Model error in data assimilation

Patrick Laloyaux

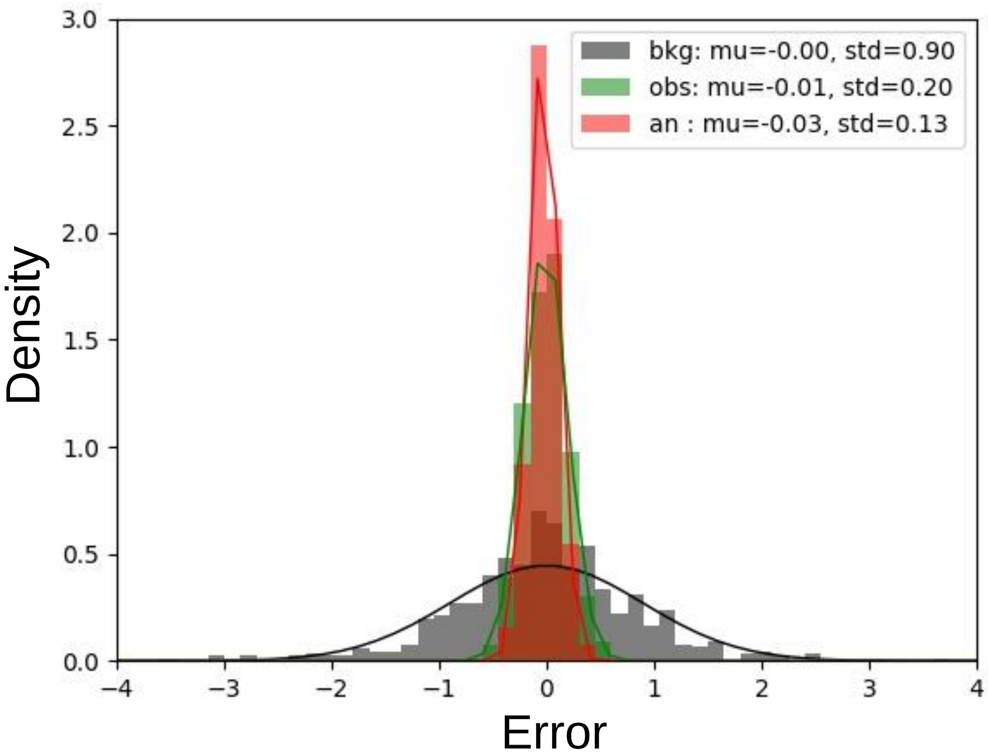
What you have learned so far

The analysis is computed by minimising 4D-Var

$$J(x_0) = \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) + \frac{1}{2}\sum_{k=0}^{K}[y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)]$$

using the model's equations

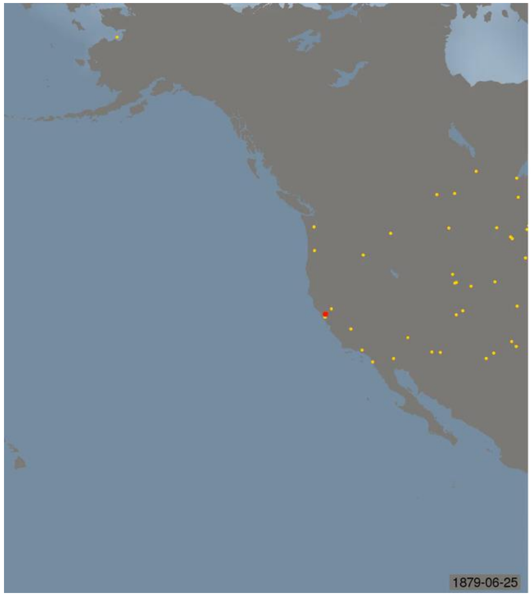
$$x_k = \mathcal{M}_k(x_{k-1})$$

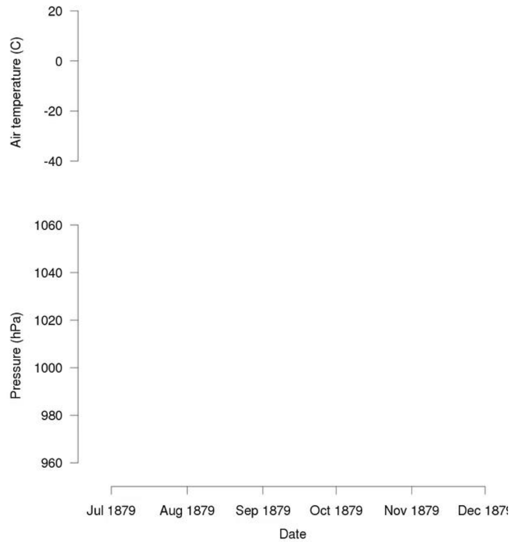


4D-Var combines model predictions with observations (**errors have to be random with zero-mean**)

Most observations have biases

The USS Jeannette (1879, Artic, 33 crew members)





SST measurements from buckets have a cold bias (~0.4C)

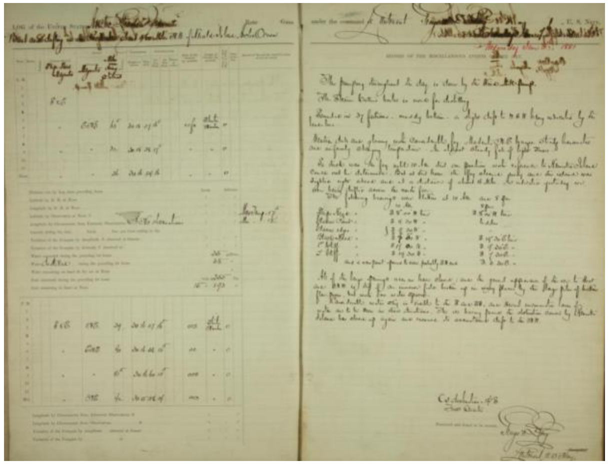


Photo # NH 52000 Steamer Jeannette sinking after being crushed by Arctic ice, June 1881

THE SINKING OF THE JEANNETTE.

Photo # NH 52002 Jeannette's crewmen drag their boats over the Arctic ice, June-August 1881

DRAGGING THE BOATS OVER THE ICE

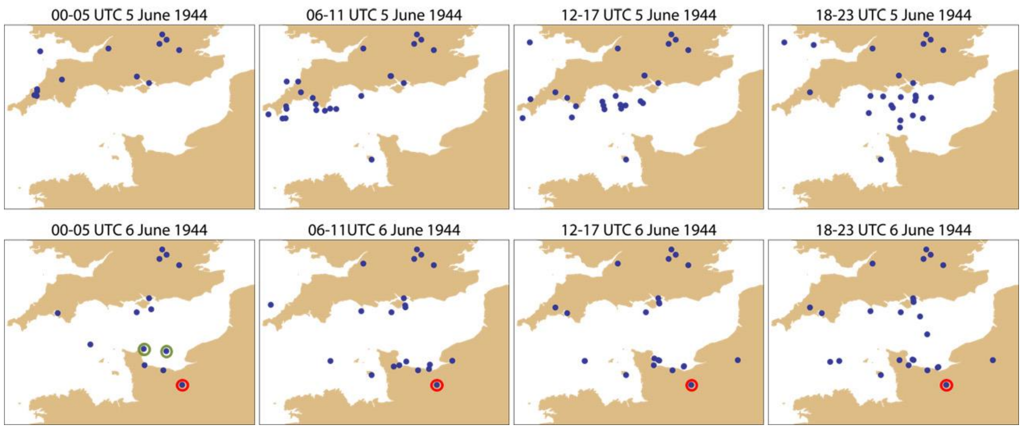
Photo # NH 92142 LCdr. DeLong and his party wading ashore on the Lena Delta, Siberia, 17 Sept. 1881

WADING ASHORE.

Courtesy of P. Brohan and G. Compo

Most observations have biases

D-Day (1944, France)

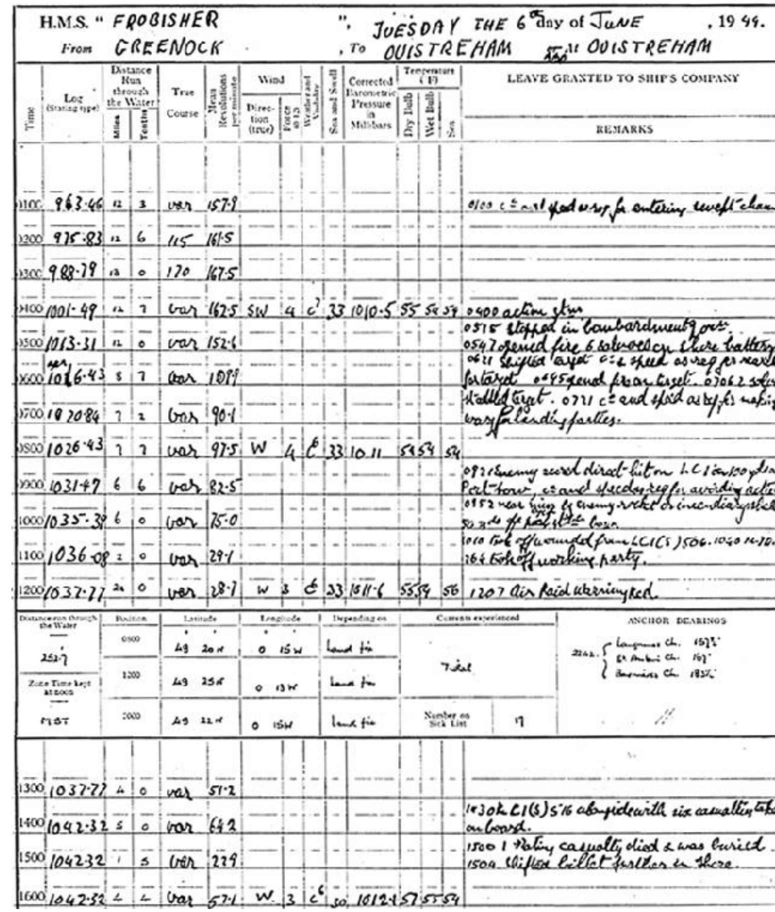


HMS Frobisher

Pressure	1010.5hPa (mb)	
Temperature	285.95K	(55°F)
Dew point (wet bulb)	284.95K	(54°F)
Wind direction	225°	(SW)
Wind speed	6.7ms^{-1}	(Force 4)
(Weather/) Visibility	Code 97	(c/7)
Sea temperature	285.35K	(54°F)

