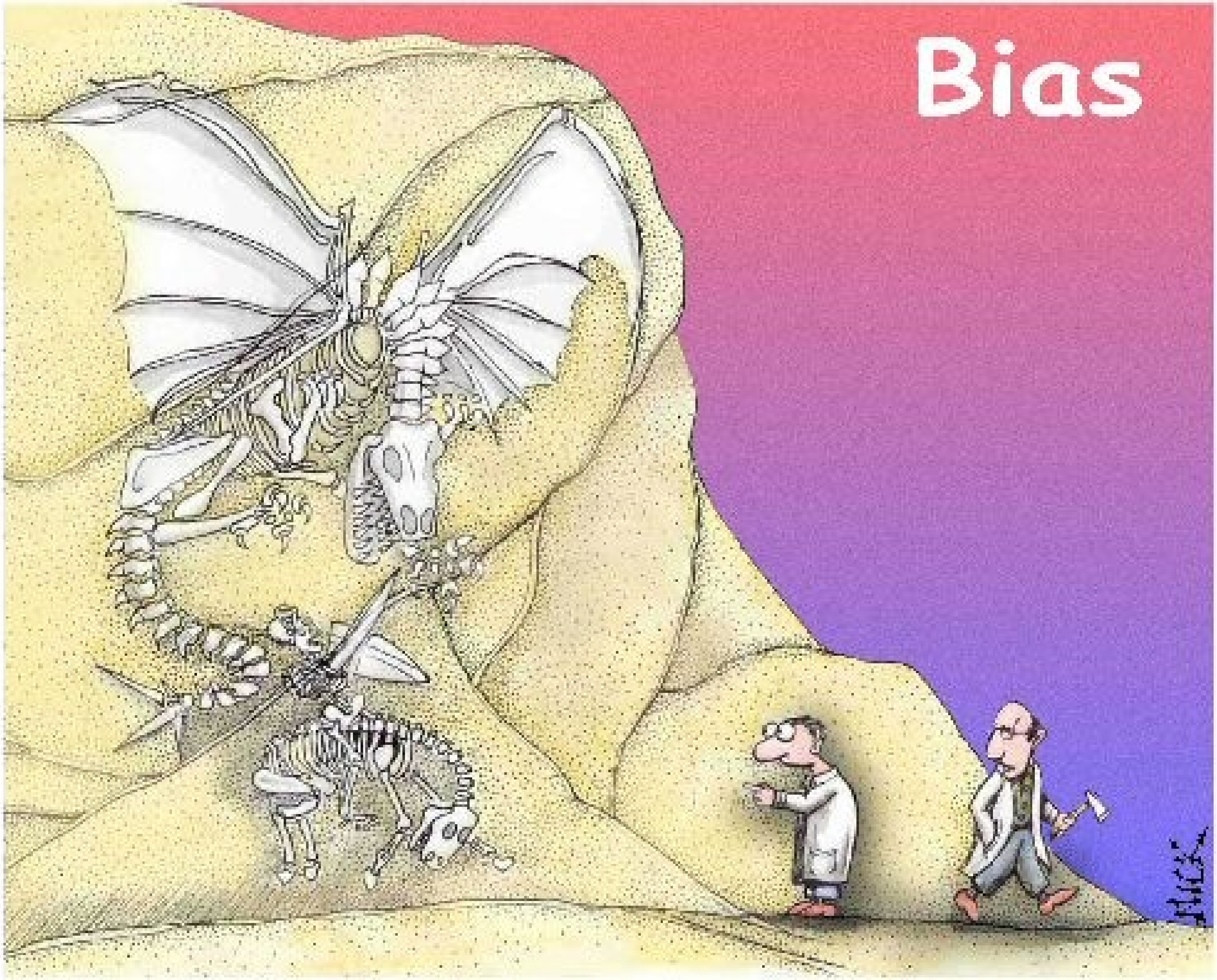
Bias correction in data assimilation

Dick Dee

ECMWF

Meteorological Training Course
Data Assimilation and Use of Satellite Data

12 May 2010



"Ignore it, Jeffries. It's unscientific."

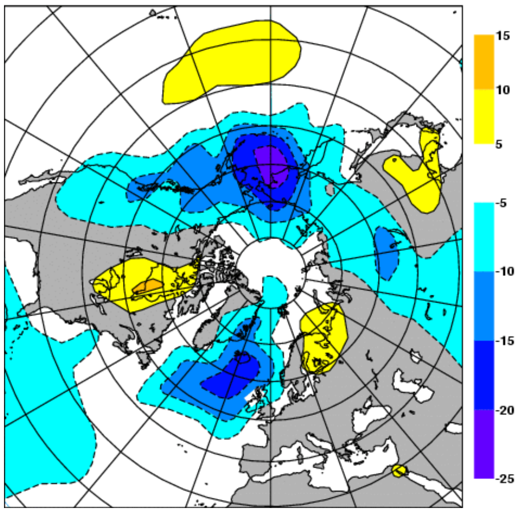
Outline

- Introduction
  - Biases in models, observations, and observation operators
  - Implications for data assimilation
- Variational analysis and correction of observation bias
  - The need for an adaptive system
  - Variational bias correction (VarBC)
- Limitations of VarBC
  - Interaction with model bias
  - Assimilation in the upper stratosphere
- Summary

Model bias:
Systematic D+3 Z500 errors in three different models

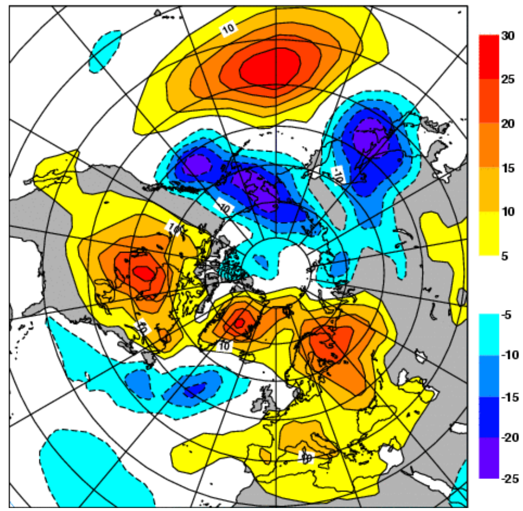
ECMWF

Systematic Z500 Error (ECMWF DJF 01-03)



