Biases in observations

Patrick Laloyaux

Many thanks to Niels Bormann, Hans Hersbach and Dick Dee

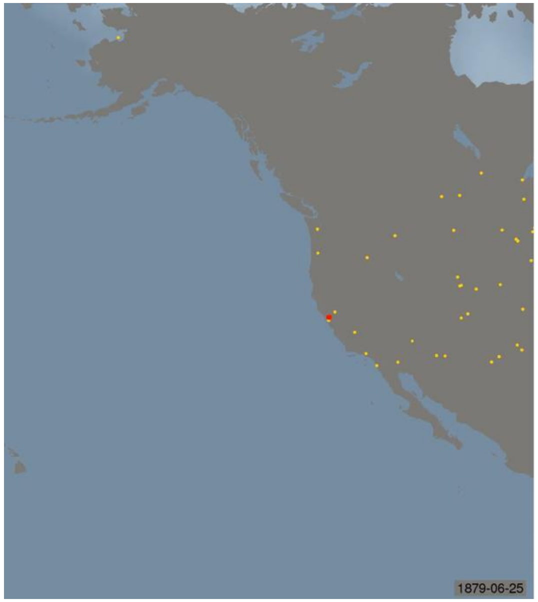
To illustrate biases in observations

To construct bias models for specific instruments

To understand the challenges of observation bias correction

Examples of biases in observations (1/4)

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



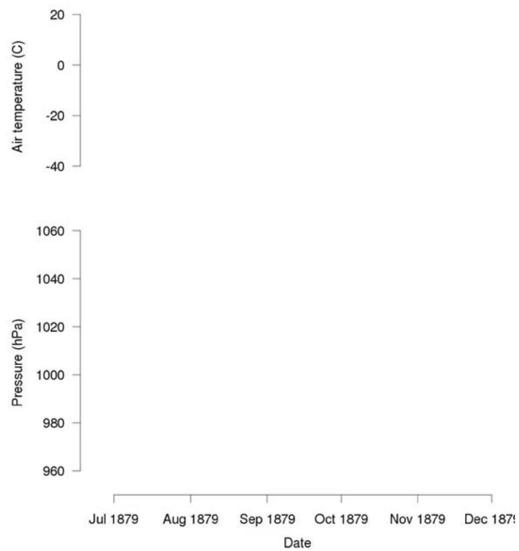
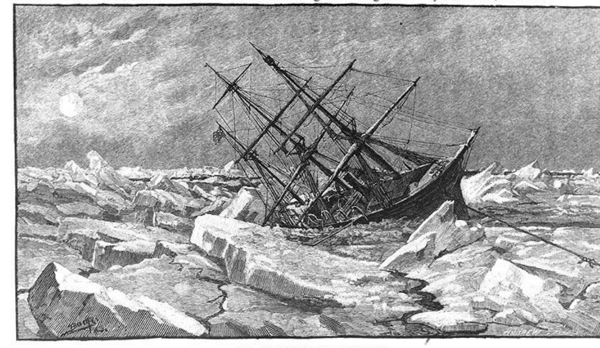


Photo # NH 52000 Steamer Jeannette sinking after being crushed by Arctic ice, June 1881



THE SINKING OF THE JEANNETTE.

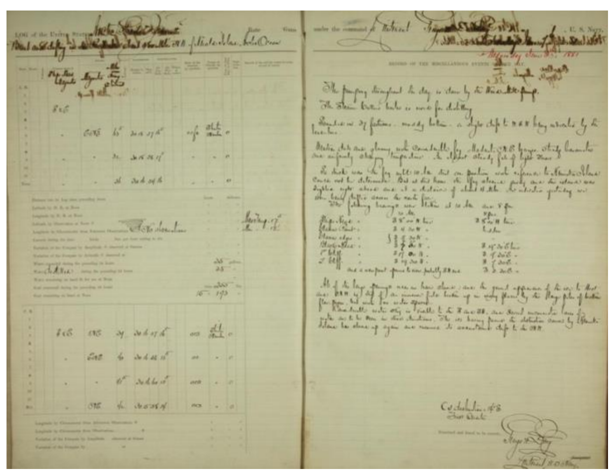
Photo # NH 52002 Jeannette's crewmen drag their boats over the Arctic ice, June-August 1881

DRAGGING THE BOATS OVER THE ICE.

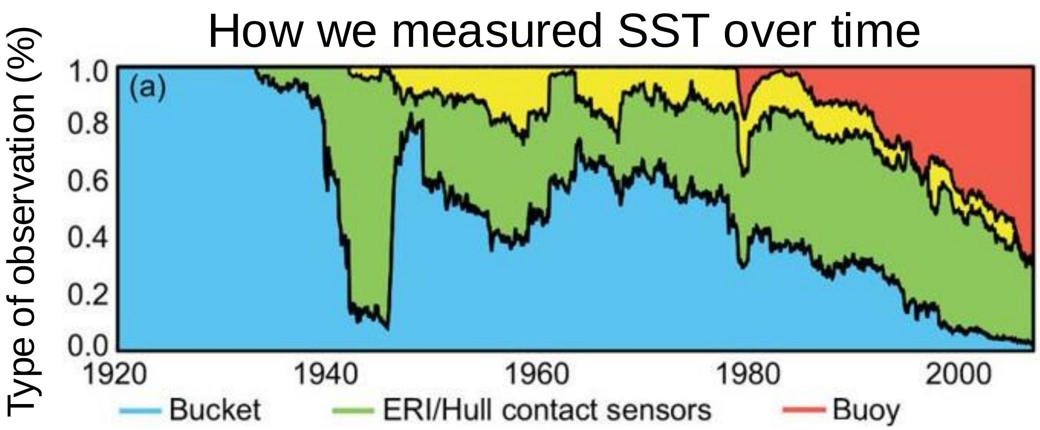
Photo # NH 92142 LCdr. DeLong and his party wading ashore on the Lena Delta, Siberia, 17 Sept. 1881

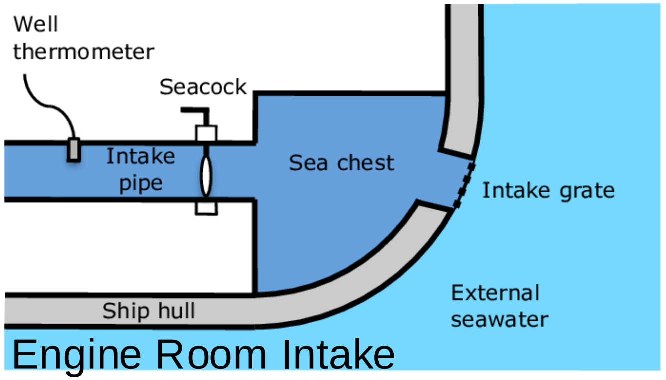
WADING ASHORE.

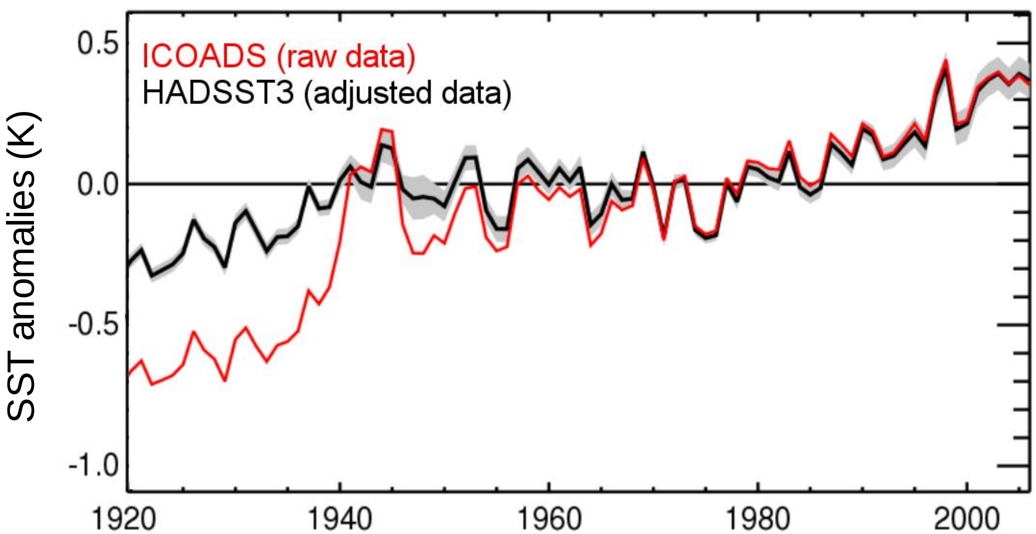
SST measurements from standard buckets have a cold bias (~0.4C)