HMS Hawkins

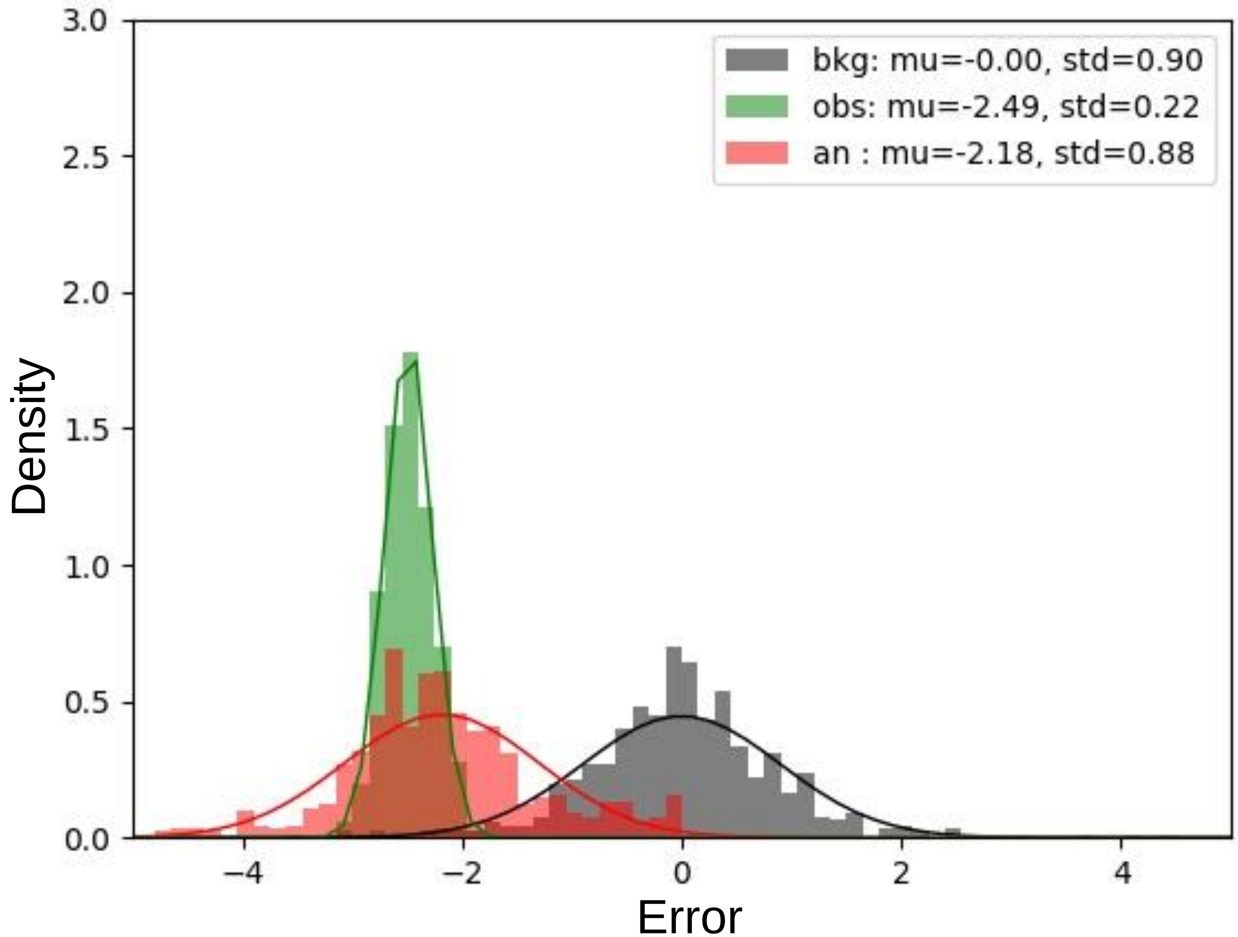
Pressure	1014.8hPa (mb)	
Temperature	285.35K	(54°F)
Dew point (wet bulb)	283.35K	(52°F)
Wind direction	270°	(W)
Wind speed	6.7ms^{-1}	(Force 4)
(Weather/) Visibility	Code 96	(c/6)
Sea temperature	284.25K	(52°F)



SST measurements from Engine Room Intake (ERI) have a warm bias (~0.2C)

Assimilation of biased observations

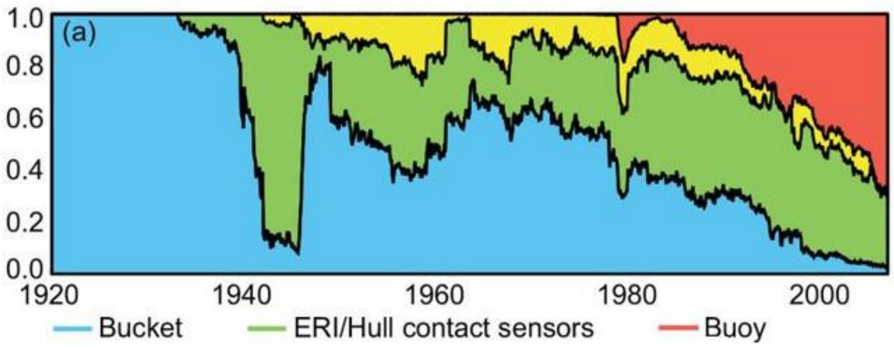
The standard 4D-Var formulation is designed to cope with random, zero-mean errors from the model and the observations

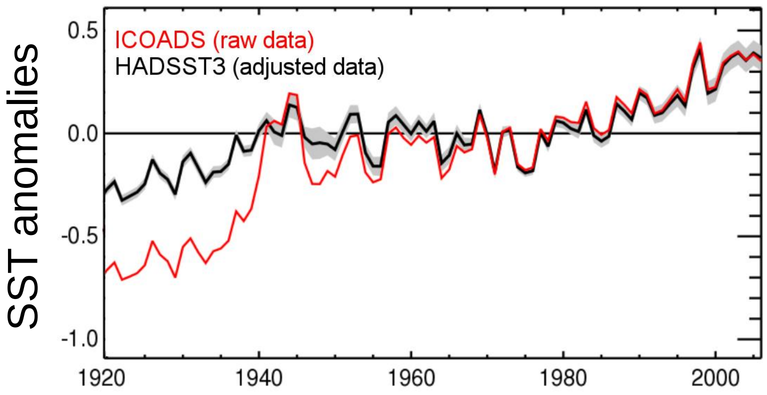


If biased observations are assimilated, the resulting analysis will be biased. In this case the background is more accurate than the analysis!

How to remove observation biases

Before the assimilation, based on instrument properties





During the assimilation, using information from the model and reference observations

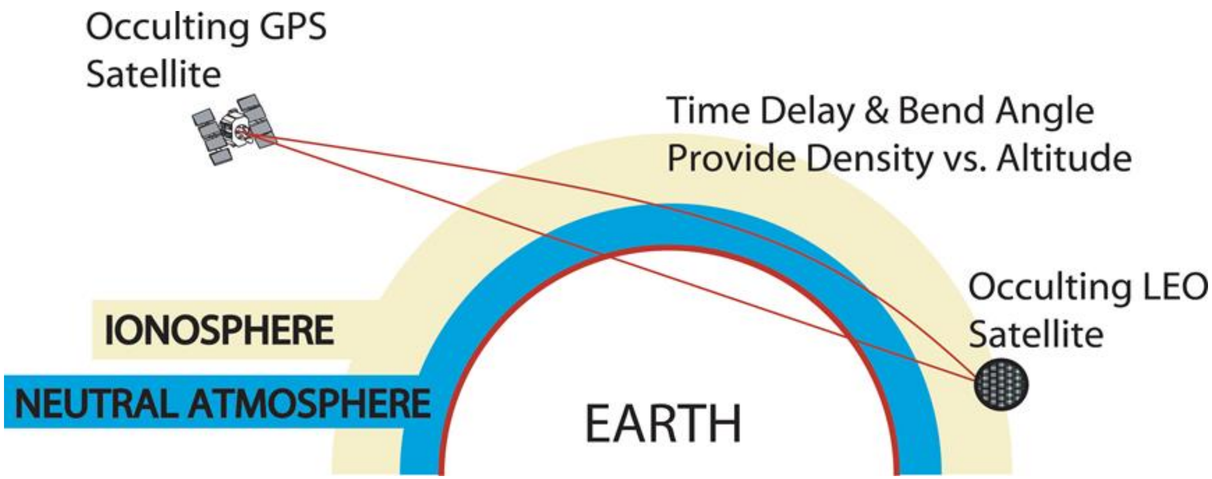
$$
\begin{aligned}
J(x_0, \beta) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\
&+ \frac{1}{2}\sum_{k=0}^{K}[y_k - \mathcal{H}(x_k) - b(x_k, \beta)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k) - b(x_k, \beta)] \\
&+ \frac{1}{2}(\beta - \beta_b)^T \mathbf{B}_\beta{}^{-1}(\beta - \beta_b)
\end{aligned}
$$

→ designed to estimate simultaneously the initial condition and parameters that represent systematic errors in the observationssystem

→ the bias model copes with instrument miscalibration (e.g. radiances systematically too warm by 1K) or systematic errors in the observation operator

How to estimate model biases

The GPS satellites are used for positioning and navigation. GPS-RO (Radio Occultation) is based on analysing the bending caused by the atmosphere along paths between a GPS satellite and a receiver placed on a low-earth-orbiting satellite.