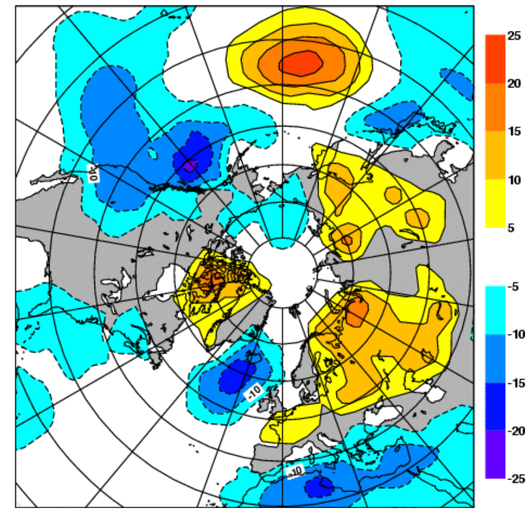
Meteo-France

Systematic Z500 Error (France DJF 01-03)



DWD

Systematic Z500 Error (DWD DJF 01-03)

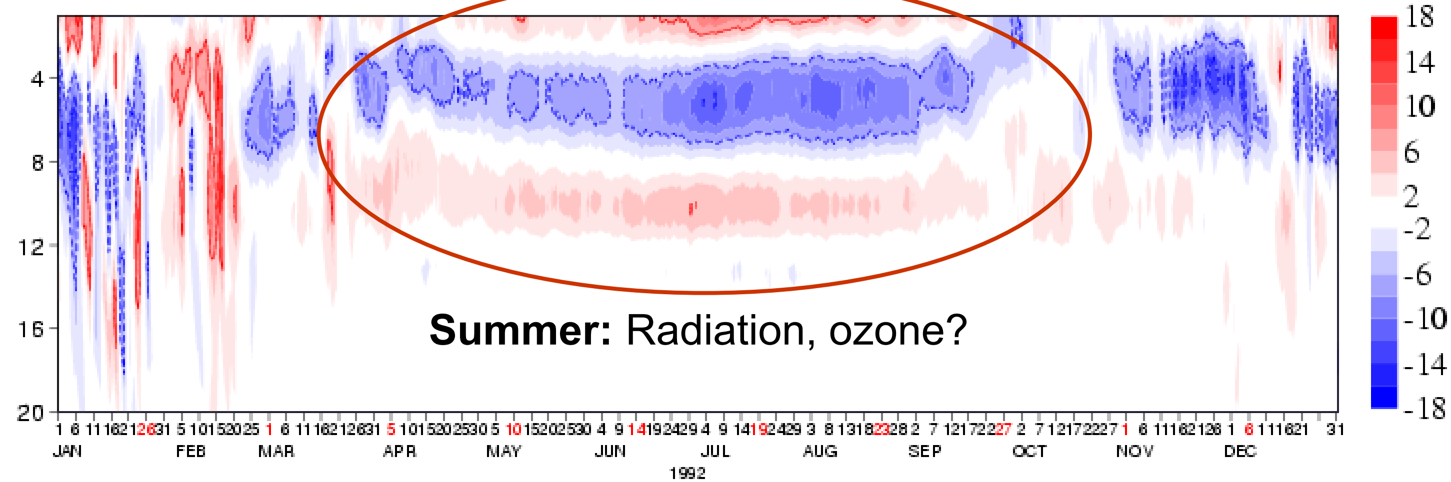


Different models often have similar error characteristics

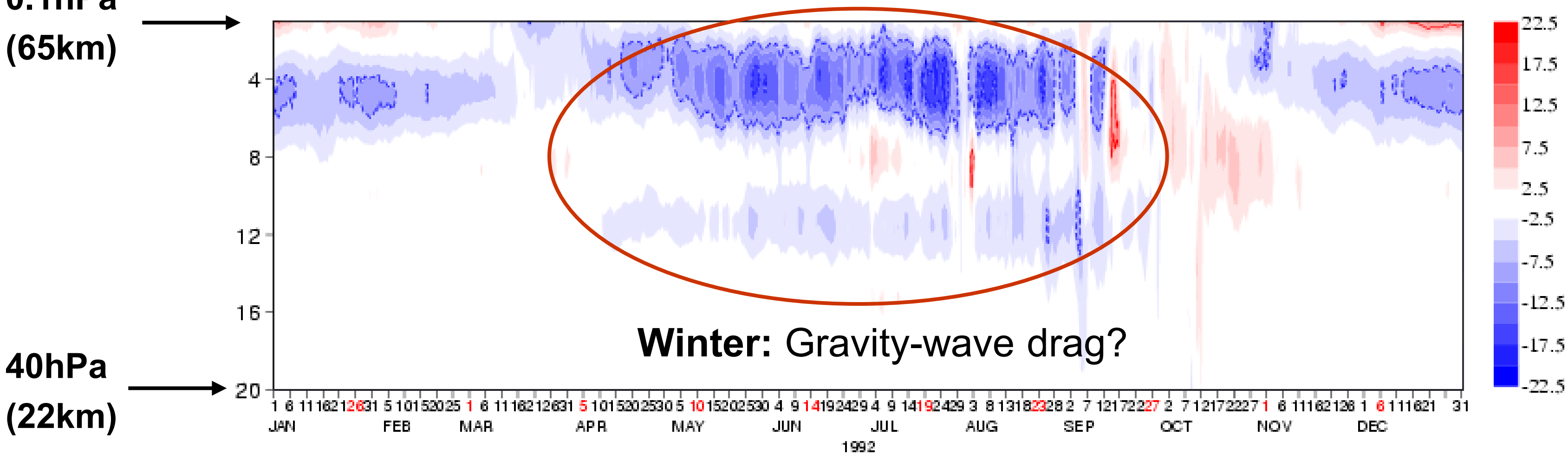
Model bias:
Seasonal variation in upper-stratospheric model errors

