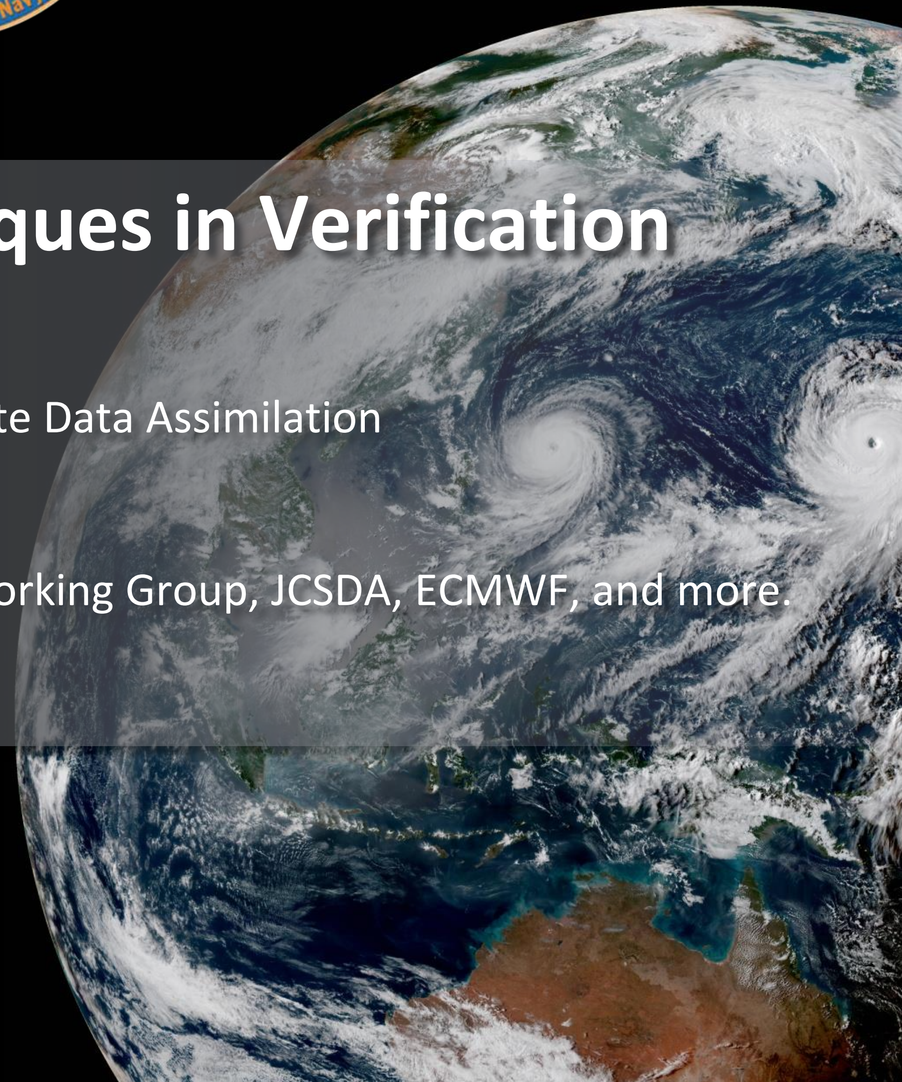
Data Assimilation Techniques in Verification

Tom Auligné, Director, Joint Center for Satellite Data Assimilation

With inputs from members of WMO DAOS Working Group, JCSDA, ECMWF, and more.



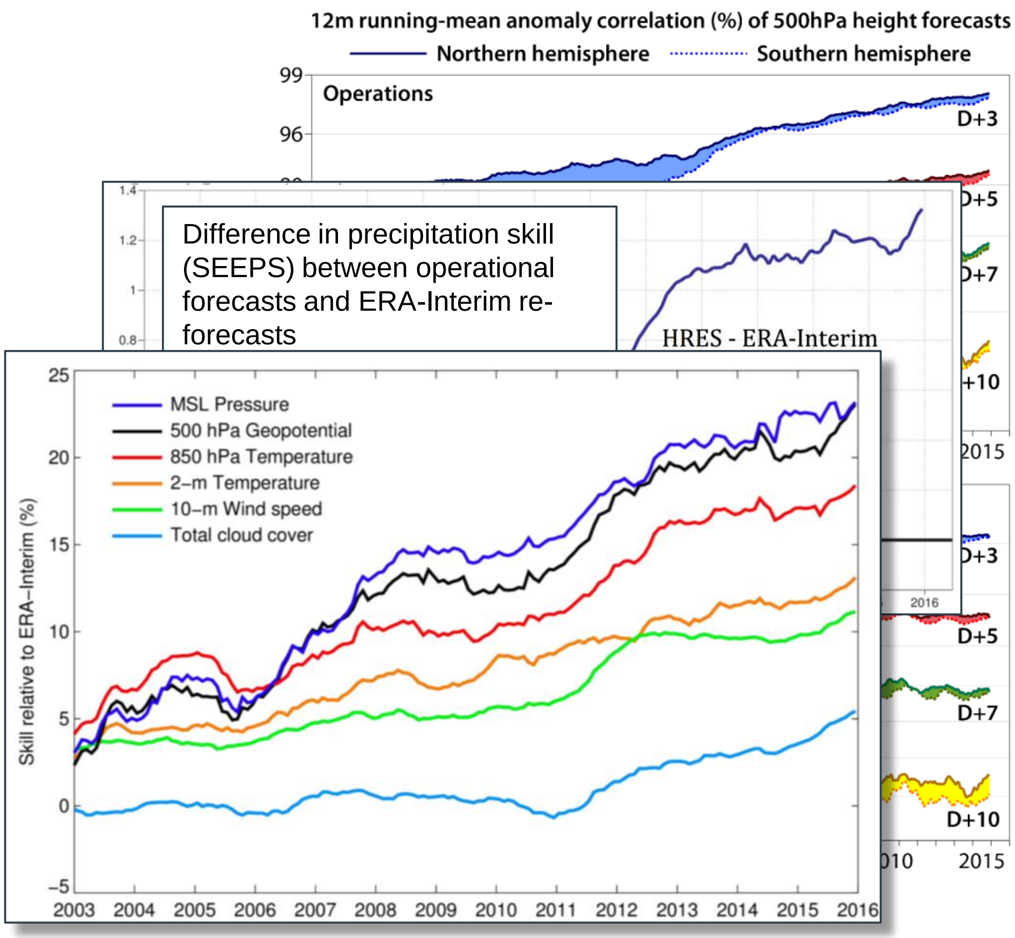
Use of Reanalysis to Evaluate Forecast

NWP forecast skill varies due to

- model and DA upgrades
- changes in the observing system
- atmospheric predictability

Comparing with re-forecasts skill can isolate the effect of model/DA upgrades

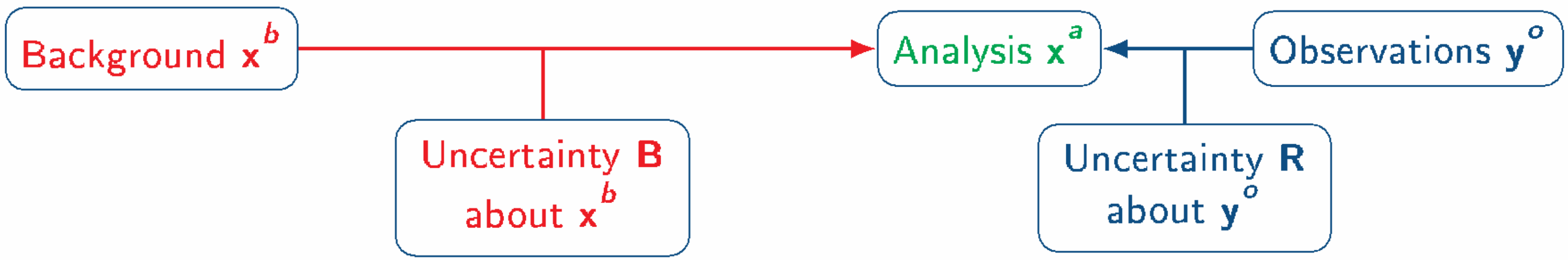
Reanalysis system must be "similar" to the NWP system



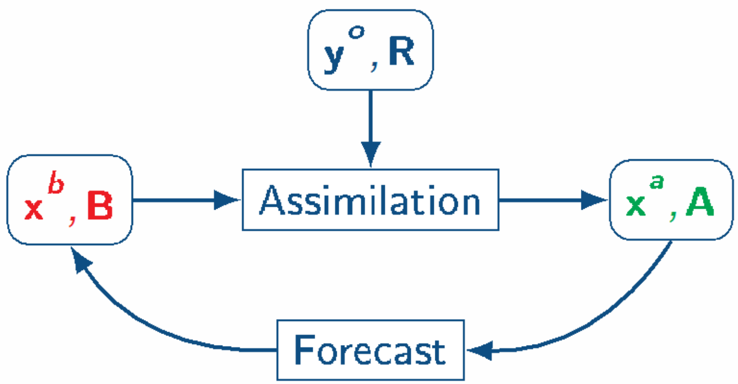
ECMWF

Data Assimilation Basics

Background and observations are combined to give the analysis:



Data assimilation is a cycled process:



Analysis increment: $x^a - x^b$

Obs-model departures: $\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k$

Credit: Ménétrier

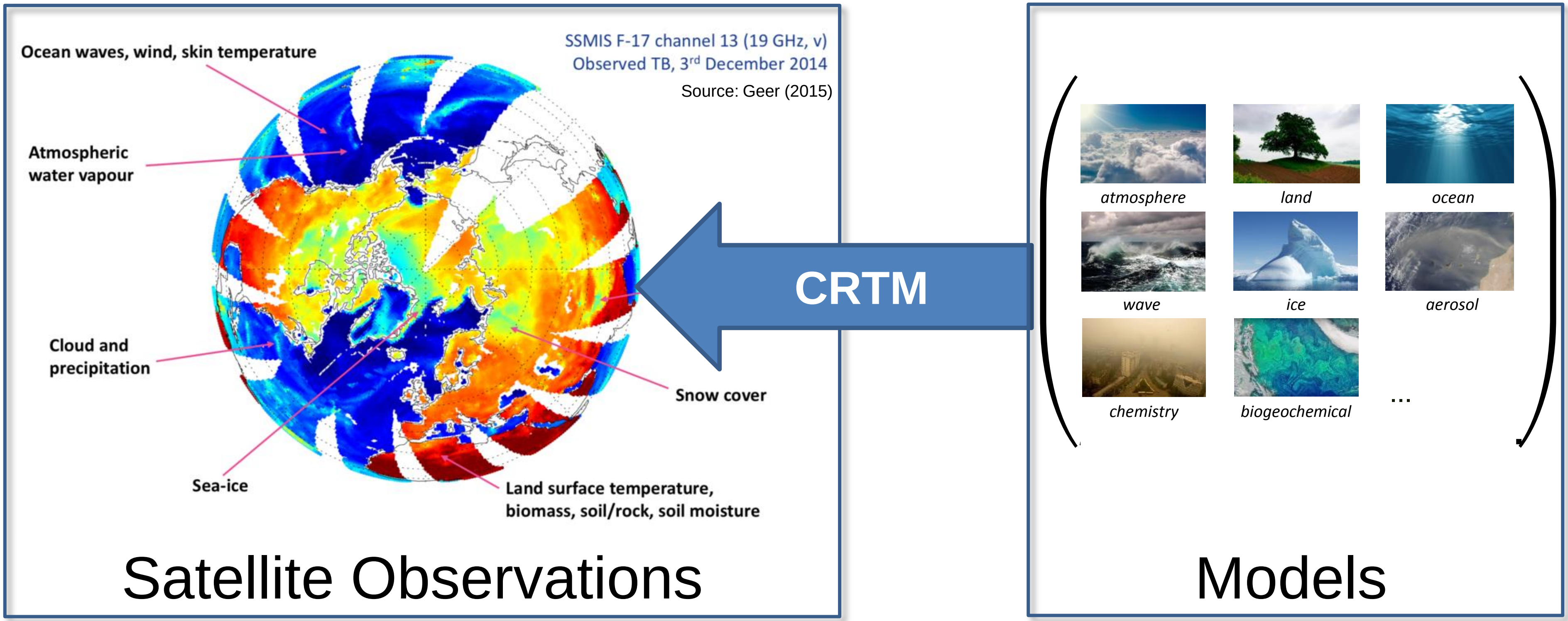
Outline

- Forward Operator
- Observation Processing
- Error and Bias Characterization
- Forecast Sensitivity - Observation impact
- Community Infrastructure

FORWARD OPERATOR

Big Data is here!