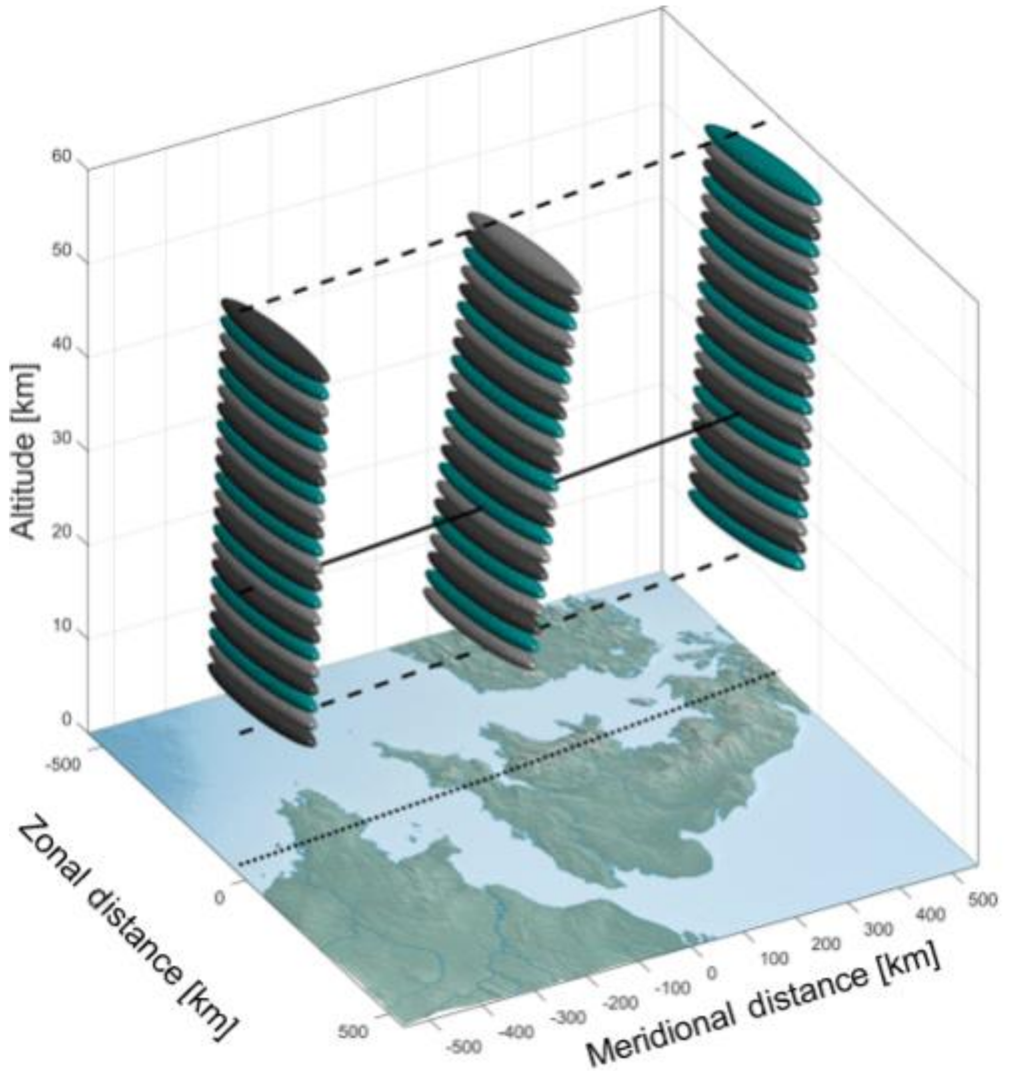
→ As the LEO moves behind the earth, we obtain a profile of bending angles

→ Temperature profiles can then be derived (a vertical interval between 10-50 km)

→ GPS-RO can be assimilated without bias correction. They are good for highlighting errors/biases

How to estimate model biases

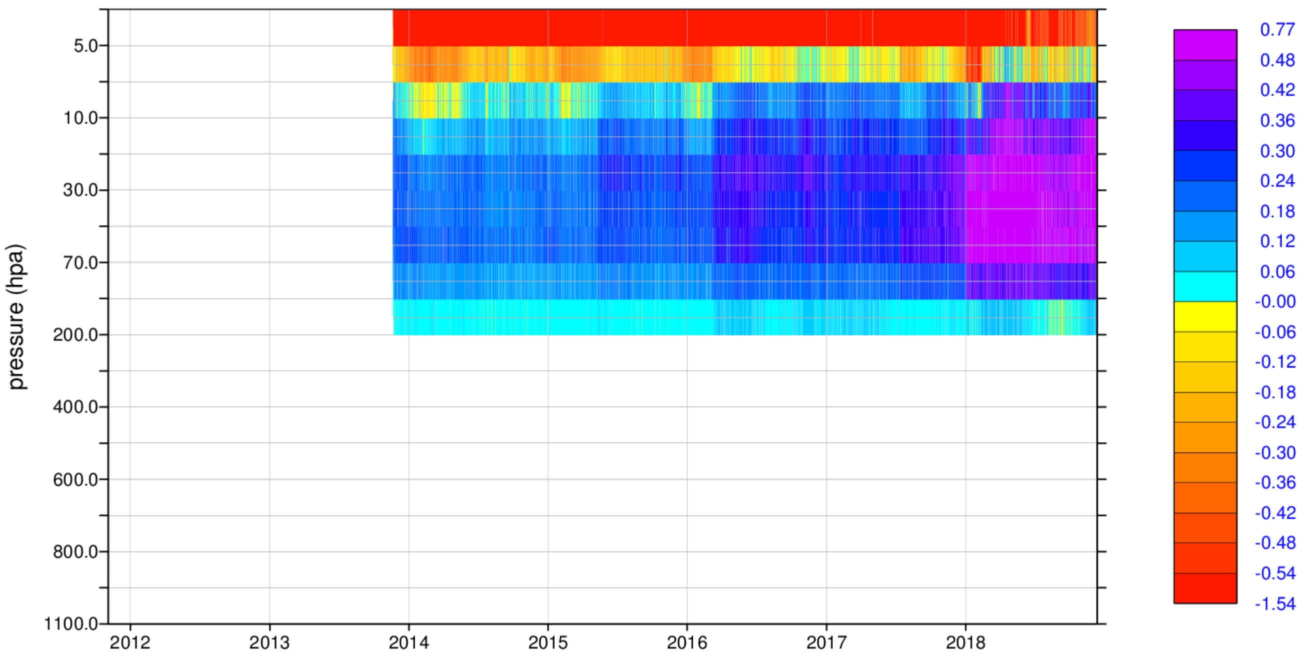
Temperature estimate from the GPS-RO measurements



Courtesy of Corwin J. Wright

How to estimate model biases

The first-guess trajectory of the model can be compared to unbiased observations



Difference between GPS-RO temperature retrievals and the IFS first-guess trajectory (O-B)

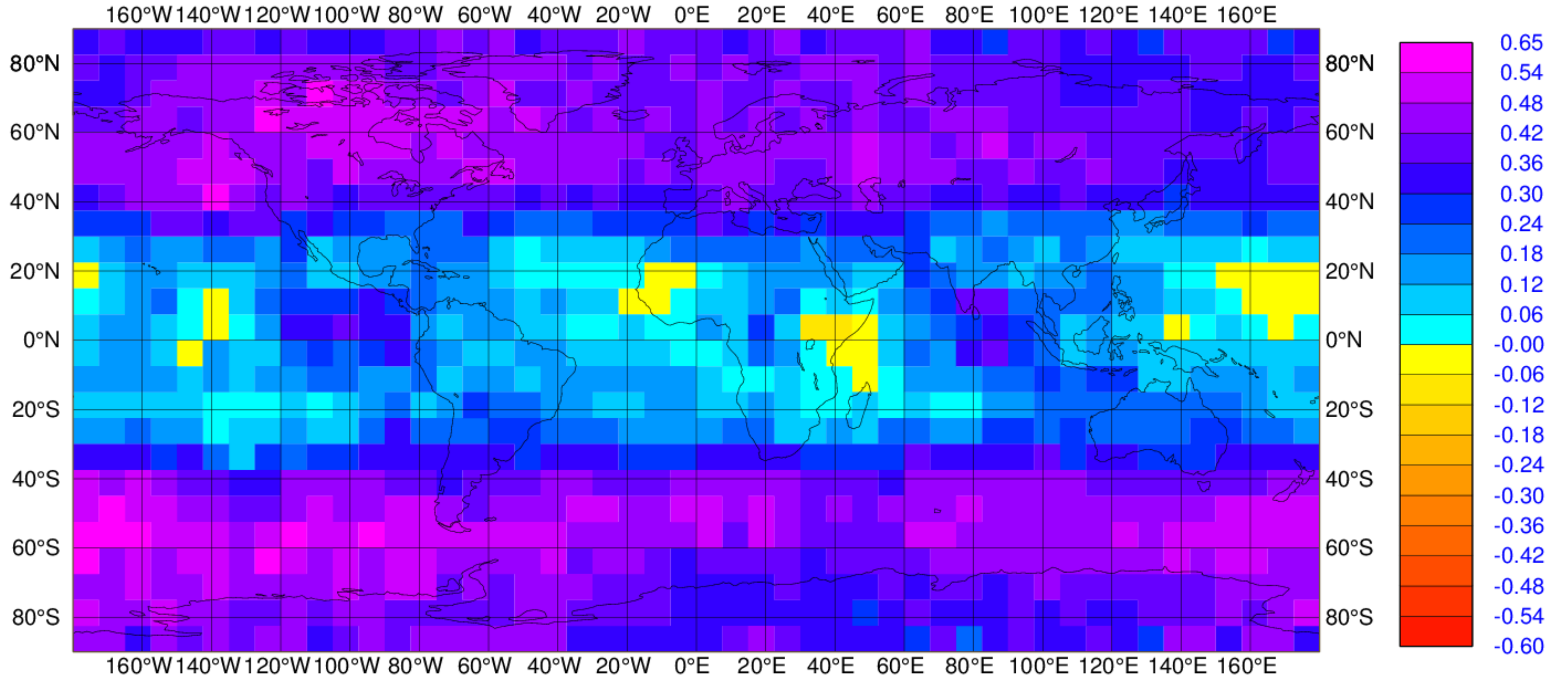
Errors in models are often systematic rather than random, zero-mean

