Model error in data assimilation

Patrick Laloyaux

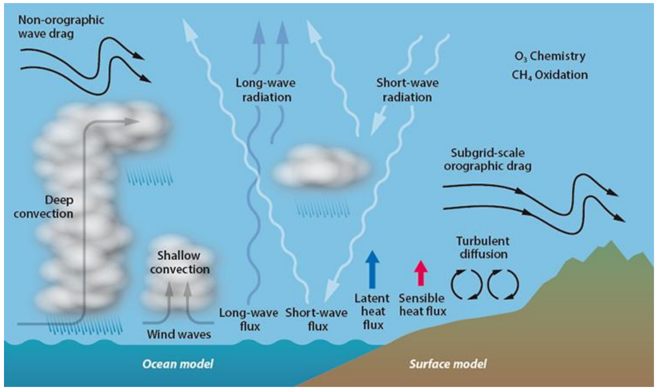
Learn about the typical issues that occur in 4D-Var

Identify systematic errors in observations and model (biases)

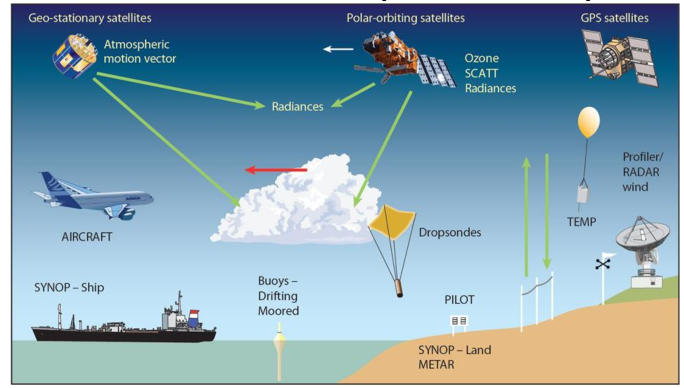
Develop bias correction methods

What you have seen so far on data assimilation

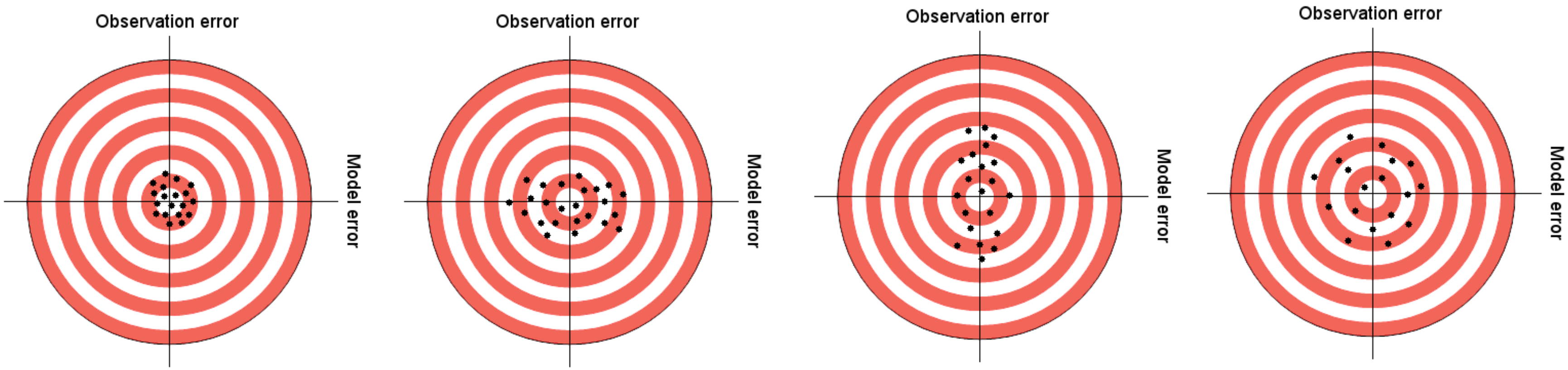
Model (with errors)



Observations (with errors)



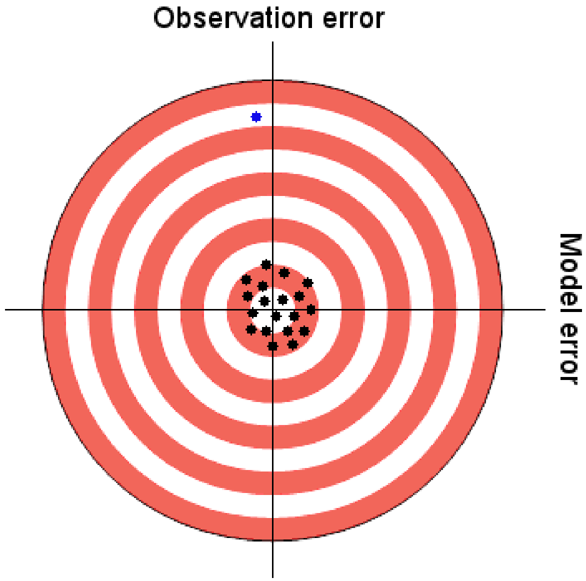
If you are lucky, model and observations are not biased

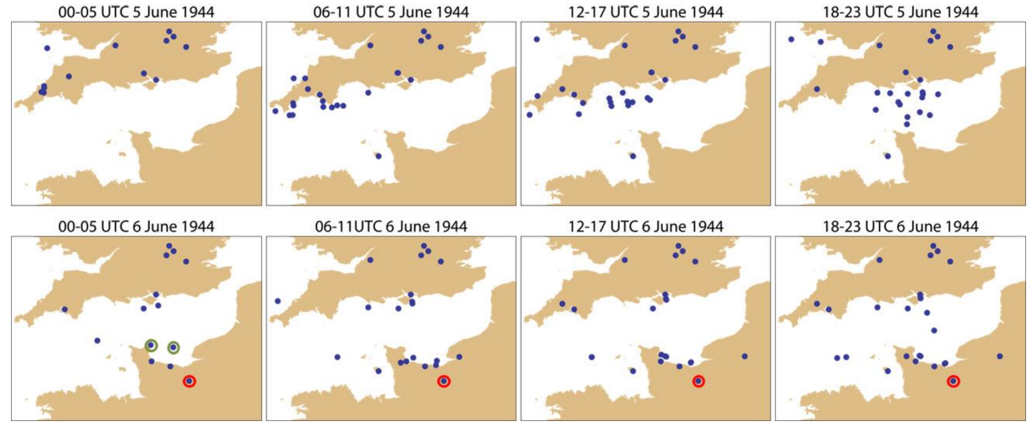


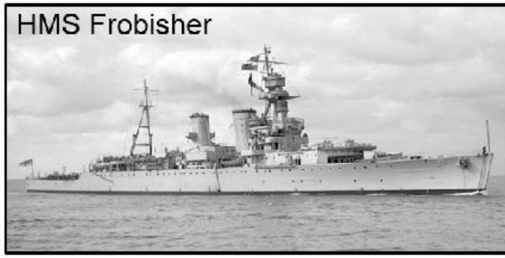
$$
\begin{aligned}
J(x_0) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\
&+ \frac{1}{2}\sum_{k=0}^{K}[y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)]
\end{aligned}
$$

What's happening if you are unlucky?

Data assimilation and gross (obvious) errors

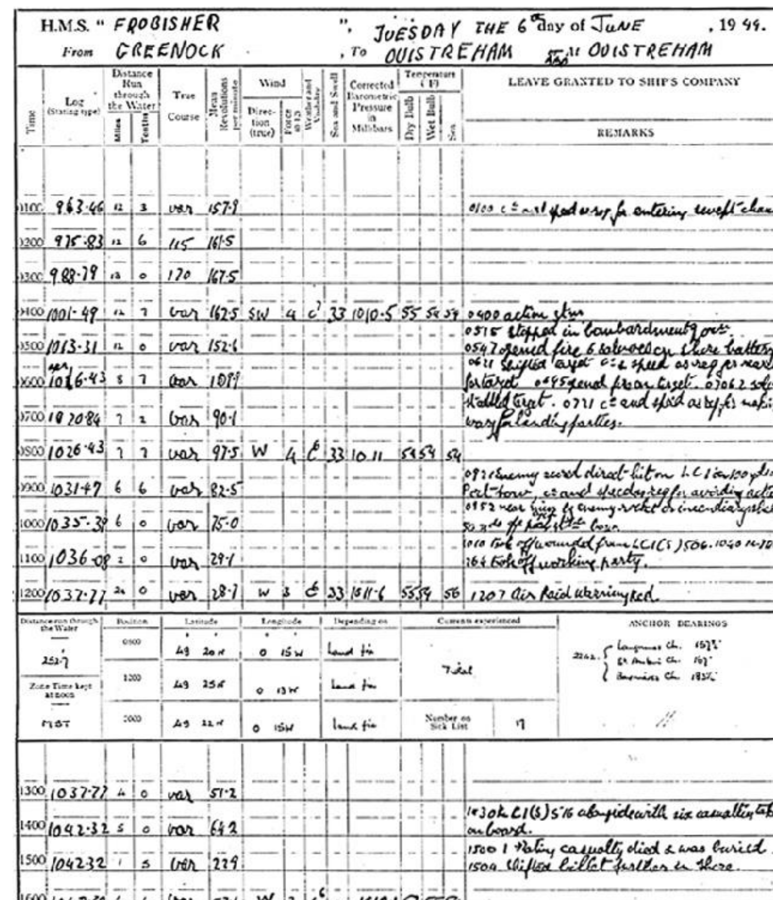






Pressure	1010.5hPa (mb)	
Temperature	285.95K	(55°F)
Dew point (wet bulb)	284.95K	(54°F)
Wind direction	225°	(SW)
Wind speed	6.7ms^{-1}	(Force 4)
(Weather/) Visibility	Code 97	(c/7)
Sea temperature	285.35K	(54°F)