Forward Operator

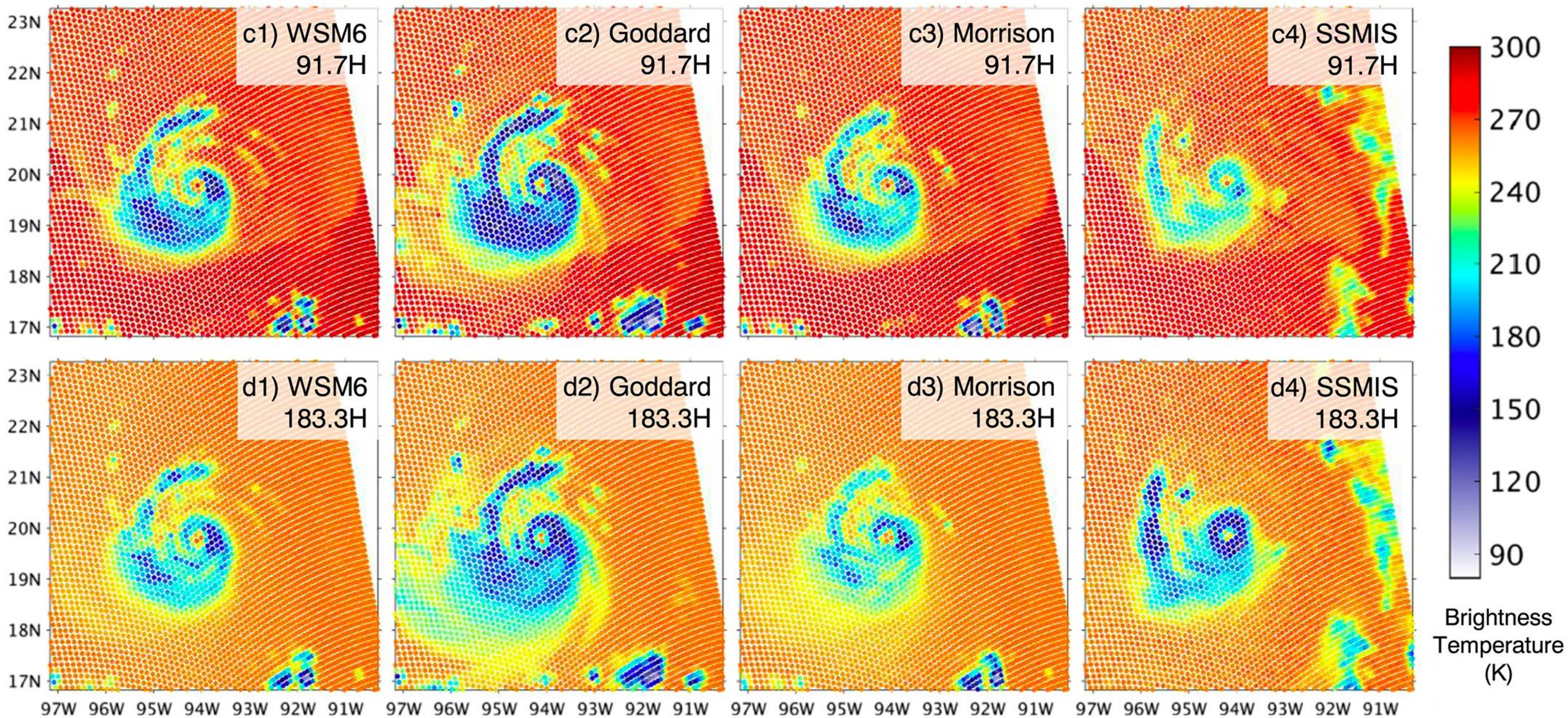


- Horizontal, vertical, and time interpolation from model to observations
- Variable transform and integration along observation geometry
- Fast Radiative Transfer Model

Community Radiative Transfer Model (CRTM)

Emulation of the Community Hydrometeor Model (CHyM) providing consistency between model microphysics and assumed CRTM cloud microphysics.

WSM-6 Goddard Morrison SSMI/S Observation

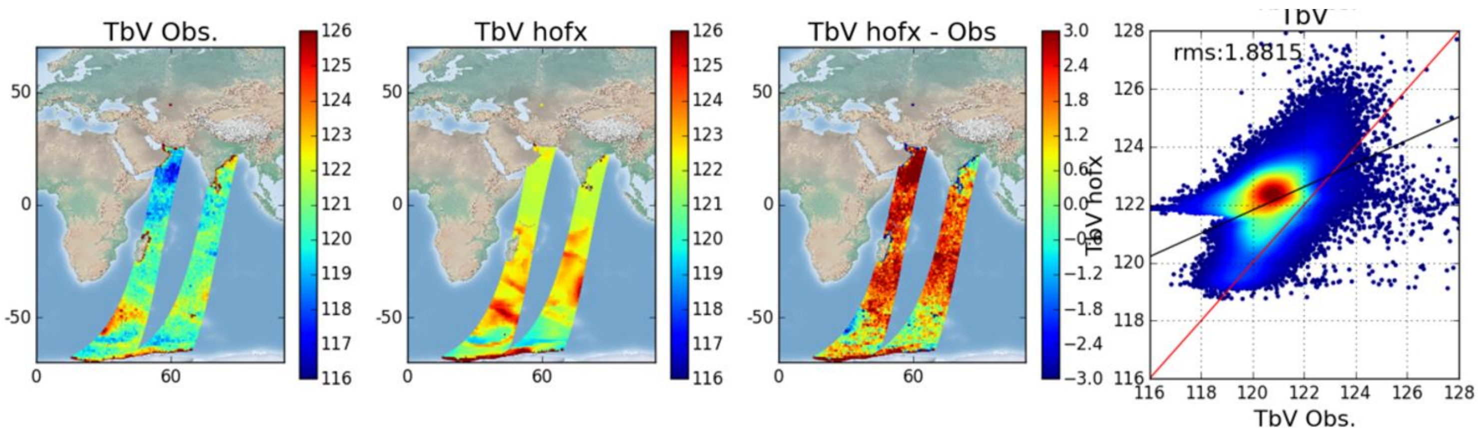


Sieron, S.B., Clothiaux, E.E., Zhang, F., Lu, Y. and Otkin, J.A., 2017. Comparison of using distribution-specific versus effective radius methods for hydrometeor single-scattering properties for all-sky microwave satellite radiance simulations with different microphysics parameterization schemes. *Journal of Geophysical Research: Atmospheres*, *122*(13), pp.7027-7046.

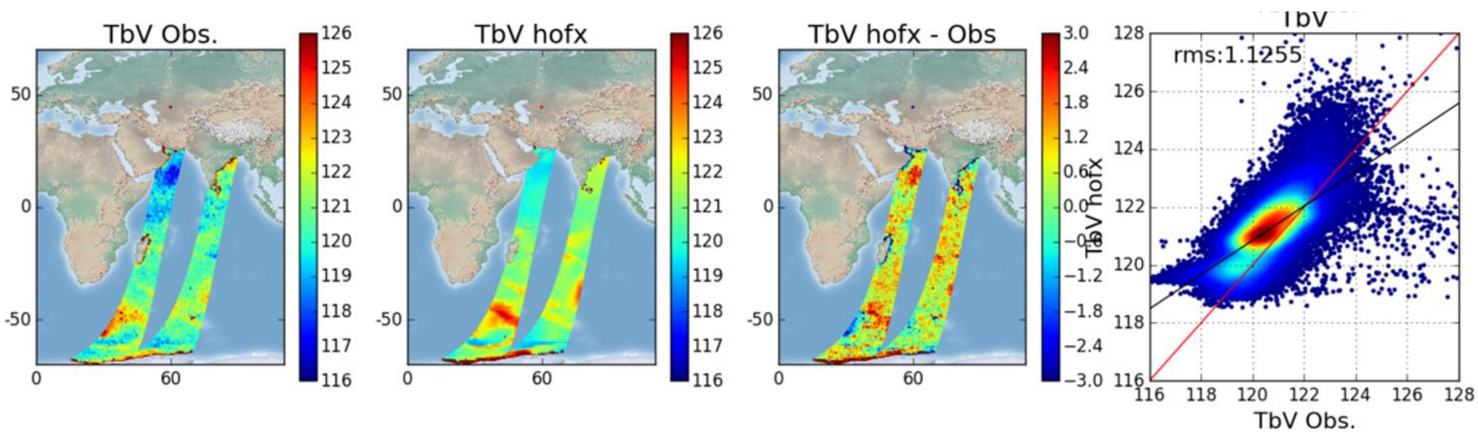
Multi-Domain Forward Operator

SMAP Observations, Simulated, and Departures

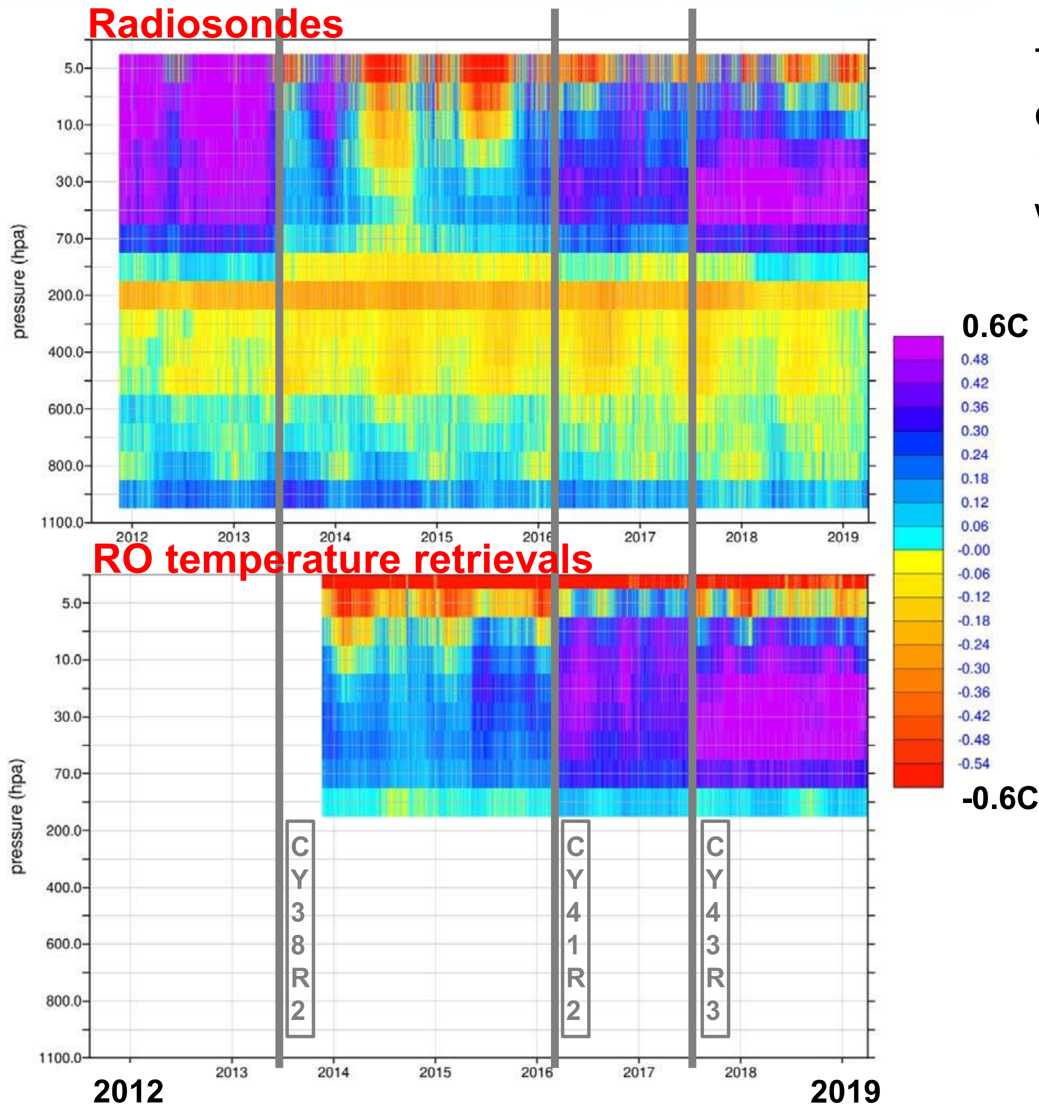
With CRTM default salinity



With MOM6 salinity background



Temperature Bias in ECMWF model



The short-term model bias is estimated by comparing the 12-hour first-guess trajectory with radiosondes and GPS-RO

Similar signal with the two types of observations:

→ bias reduced with new vertical resolution (L137 in CY38R2)

→ bias increased with new horizontal resolution (Tco1279 in CY41R2)

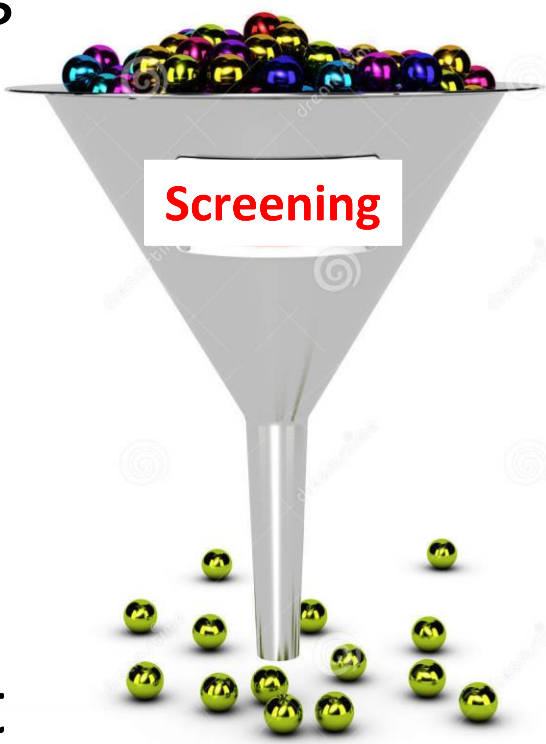
→ bias increased with new radiative scheme (CY43R3)