→ Model has a temperature cold bias in the lower/mid stratosphere

→ Model has a warm bias in the upper stratosphere

How to estimate model biases

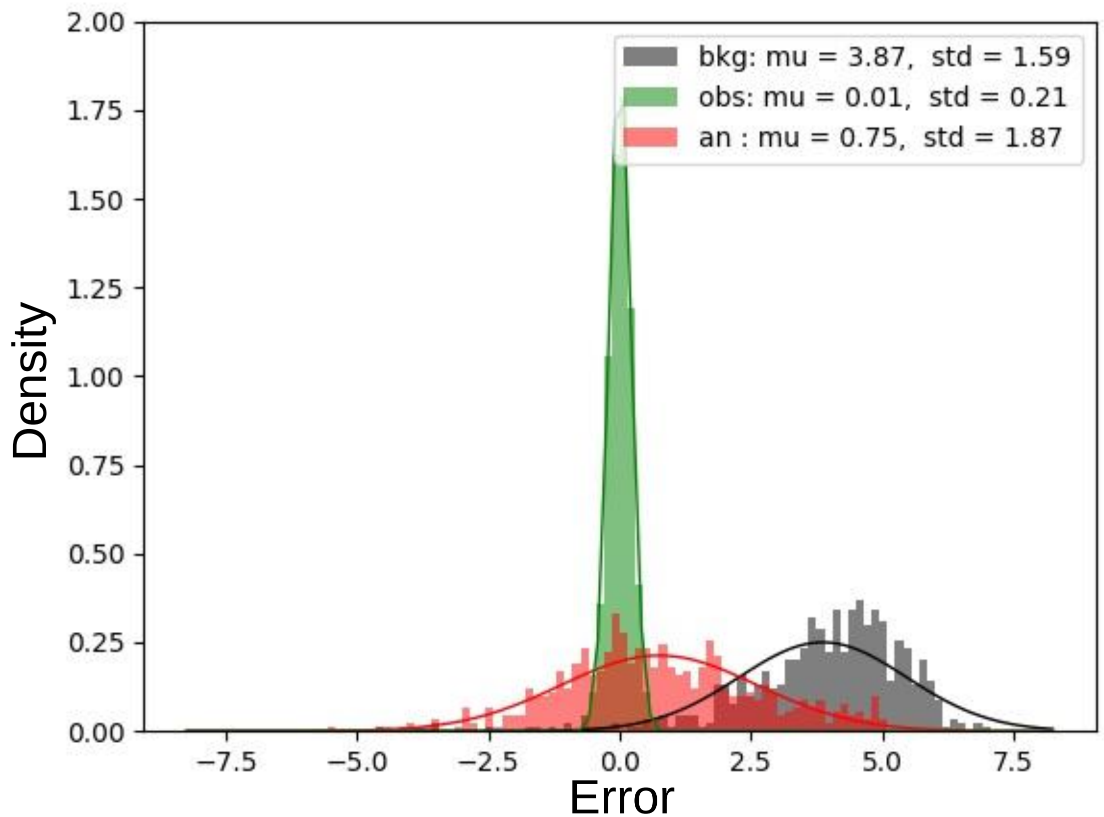
GPS-RO temperature retrievals provide an homogeneous observing system that can be used to study the spatial distribution of the model error



→ The IFS model shows very large structures in the temperature model bias

Assimilation with a biased model

The standard 4D-Var formulation is designed to cope with random, zero-mean errors from the model and the observations



The model produces a biased first-guess trajectory. Even if observations are unbiased and accurate, the final analysis will be biased.

Weak constraint 4D-Var

We assume that the model is not perfect, adding an error term η in the model equation

$$x_k = \mathcal{M}_k(x_{k-1}) + \eta \quad \text{for } k = 1, 2, \cdots, K$$

The model error estimate η contains 3 physical fields

- temperature
- vorticity
- Divergence

Constant model error forcing over the assimilation window to correct the model bias

$$\begin{aligned} J(x_0, \beta, \eta) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\ &+ \frac{1}{2}\sum_{k=0}^{K}[y_k - \mathcal{H}(x_k) - b(x_k, \beta)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k) - b(x_k, \beta)] \\ &+ \frac{1}{2}(\beta - \beta_b)^T \mathbf{B}_\beta{}^{-1}(\beta - \beta_b) \\ &+ \frac{1}{2}(\eta - \eta_b)^T \mathbf{Q}^{-1}(\eta - \eta_b) \end{aligned}$$

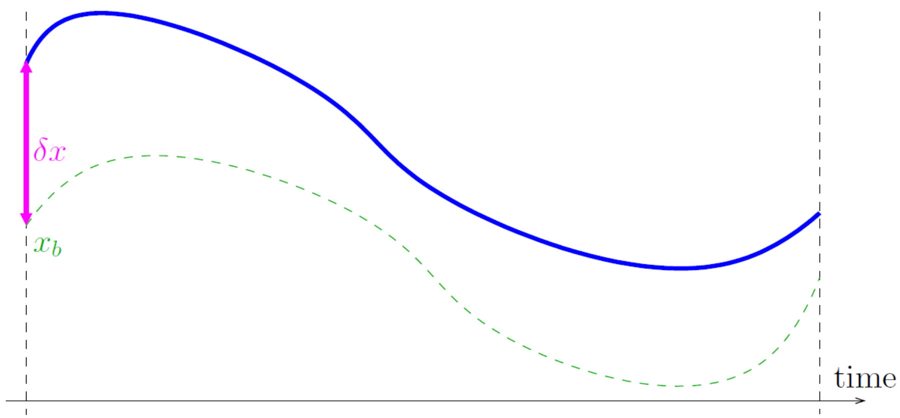
Model state → x_0
Observation bias parameters → β
Model error → η

→ Introduce additional controls to target an unbiased analysis

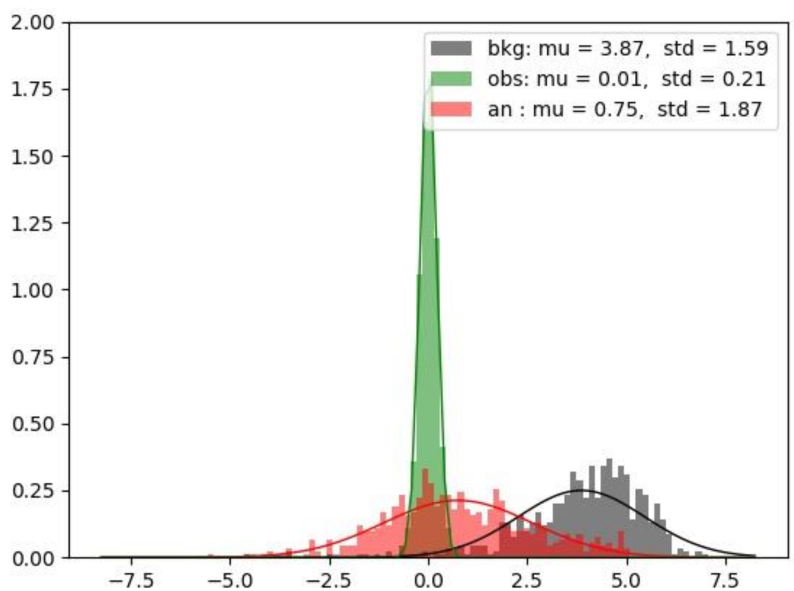
→ The model error covariance matrix Q constrains the model error field (22 millions of parameters)