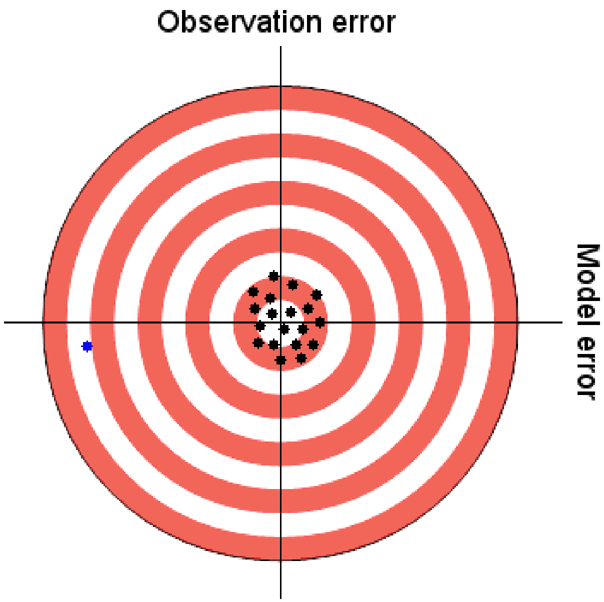
Pressure	1014.8hPa (mb)	
Temperature	285.35K	(54°F)
Dew point (wet bulb)	283.35K	(52°F)
Wind direction	270°	(W)
Wind speed	6.7ms^{-1}	(Force 4)
(Weather/) Visibility	Code 96	(c/6)
Sea temperature	284.25K	(52°F)


H.M.S. " FROBISHER ", TUESDAY THE 6 day of JUNE , 19 44.
From GREENOCK , To OUISTREHAM At OUISTREHAM
LEAVE GRANTED TO SHIP'S COMPANY
REMARKS

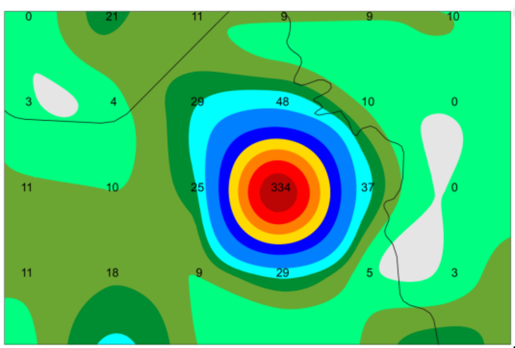
→ Preliminary analysis (blacklist,…)

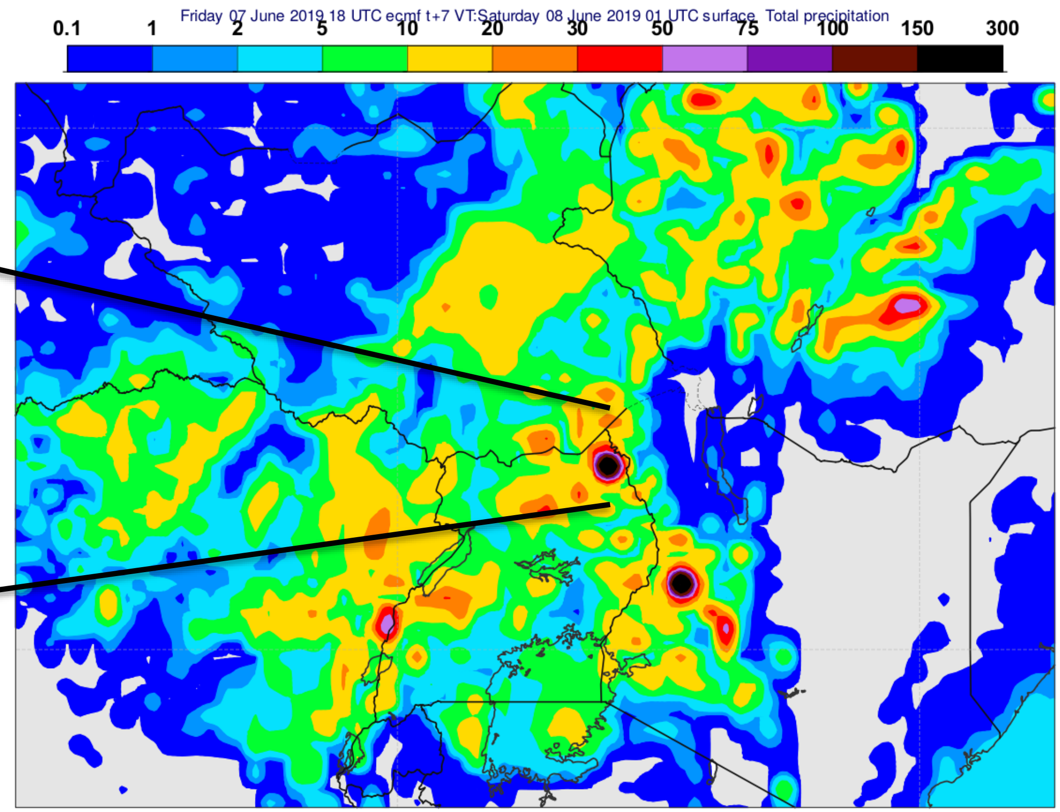
→ Online Quality Control

Data assimilation and gross (obvious) errors



Total precipitation on 07 June 2019
(accumulated over 6 hours)

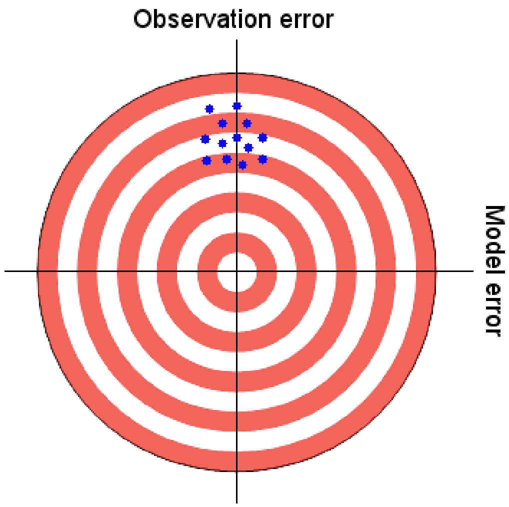




→ Continuous monitoring

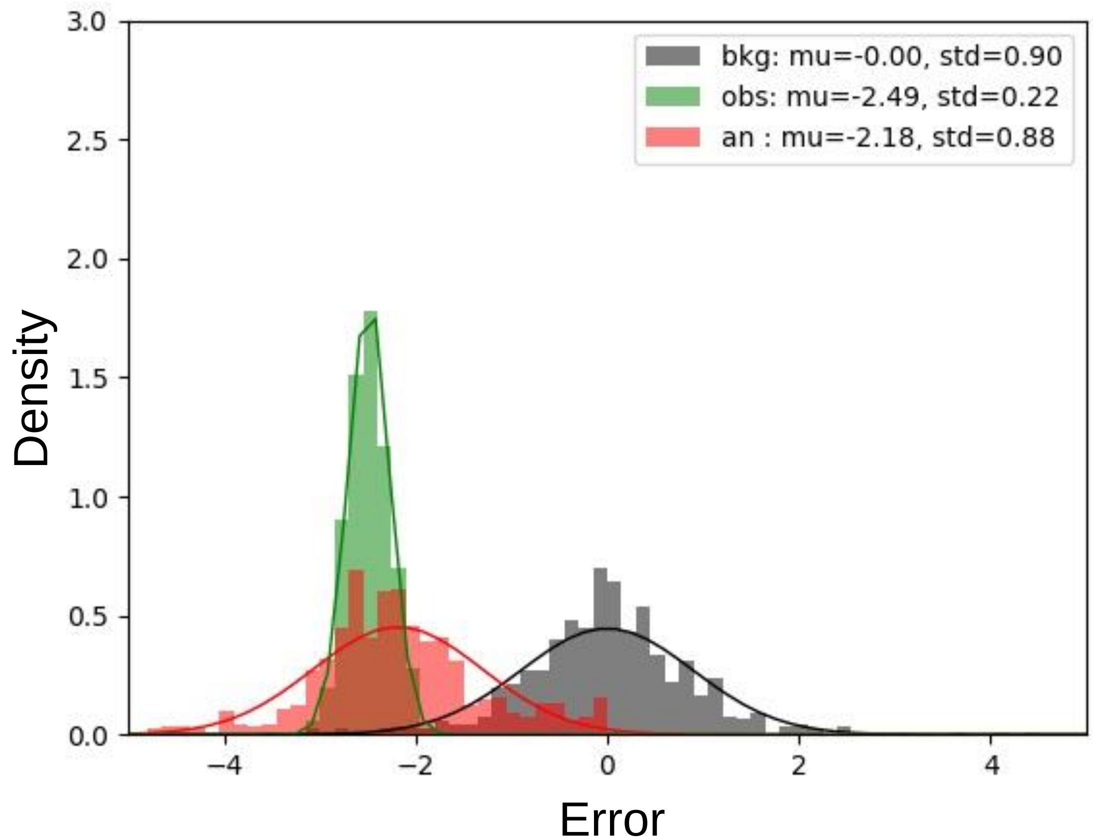
→ Keep improving the model

Data assimilation and biased observations



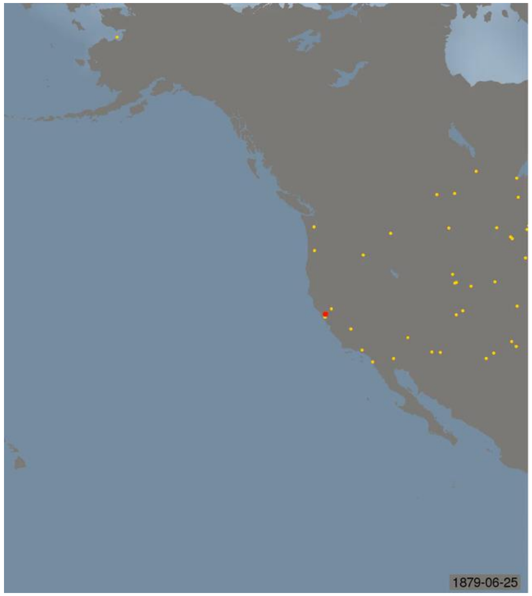
If standard 4D-Var is used to assimilate biased observations (systematic errors), the resulting analysis will be biased.

In this case the background is more accurate than the analysis!