T255L60 model currently used for ERA-Interim

Mean temperature [K] 120-hour forecast errors for experiment 1112 : Arctic

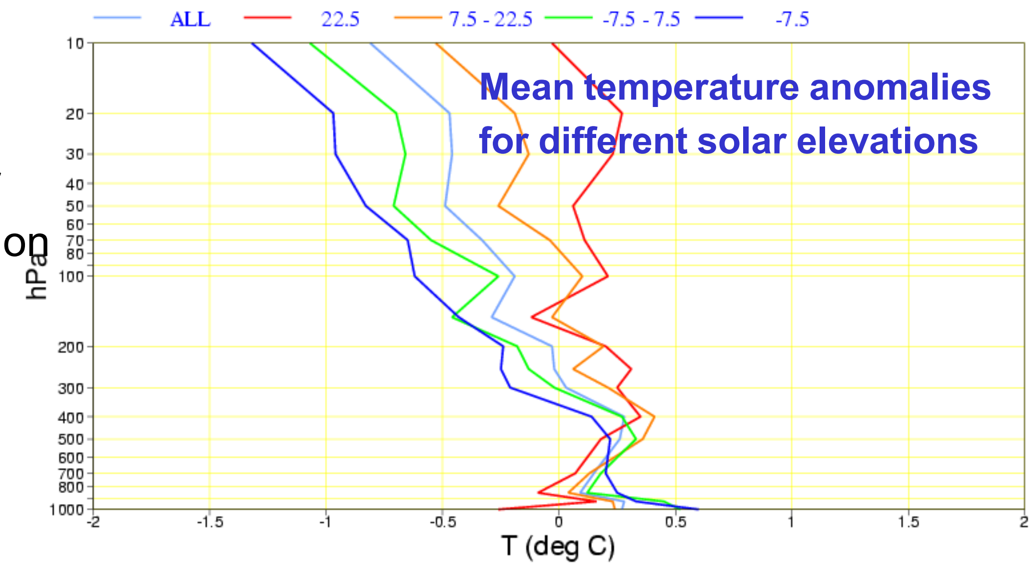


Mean temperature [K] 120-hour forecast errors for experiment 1112 : Antarctica

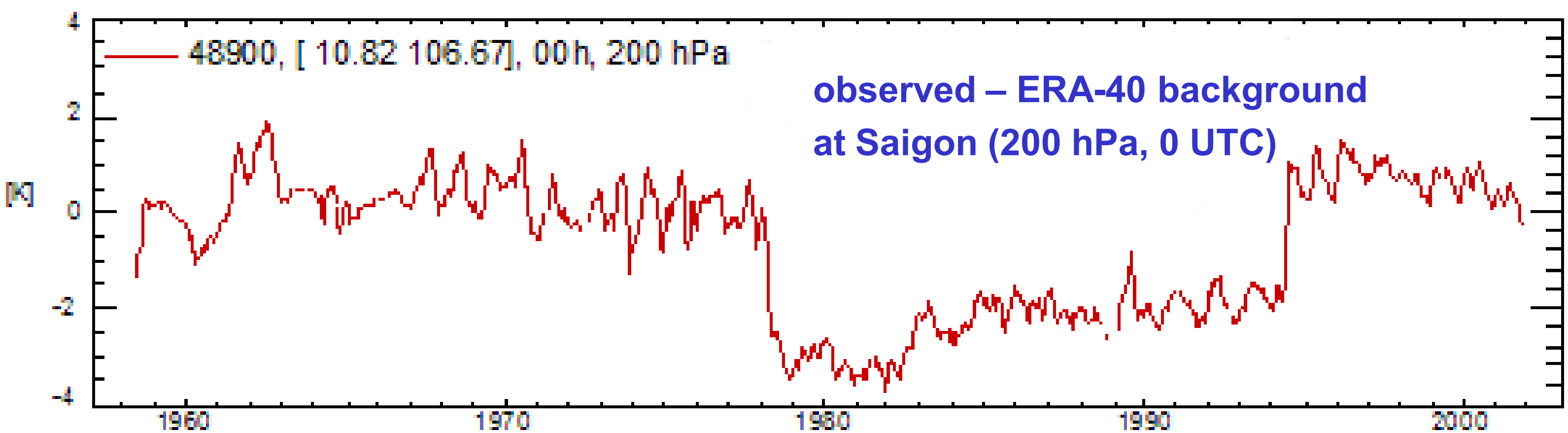


Observation bias:
Radiosonde temperature observations

Daytime warm bias due to radiative heating of the temperature sensor (depends on solar elevation and equipment type)



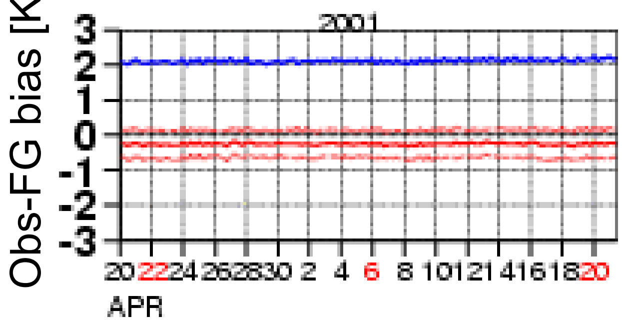
Bias changes due to change of equipment



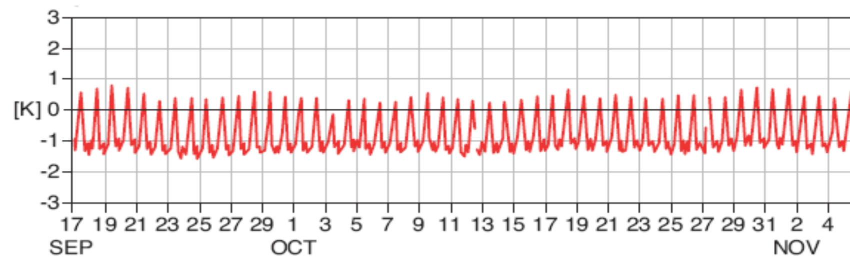
Observation and observation operator bias:
Satellite radiances

Monitoring the background departures (averaged in time and/or space):

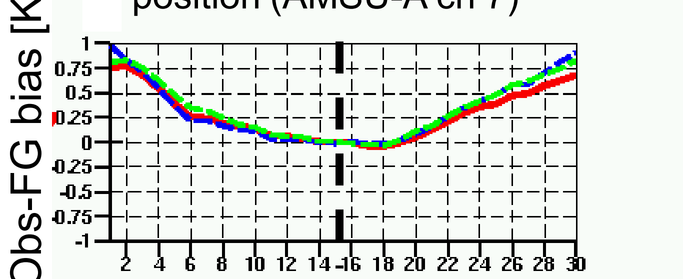
Constant bias (HIRS channel 5)