Weak constraint 4D-Var

Strong constraint 4D-Var

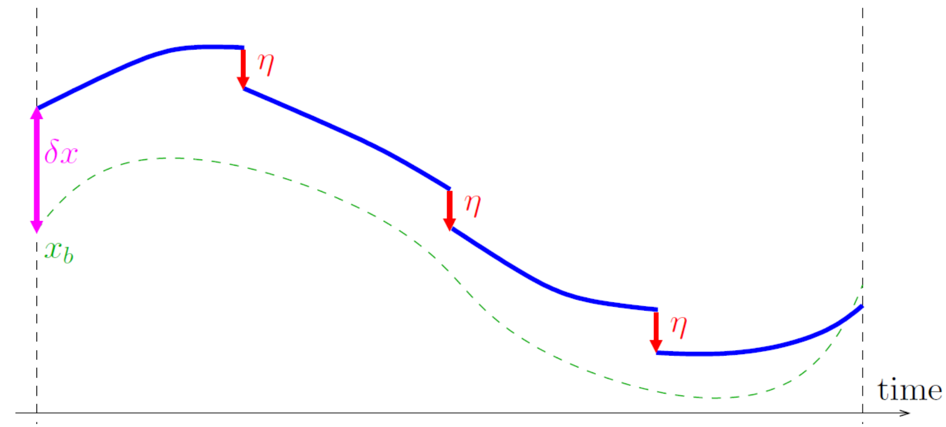


$$x_k = \mathcal{M}_k(x_{k-1})$$

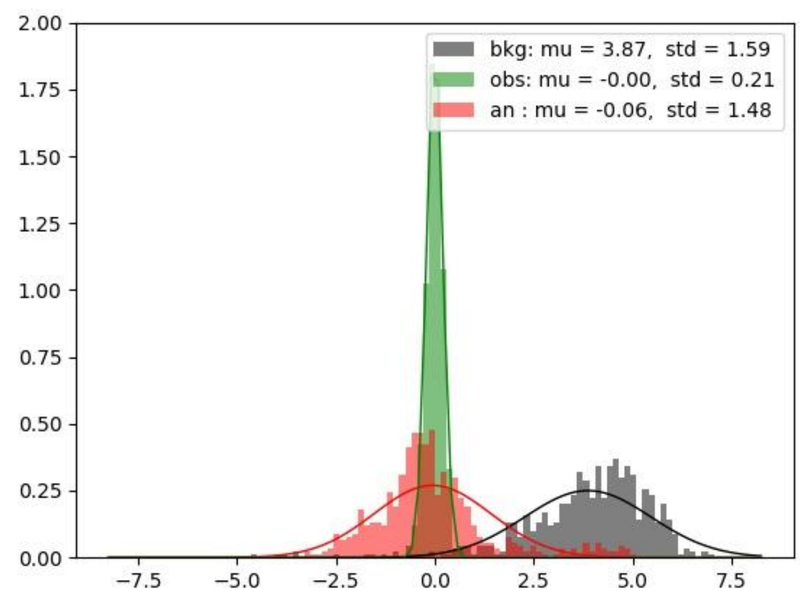


→ Large bias and standard deviation in the analysis

Weak constraint 4D-Var

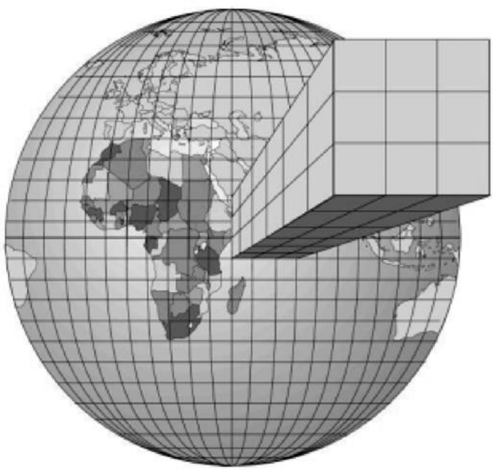


$$x_k = \mathcal{M}_k(x_{k-1}) + \eta \quad \text{for } k = 1, 2, \cdots, K$$



→ Bias in the analysis has been reduced, standard deviation as well

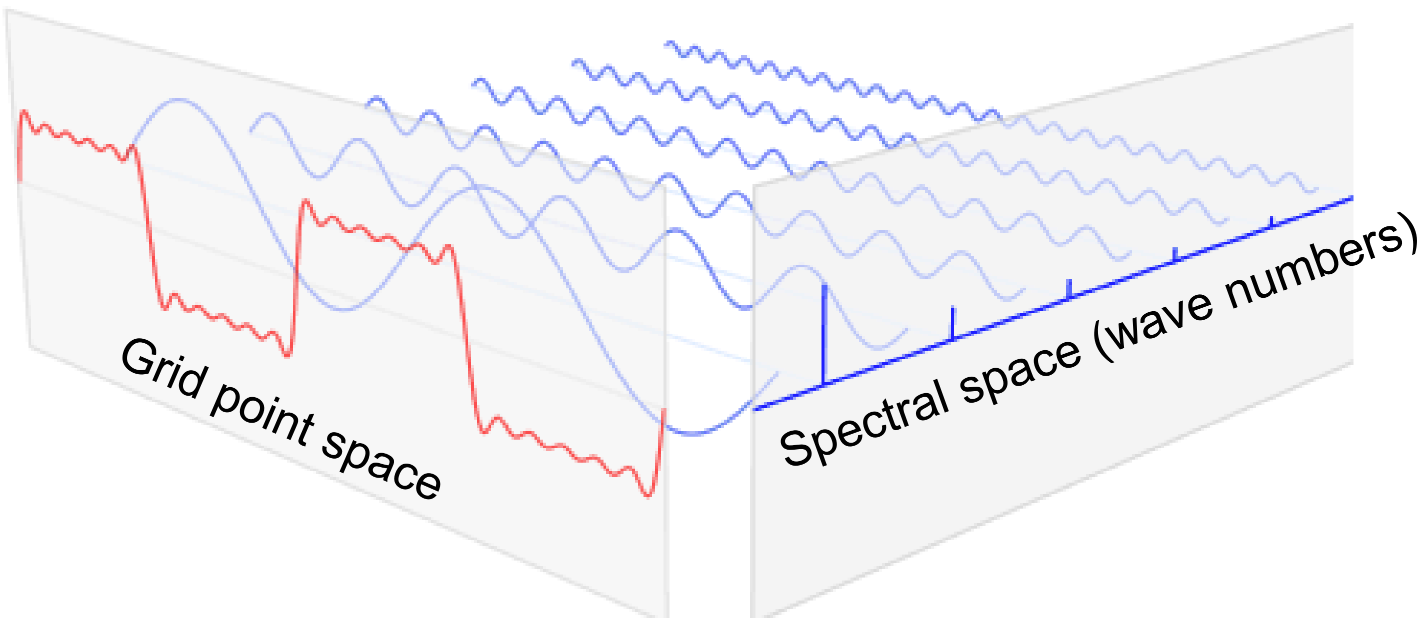
Specification of B and Q



IFS is based on a hybrid spectral/grid-point model

The spectral transform

$$A(\lambda,\mu,\eta,t) = \sum_{l=0}^{T}\sum_{|m|\le l}^{T}\psi_{lm}(\eta,t)Y_{lm}(\lambda,\mu)$$



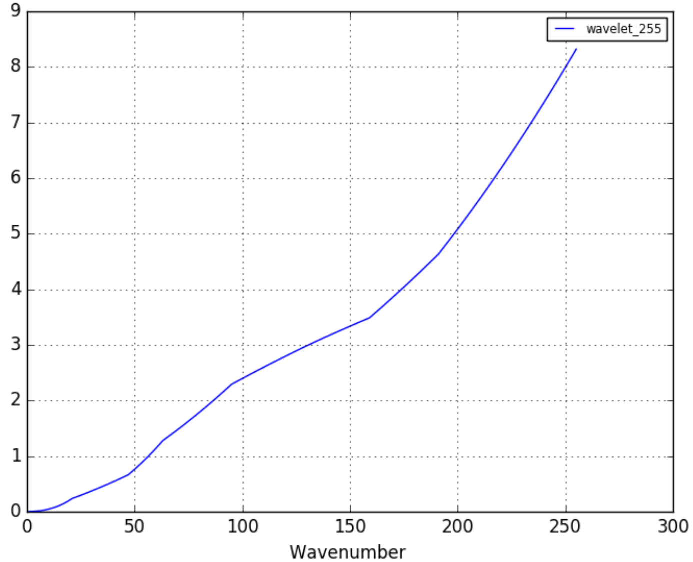
→ First wave numbers contain large scale signals