Examples of biases in observations (2/4)







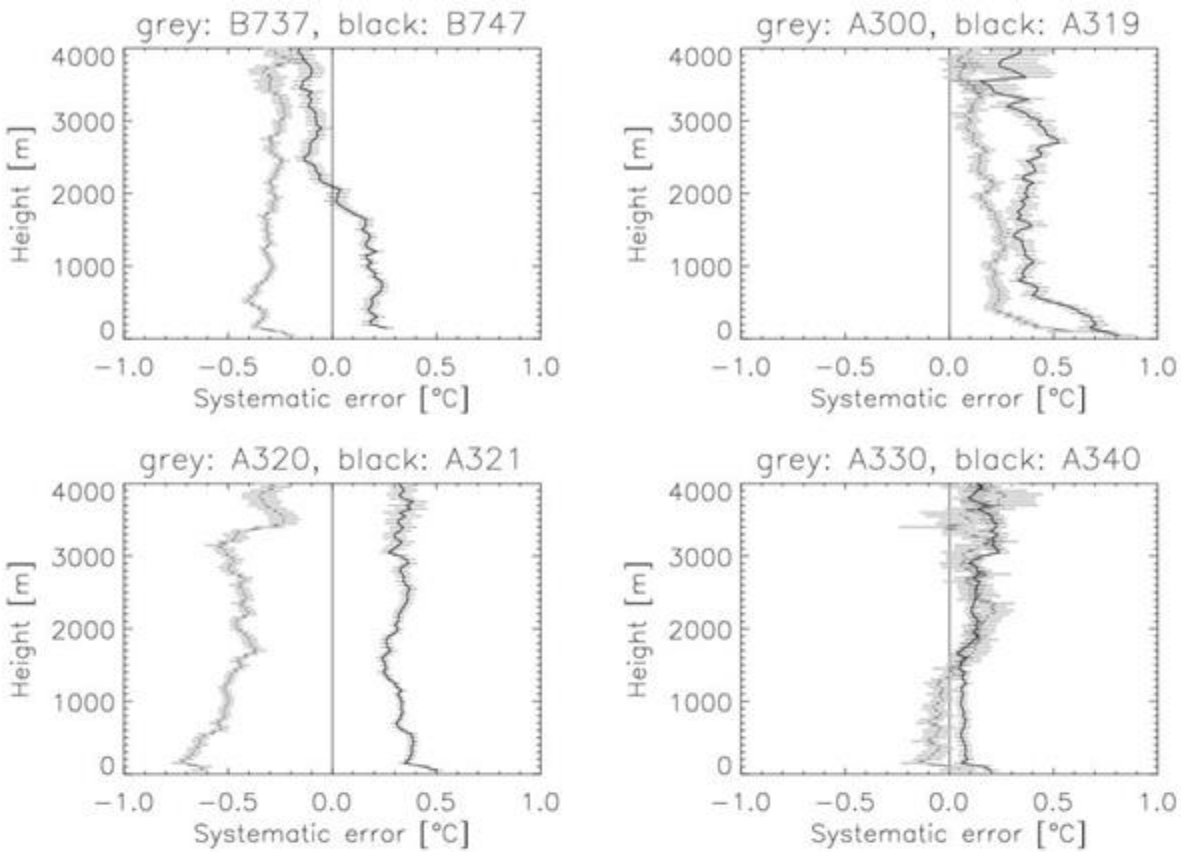
Estimation of observation biases done by inter-comparison between instruments

➔ Involve experts knowing the instruments

➔ Not straightforward as incomplete metadata

Examples of biases in observations (3/4)

One year of measurements from aircrafts landing at Frankfurt



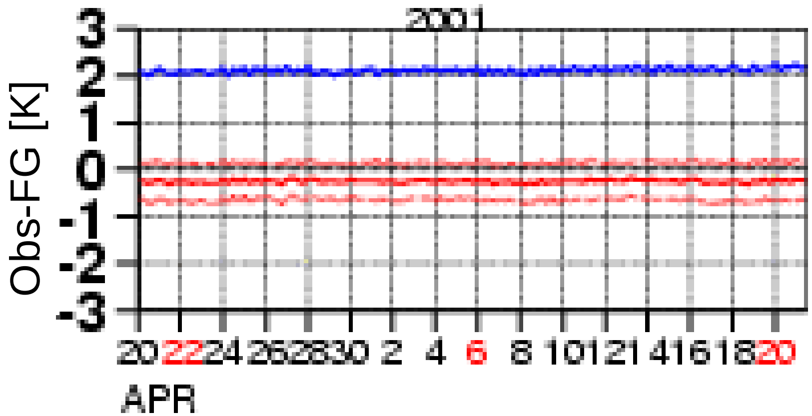
Estimation of observation biases done by inter-comparison between instruments

➔ Involve experts knowing the instruments

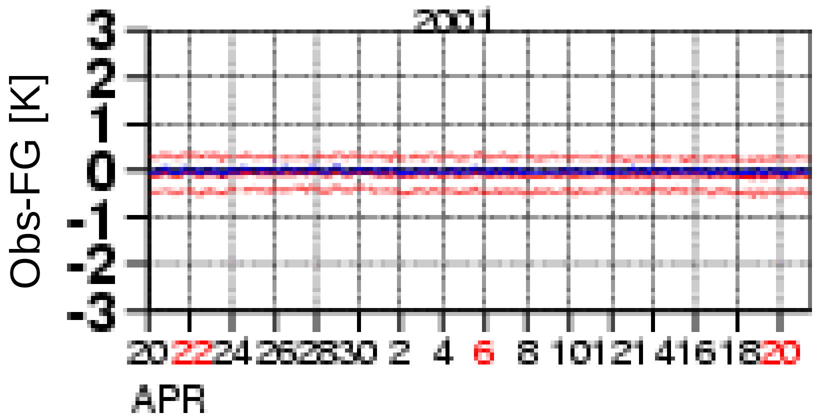
➔ observation bias is estimated using the hourly mean of all measured profiles

Examples of biases in observations (4/4)

High-Resolution Infrared Radiation Sounder (HIRS) measures temperature profiles, moisture content, cloud height and surface albedo. Channel 5 peaks around 600hPa



HIRS channel 5 on NOAA-14 satellite has +2.0K radiance bias against FG (blue line)



HIRS channel 5 on NOAA-16 satellite has no significant bias against FG (blue line)

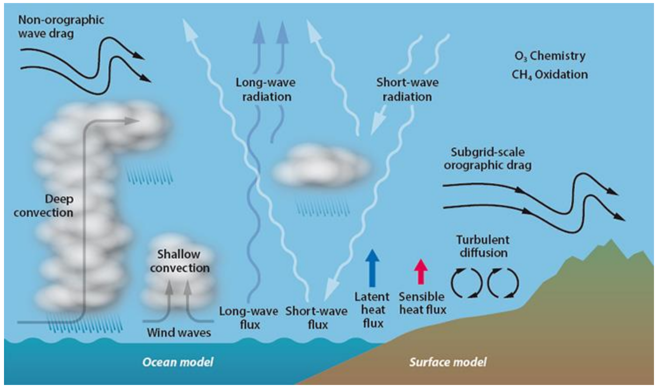
Estimation of observation biases done by inter-comparison between instruments

➔ Involve experts knowing the instruments

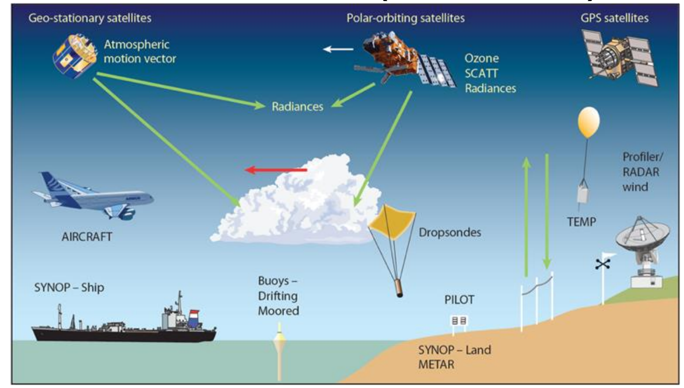
➔ observation bias is estimated comparing obs with the model (time/space average)