Data assimilation and biased observations

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



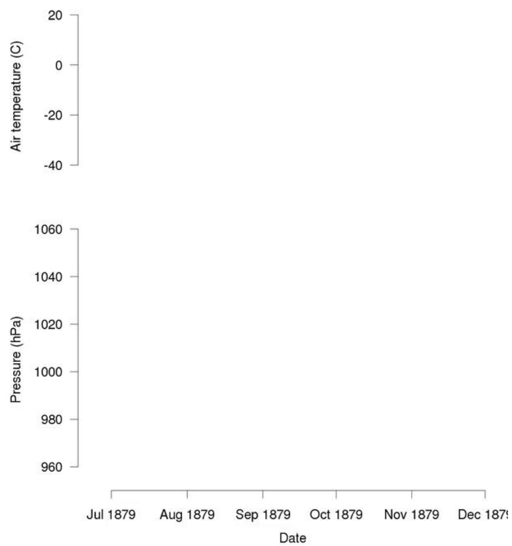
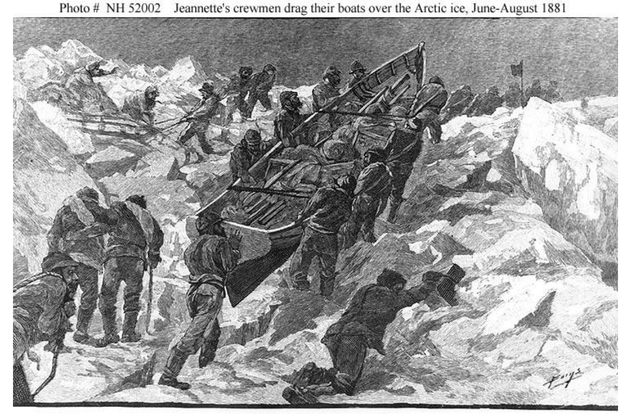


Photo # NH 52000 Steamer Jeannette sinking after being crushed by Arctic ice, June 1881

THE SINKING OF THE JEANNETTE.

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

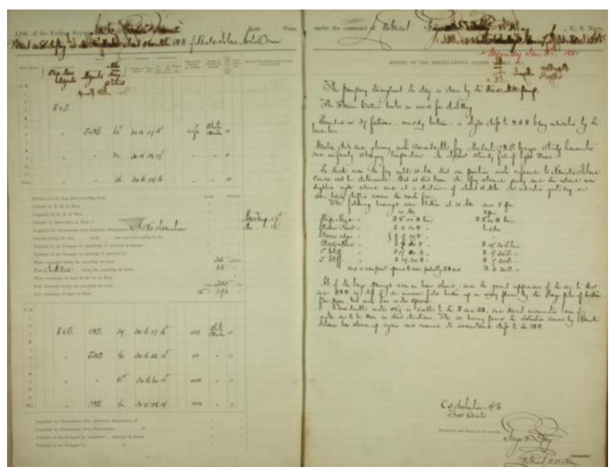


DRAGGING THE BOATS OVER THE ICE

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

WADING ASHORE.

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

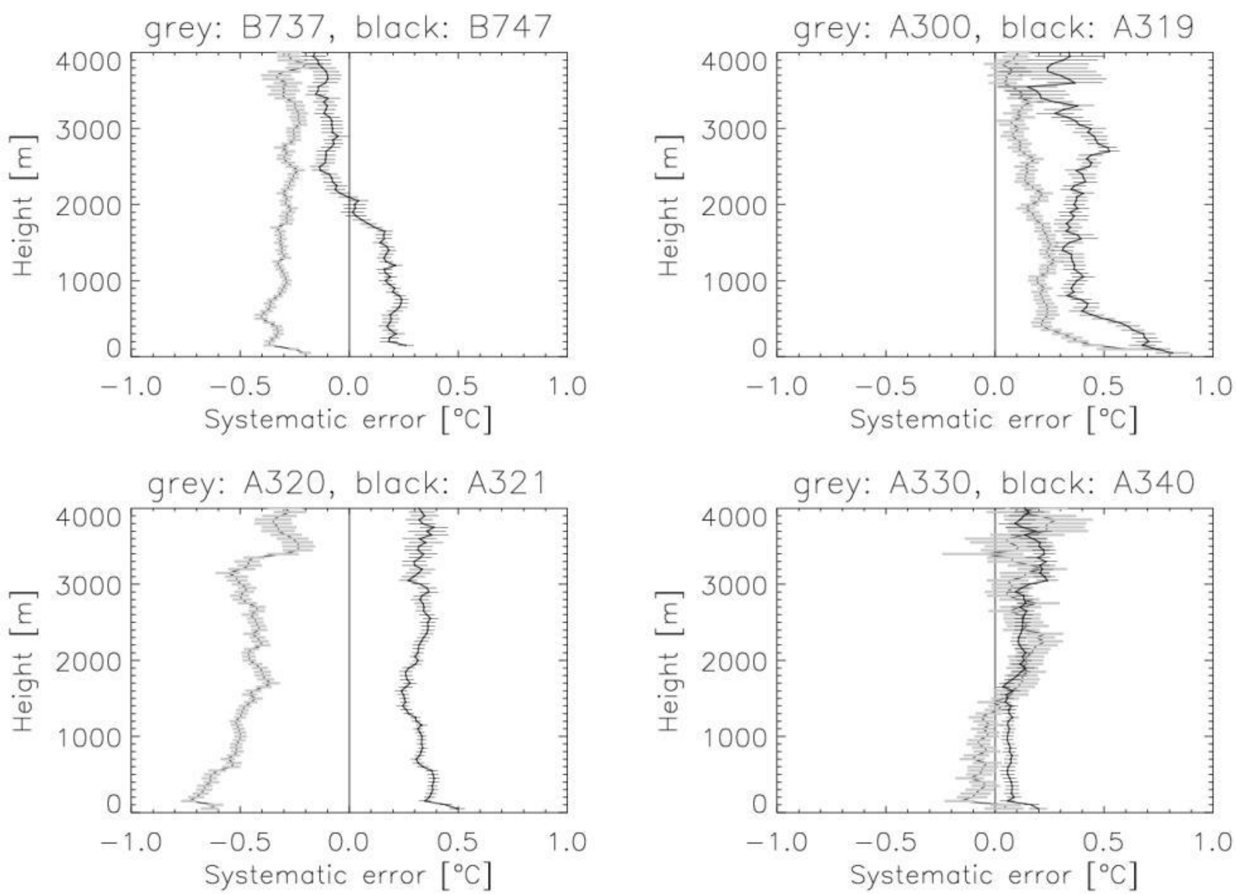


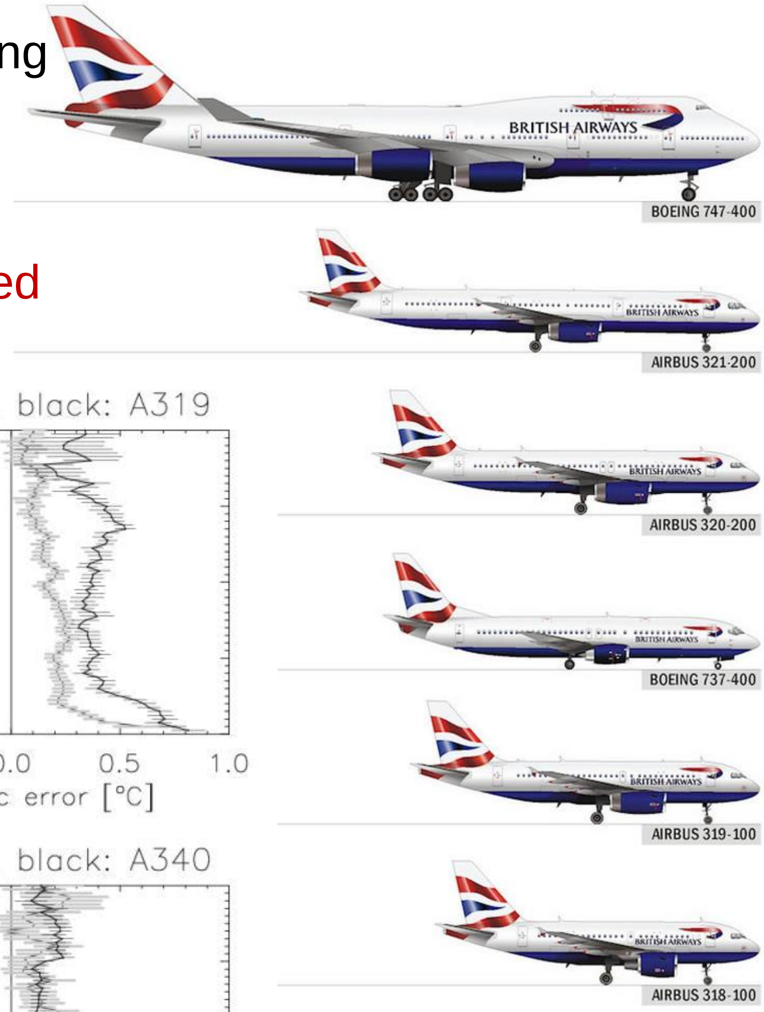
Observation bias is estimated using other (more accurate) instruments

Data assimilation and biased observations

One year of measurements from aircrafts landing at Frankfurt

No external reference, observation bias is estimated using the hourly mean of all measured profiles