→ Last wave numbers contain small scale signals

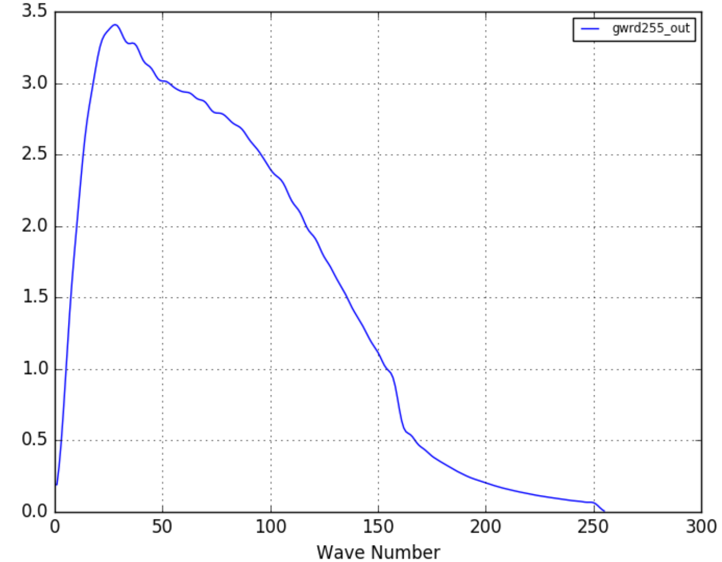
Specification of B and Q

Assimilation is computed in spectral space

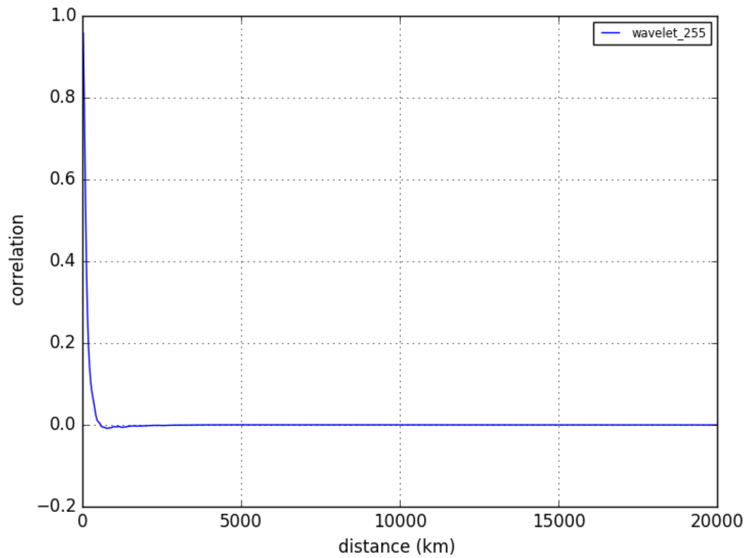
Standard deviation in B



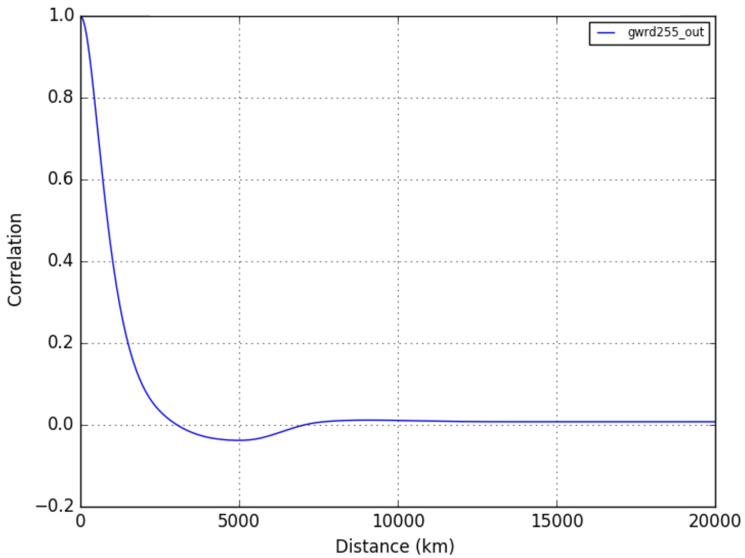
Standard deviation in Q



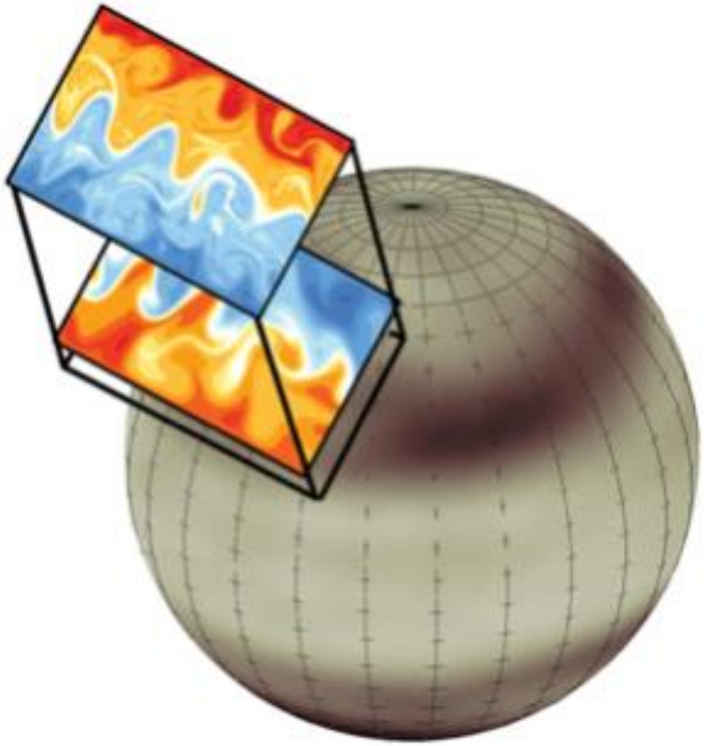
Information spread in B



Information spread in Q

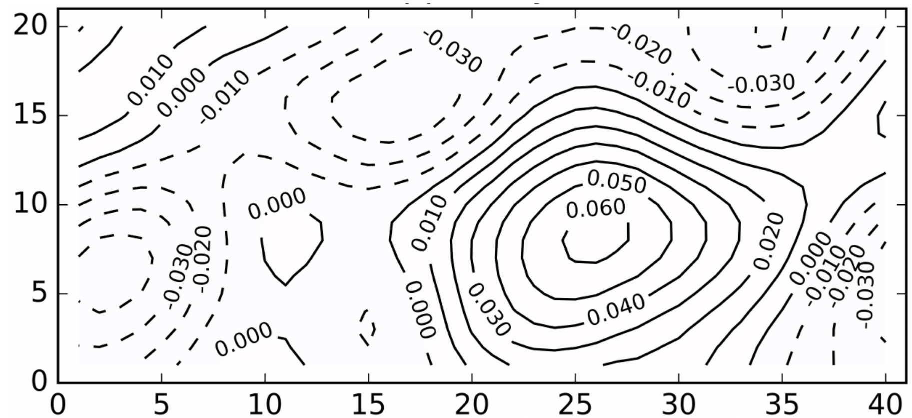


Weak constraint 4D-Var results with the QG model



The Quasi-Geostrophic (QG) model is very important in geophysical fluid dynamics as it describes some aspects of flows in the oceans and atmosphere very well

Model bias for the upper level



Experiment framework

→ A bias is introduced in the model

$$x_k = \mathcal{M}_k(x_{k-1}) + \eta \quad \text{for } k = 1, 2, \cdots, K$$

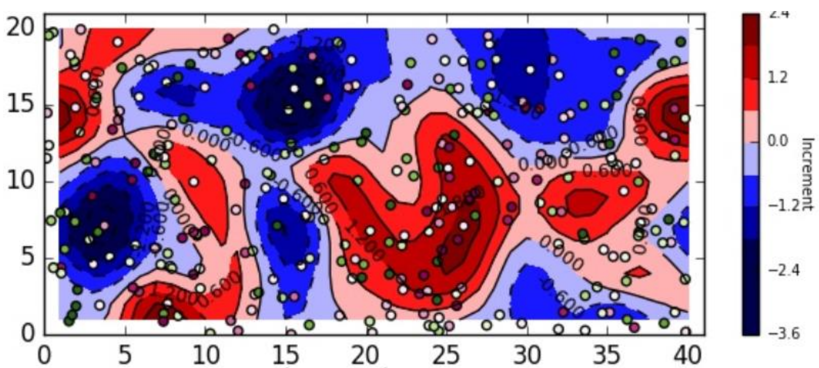
→ Observations are generated

→ Can weak constraint 4D-Var estimate correctly the model bias?

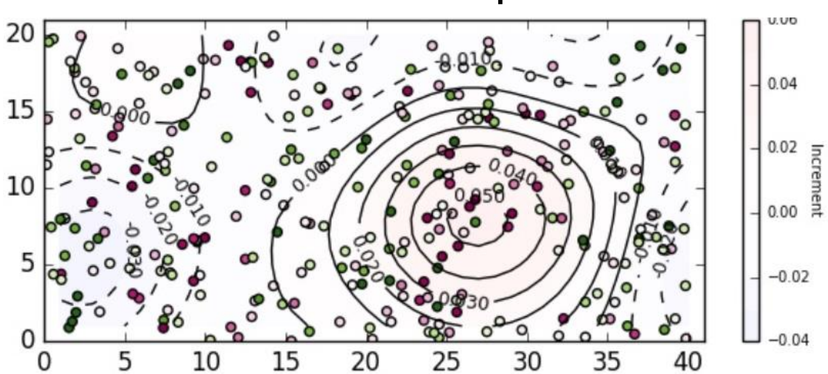
Weak constraint 4D-Var results with the QG model

Weak constraint 4D-Var

Model state increment δx



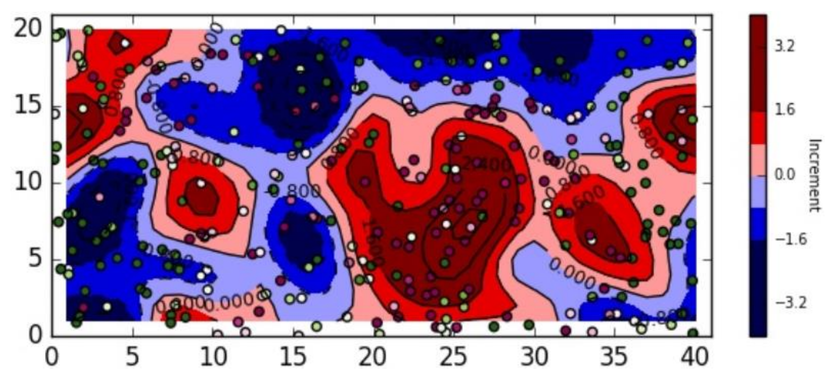
Model bias estimate η



→ Good separation between δx correcting the small scale random errors and η correcting the large scale systematic errors

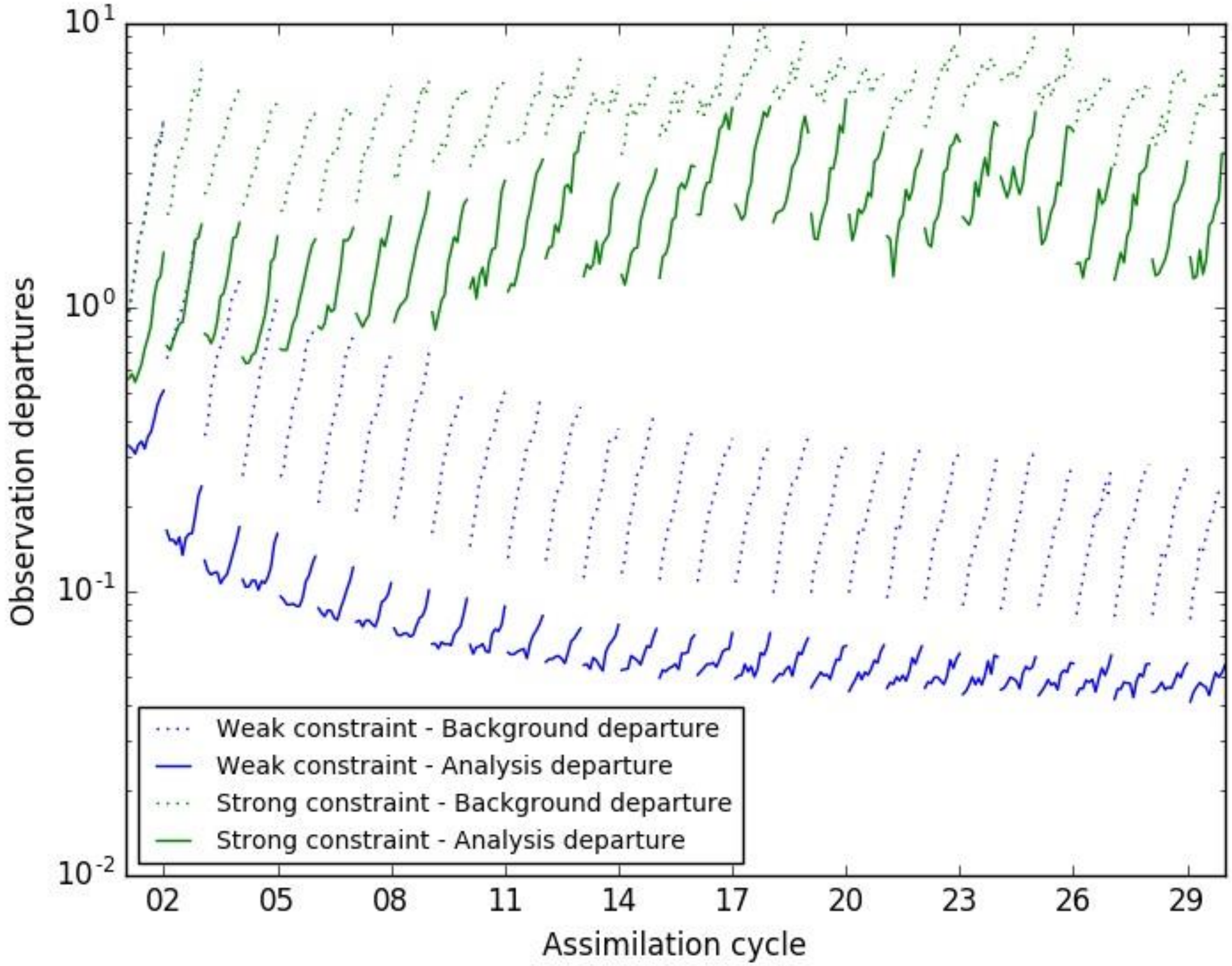
Strong constraint 4D-Var

Model state increment δx



→ Model state increment δx includes the two types of error, poor fit to the observations