Diurnal bias variation in a geostationary satellite

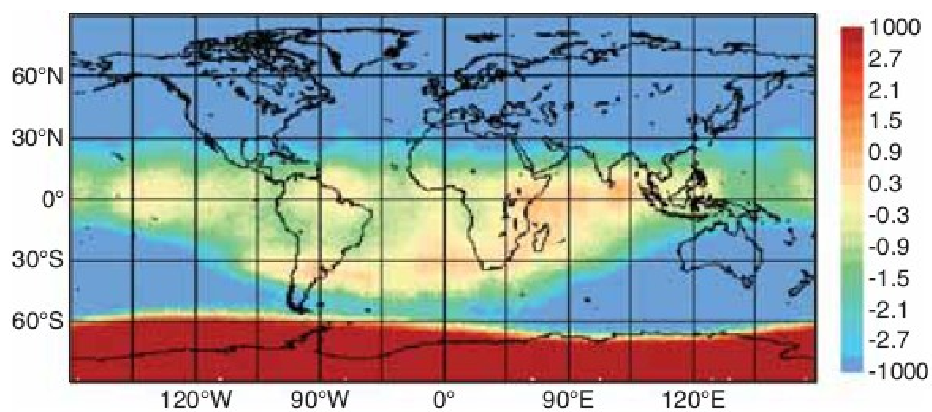


Bias depending on scan position (AMSU-A ch 7)



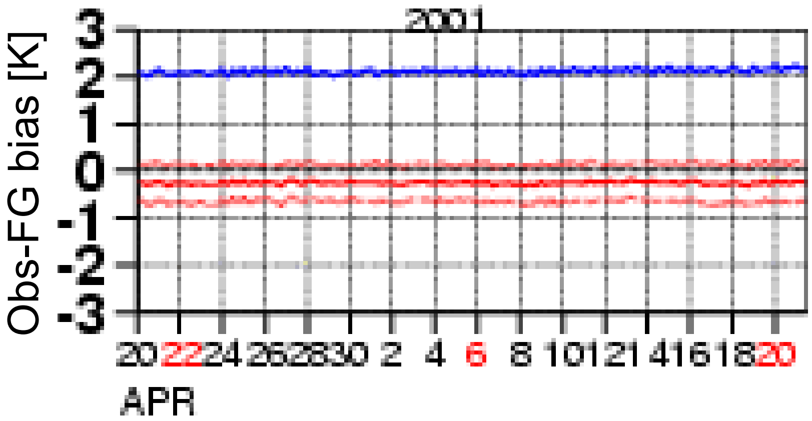
high zenith angle **nadir** **high zenith angle**

Air-mass dependent bias (AMSU-A channel 14)

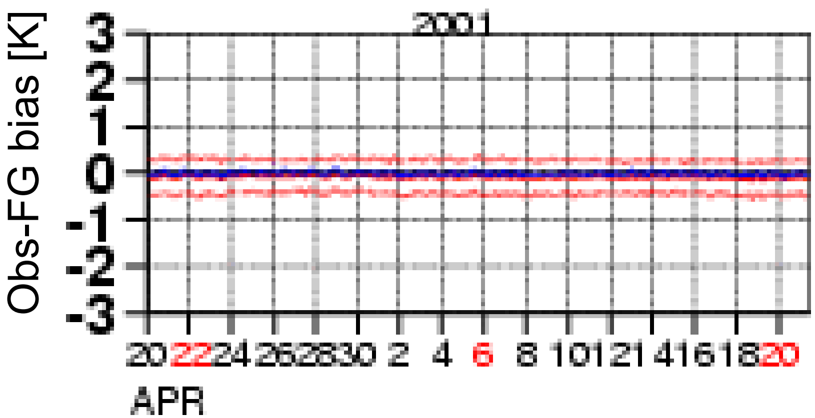


Observation and observation operator bias:
Satellite radiances

Monitoring the background departures (averaged in time and/or space):



HIRS channel 5 (peaking around 600hPa) on NOAA-14 satellite has +2.0K radiance bias against FG.



Same channel on NOAA-16 satellite has no radiance bias against FG.

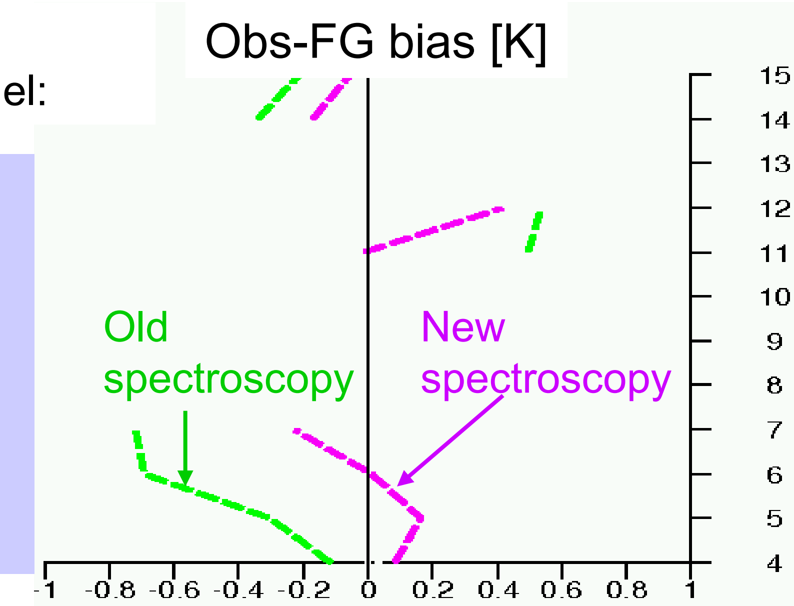
→ **NOAA-14 channel 5 has an instrument bias.**