What you have seen so far on data assimilation

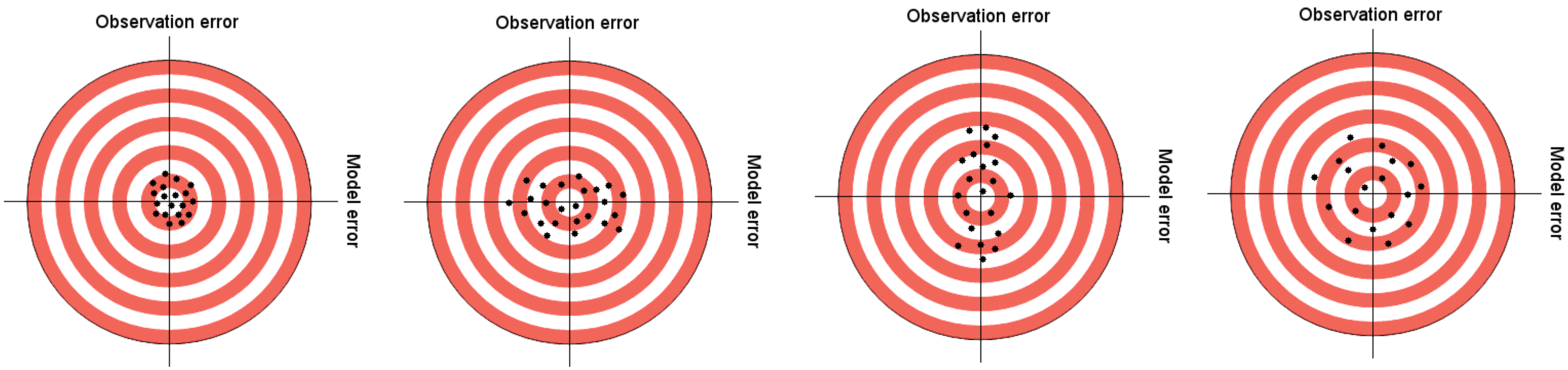
Model (with errors)



Observations (with errors)



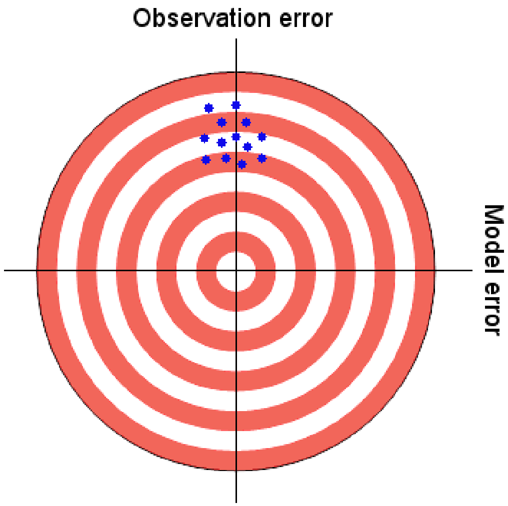
If you are lucky, model and observations are **accurate** (no biases, mean error is zero)



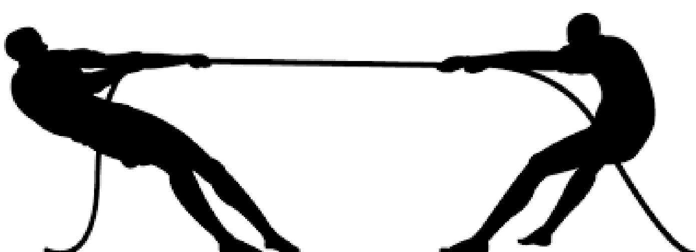
$$
\begin{aligned}
J(x_0) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\
&+ \frac{1}{2}\sum_{k=0}^{K}[y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)]
\end{aligned}
$$

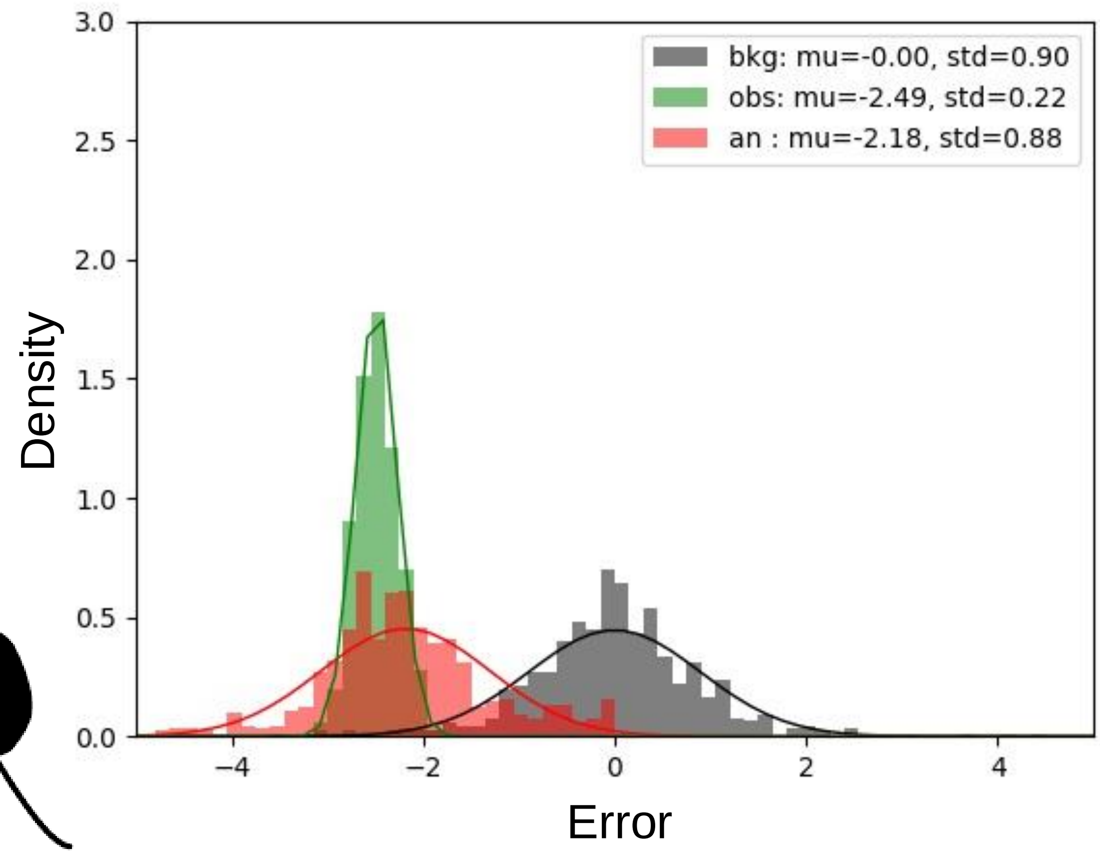
Most of the time, we are unlucky!

Observation biases matter



- 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!





Changing the 4D-Var formulation (introducing VarBC)

$$
\begin{aligned}
J(x_0, \beta) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\
&+ \frac{1}{2}(\beta - \beta_b)^T \mathbf{B}_\beta{}^{-1}(\beta - \beta_b) \\
&+ \frac{1}{2}\sum_{k=0}^{\text{Radiosonde}} [y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \\
&+ \frac{1}{2}\sum_{k=0}^{\text{NOAA-16}} [y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \\
&+ \frac{1}{2}\sum_{k=0}^{\text{NOAA-14}} [y_k - \beta - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \beta - \mathcal{H}(x_k)]
\end{aligned}
$$

Model state

Observation bias parameters

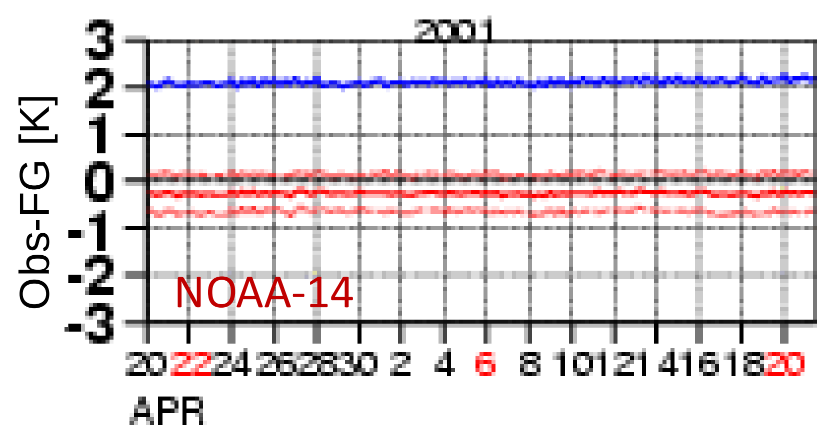
Unbiased observations (anchor)