Source: Laloyaux et al. (2019)

OBSERVATION PROCESSING

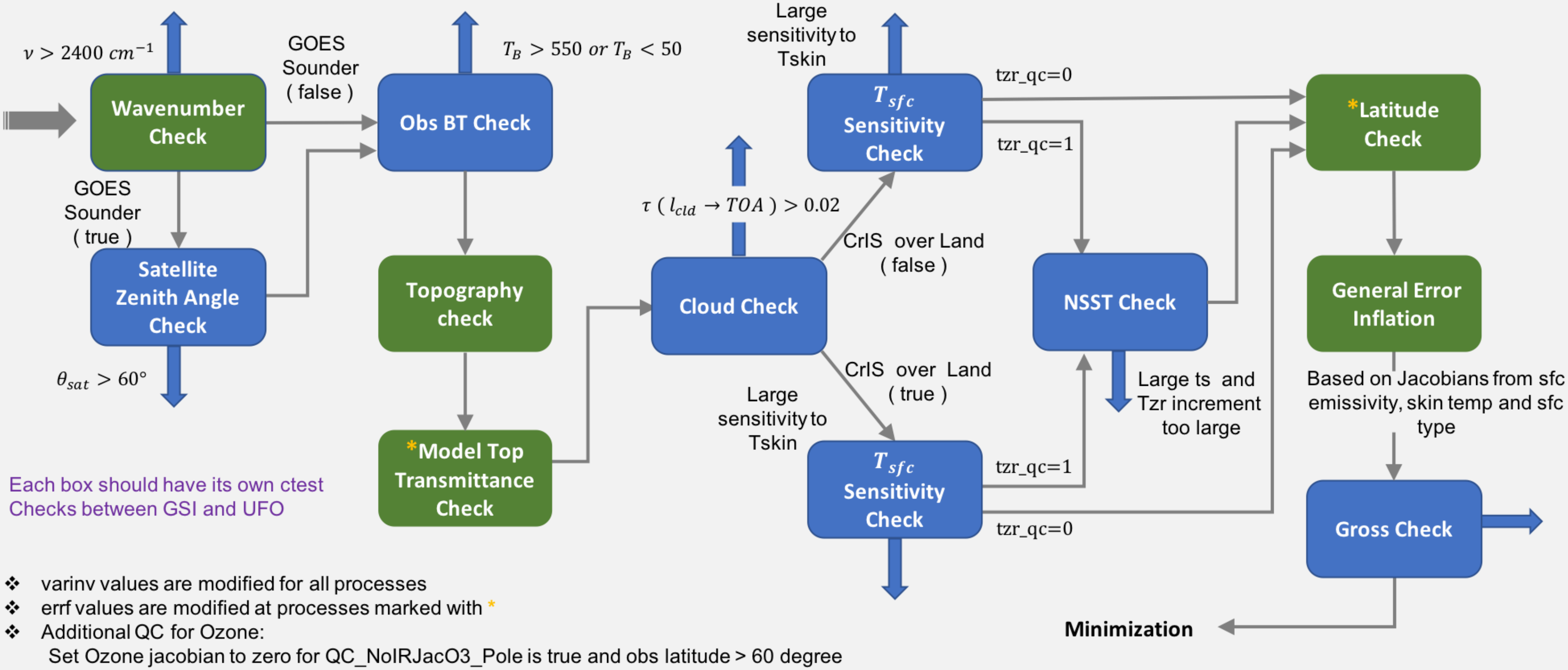
Observation Processing

- Channel selection, blacklisting
- Thinning or super-obbing
- Forward operator
- Observation bias correction
- Quality control
  - Cloud detection, Variational QC
- Observation error assignment



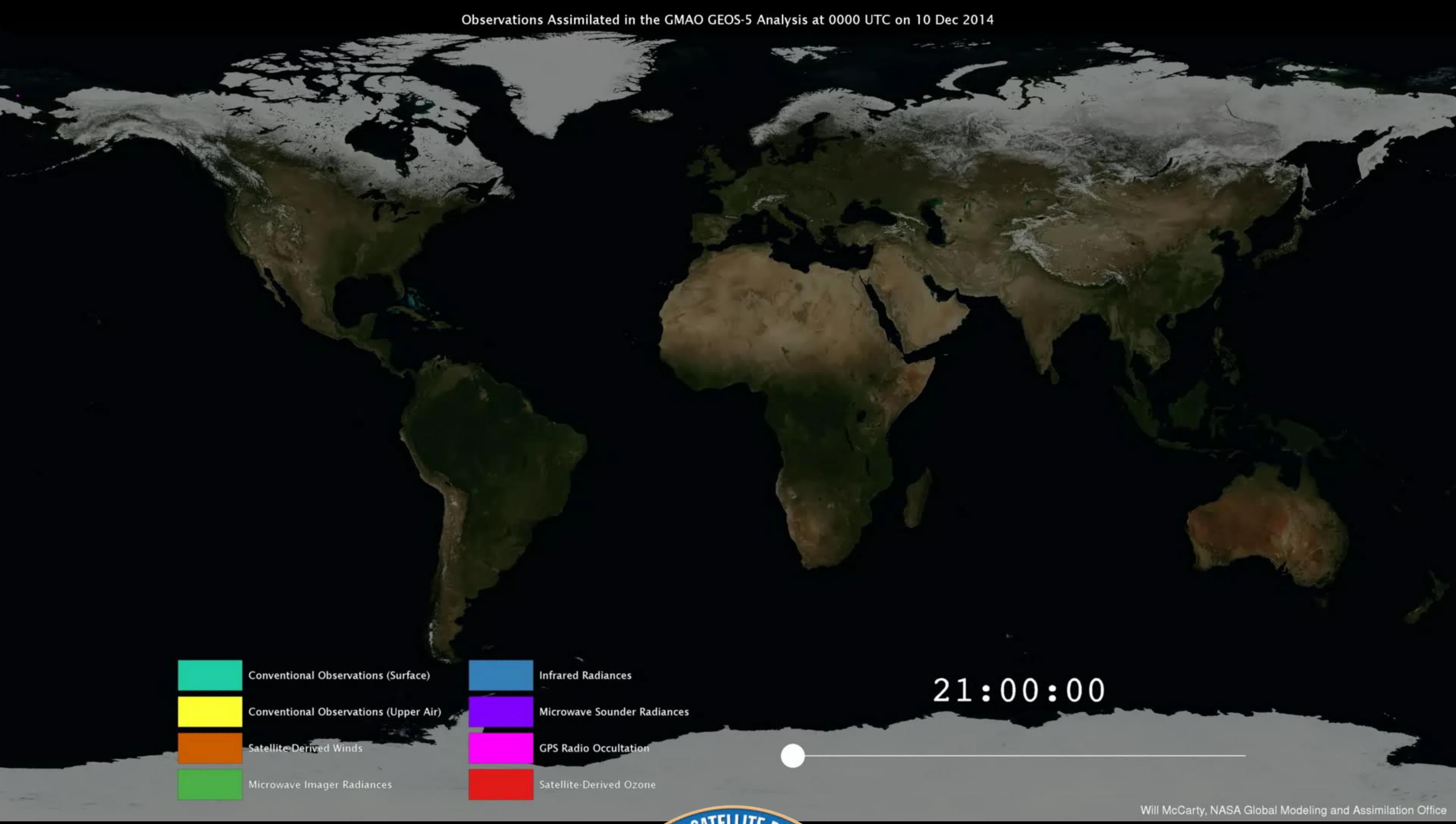
Metric of fit to observations: $J_o = \frac{1}{2}\sum_{k=0}^{N}\left(\mathcal{H}_k\mathcal{M}_{k,0}(\mathbf{x}_0)-\mathbf{y}_k\right)^{\mathrm{T}}\mathbf{R}_k^{-1}\left(\mathcal{H}_k\mathcal{M}_{k,0}(\mathbf{x}_0)-\mathbf{y}_k\right)$

Complex Quality Control Procedures



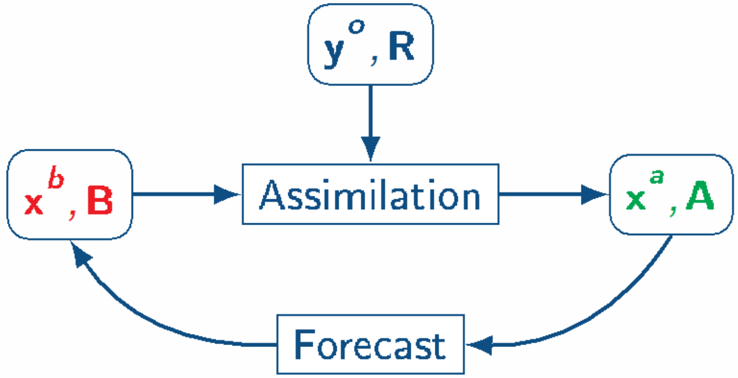
- varinv values are modified for all processes
- errf values are modified at processes marked with *
- Additional QC for Ozone:
  Set Ozone jacobian to zero for QC_NoIRJacO3_Pole is true and obs latitude > 60 degree

NOAA Operational QC Flowchart for Infrared Sounders


Observations Assimilated in the GMAO GEOS-5 Analysis at 0000 UTC on 10 Dec 2014
Conventional Observations (Surface)
Conventional Observations (Upper Air)
Satellite-Derived Winds
Microwave Imager Radiances
Infrared Radiances
Microwave Sounder Radiances
GPS Radio Occultation
Satellite-Derived Ozone
21:00:00
Will McCarty, NASA Global Modeling and Assimilation Office

ERROR AND BIAS CHARACTERIZATION

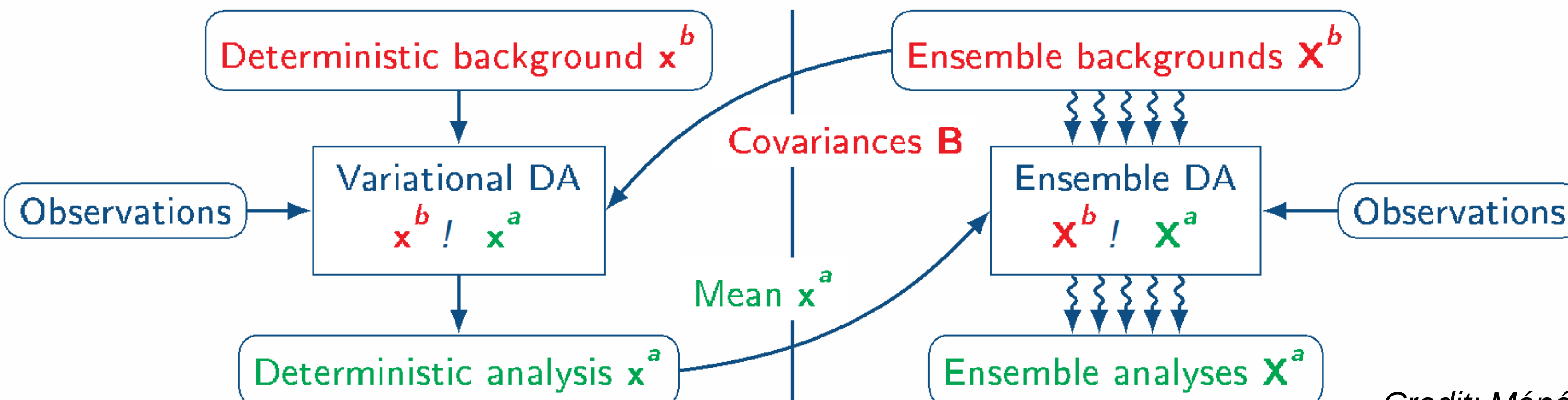
Ensemble Data Assimilation



Cycling $\mathbf{x}^b$ and $\mathbf{x}^a$ is easy ... but $\mathbf{B}$ and $\mathbf{A}$ are way too big!