Observation and observation operator bias:
Satellite radiances

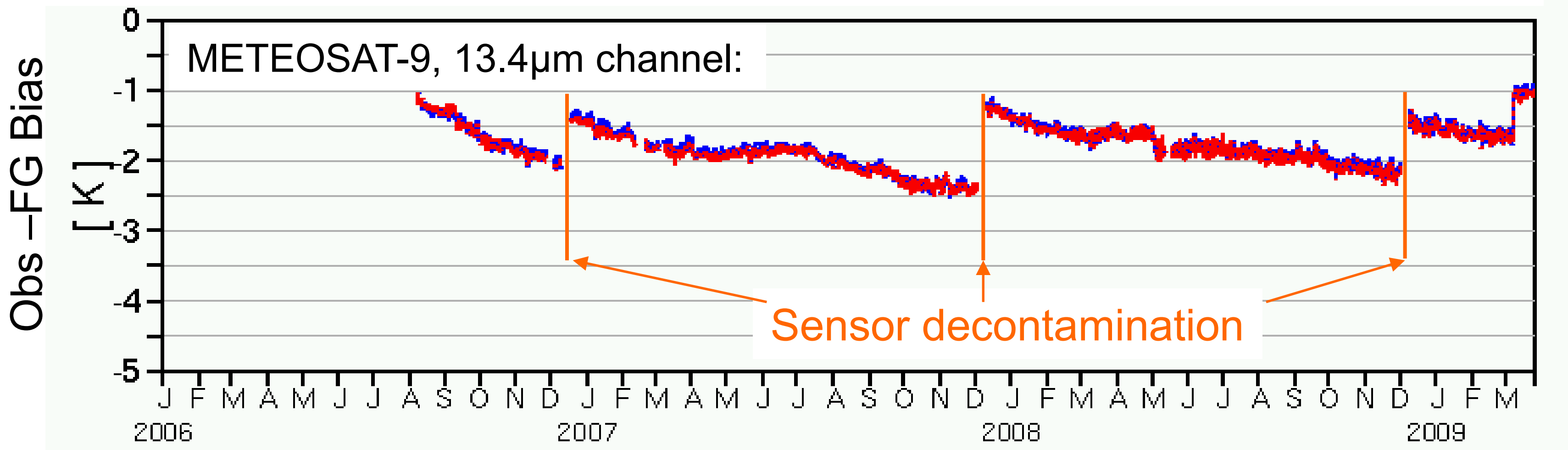
Different bias for HIRS due to different spectroscopy in the radiative transfer model:

Other common causes for biases in radiative transfer:

- Bias in assumed concentrations of atmospheric gases (e.g., CO_2)
- Neglected effects (e.g., clouds)
- Incorrect spectral response function
- ….



Drift in bias due to ice-build up on sensor:

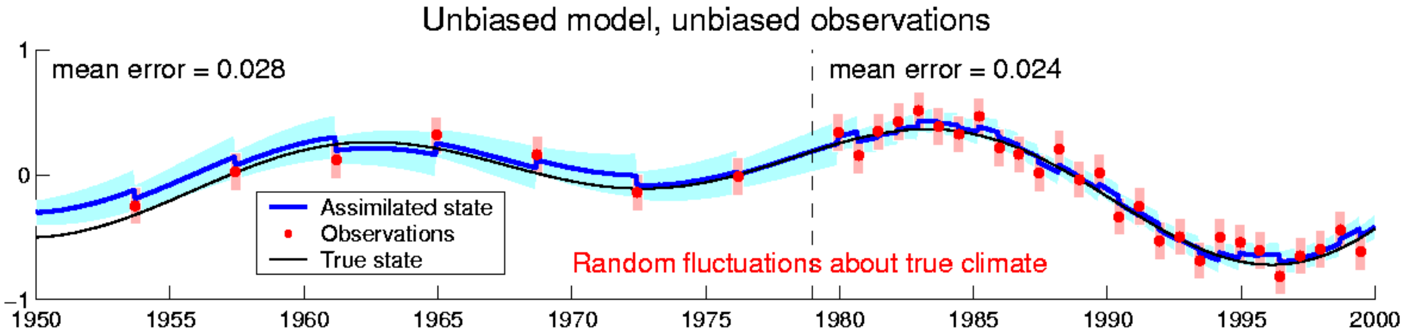


Implications for data assimilation:
Bias problems in a nutshell

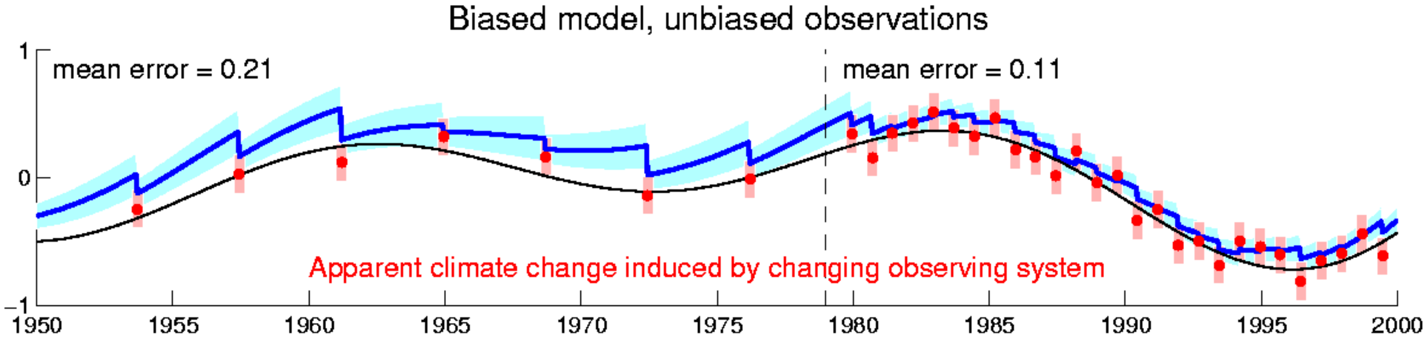
- **Observations and observation operators have biases, which may change over time**
    - Daytime warm bias in radiosonde measurements of stratospheric temperature; radiosonde equipment changes
    - Biases in cloud-drift wind data due to problems in height assignment
    - Biases in satellite radiance measurements and radiative transfer models
- **Models have biases, and changes in observational coverage over time may change the extent to which observations correct these biases**
    - Stratospheric temperature bias modulated by radiance assimilation
    - This is especially important for reanalysis (trend analysis)
- **Data assimilation methods are primarily designed to correct small (random) errors in the model background**
    - Large corrections generally introduce spurious signals in the assimilation
    - Likewise, inconsistencies among different parts of the observing system lead to all kinds of problems