Weak constraint 4D-Var results with the QG model



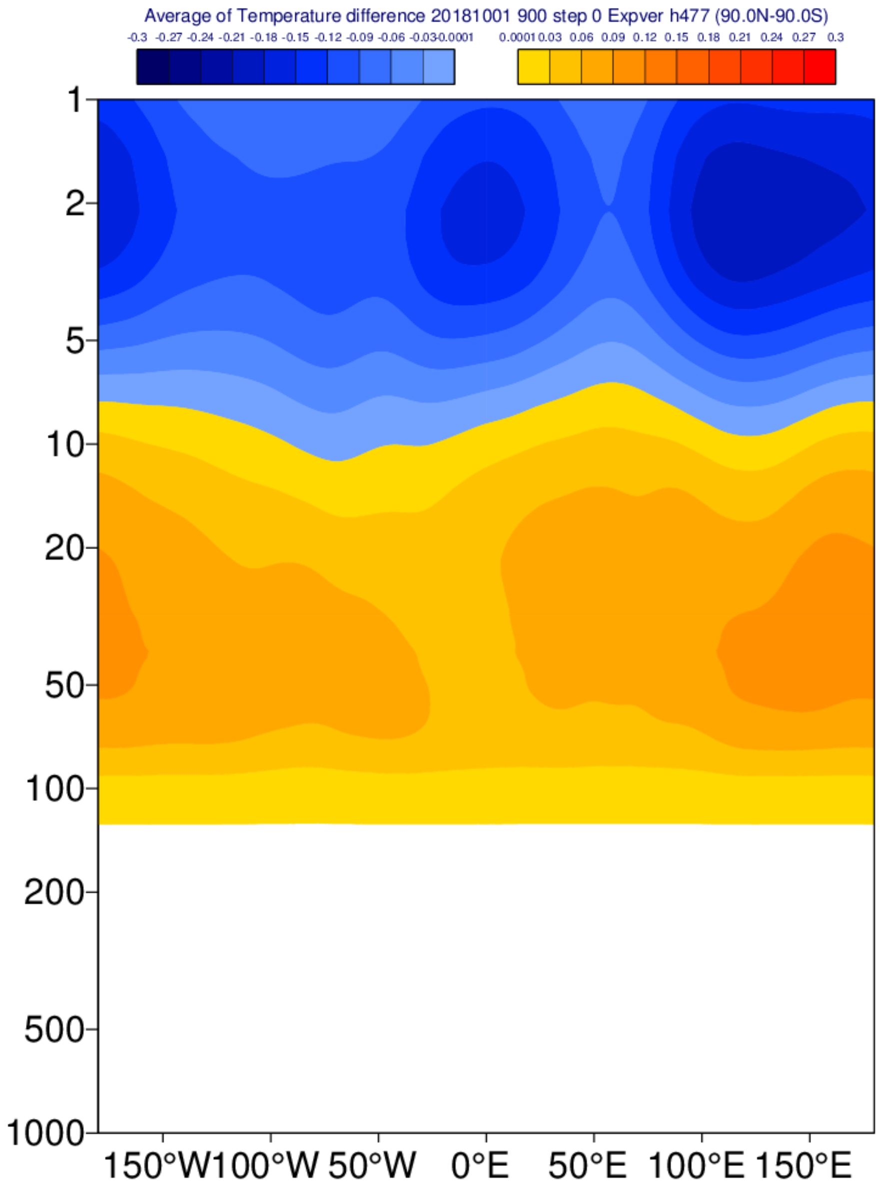
Information is cycle in the weak constraint 4D-Var

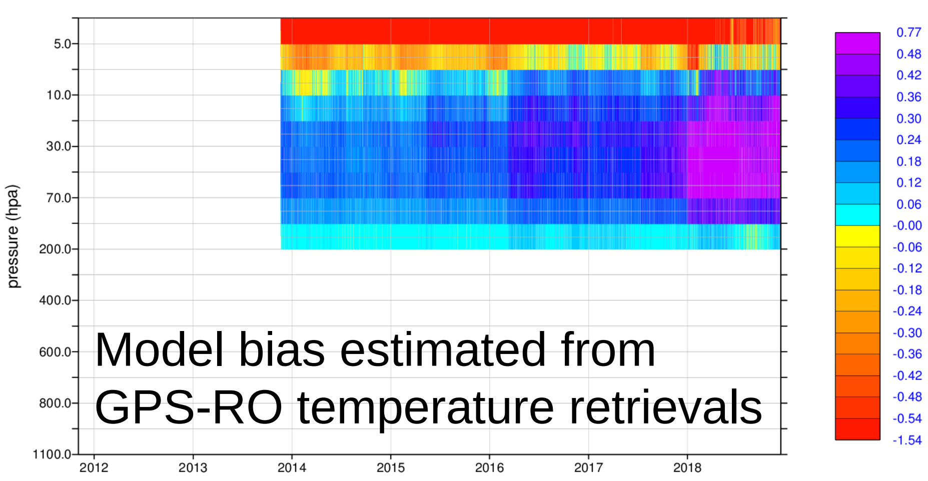
→ First-guess trajectory fir better and better the observations

→ More accurate analysis is produced

Weak constraint 4D-Var results with IFS

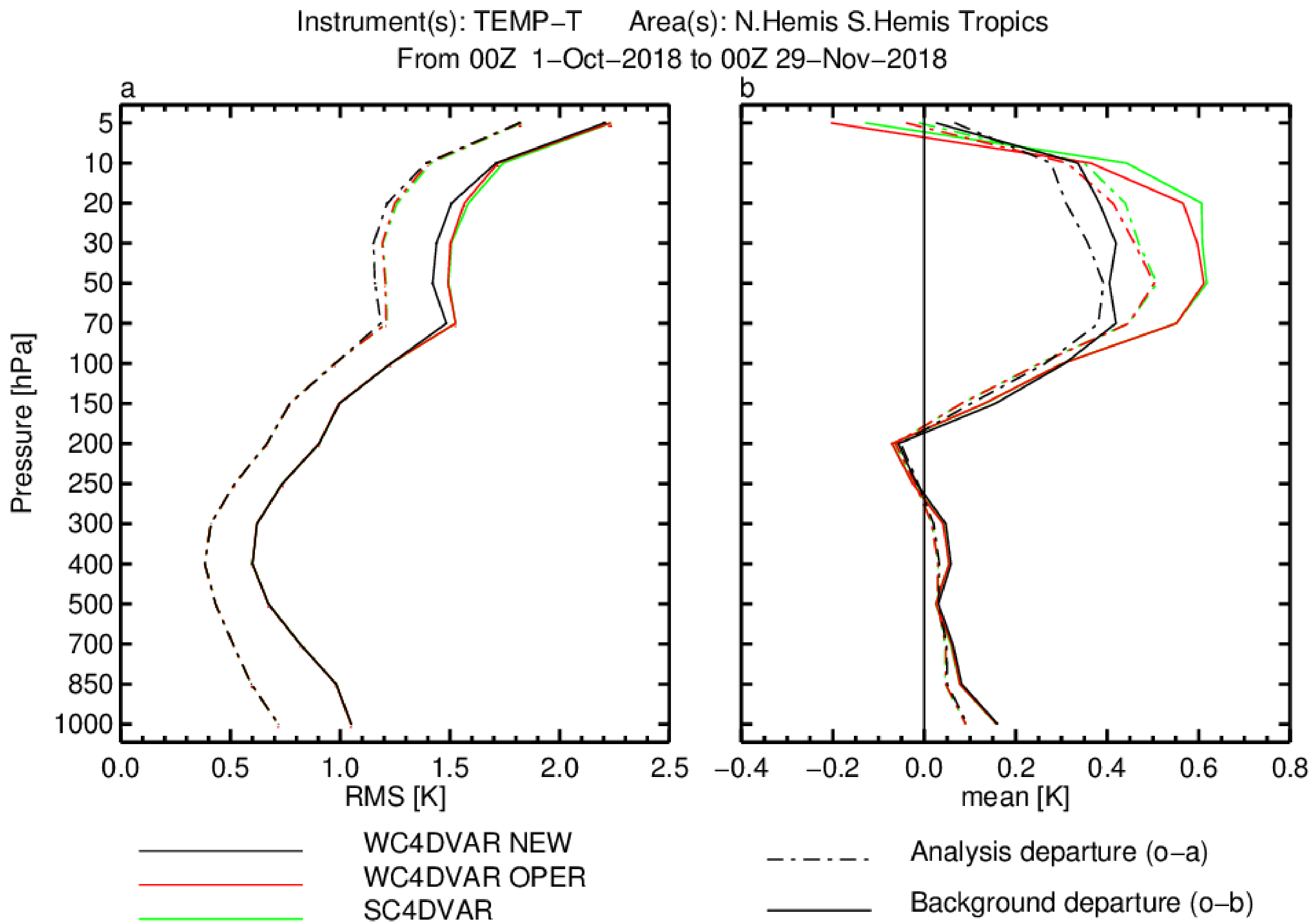
Bias estimated by the weak constraint 4D-Var