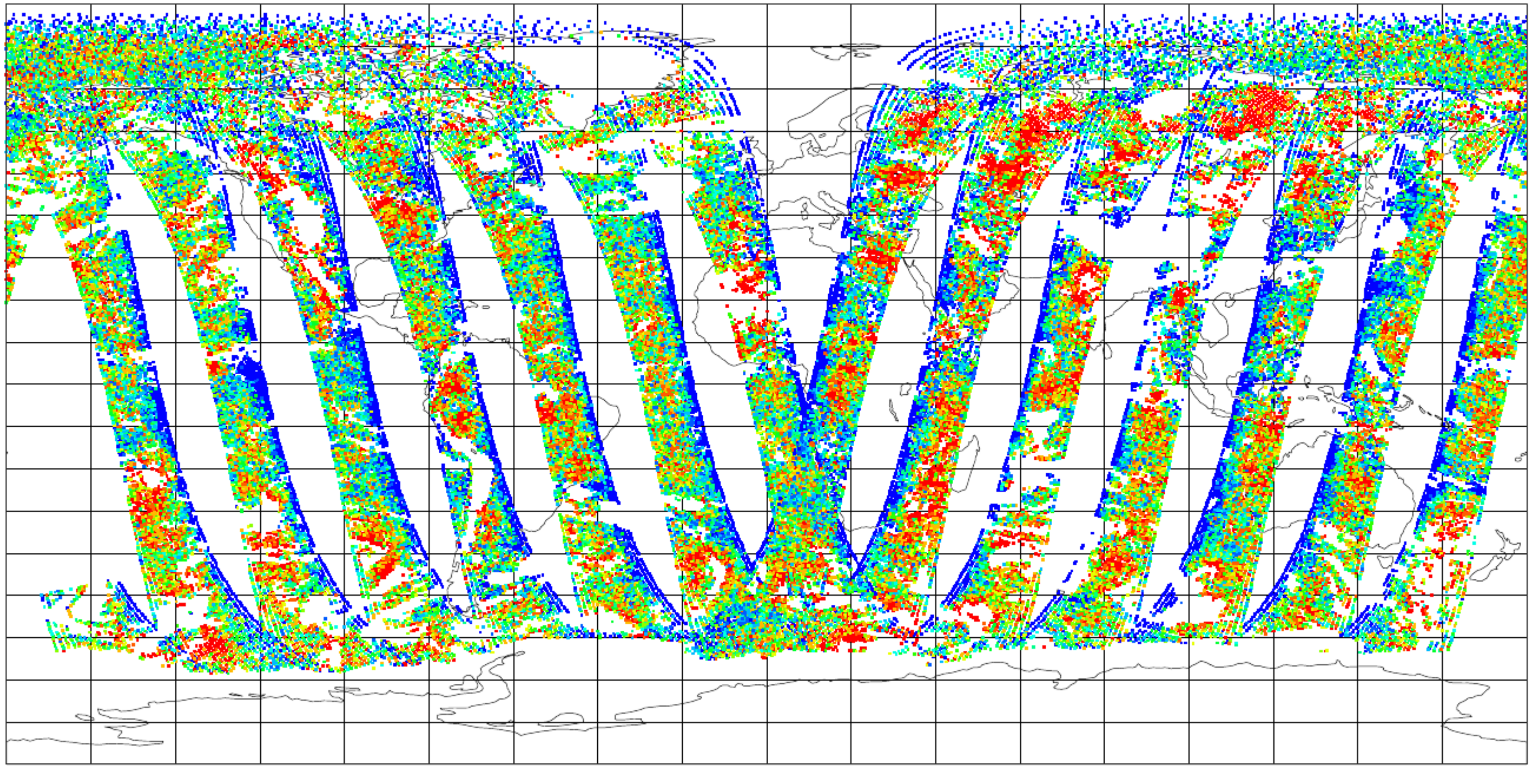


Data assimilation and biased observations

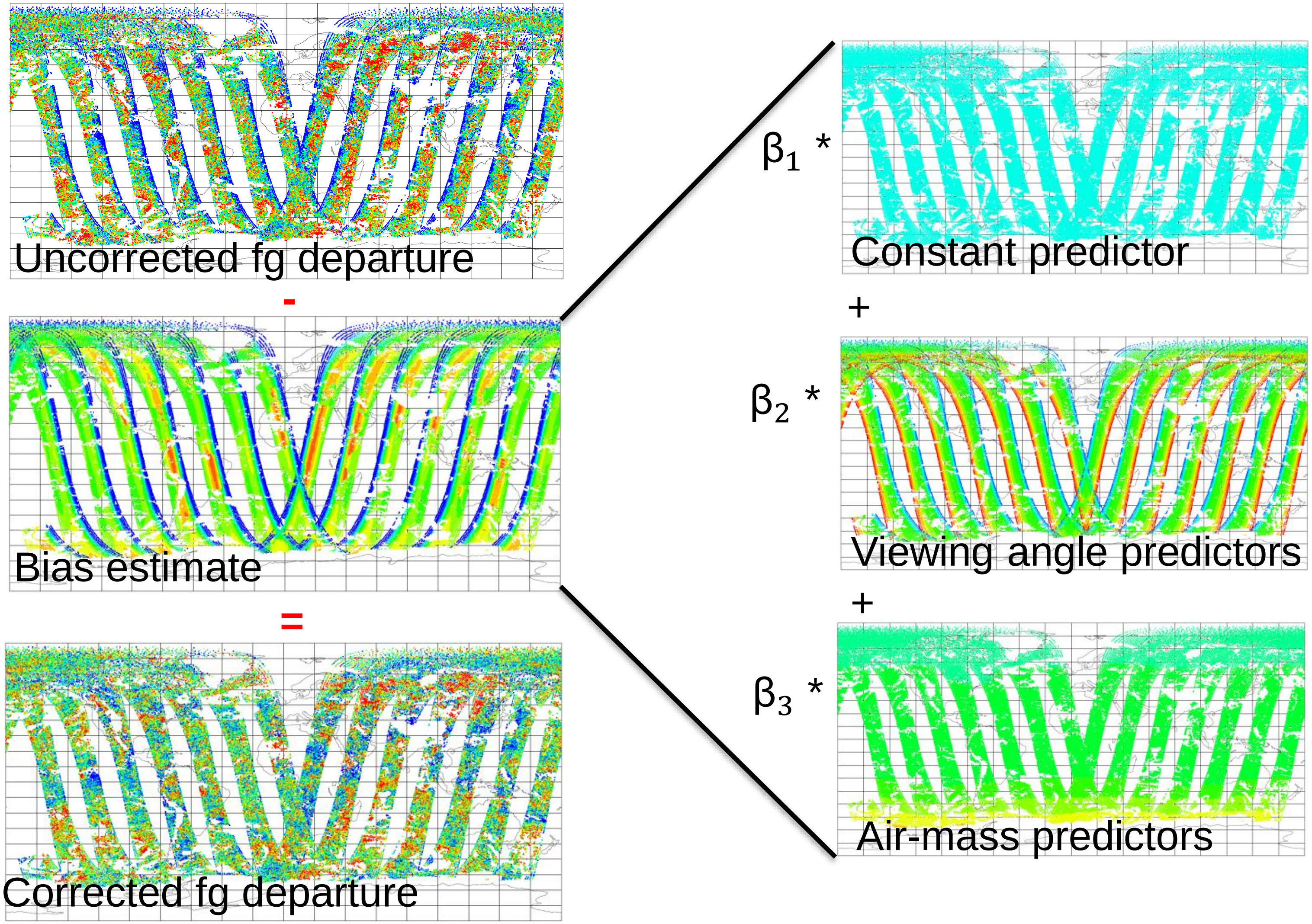
ECMWF is assimilating polar-orbiting Metop-C satellite (launched on 7 November 2018)

Observation bias is estimated inside 4D-Var
➔ comparing measurements with model
➔ specifying the structure of the model bias

Metop-C AMSUA-A Channel 5 (obs-model)



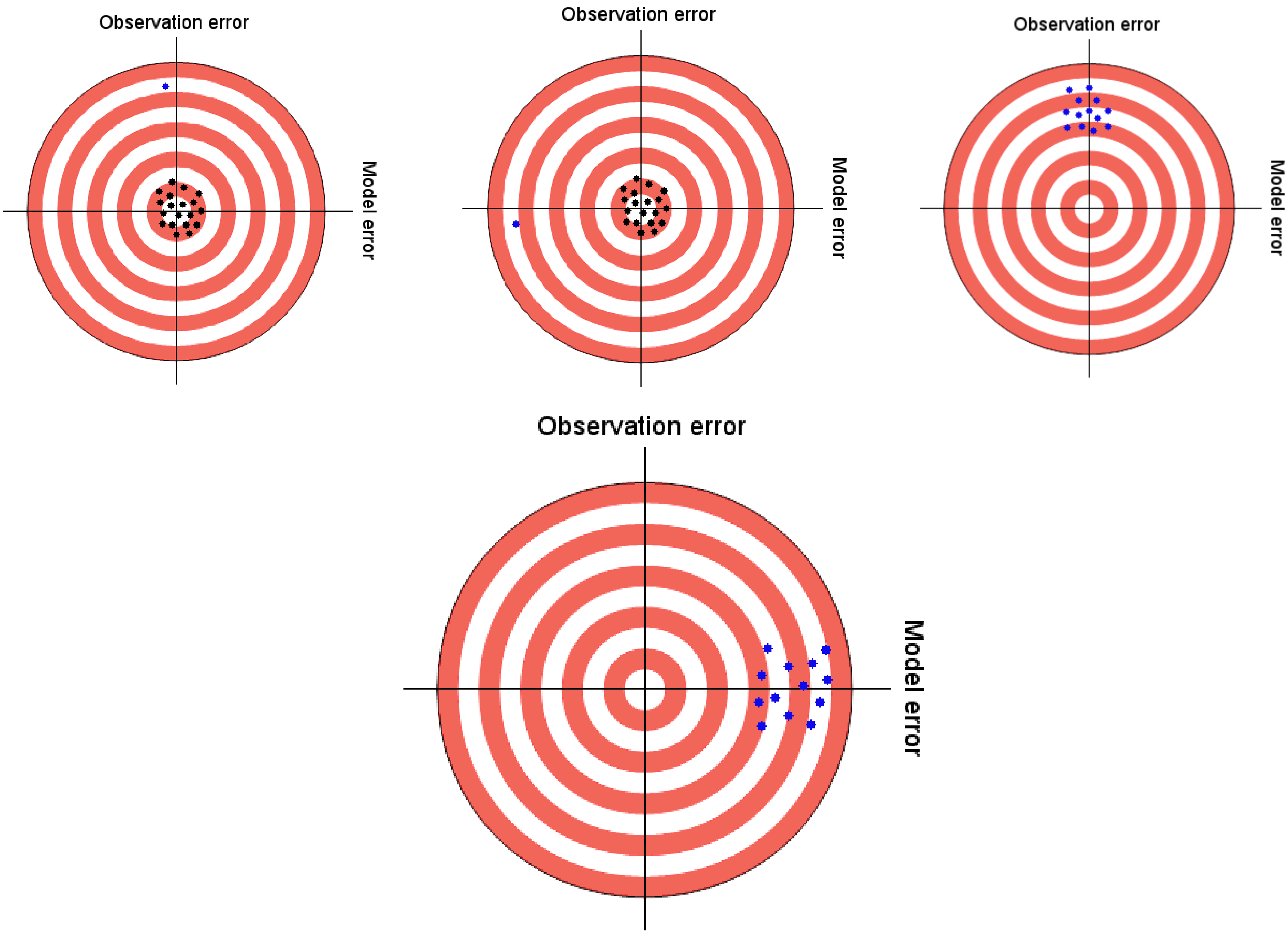
Data assimilation and biased observations