Implications for data assimilation:
The effect of model bias on trend estimates

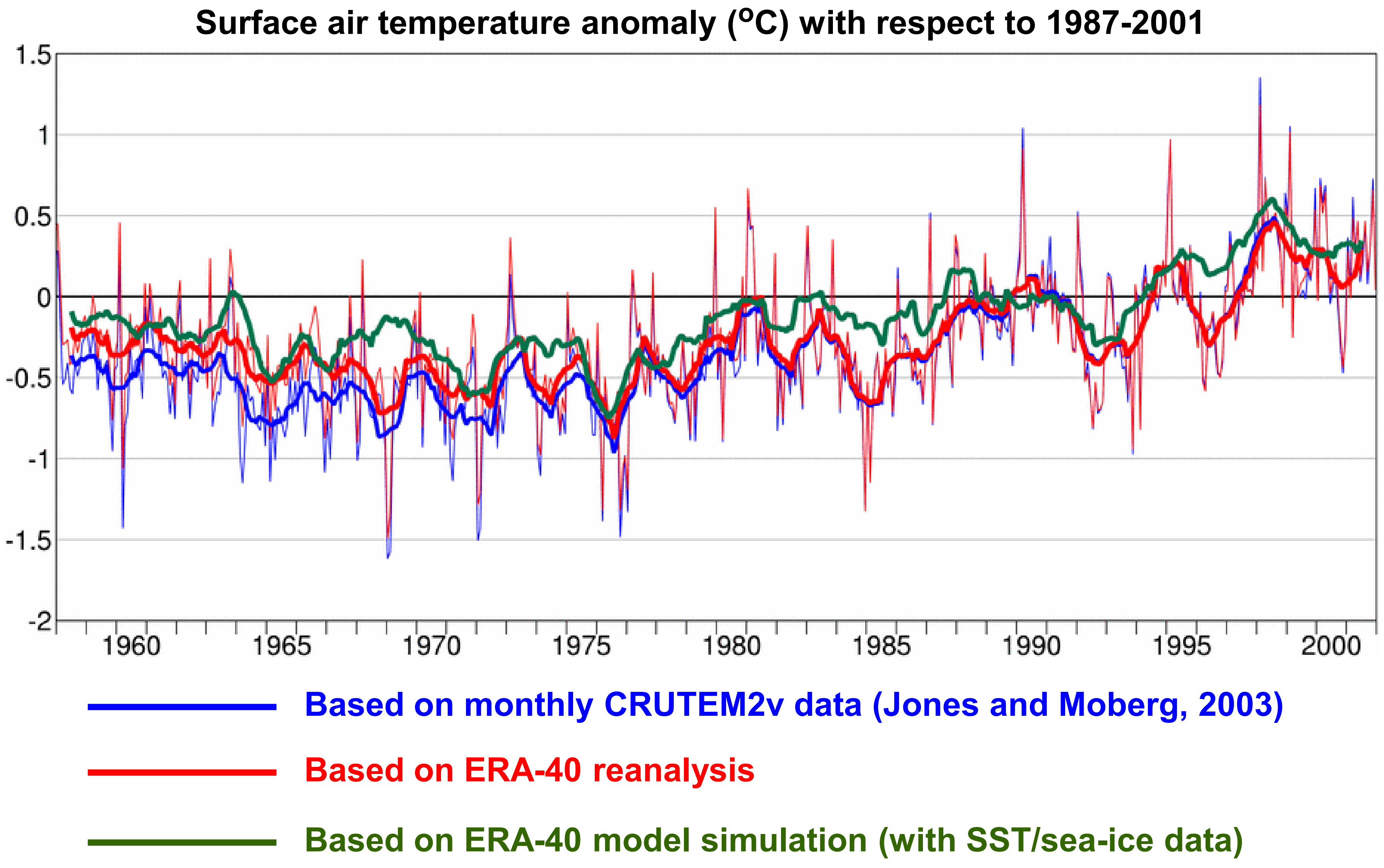
Most assimilation systems assume unbiased models and unbiased data



Biases in models and/or data can induce spurious trends in the assimilation



Implications for data assimilation:
ERA-40 surface temperatures compared to land-station values



Outline

- Introduction
  - Biases in models, observations, and observation operators
  - Implications for data assimilation
- Variational analysis and correction of observation bias
  - The need for an adaptive system
  - Variational bias correction (VarBC)
- Limitations of VarBC
  - Interaction with model bias
  - Assimilation in the upper stratosphere
- Summary

Variational analysis and bias correction:
A brief review of variational data assimilation

Minimise $$J(x) = (x_b - x)^T B^{-1} (x_b - x) + [y - h(x)]^T R^{-1} [y - h(x)]$$

background constraint (J_b) observational constraint (J_o)

- The input x_b represents past information propagated by the forecast model
  (the model background)

- The input $[y – h(x_b)]$ represents the new information entering the system
  (the background departures - sometimes called the innovation)

- The function h(x) represents a model for simulating observations
  (the observation operator)

- Minimising the cost function J(x) produces an adjustment to the model background based on all used observations
  (the analysis)

Variational analysis and bias correction:
Error sources in the input data

Minimise $$J(x) = (x_b - x)^T B^{-1} (x_b - x) + [y - h(x)]^T R^{-1} [y - h(x)]$$

background constraint (J_b) observational constraint (J_o)

- **Errors in the input $[y - h(x_b)]$ arise from:**
  - errors in the actual observations
  - errors in the model background
  - errors in the observation operator

- **There is no general method for separating these different error sources**
  - we only have data about differences
  - there is no true reference in the real world

- **The analysis does not respond well to contradictory input information**

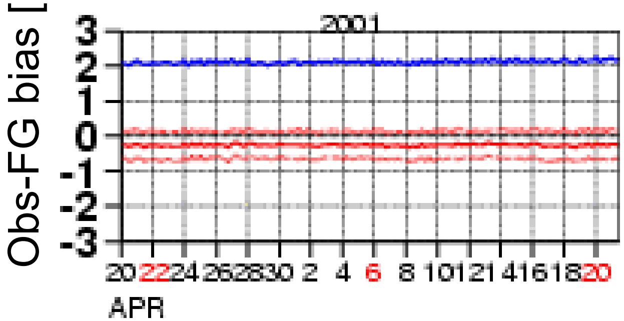
  A lot of work is done to remove biases prior to assimilation:
  - ideally by removing the cause
  - in practise by careful comparison against other data

The need for an adequate bias model

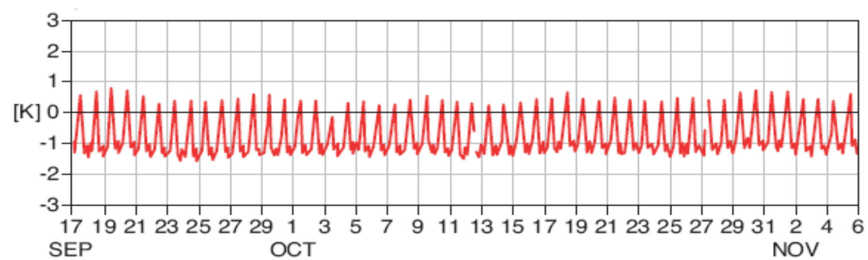
Prerequisite for any bias correction is a good model for the bias (b(x,β)):

- Ideally, “corrects only what we want to correct”.
- If possible, bias model is guided by the physical origins of the bias.
- Usually, bias models are derived empirically from observation monitoring.