! solution: ensemble data assimilation

Ensemble backgrounds are updated into ensemble analyses:

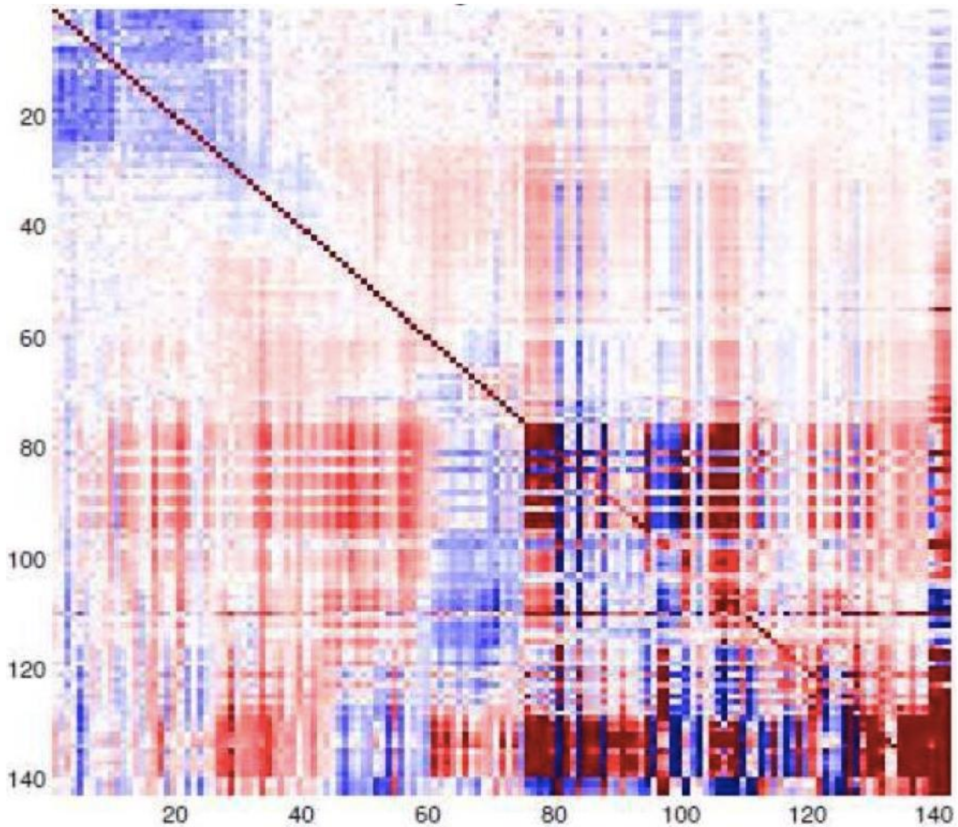


Credit: Ménétrier

Error Covariances Diagnostics

$$E[\mathbf{d}_b^a(\mathbf{d}_b^o)^T] = \mathbf{HBH}^T, \qquad E[\mathbf{d}_a^o(\mathbf{d}_b^o)^T] = \mathbf{R},$$
$$E[\mathbf{d}_b^o(\mathbf{d}_b^o)^T] = \mathbf{HBH}^T + \mathbf{R}, \qquad E[\mathbf{d}_b^a(\mathbf{d}_a^o)^T] = \mathbf{HAH}^T$$

Desroziers et al., 2005



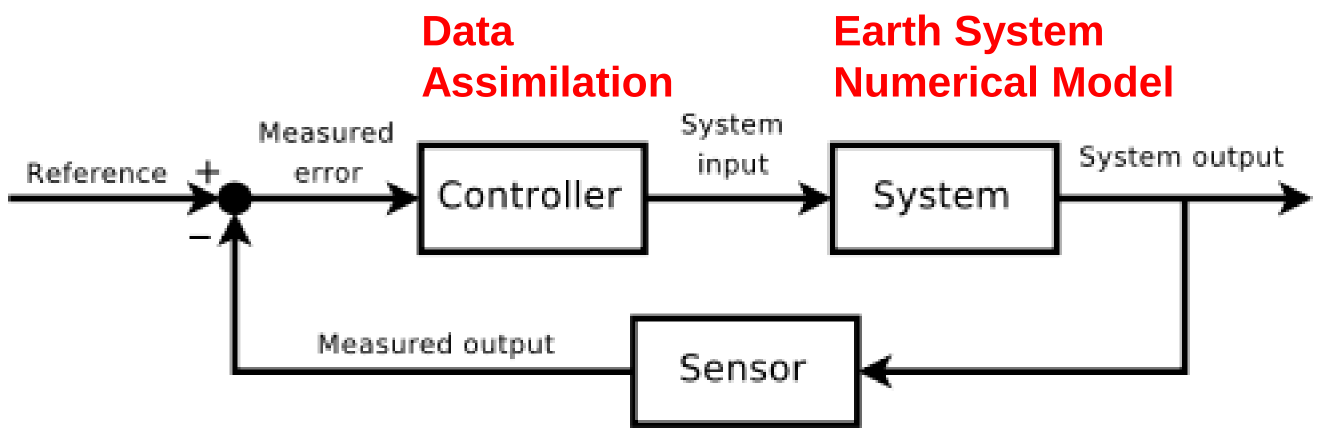
AIRS Diagnostic R Matrix

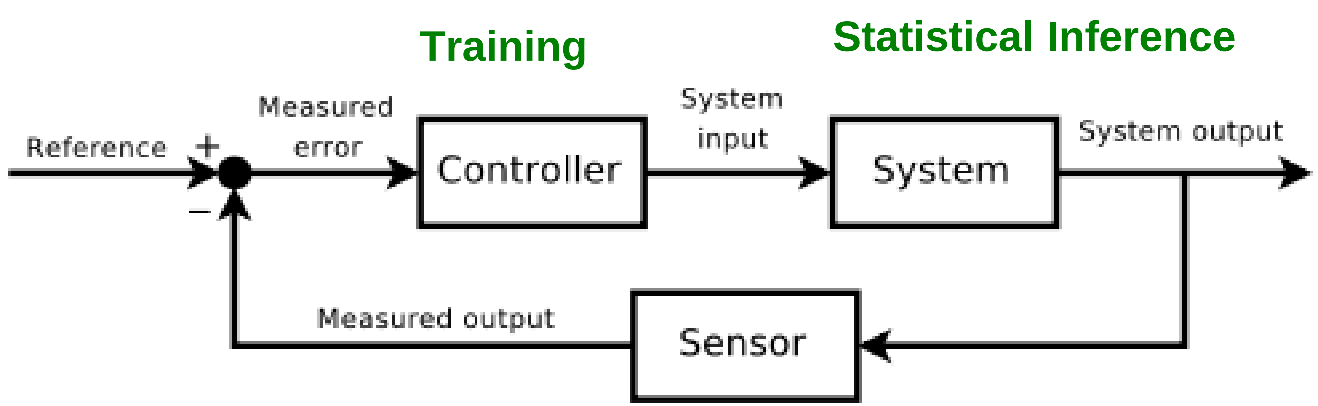
Correlated errors
(esp. for moisture channels)

At least partly due to
representativeness error
(Waller et al. 2014)

Source: Weston (2011)

Control System Theory

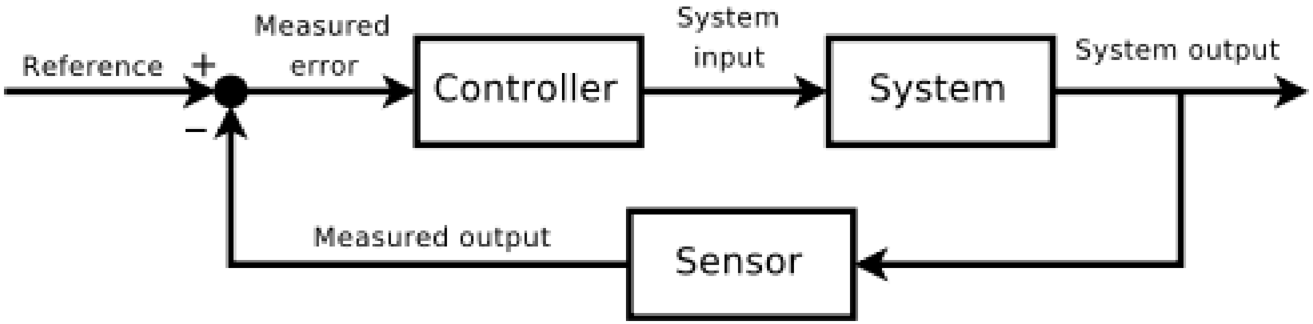




Control System Theory

Data Assimilation
+ Training

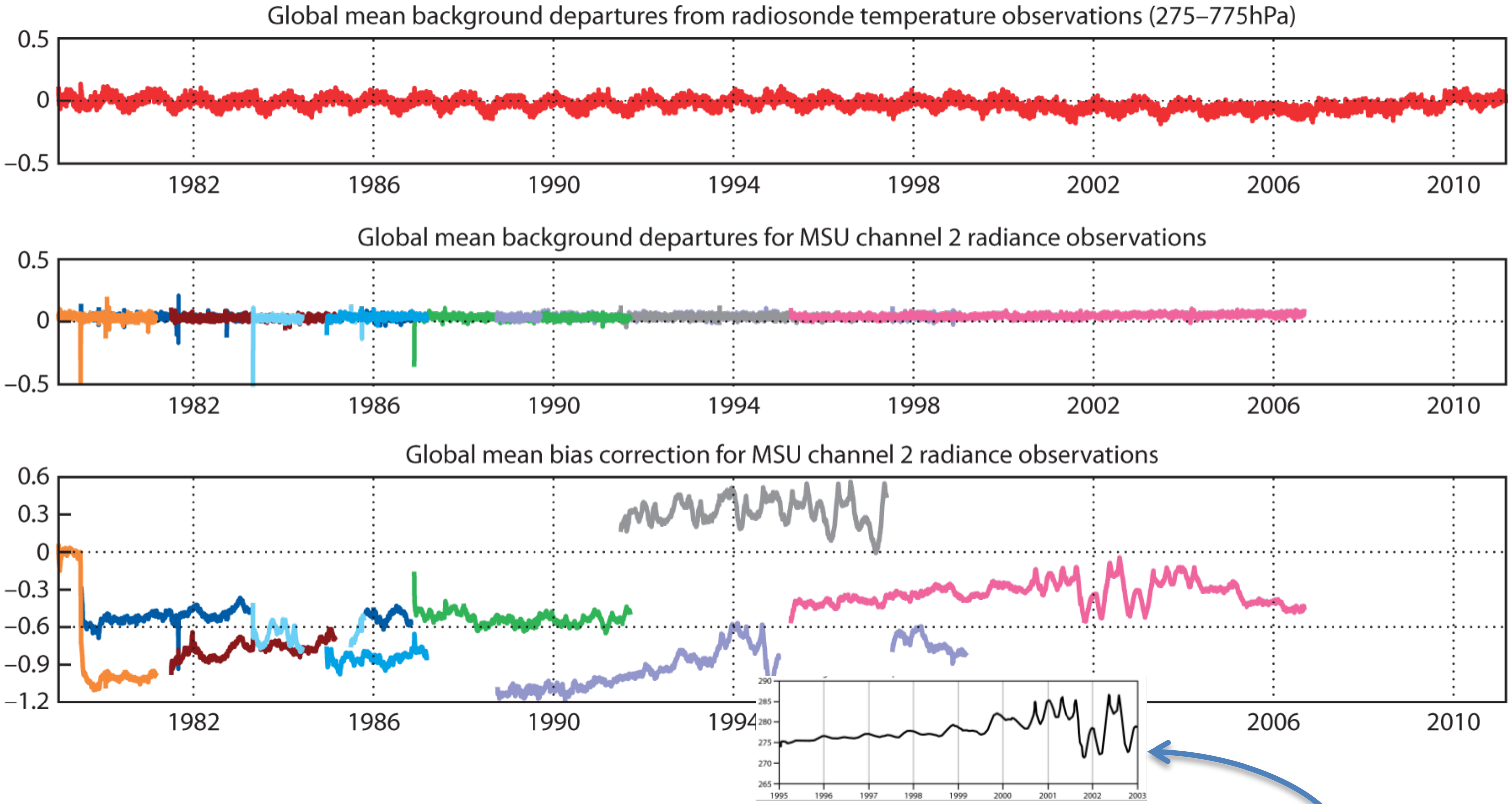
Earth System
Numerical + Statistical Inference



Example 1: Observation bias

Example 2: Systematic model error

Satellite Radiance Bias Correction



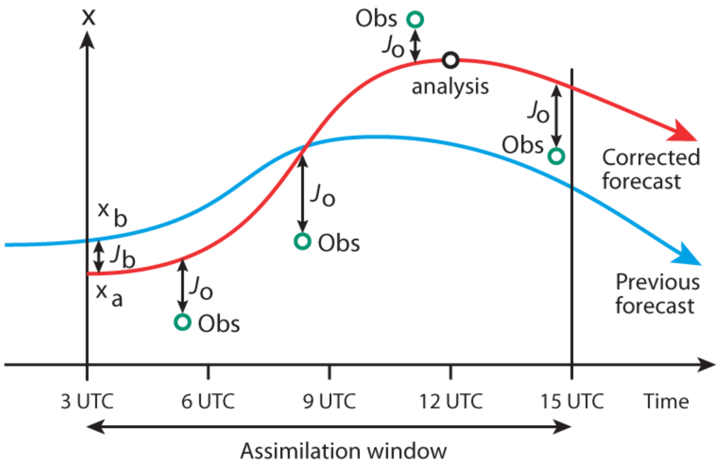
NOAA-14 recorded warm-target temperature changes, due to orbital drift (Grody et al. 2004)

Variational Bias Correction

$$\tilde{H}(\mathbf{x}, \boldsymbol{\beta}) = H(\mathbf{x}) + \sum_{i=0}^{N} \beta_i \, p_i(\mathbf{x})$$

Source: Dick Dee (ECMWF)