Biased observations

Bias model

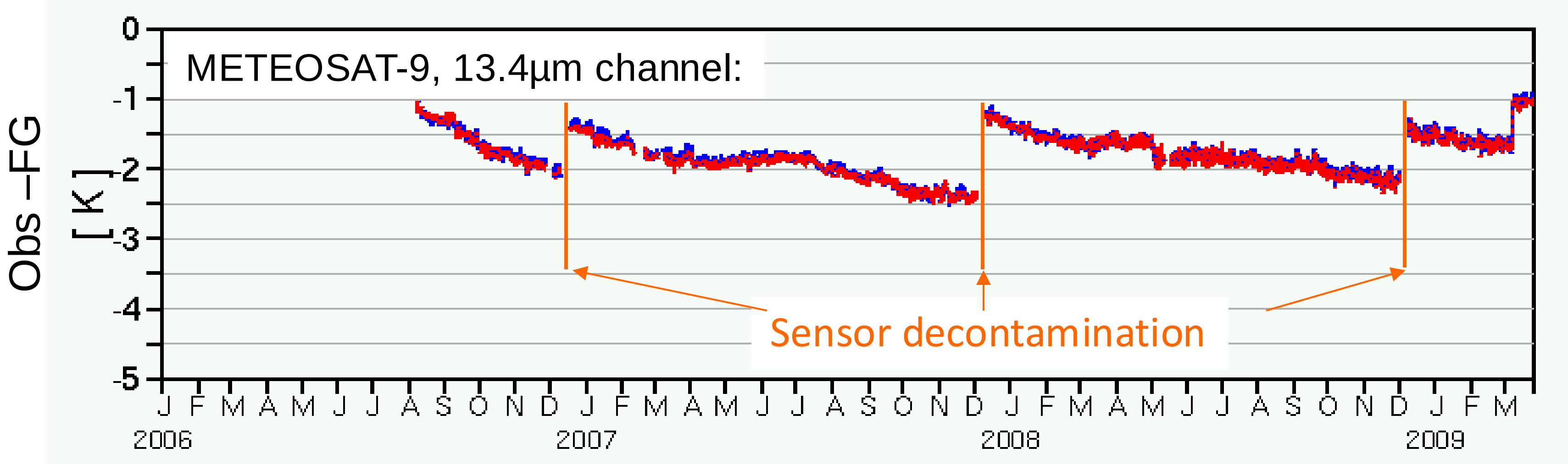
Variational Bias Correction (VarBC)

- We choose which observations we want to correct and which observations we trust
- We choose the bias model b(β) = β
- 4D-Var minimization estimates the value of the VarBC parameters



Changing the 4D-Var formulation (introducing VarBC)

Drift in bias due to ice building up on sensor



VarBC needs to correct for observation bias changing over time

- Bias model = b(β) = β
- β is evolving over time depending how much ice is building up

Changing the 4D-Var formulation (introducing VarBC)

$$
\begin{aligned}
J(x_0, \beta) \;=\;& \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\
+\;& \frac{1}{2}(\beta - \beta_b)^T \mathbf{B}_\beta^{-1}(\beta - \beta_b) \\
+\;& \frac{1}{2}\sum_{k=0}^{\text{Radiosonde}} [y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \\
+\;& \frac{1}{2}\sum_{k=0}^{\text{NOAA-16}} [y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \\
+\;& \frac{1}{2}\sum_{k=0}^{\text{NOAA-14}} [y_k - \beta - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \beta - \mathcal{H}(x_k)]
\end{aligned}
$$

Model state

Observation bias parameters

Parameter estimates from previous analysis

Background covariance matrix for VarBC parameters

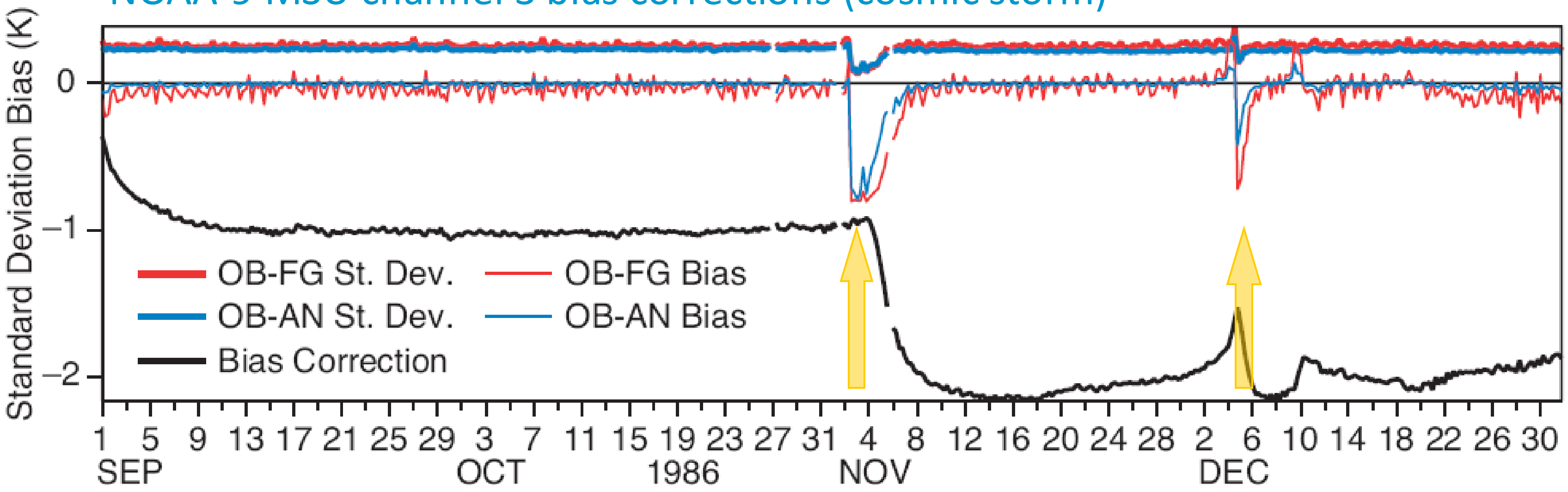
Variational Bias Correction (VarBC)

- A cycling scheme for updating the bias parameter estimates
- Specification of the background covariance matrix $\mathbf{B}_\beta$ (large value ➔ fast adaptation, small value ➔ slow adaptation)

$$
\mathbf{B}_\beta = \begin{bmatrix} \mathbf{B}_\beta^{(1)} & & 0 \\ & \ddots & \\ 0 & & \mathbf{B}_\beta^{(J)} \end{bmatrix}
$$

Changing the 4D-Var formulation (introducing VarBC)

NOAA-9 MSU channel 3 bias corrections (cosmic storm)

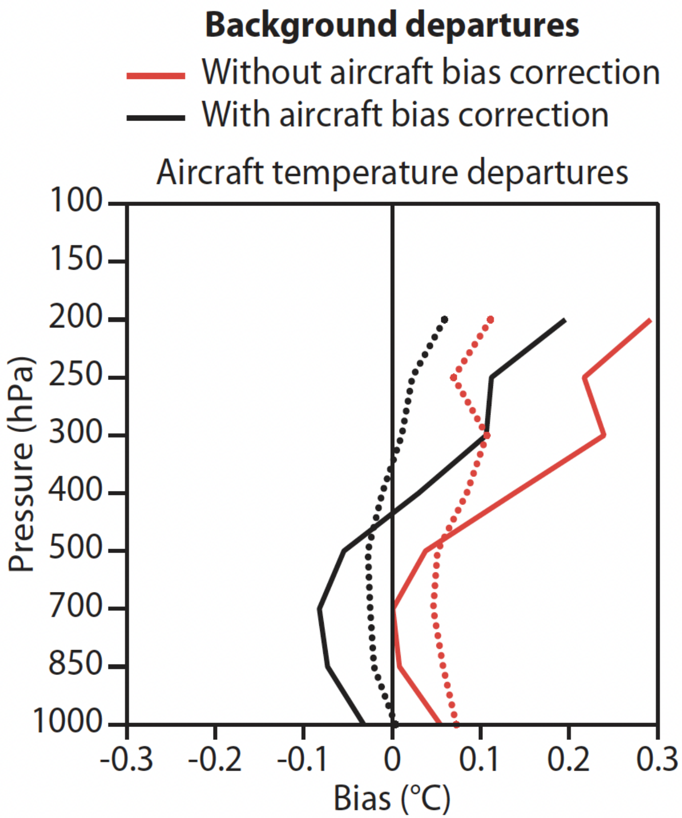


Two cosmic storms trigger large observation biases, but the whole 4D-Var system handles this automatically (thanks to VarBC and its online learning)

1. Initially QC rejects most data from this channel
2. VarBC adjusts the bias estimates
3. Bias-corrected data are gradually assimilated again

No shocks to the system!