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

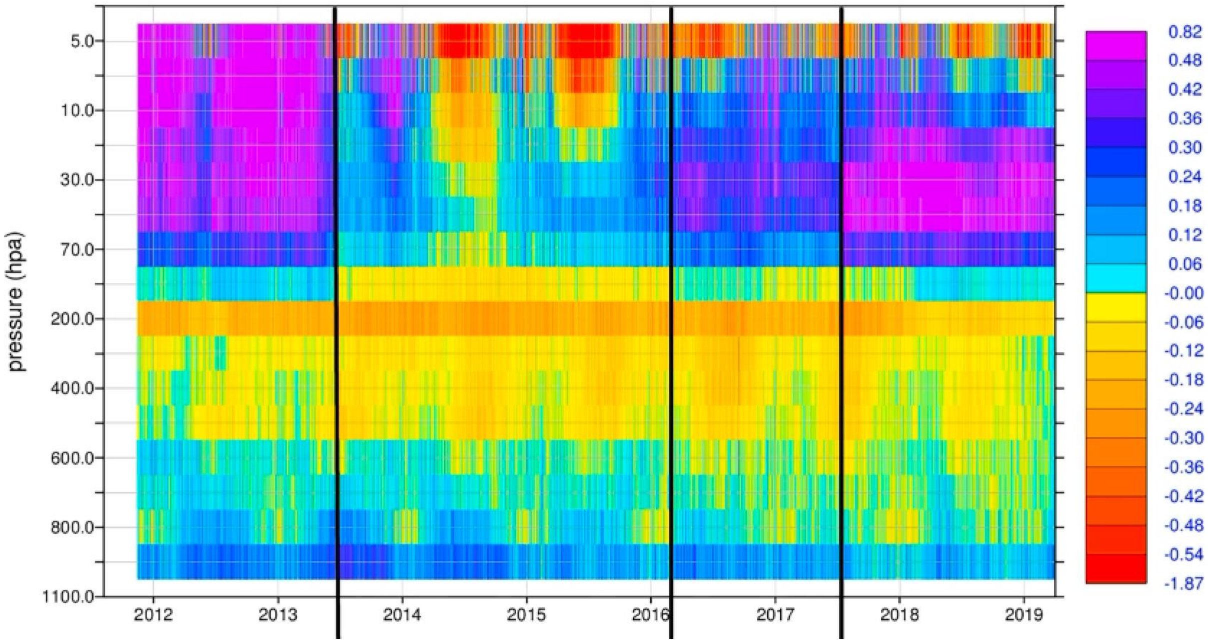
We choose the predictors and 4D-Var estimate the parameters β

Data assimilation and biased model



How to estimate model biases

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



Difference between radiosonde temperature observations and the IFS first-guess trajectory (O-B)

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

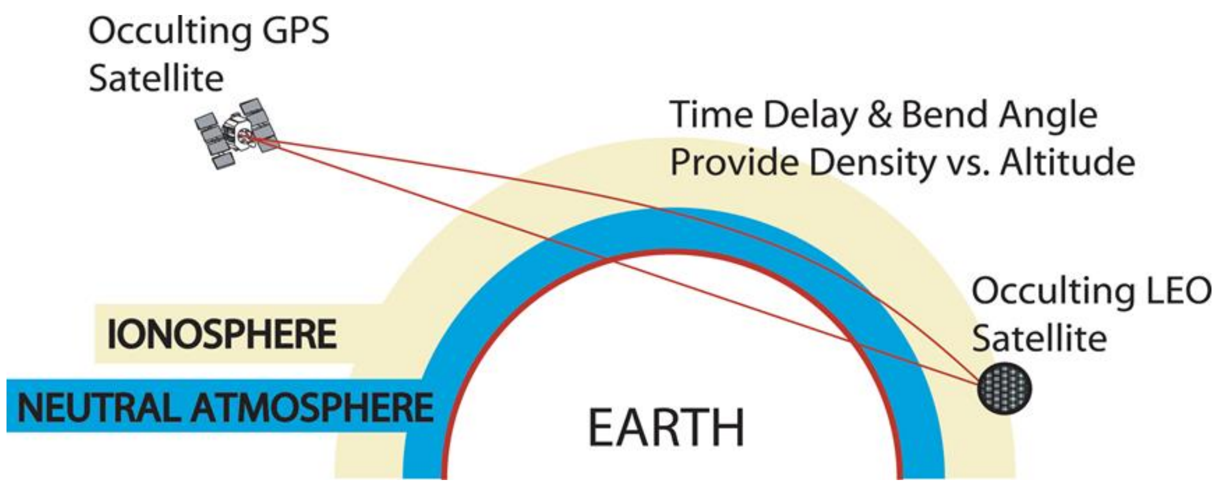
→ Largest bias in the stratosphere

→ 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

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.



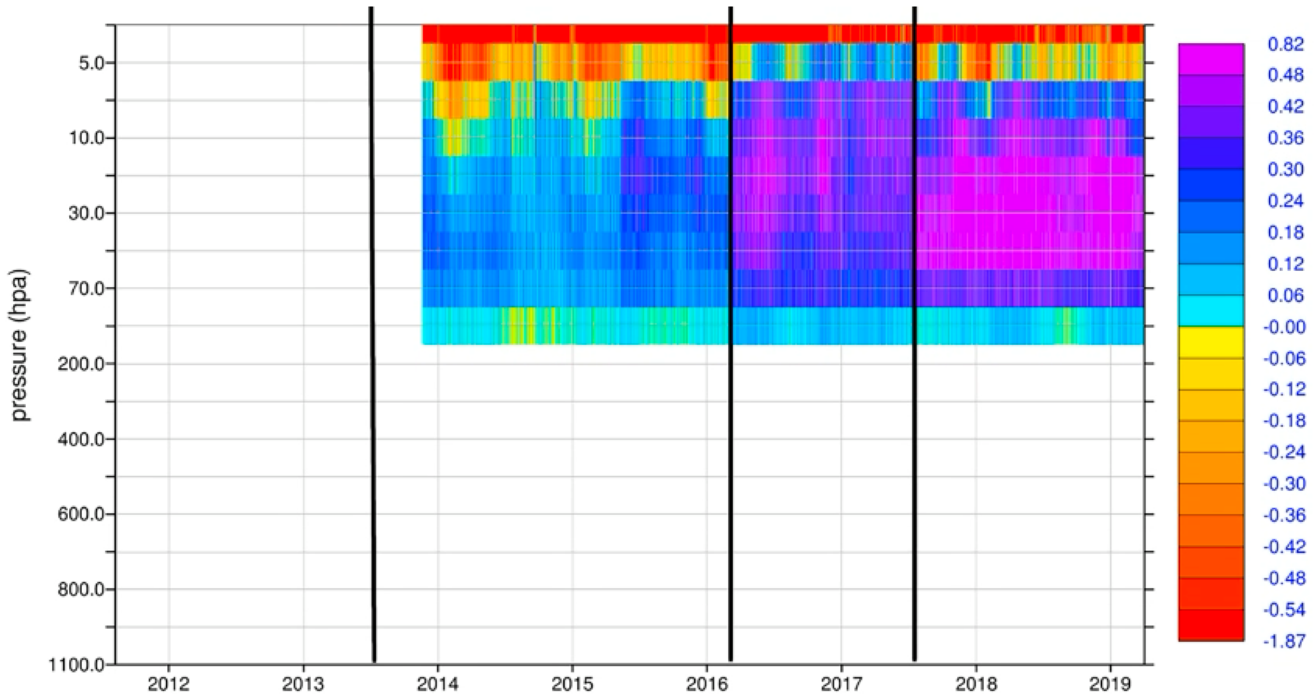
→ As the LEO moves behind the earth, we obtain a profile of bending angles

→ Temperature profiles can then be derived

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

How to estimate model biases

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



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

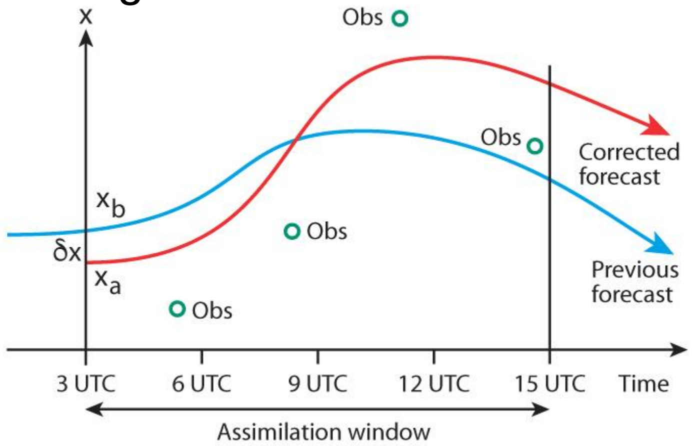
Errors in models are often systematic rather than random, zero-mean

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

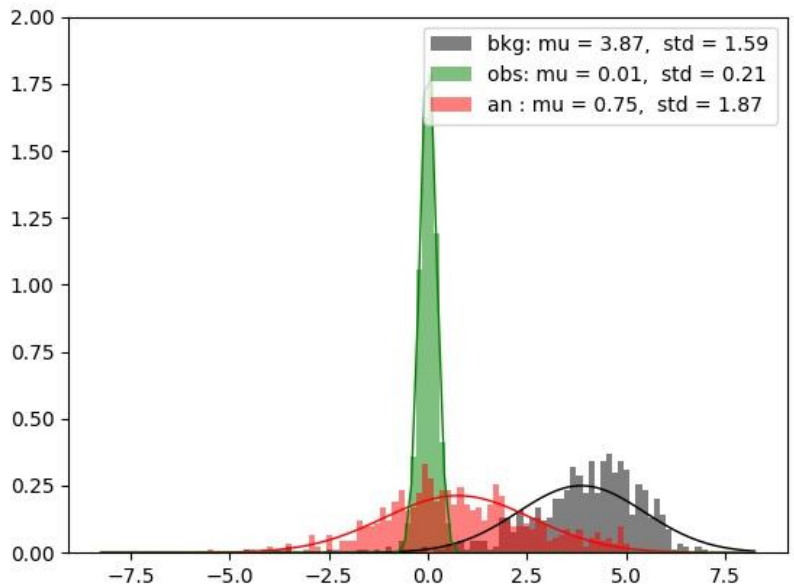
→ Model has a warm bias in the upper stratosphere