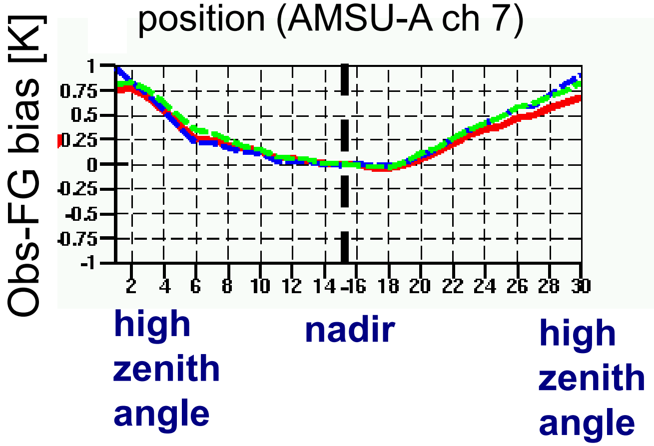
Constant bias (HIRS channel 5)



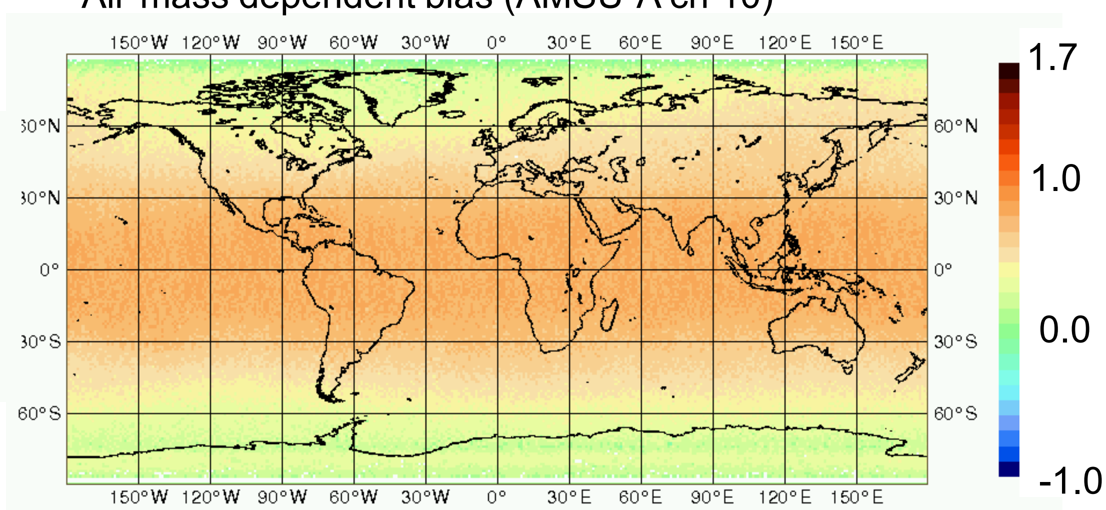
Diurnal bias variation in a geostationary satellite



Bias depending on scan position (AMSU-A ch 7)



Air-mass dependent bias (AMSU-A ch 10)



Past* scheme for radiance bias correction at ECMWF

Scan bias and **air-mass dependent bias** for each sensor/channel were estimated off-line from background departures, and stored on files **(Harris and Kelly 2001)**

Error model for brightness temperature data:

$$y = h(x) + b^{scan} + b^{air}(x) + e^{obs}$$

where

$$b^{scan} = b^{scan}(\text{latitude, scan position})$$

$$b^{air} = \beta_0 + \sum_{i=1}^{N} \beta_i p_i(x)$$

$$e^{obs} = \text{random observation error}$$

Predictors, for instance:

- 1000-300 hPa thickness
- 200-50 hPa thickness
- surface skin temperature
- total precipitable water

Average the background departures:

$$\langle y - h(x_b) \rangle = b^{scan} + b^{air}(x)$$

Periodically estimate scan bias and predictor coefficients:

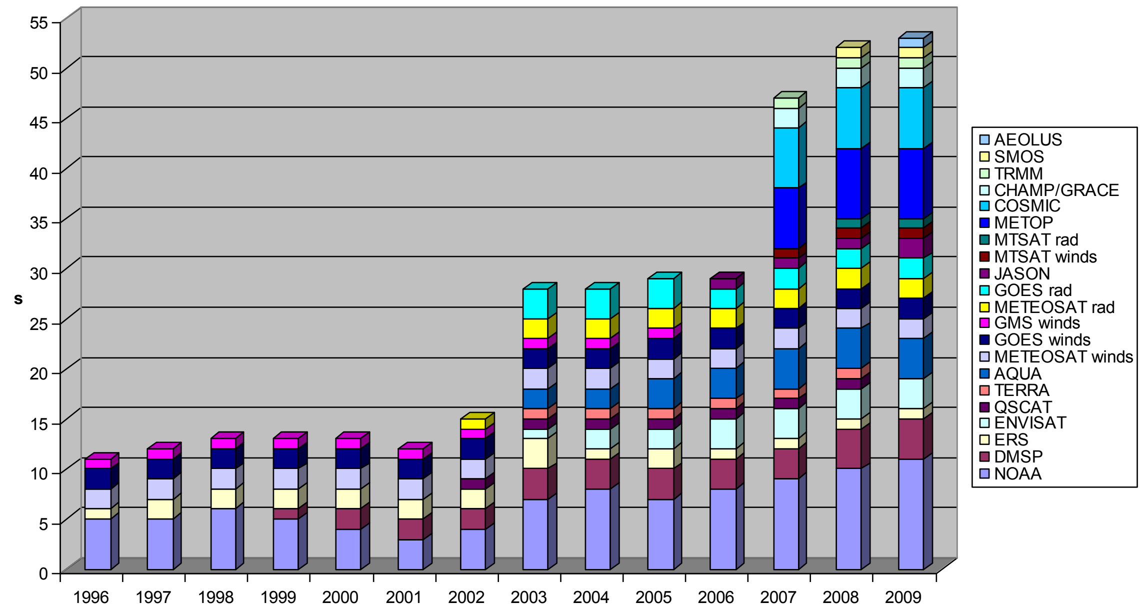
- typically 2 weeks of background departures
- 2-step regression procedure
- careful masking and data selection

***Replaced in operations September 2006 by VarBC (Variational Bias Correction)**

The need for an adaptive bias correction system

- **The observing system is increasingly complex and constantly changing**
- **It is dominated by satellite radiance data:**