Building models of observation biases (aircraft)



For each aircraft separately (~5000 distinct aircraft)

Bias model = $b(\beta) = b(\beta_0, \beta_1, \beta_2) = \beta_0 + \beta_1 *$ ascent rate $+ \beta_2 *$ descent rate

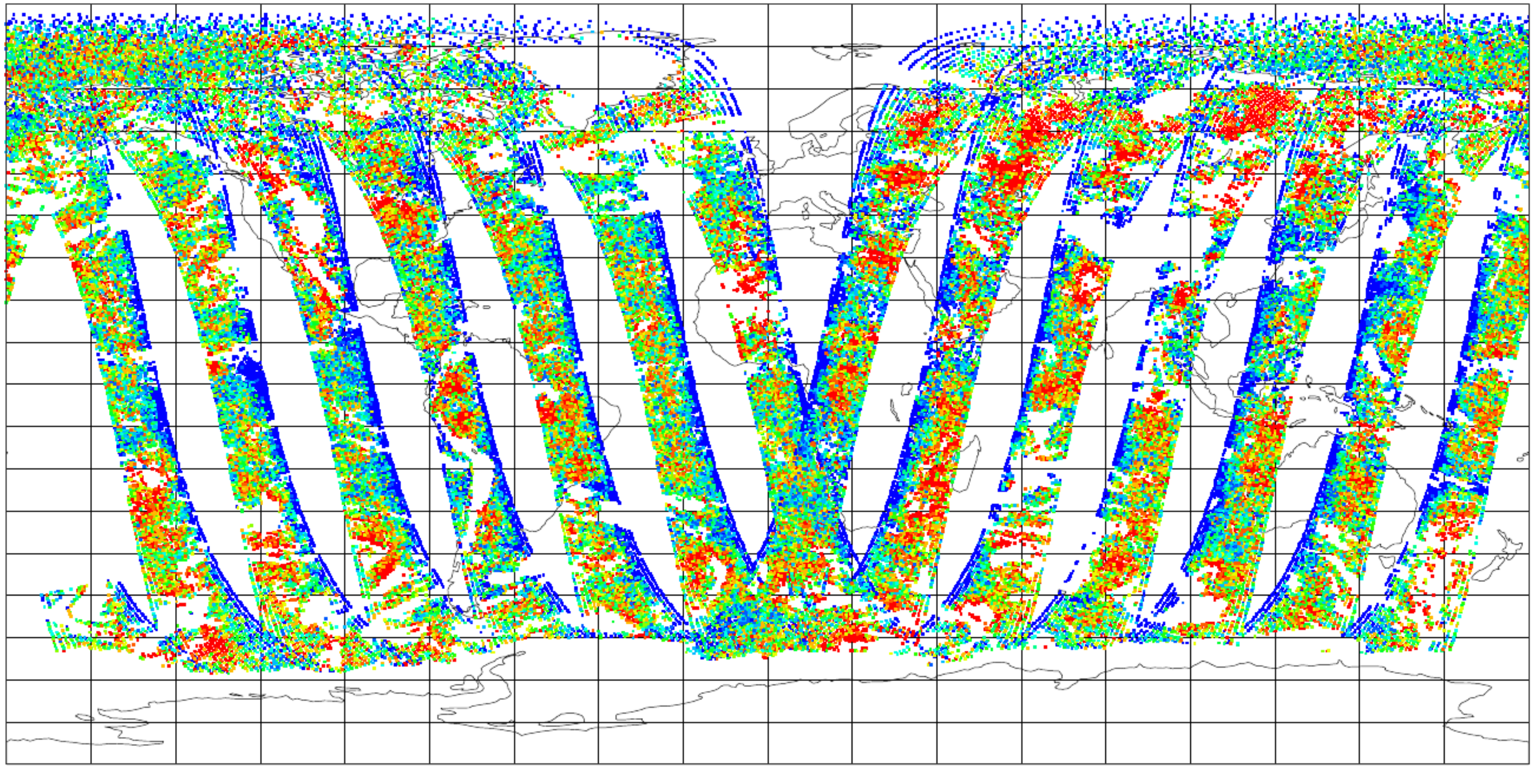
the parameters **the predictors**

Building models of observation biases (a more complex case)

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

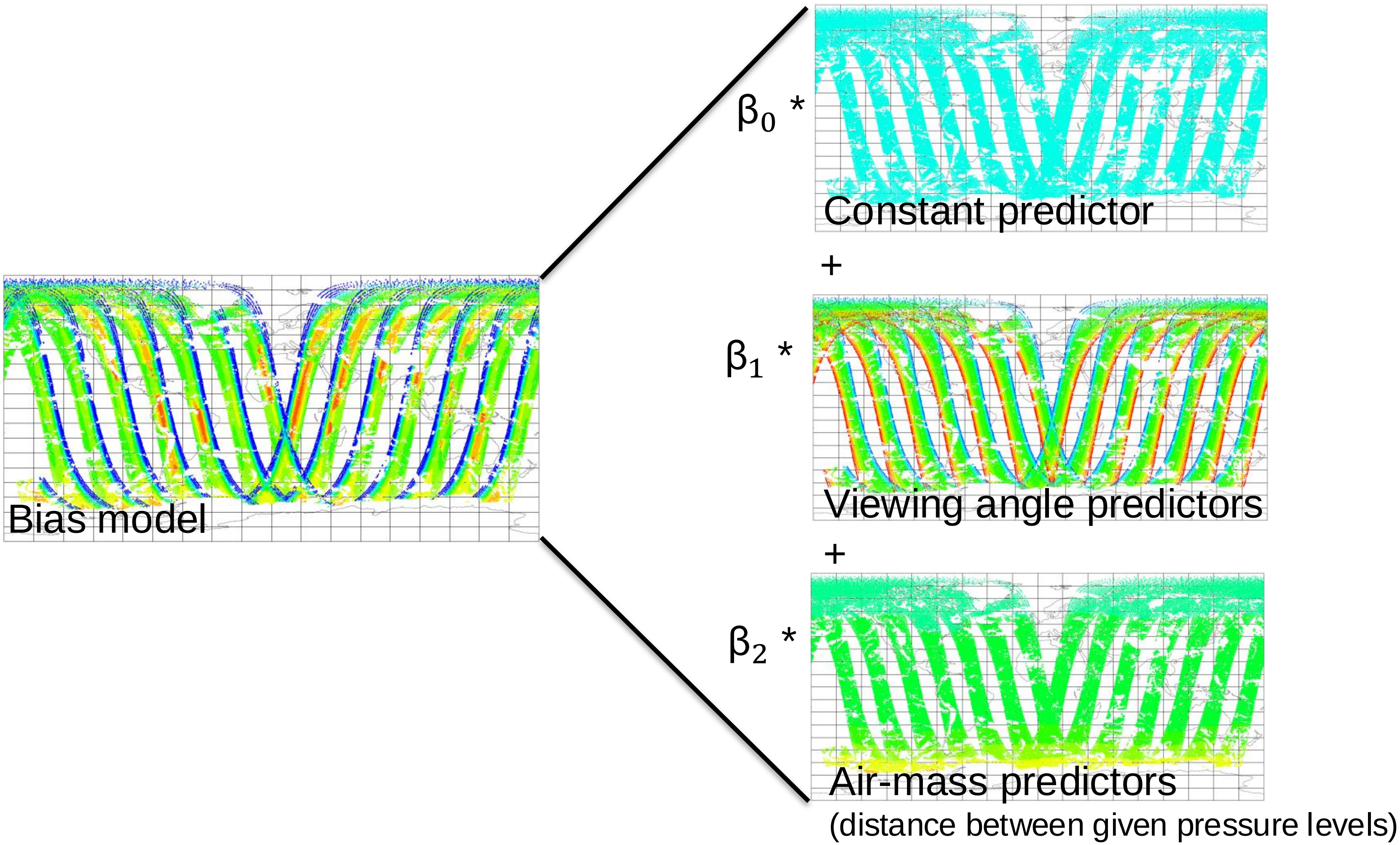
Observation bias is estimated inside 4D-Var
➔ comparing measurements with model
➔ specifying the structure of the model bias