How to deal with model biases in data assimilation

Strong constraint 4D-Var

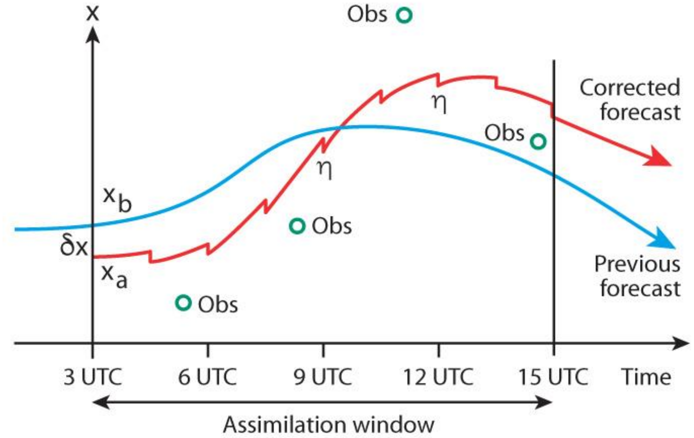


$$x_k = \mathcal{M}_k(x_{k-1})$$

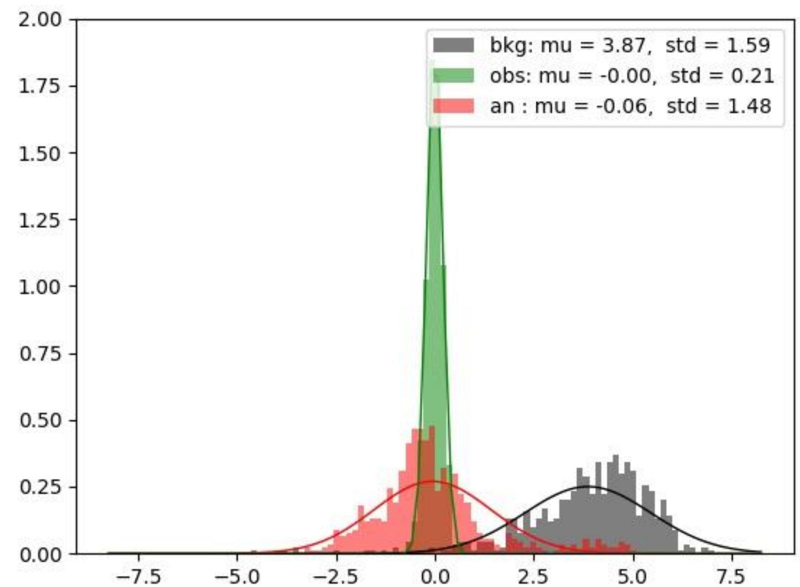


→ Large bias and standard deviation in the analysis

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

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 → β
Model bias → η

→ Introduce additional controls to target an unbiased analysis

→ The model error covariance matrix Q constrains the model error field

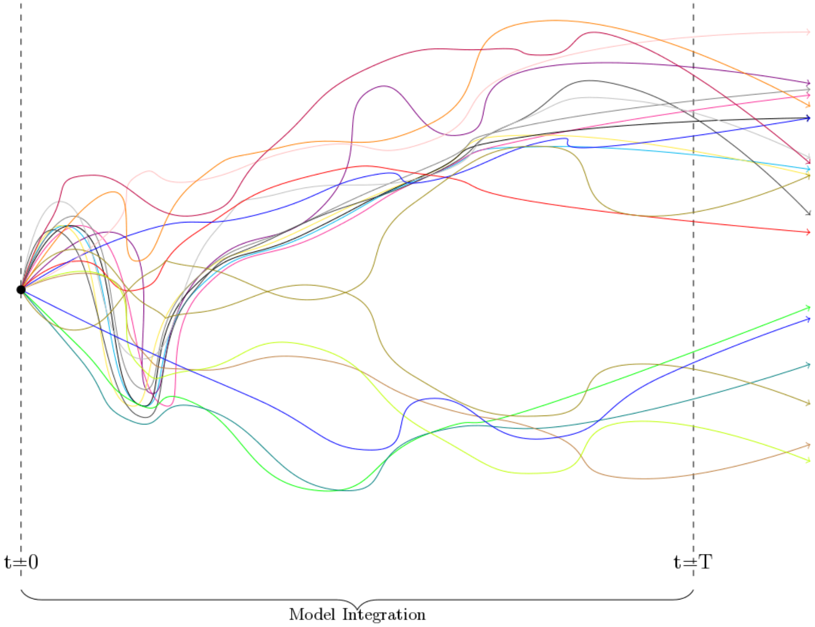
How to estimate the model error covariance matrix (Q)

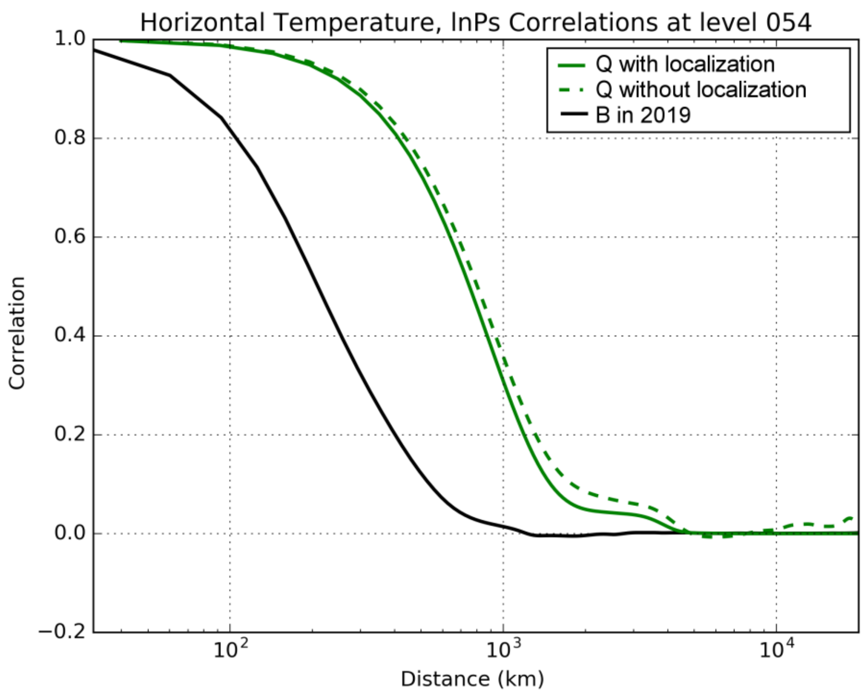
Estimate the model error covariance matrix

➔ run the ensemble forecasting system (ENS) with perturbed physics (51 members with the same initial condition for different days)

➔ differences after 12 hours are used to compute Q

$$Q_f = \frac{1}{N-1} \sum_{i=1}^{N} \left(f_i^{12} - f_{i+1}^{12}\right) \left(f_i^{12} - f_{i+1}^{12}\right)^T$$