Model Error in Weak Constraint 4DVar



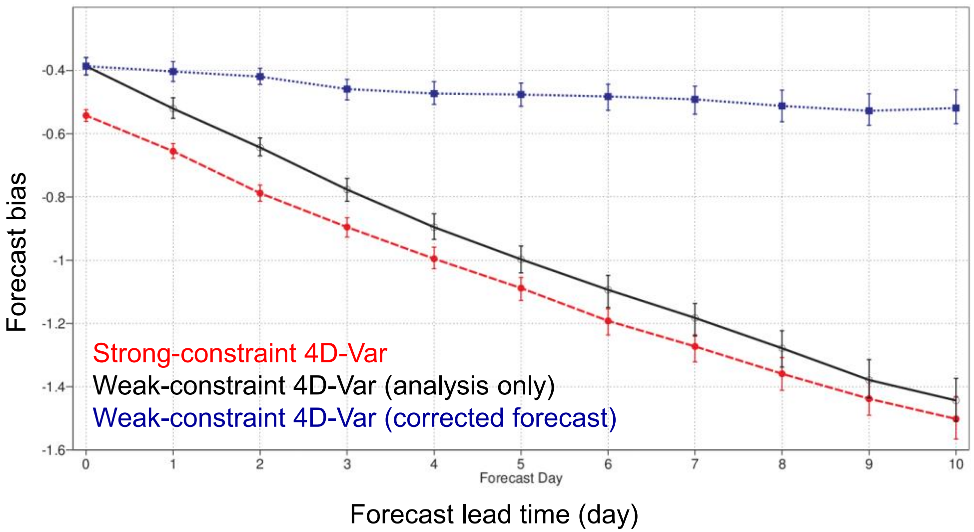
Weak-constraint

- First-guess trajectory
- Observations
- Compute a correction at initial time
- Compute a model forcing η
- Analysis trajectory

Weak Constraint 4DVar

$$\mathbf{x}_k = \mathcal{M}_{k,k-1}(\mathbf{x}_{k-1}) + \eta$$

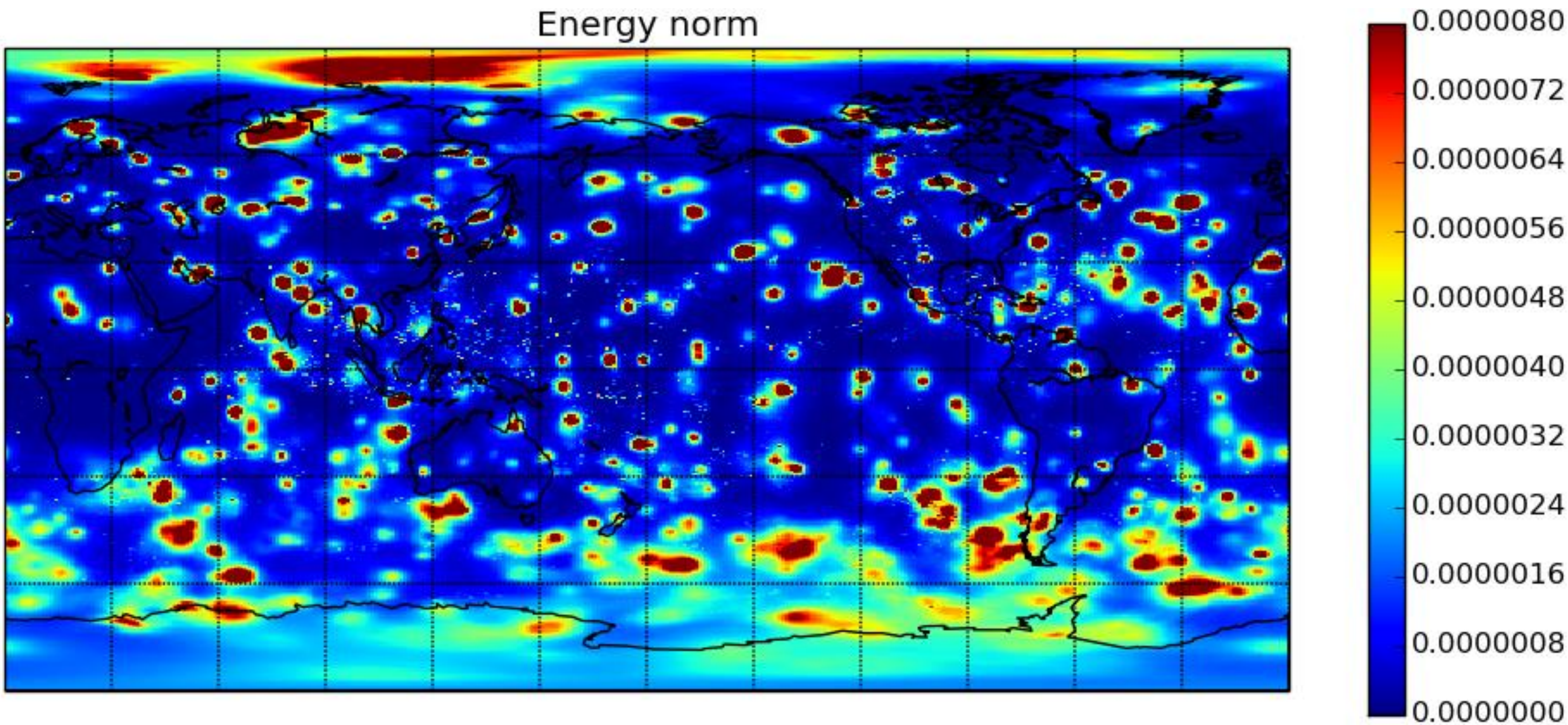
Data assimilation cost function depends on initial conditions and model forcing



Source: Laloyaux et al. (2019)

FORECAST SENSITIVITY OBSERVATION IMPACT

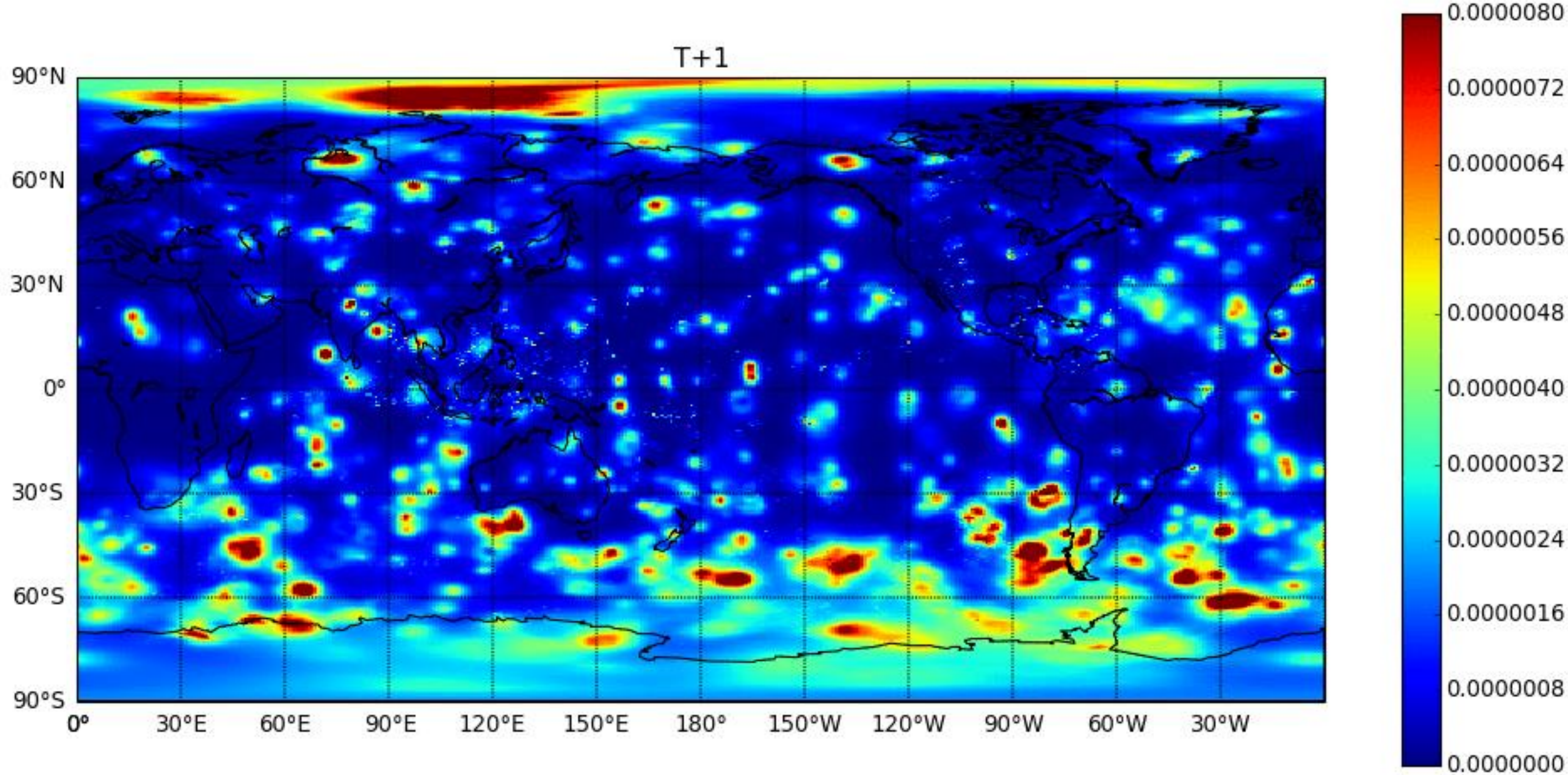
Where does GNSS-RO Influence the Forecast?



Vertically-integrated moist energy norm of the difference between analysis increments with and without radio occultation. Each blob corresponds to a GNSS-RO observation.

Source: Chris Burrows (2016)

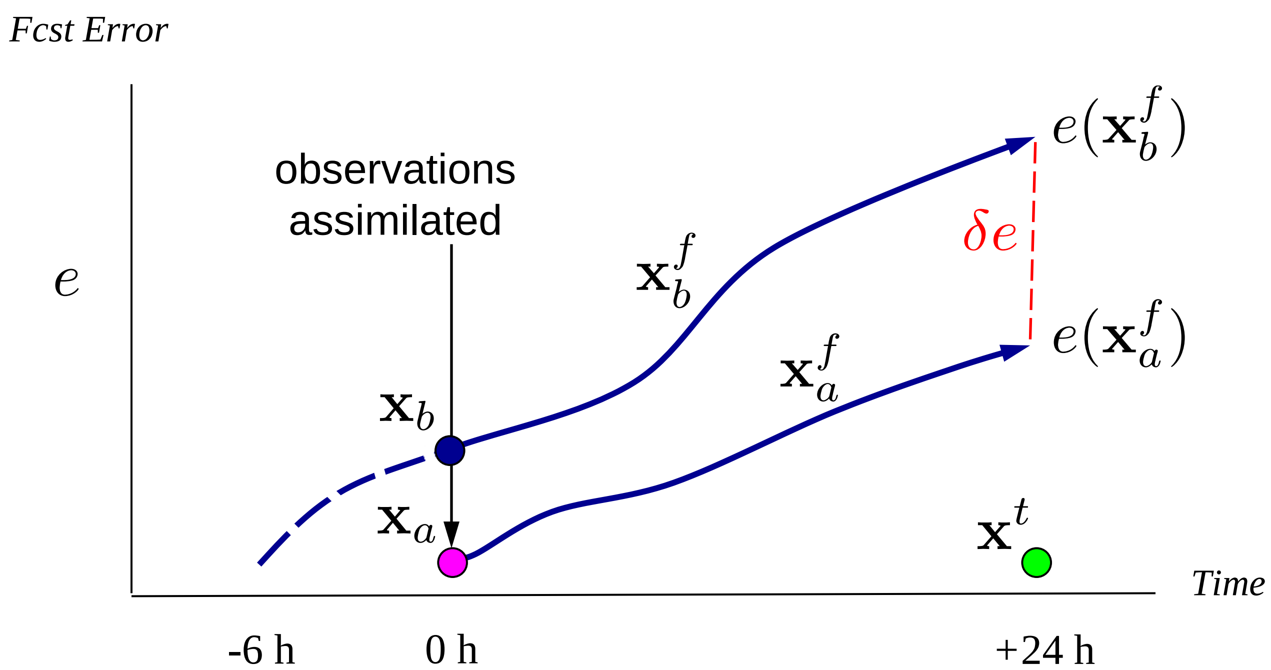
Where does GNSS-RO Influence the Forecast?



Vertically-integrated moist energy norm of the difference between analysis increments with and without radio occultation. Each blob corresponds to a GNSS-RO observation.