Weak constraint 4D-Var results with IFS

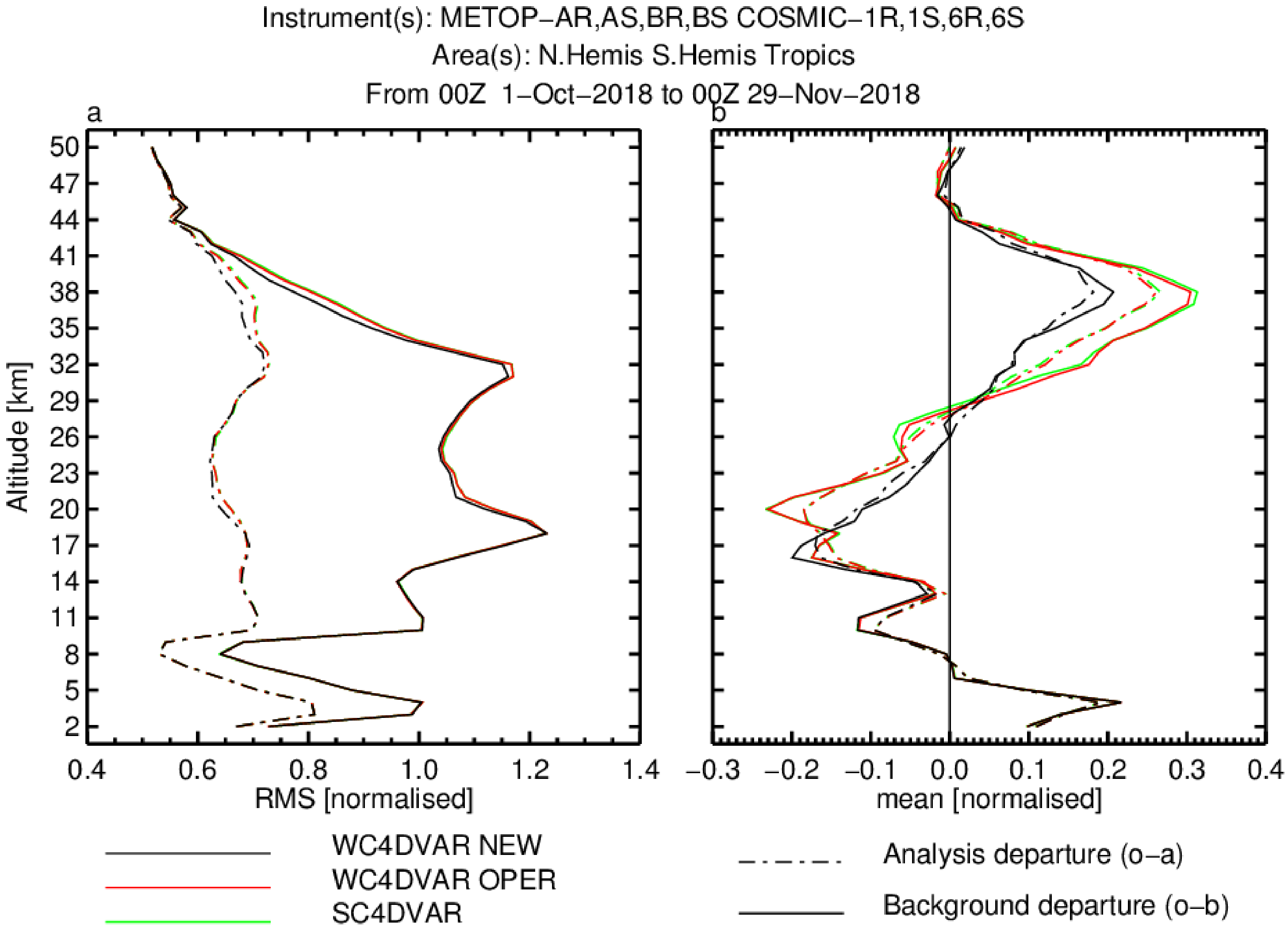
Departure statistics with radiosondes



→ Stratospheric bias in the analysis and in the first guess is reduced

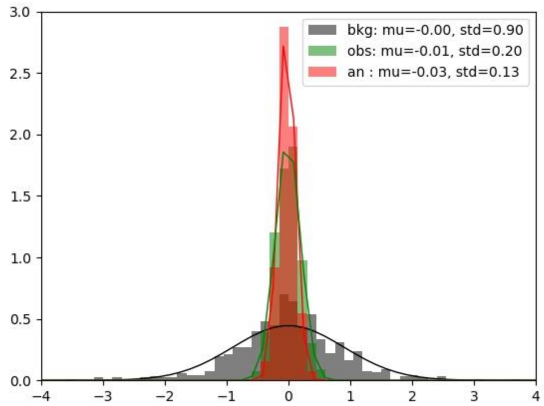
Weak constraint 4D-Var results with IFS

Departure statistics with GPS-RO

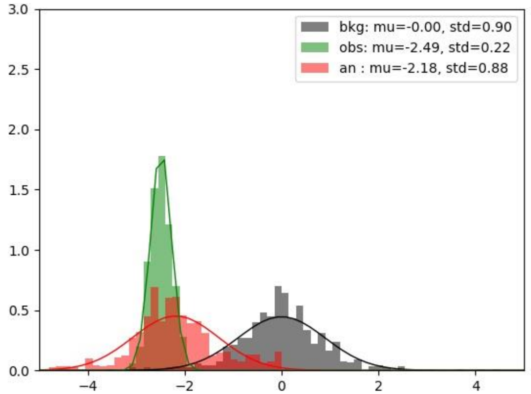


→ Stratospheric bias in the analysis and in the first guess is reduced

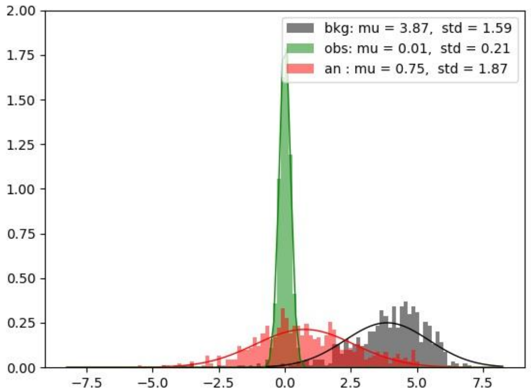
Summary



Background: unbiased (only random errors)
Observation: unbiased (only random errors)
Standard 4D-Var



Background: unbiased (only random errors)
Observation: biased
Standard 4D-Var & Variational Bias Control (VarBC)



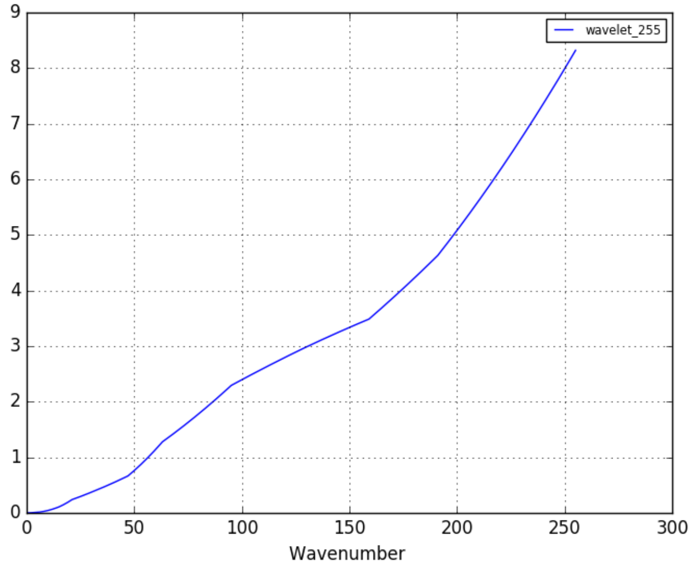
Background: biased
Observation: unbiased (only random errors)
Weak constraint 4D-Var

Summary

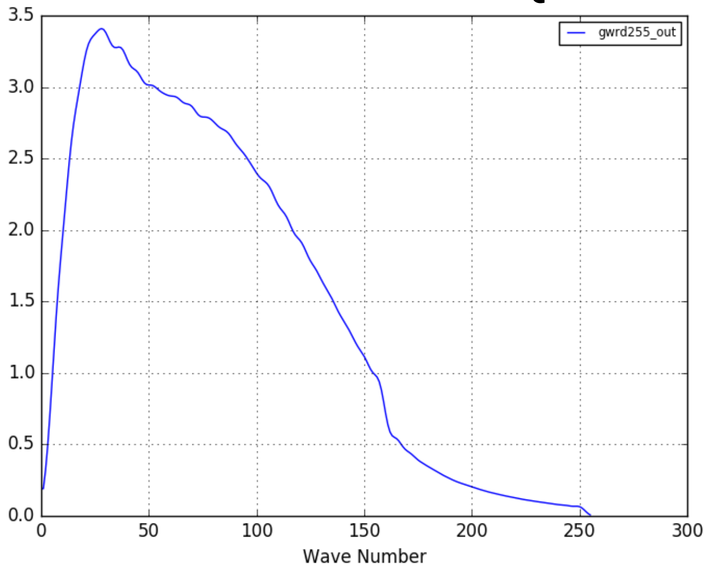
How can we separate background and model error?
How can we specify B and Q covariance matrices?

Random errors tend to be small scales while systematic errors tend to be large scale

Standard deviation in B



Standard deviation in Q



Weak constraint 4D-Var with scale separation reduces the stratospheric temperature bias in the analysis

In real world applications

How do I know if my observations are biased?
How do I know if my model is biased?
I'm not running twin experiments, I don't know the truth

Reference observations are used

Radiosondes

GPS-RO

In real world applications

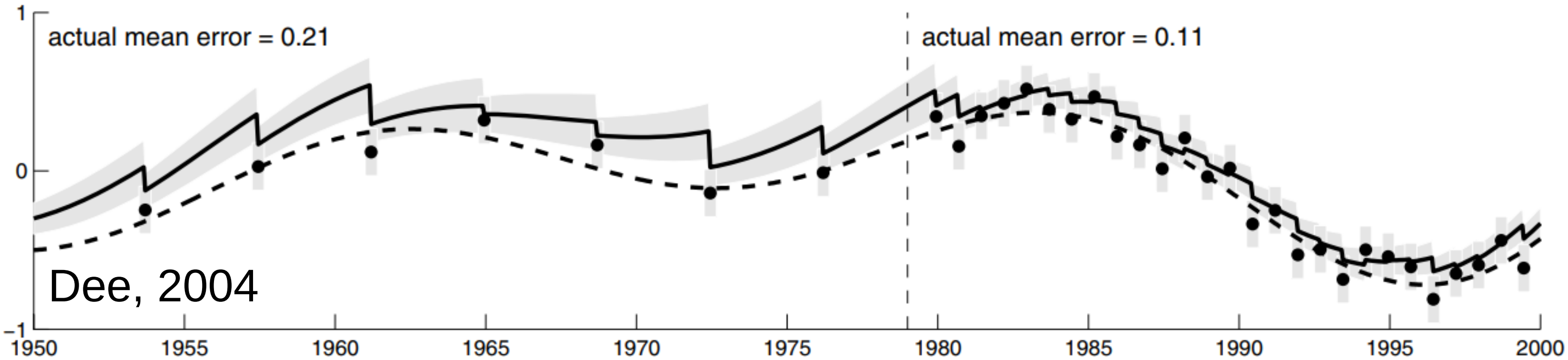
From bias-blind to bias-aware data assimilation

$$
\begin{aligned}
J(x_0, \beta, \eta) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\
&+ \frac{1}{2}\sum_{k=0}^{\text{Radiosonde}} [y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \\
&+ \frac{1}{2}\sum_{k=0}^{\text{GPSRO}} [y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \\
&+ \frac{1}{2}\sum_{k=0}^{\text{Others}} [y_k - \mathcal{H}(x_k) - b(x_k, \beta)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k) - b(x_k, \beta)] \\
&+ \frac{1}{2}(\beta - \beta_b)^T \mathbf{B}_\beta{}^{-1}(\beta - \beta_b) \\
&+ \frac{1}{2}(\eta - \eta_b)^T \mathbf{Q}^{-1}(\eta - \eta_b)
\end{aligned}
$$

Future of weak constraint 4D-Var

In Numerical Weather Prediction, weak constraint 4D-Var could be investigated to correct other variables (e.g. humidity) or other areas (e.g. troposphere)

In climate reanalysis, weak constraint 4D-Var could be applied over the whole period to reduce spurious climate change due to changes in the observing system



In other component of the Earth system, weak constraint 4D-Var could be applied to the ocean to deal with temperature/ salinity biases

Any questions? Feel free to contact me patrick.laloyaux@ecmwf.int