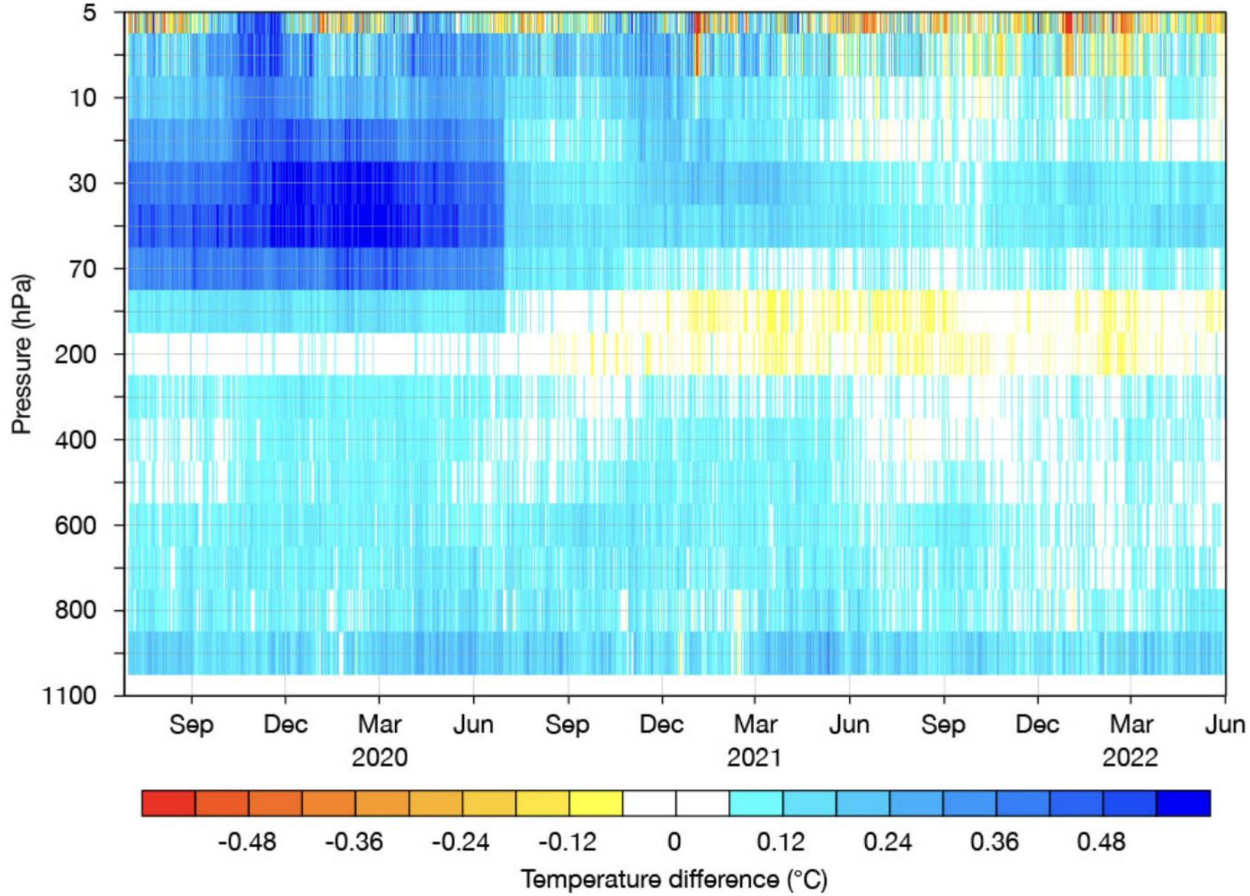




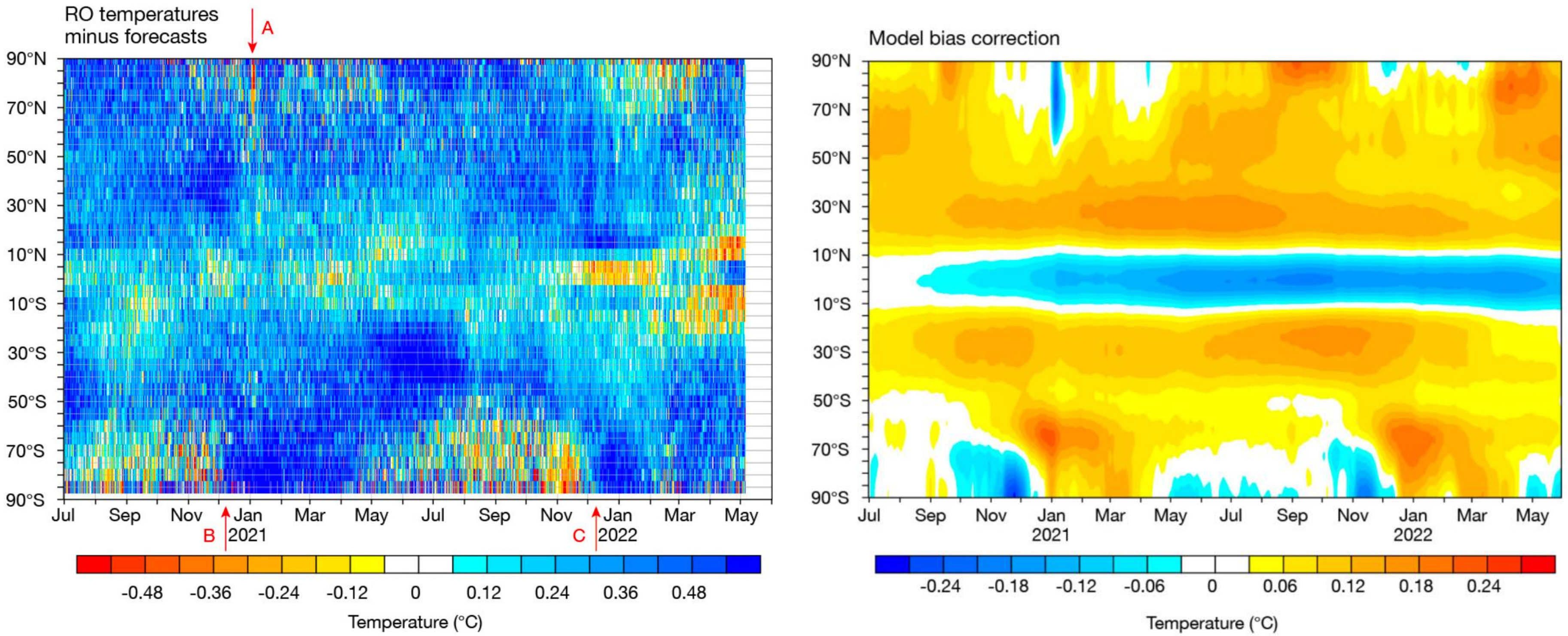
4D-Var corrects small scale errors (background errors) by changing the initial condition and large scale errors (model errors) by changing the model forcing

Weak-constraint 4D-Var in operations for the stratosphere

Time series of the difference between radiosonde temperature observations and model first-guess (47r1 implemented on 30 June 2020)



Weak-constraint 4D-Var in operations for the stratosphere

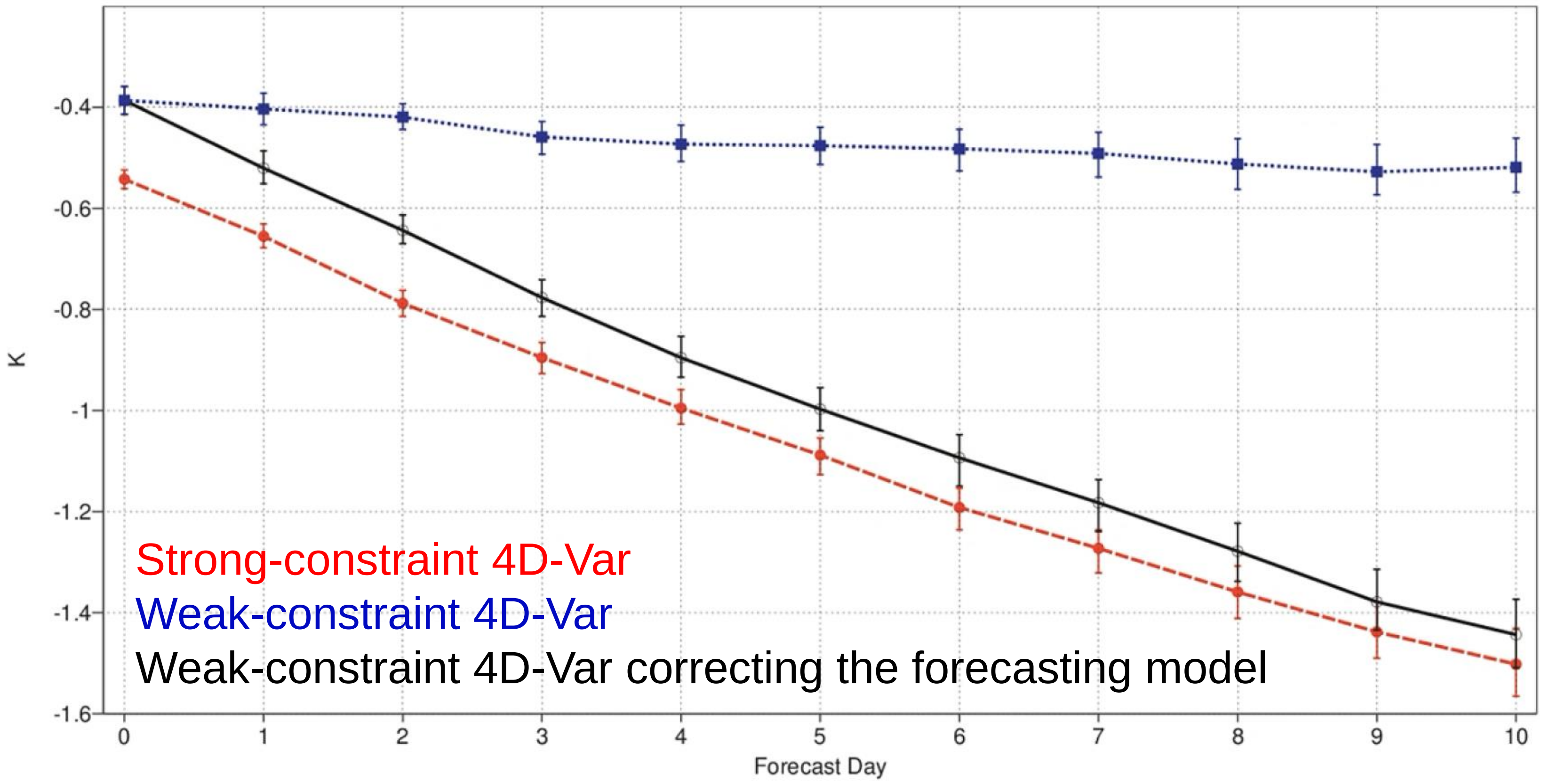


A) On 31 December 2020, a Sudden Stratospheric Warming (SSW) event started over the northern hemisphere

B&C) Clear seasonal cycle in the model bias over the southern hemisphere with a sharp transition in early December 2020 and 2021

Weak-constraint 4D-Var and medium-range forecast

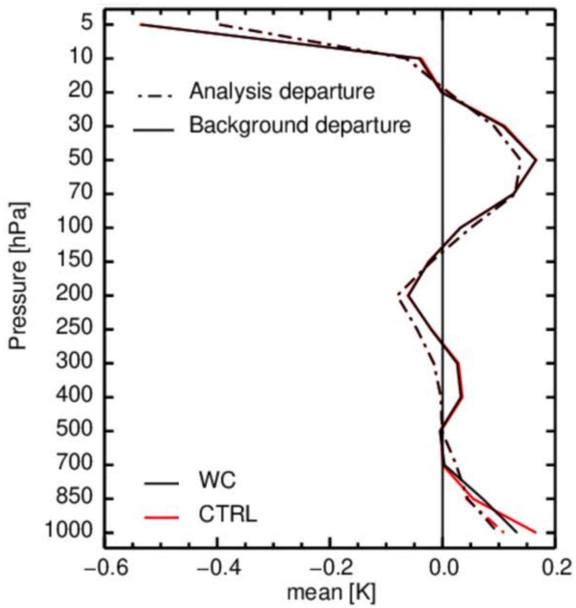
Mean error of the 10-day forecast at 50hPa with respect to the radiosonde observations



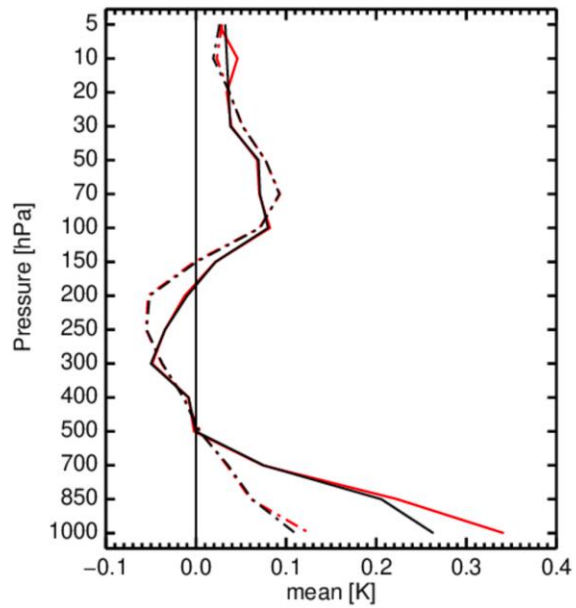
Weak-constraint 4D-Var in the boundary layer