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

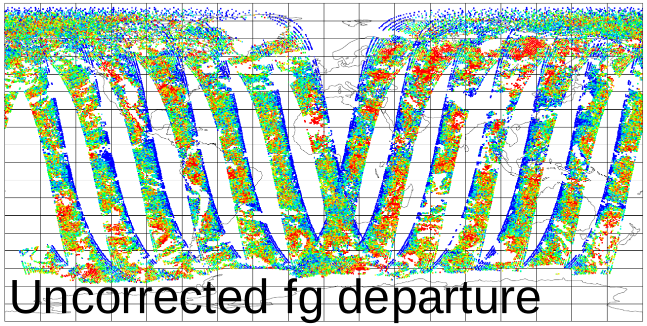


Building models of observation biases (a more complex case)

Bias model = $b(\beta) = b(\beta_0\ , \beta_1\ , \beta_2) = \beta_0 + \beta_1$ * viewing angle + β_2 * air-mass

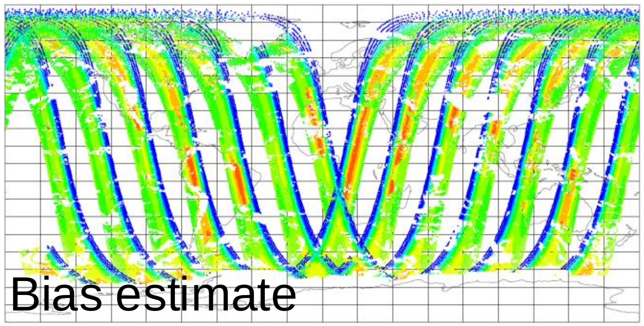


Building models of observation biases (a more complex case)



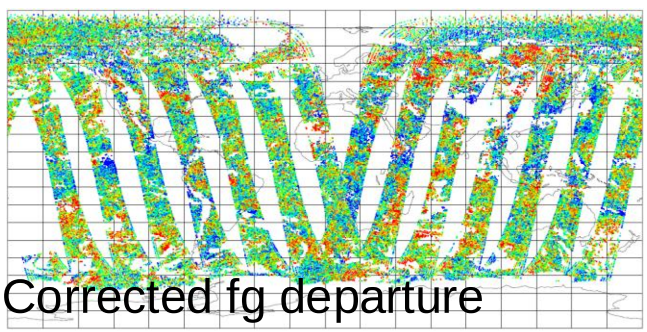
Uncorrected fg departure

-



Bias estimate

=



Corrected fg departure

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

Do not include too many predictors in the bias correction models

➔ to avoid correcting for other sources of errors (background errors/model error)

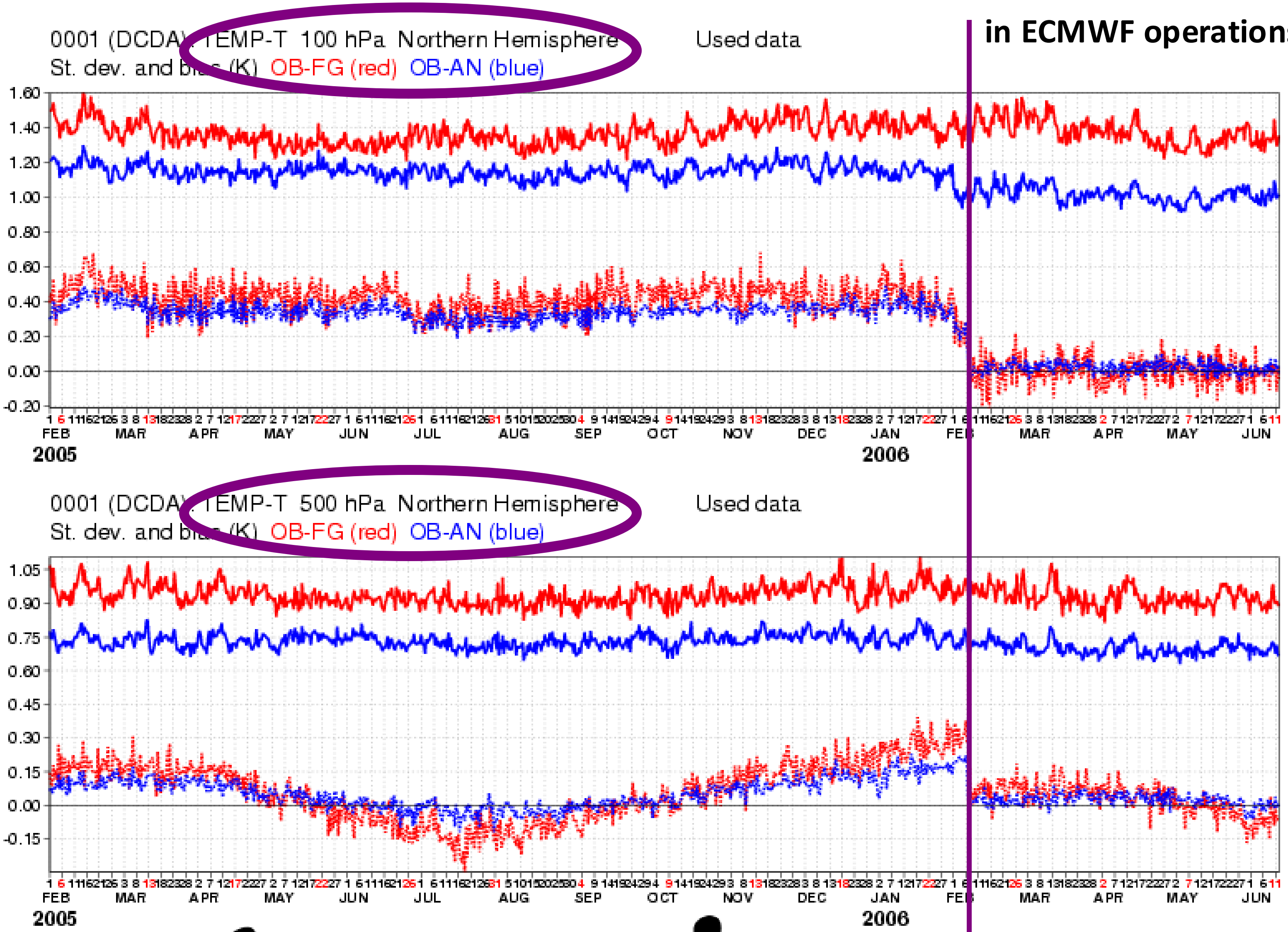
➔ corrected fg departure should still contain some information to constrain x_0