Source: Chris Burrows (2016)

Observation Impact Estimation

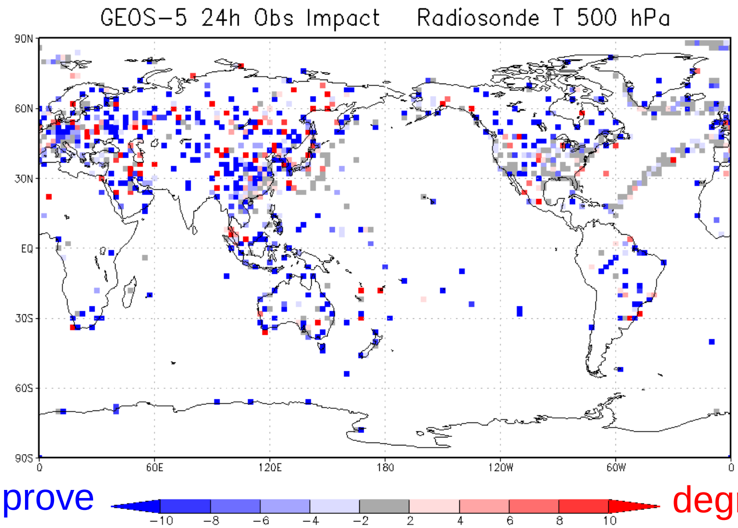


The difference $\delta e = e(\mathbf{x}_a^f) - e(\mathbf{x}_b^f)$ measures the collective impact at 24 h of **all observations** assimilated at 0 h. (model space)

From Langland and Baker (2004)

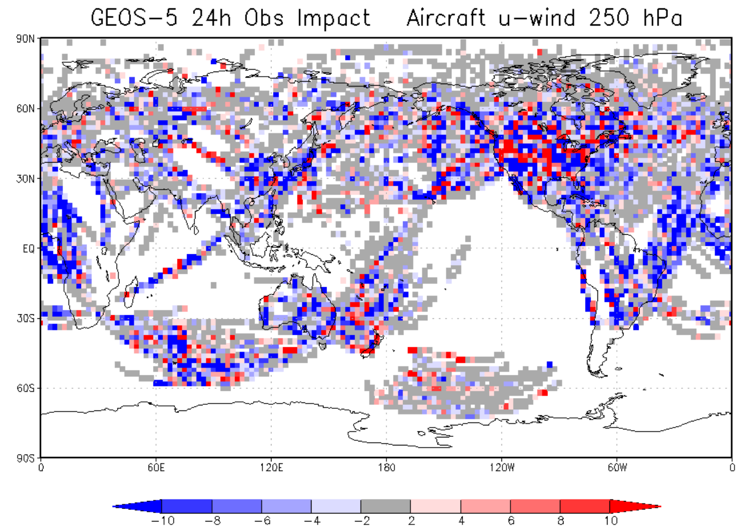
Observation Impact Estimation

Raob T-500



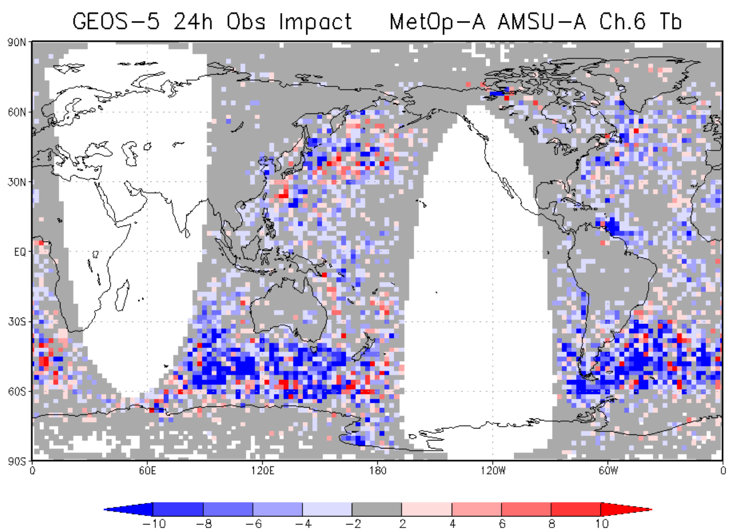
improve

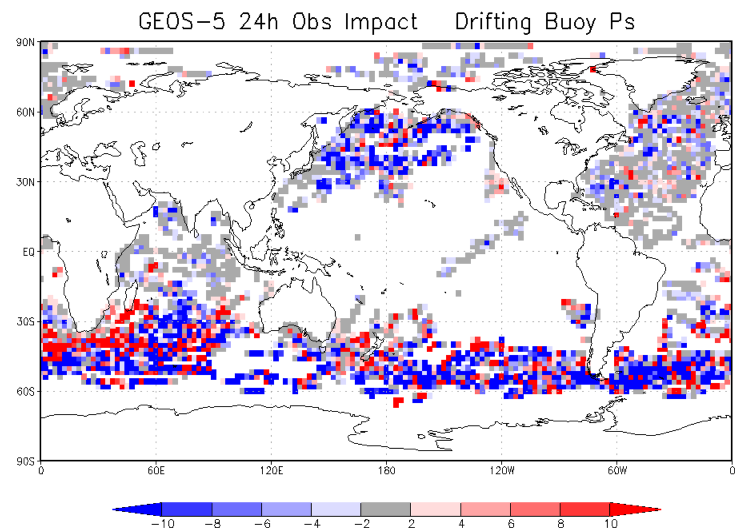
degrade



Aircraft u-250

AMSU-A Ch.6 Tb





Drift Buoy Ps

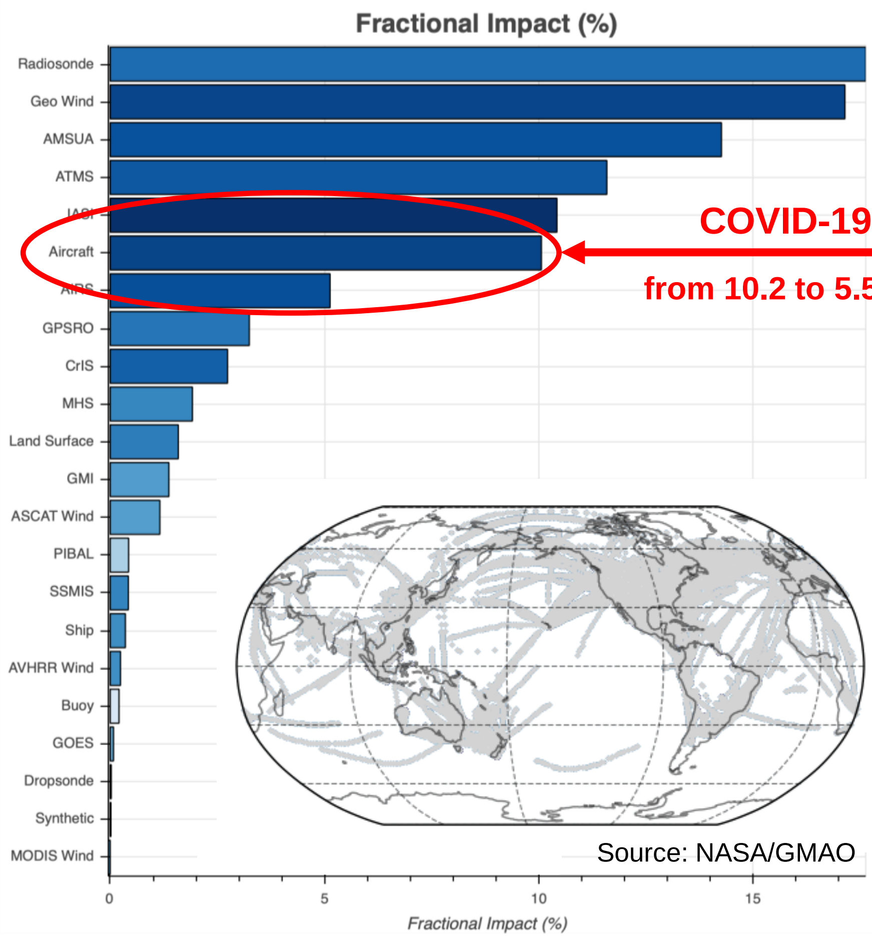
3-Month Average (MJJ 2015) for 24h Forecasts Initialized at 00z