Activate WC4D-Var in the boundary layer (up to ~ 900hPa) to correct temperature

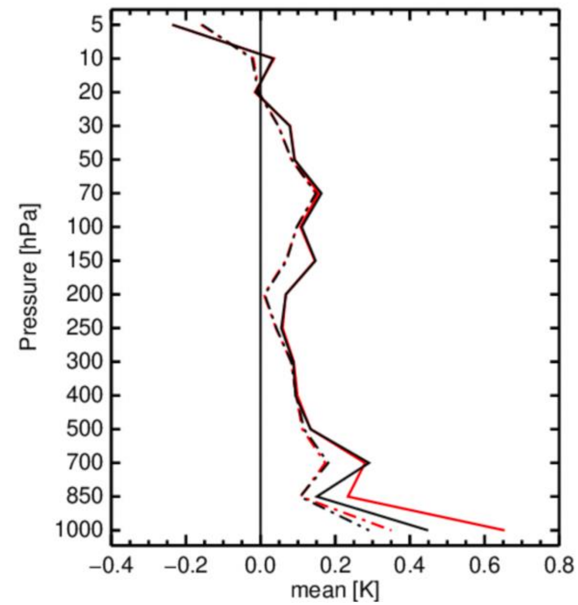
NH

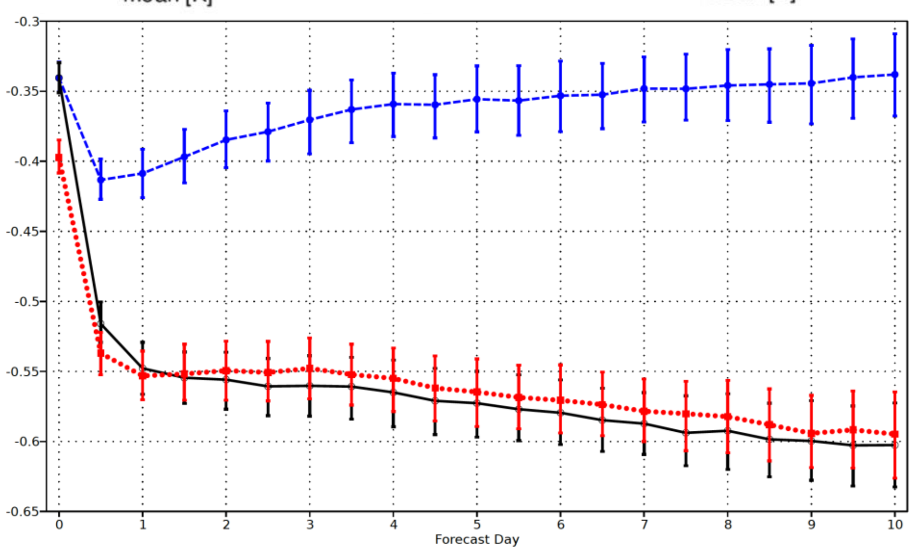


SH



Tropics





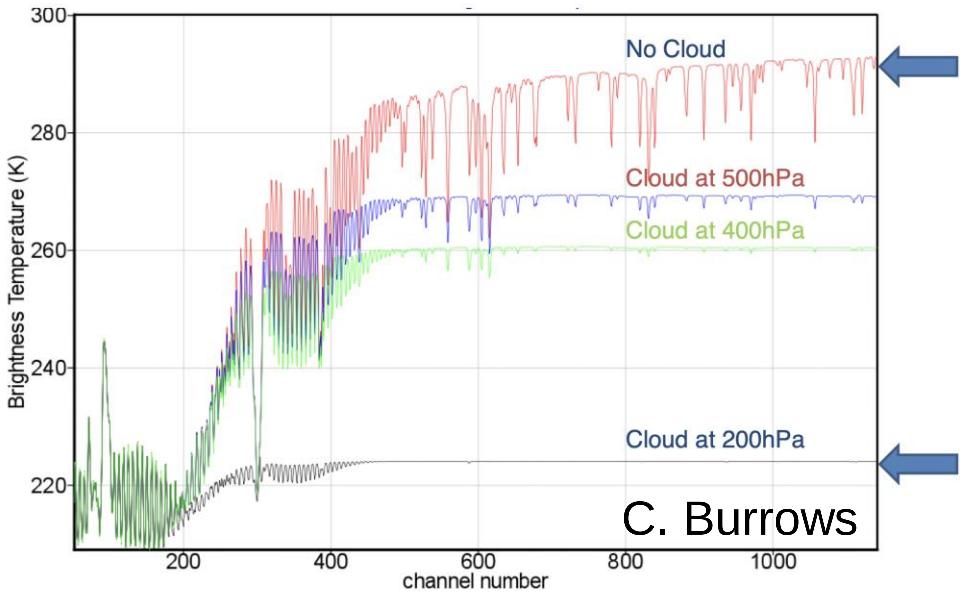
T2m mean error (Tropics)

CTRL

WC4D-Var correcting IC

WC4D-Var correcting IC and forecast model

Positive feedback for satellite infrared assimilation



Channels sensitive to the surface in clear scenes.

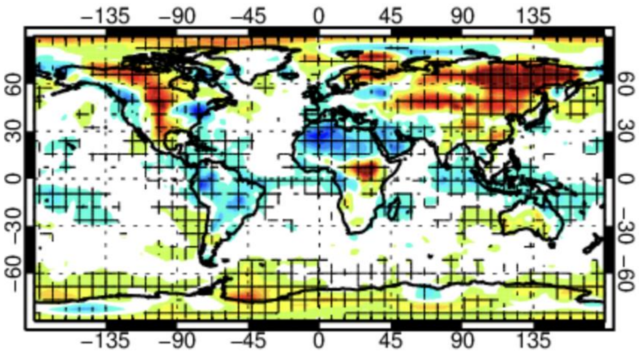
Same channels are strongly sensitive to cold high clouds!

Clouds have a very strong impact on infrared radiance measurements

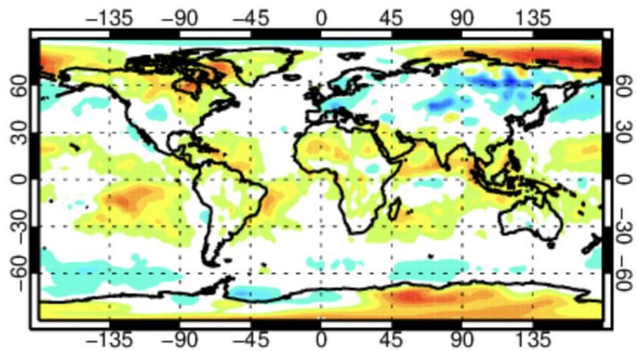
$$\Delta BT_{thresh1} < \left(y_{obs} - H(x_{clear})\right) < \Delta BT_{thresh2}$$

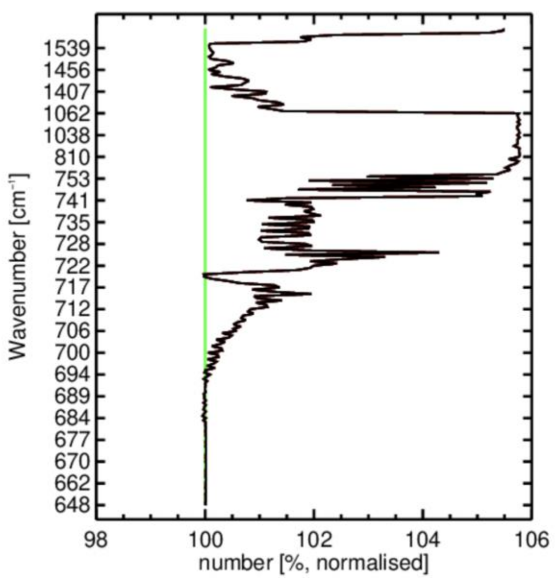
RT operator

Temperature model bias at 1000hPA

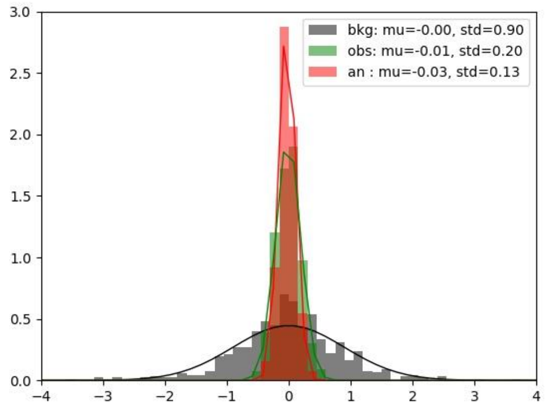


Model bias correction at 1000hPA

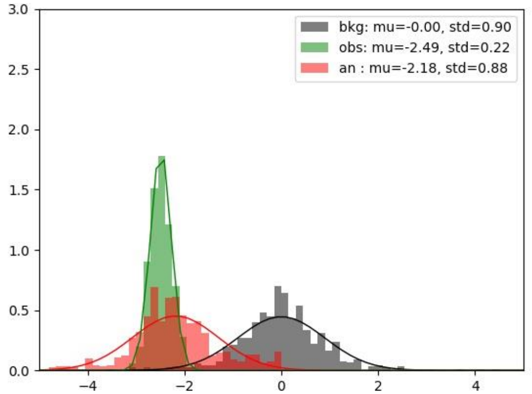




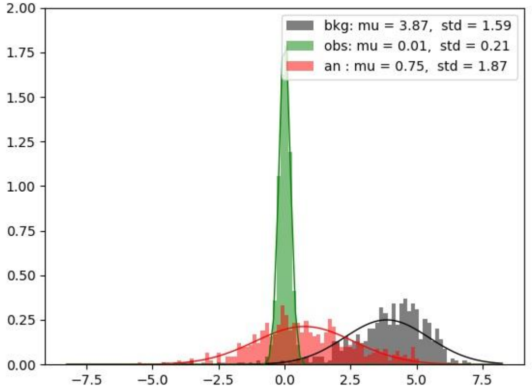
Summary 1/3



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 2/3

How do I know if my observations are biased?
How do I know if my model is biased?
You don't know the truth, but you have to trust something

Reference observations are used

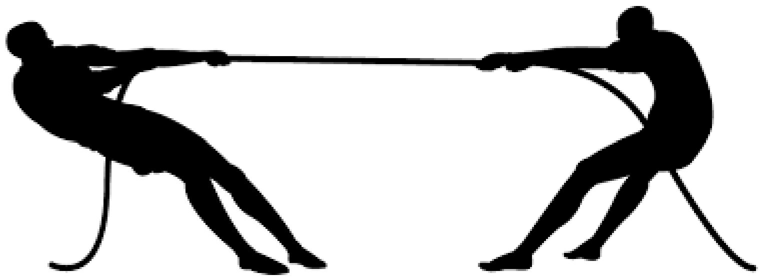
Radiosondes

GPS-RO

Summary 3/3

From bias-blind to bias-aware data assimilation

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



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