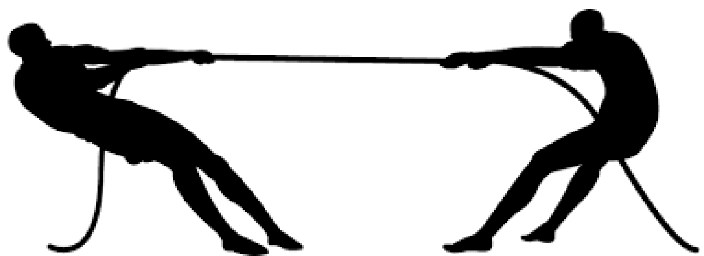
Generic VarBC formulation

$$b(\beta, x_k) = \beta_0 + \sum_{i=0}^{N}\beta_i p_i(x_k)$$

VarBC introduced in operations at ECMWF

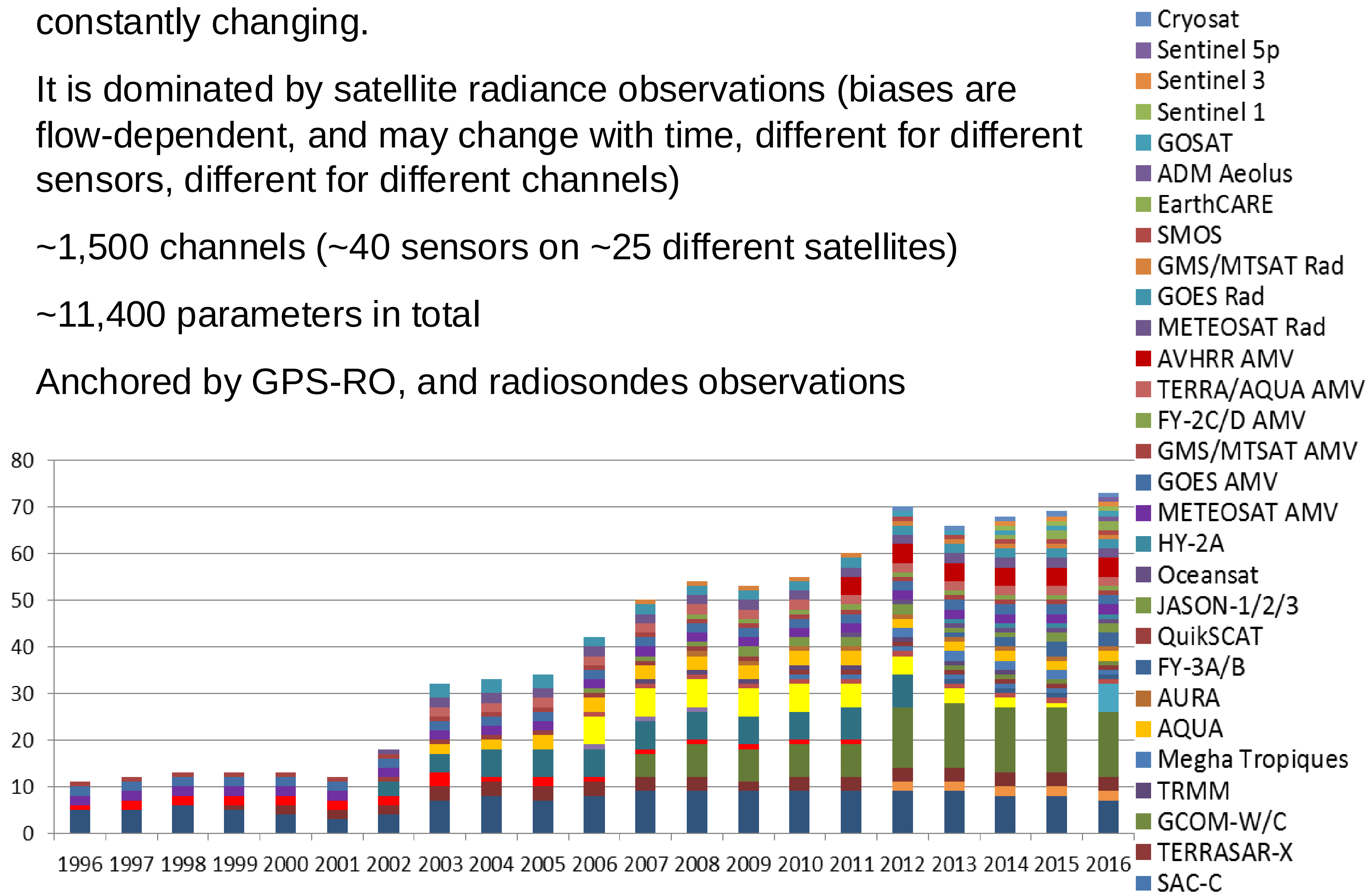
Introduction of VarBC in ECMWF operations





The power of VarBC

- The global observing system is increasingly complex and constantly changing.
- It is dominated by satellite radiance observations (biases are flow-dependent, and may change with time, different for different sensors, different for different channels)
- ~1,500 channels (~40 sensors on ~25 different satellites)
- ~11,400 parameters in total
- Anchored by GPS-RO, and radiosondes observations

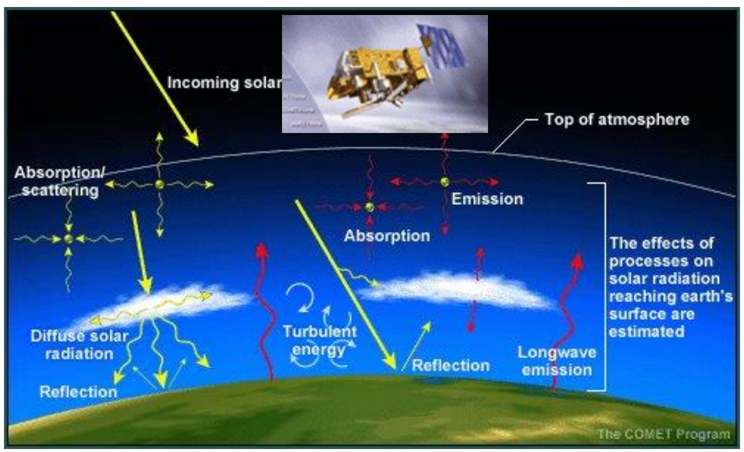


Biases in the observation operators

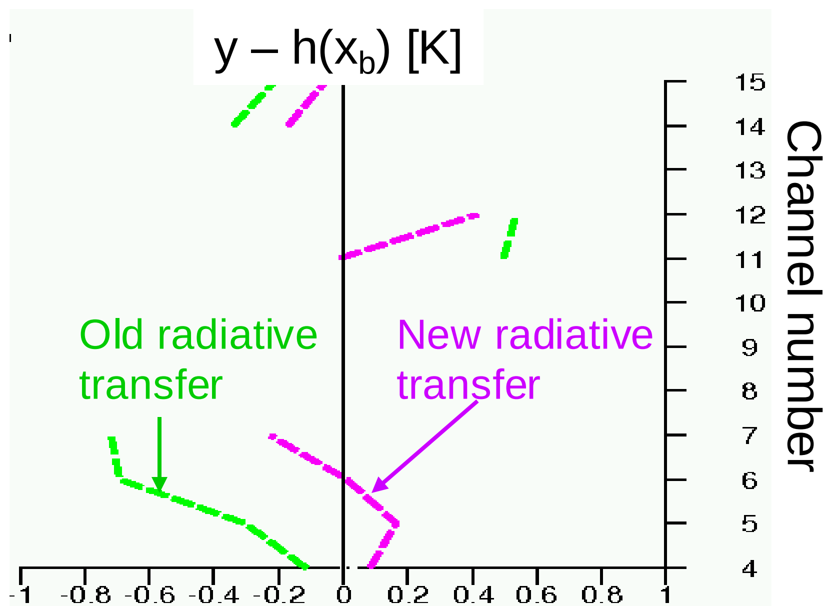
Examples of causes for biases in radiative transfer $y - h(x_b)$:

- Bias in assumed concentrations of atmospheric gases (e.g., CO_2, aerosols)
- Neglected effects (e.g., clouds)

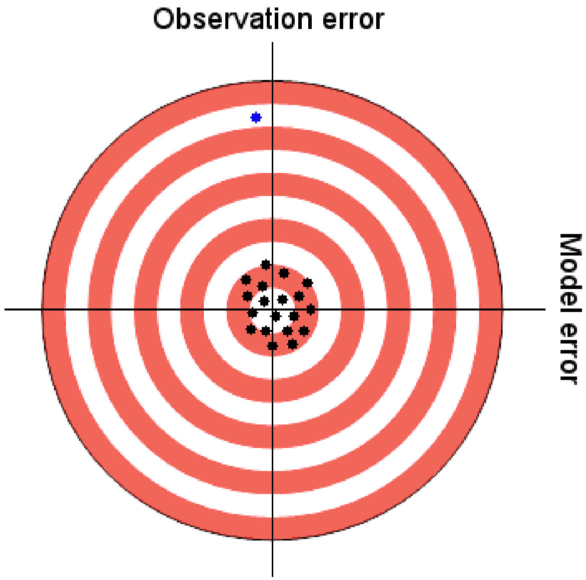
VarBC needs to handle these biases in its model