Credit: NASA/GMAO

Observation Impact Estimation

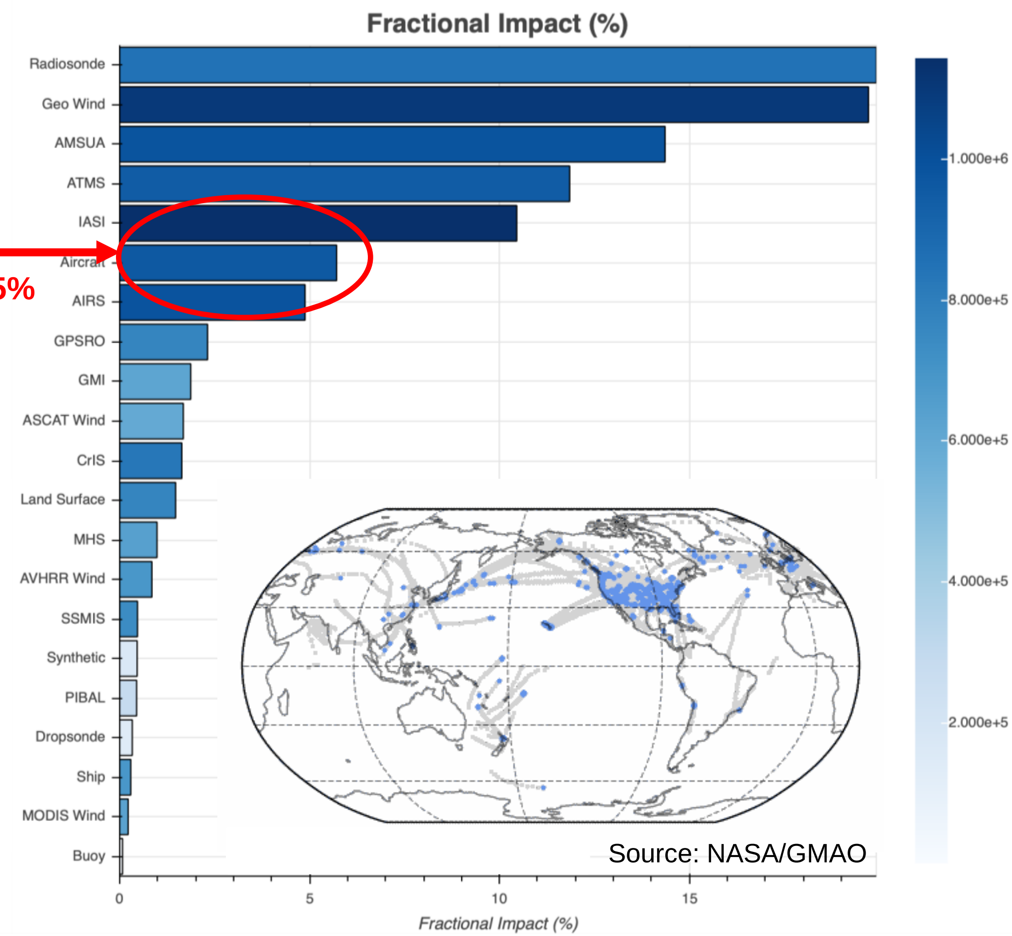
Mar 10 - May 10, 2019

Mar 10 - May 10, 2020



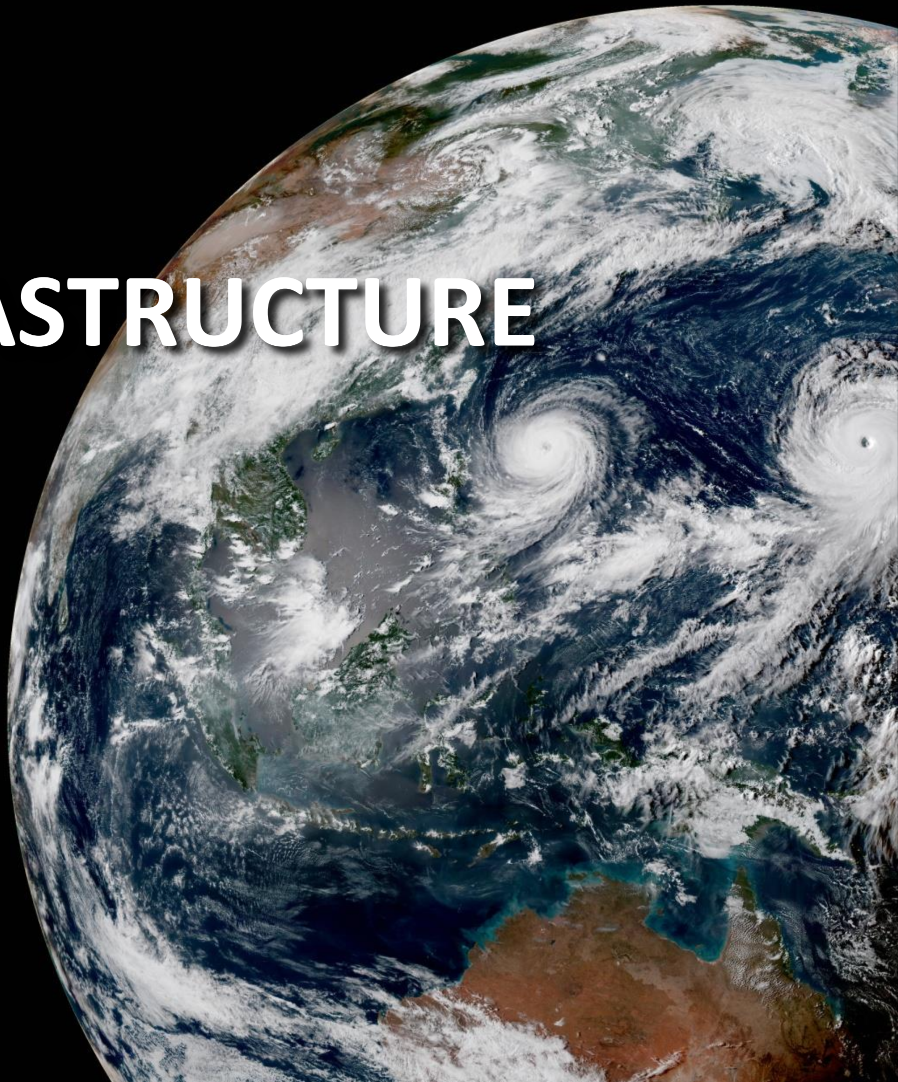
COVID-19

from 10.2 to 5.5%



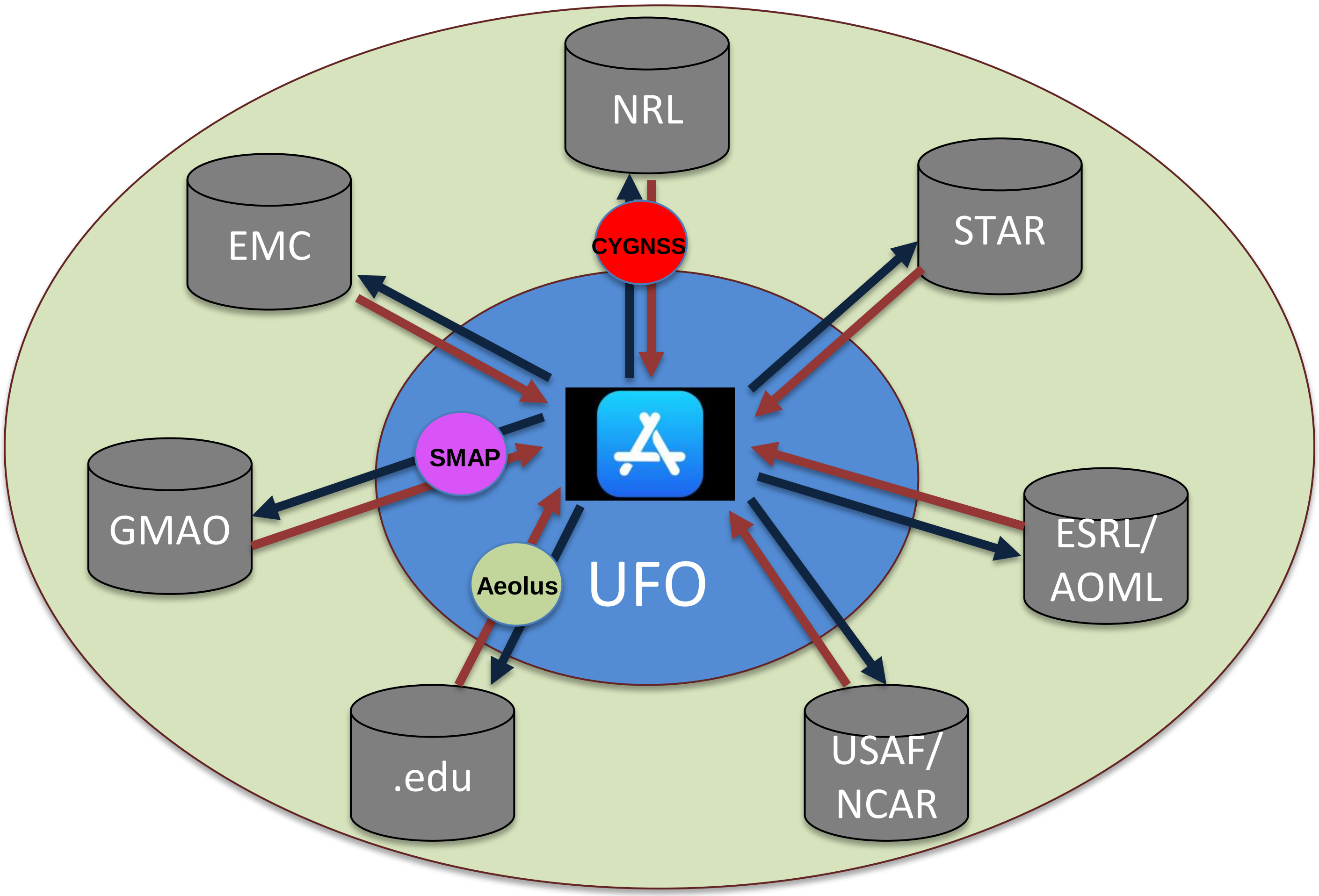
JCSDA processing and distributing near-real-time data from **NASA, NRL** and **Met Office** on https://ios.jcsda.org

COMMUNITY INFRASTRUCTURE



Unified Forward Operator (UFO)

Model-agnostic operators and abstract filters that can be shared and compared



Joint Effort for Data assimilation Integration (JEDI)

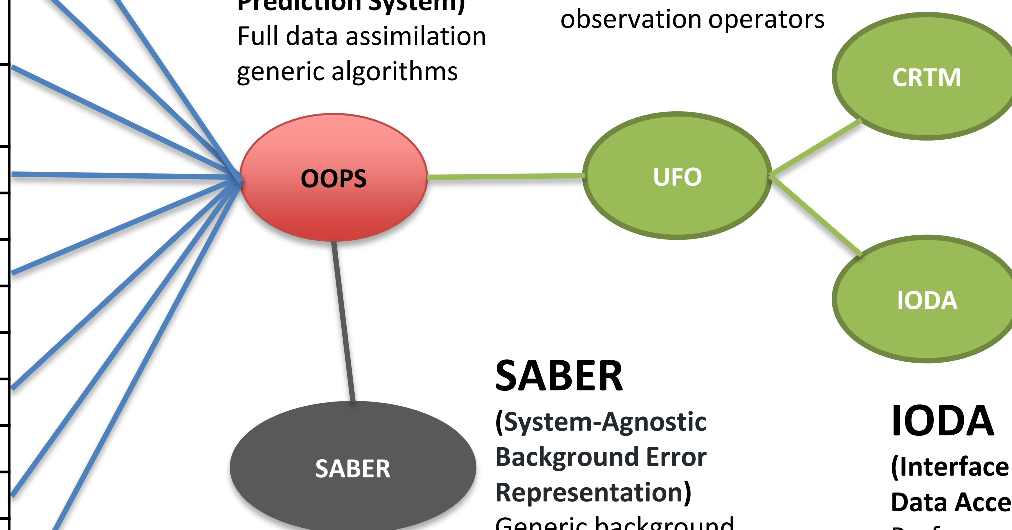
JEDI: One system with multiple configurations for Earth system data assimilation

MODEL	TYPE	CENTER
FV3GFS	Atmosphere	NOAA-EMC
GEOS	Atmosphere	NASA-GMAO
FV3GFS GSDChem	Atmospheric chemistry	NOAA-ESRL
GEOS-AERO	Atmospheric aerosols	NASA-GMAO
MPAS	Atmosphere	NCAR
WRF	Atmosphere	NCAR
LFRic	Atmosphere	Met Office (UK)
MOM6	Ocean	NOAA-EMC
SIS2	Sea ice	NOAA-EMC
CICE6	Sea ice	NOAA-EMC
NEPTUNE	Atmosphere	NRL
QG	Toy model	ECMWF
Lorenz 95	Toy model	ECMWF
ShallowWater	Toy model	NOAA-ESRL

OOPS
(Object Oriented Prediction System)
Full data assimilation generic algorithms

UFO
(Unified Forward Operator)
The 'app-store' of model-agnostic observation operators

CRTM
(Community Radiative Transfer Model)
Accurately and efficiently simulate satellite radiances

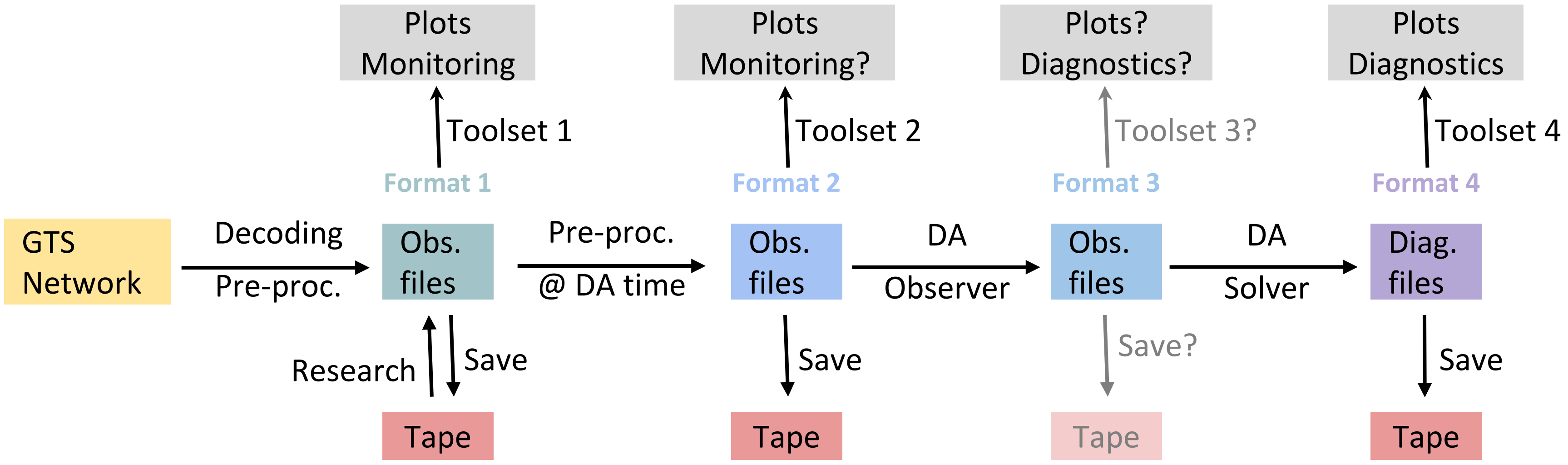


SABER
(System-Agnostic Background Error Representation)
Generic background error covariance modeling (incl. BUMP)

IODA
(Interface for Observation Data Access)
Performs all the I/O of the observations

Abstract interfaces and separation of concerns
The end of the monolithic gigantic jumble of code

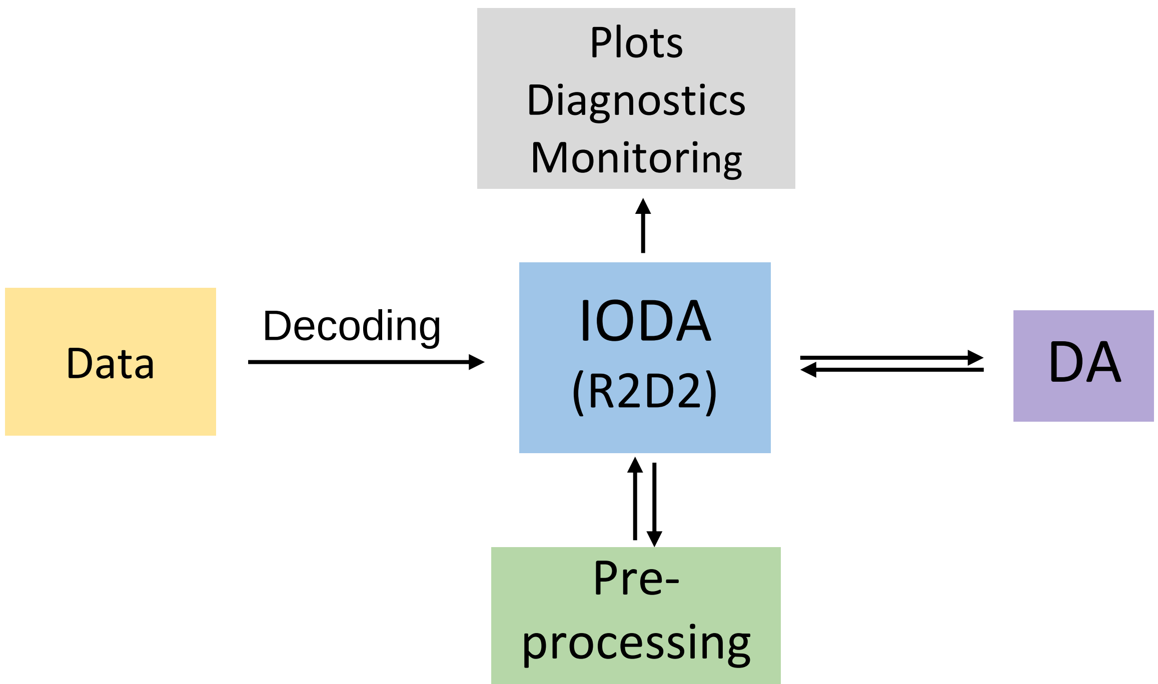
Typical Observation Workflow



This is not sustainable:

- maintenance and evolution are impossible
- does not fit possible evolution of DA methods (continuous assimilation)

Research Repository for Data & Diagnostics

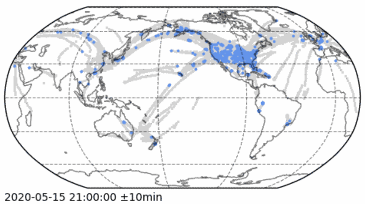


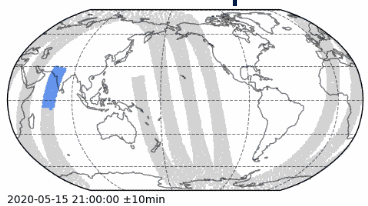
IODA: One observation data handling interface across the whole NWP chain

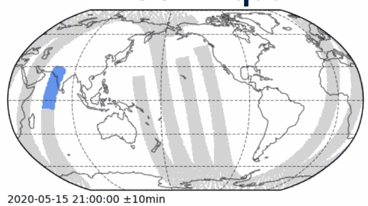
R2D2: cloud-based repository to host a comprehensive collection of Earth observations and model output, along with tools and workflows needed for DA development, model forecast verification and process-level model diagnostics

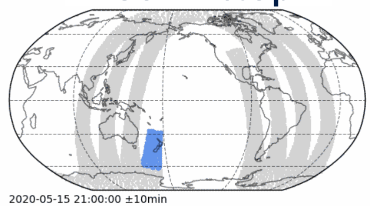
Near Real-Time Monitoring Applications

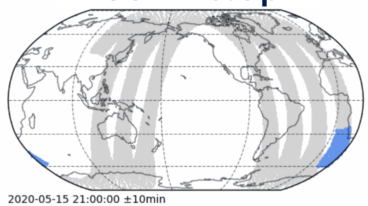
nrt.jcsda.org **soca.jcsda.org**

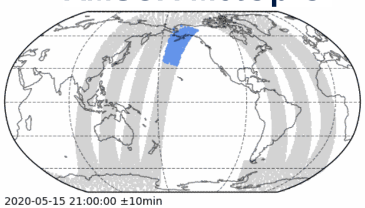
Aircraft


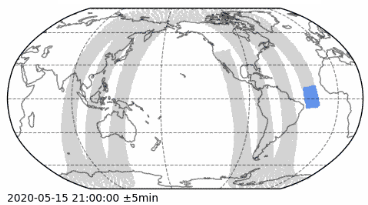
AIRS-Aqua


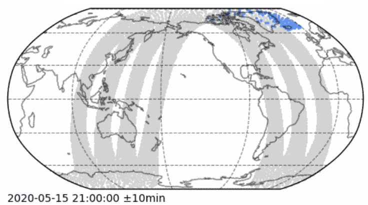
AMSUA Aqua


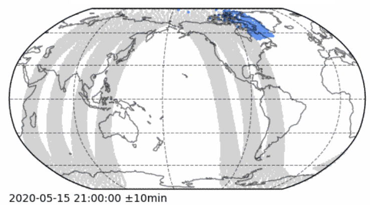
AMSUA MetOp-A


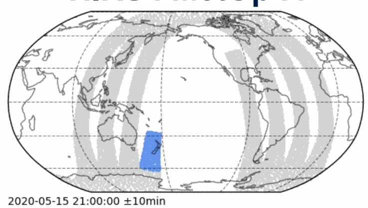
AMSUA MetOp-B


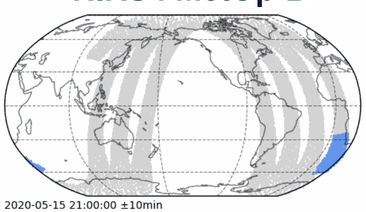
AMSUA MetOp-C


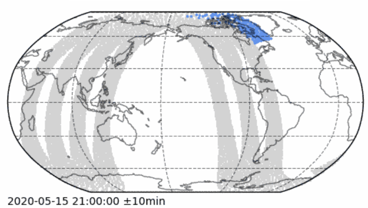
AMSUA NOAA15


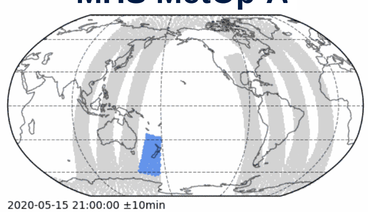
AMSUA NOAA18


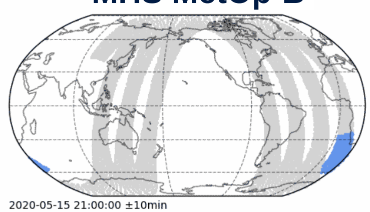
AMSUA NOAA19


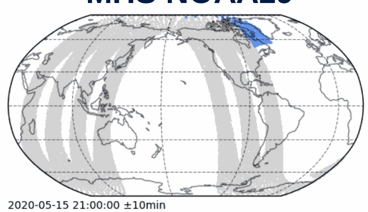
HIRS4 MetOp-A


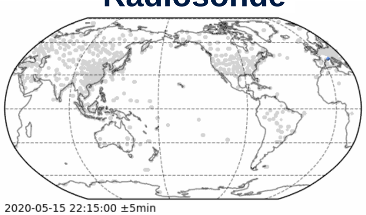
HIRS4 MetOp-B


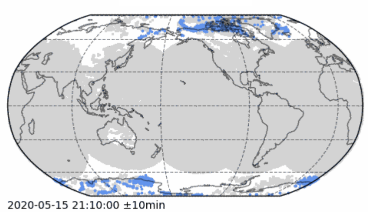
HIRS4 NOAA19


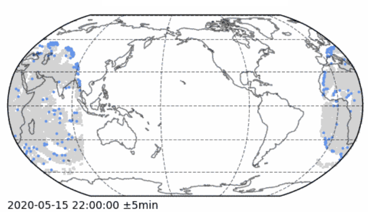
MHS MetOp-A


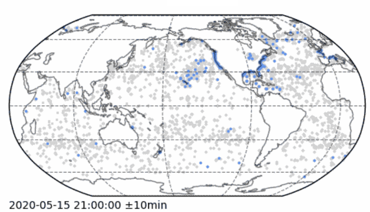
MHS MetOp-B


MHS NOAA19


Radiosonde


Satwind


SEVIRI M08


SST


Website is served by community code on GitHub.

Final Remarks and Integration Perspectives

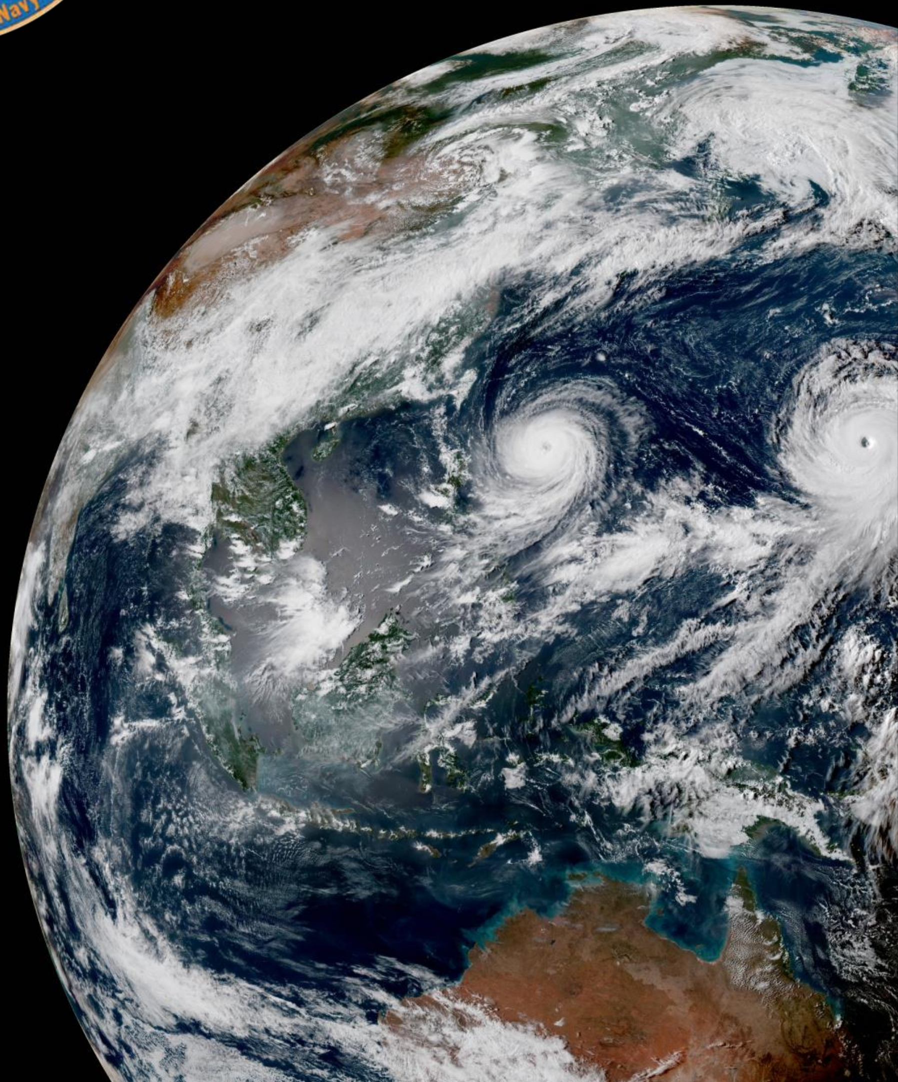
Byproducts of Data Assimilation useful for verification:

- Obs-model departures (4D) for many observing system
- J_o = Metric of fit to observations
- Analysis increments and spatial displacement increments
- Some ability to disentangle obs vs. model bias

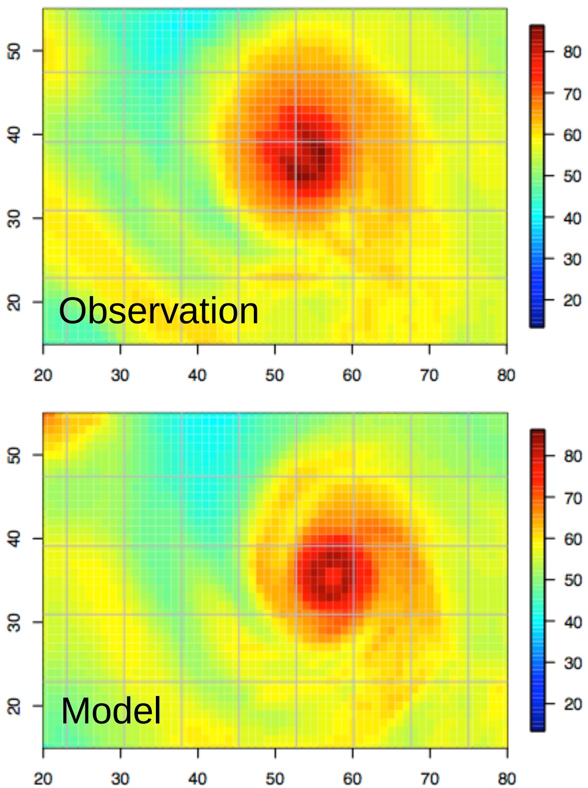
Verification practices ingestible into Data Assimilation:

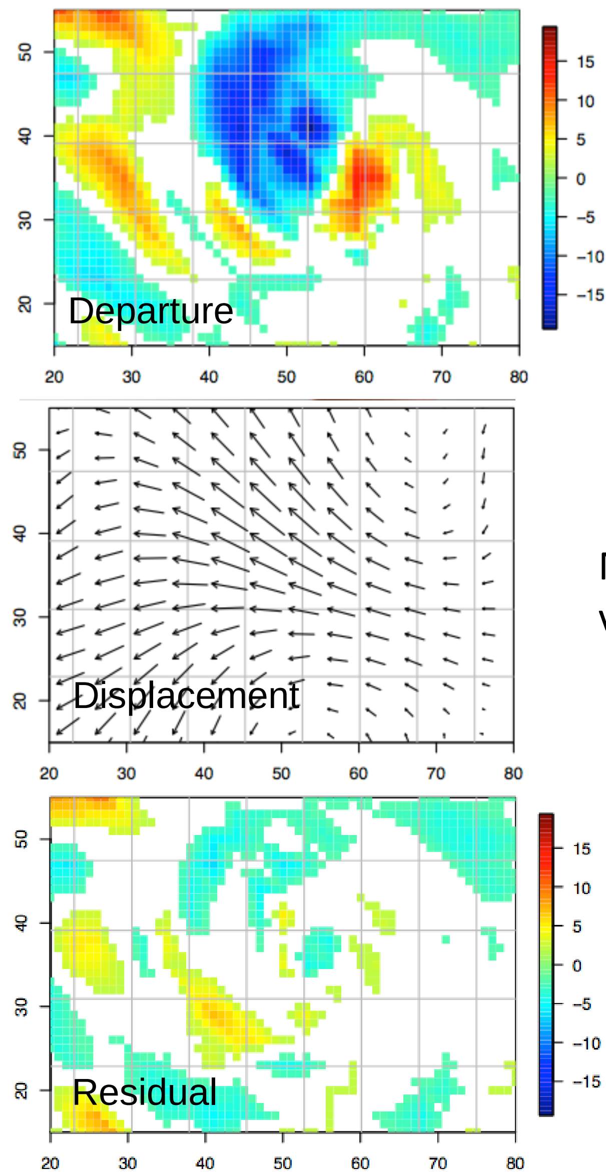
- Metrics for Observation Impact Estimation
- Parametric representation of systematic errors
- Obs-space diagnostics in community repository

Discussion



Displacement Analysis





Metric for spatial verification

Nehrkorn, Woods, Aulігné and Hoffman (MWR 2014a)