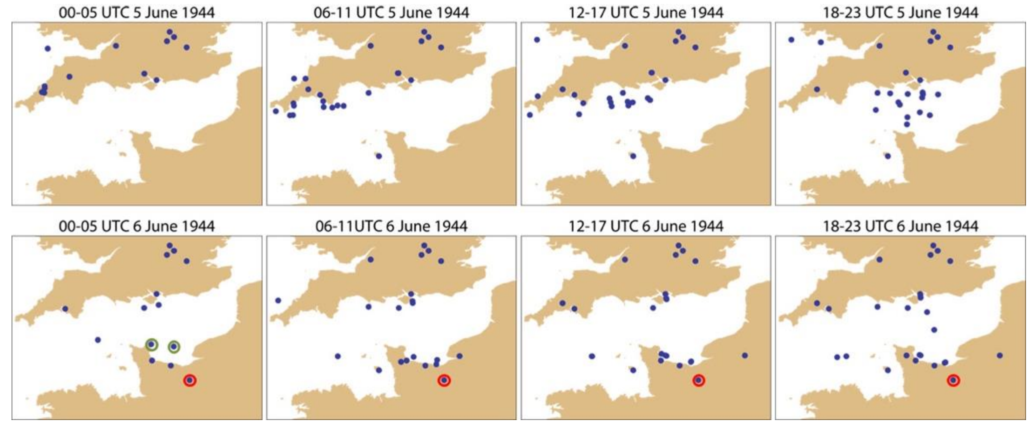


Change in bias for HIRS resulting from an update of the Radiative Transfer model:



Not the job of VarBC



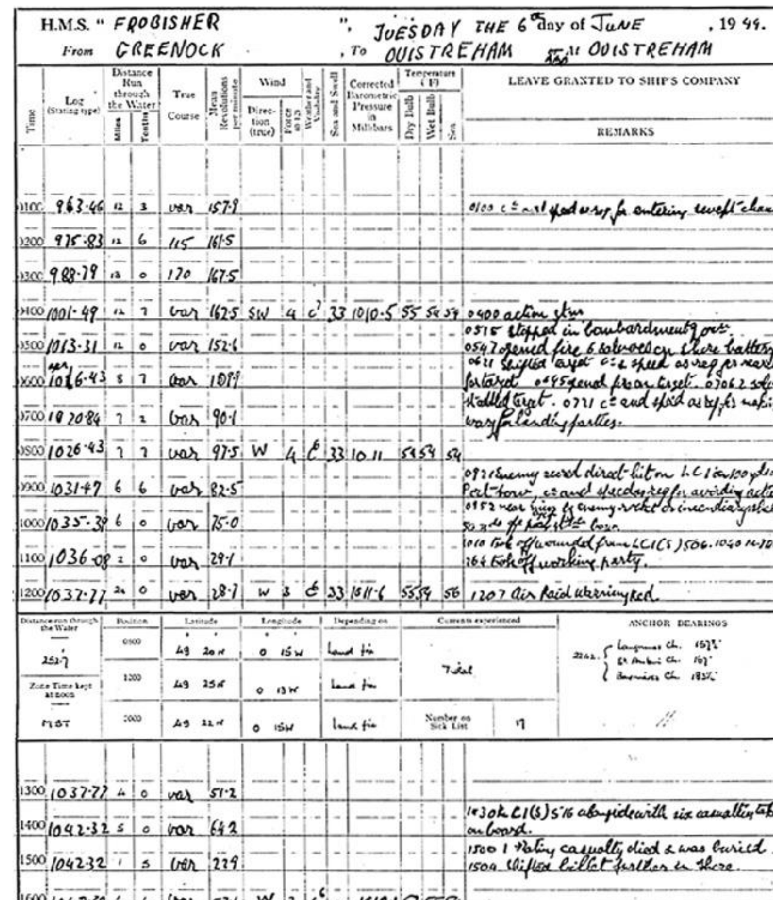


HMS Frobisher

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)


H.M.S. " FROBISHER ", TUESDAY THE 6th day of JUNE , 19 44.
From GREENOCK , To OUISTREHAM At OUISTREHAM

LEAVE GRANTED TO SHIP'S COMPANY

REMARKS

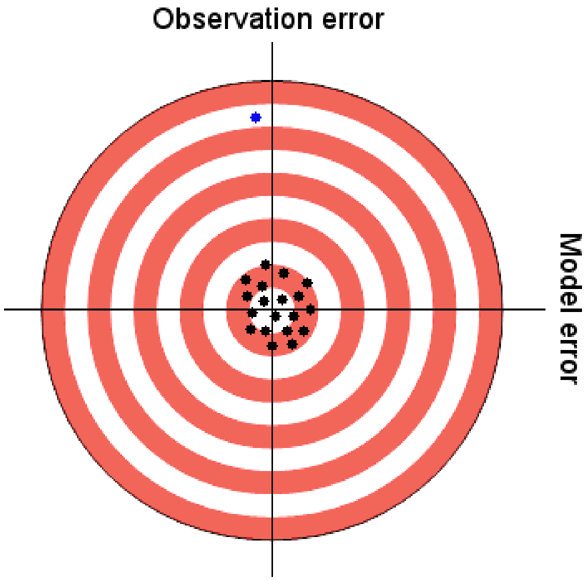
Time	Log	Miles	Tenths	True Course	Wind Direction	Force	Weather	Corrected Barometric Pressure	Dry Bulb	Wet Bulb	Sea
0100	963·46	12	3	var	157·9						
0200	975·83	12	6	115	161·5						
0300	988·19	12	0	170	167·5						
0400	1001·49	12	7	var	162·5	SW 4	c	1010·5	55	54	54
0500	1013·31	12	0	var	152·6						
0600	1016·43	3	7	var	109·9						
0700	1020·84	7	2	var	90·1						
0800	1026·43	7	7	var	97·5	W 4	c	1011	54	54	54
0900	1031·47	6	6	var	82·5						
1000	1035·39	6	0	var	75·0						
1100	1036·08	2	0	var	29·1						
1200	1037·77	2	0	var	28·1	W 3	c	1011·6	55	54	56
1300	1037·77	4	0	var	51·2						
1400	1042·32	5	0	var	64·2						
1500	1042·32	1	5	var	22·9						
1600	1042·32	4	4	var	57·1	W 3	c	1012·1	57	55	54

ANCHOR BEARINGS

Position: 0800 49 20 N 0 15 W land fix; 1200 49 25 N 0 13 W land fix; 2000 49 22 N 0 15 W land fix

Zone Time kept: B.S.T.

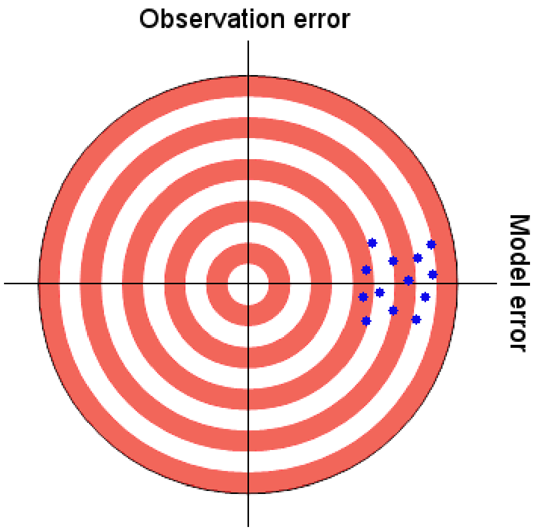
Not the job of VarBC



Gross (obvious) errors

→ Preliminary analysis (blacklist,…)

→ Variational Quality Control (VarQC)



Bias in the model

→ Tomorrow's lecture

Take-away messages (1/3)

To illustrate biases in observations

To construct bias models for specific instruments

To understand the challenges of observation bias correction

Take-away messages (2/3)

To illustrate biases in observations

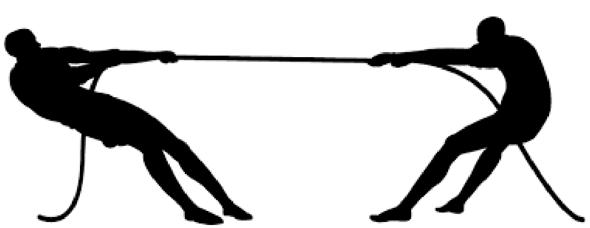
To construct bias models for specific instruments

To understand the challenges of observation bias correction

From bias-blind to bias-aware data assimilation

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

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



Take-away messages (3/3)

To illustrate biases in observations

To construct bias models for specific instruments

To understand the challenges of observation bias correction

➔ we only have information about differences $\mathbf{y - h(x_b)}$

➔ there is no true reference in the real world!

➔ the success of VarBC relies on ***anchoring*** and ***redundancy***

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