  - biases are flow-dependent, and may change with time
  - they are different for different sensors
  - they are different for different channels



- **How can we manage the bias corrections for all these different components?**
- **This requires a consistent approach and a flexible, automated system**

Variational bias correction:
The general idea

The **bias** in a given instrument/channel is described by (a few) **bias parameters:**
typically, these are functions of air-mass and scan-position (the **predictors**)

These parameters can be estimated in a variational analysis along with the model state **(Derber and Wu, 1998 at NCEP, USA)**

The **standard variational analysis** minimizes

$$J(x) = (x_b - x)^T B_x^{-1} (x_b - x) + [y - h(x)]^T R^{-1} [y - h(x)]$$

Modify the observation operator to account for bias: $\tilde{h}(z) = \tilde{h}(x, \beta)$

Include the bias parameters in the control vector: $z^T = [\, x^T \;\; \beta^T \,]$

Minimize instead $J(z) = (z_b - z)^T B_z^{-1} (z_b - z) + [y - \tilde{h}(z)]^T R^{-1} [y - \tilde{h}(z)]$

What is needed to implement this:

1. The modified operator $\tilde{h}(x, \beta)$ and its TL + adjoint
2. A cycling scheme for updating the bias parameter estimates
3. An effective preconditioner for the joint minimization problem

Variational bias correction:
The modified analysis problem

The original problem:

J_b: background constraint

$$J(\mathbf{x}) = (\mathbf{x}_b - \mathbf{x})^T \mathbf{B}^{-1} (\mathbf{x}_b - \mathbf{x}) + [\mathbf{y} - \mathbf{h}(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - \mathbf{h}(\mathbf{x})]$$

J_o: observation constraint

The modified problem:

J_b: background constraint for x

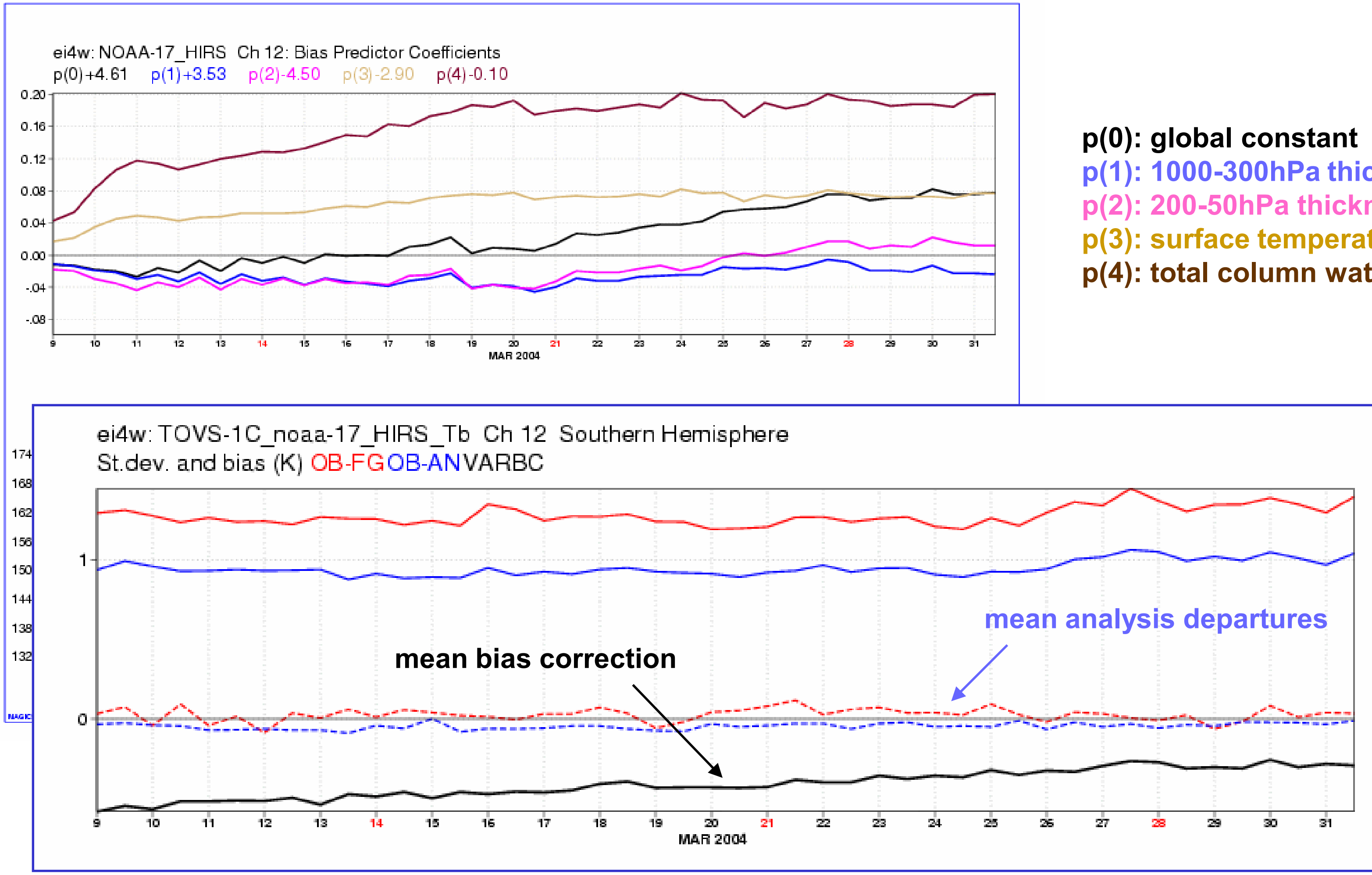
J_β: background constraint for β

$$\mathbf{J}(\mathbf{x}, \boldsymbol{\beta}) = (\mathbf{x}_b - \mathbf{x})^T \mathbf{B}_x^{-1} (\mathbf{x}_b - \mathbf{x}) + (\boldsymbol{\beta}_b - \boldsymbol{\beta})^T \mathbf{B}_\beta^{-1} (\boldsymbol{\beta}_b - \boldsymbol{\beta})$$
$$+ [\mathbf{y} - \mathbf{b}_o(\mathbf{x}, \boldsymbol{\beta}) - \mathbf{h}(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - \mathbf{b}_o(\mathbf{x}, \boldsymbol{\beta}) - \mathbf{h}(\mathbf{x})]$$

Parameter estimate from previous analysis

J_o: bias-corrected observation constraint

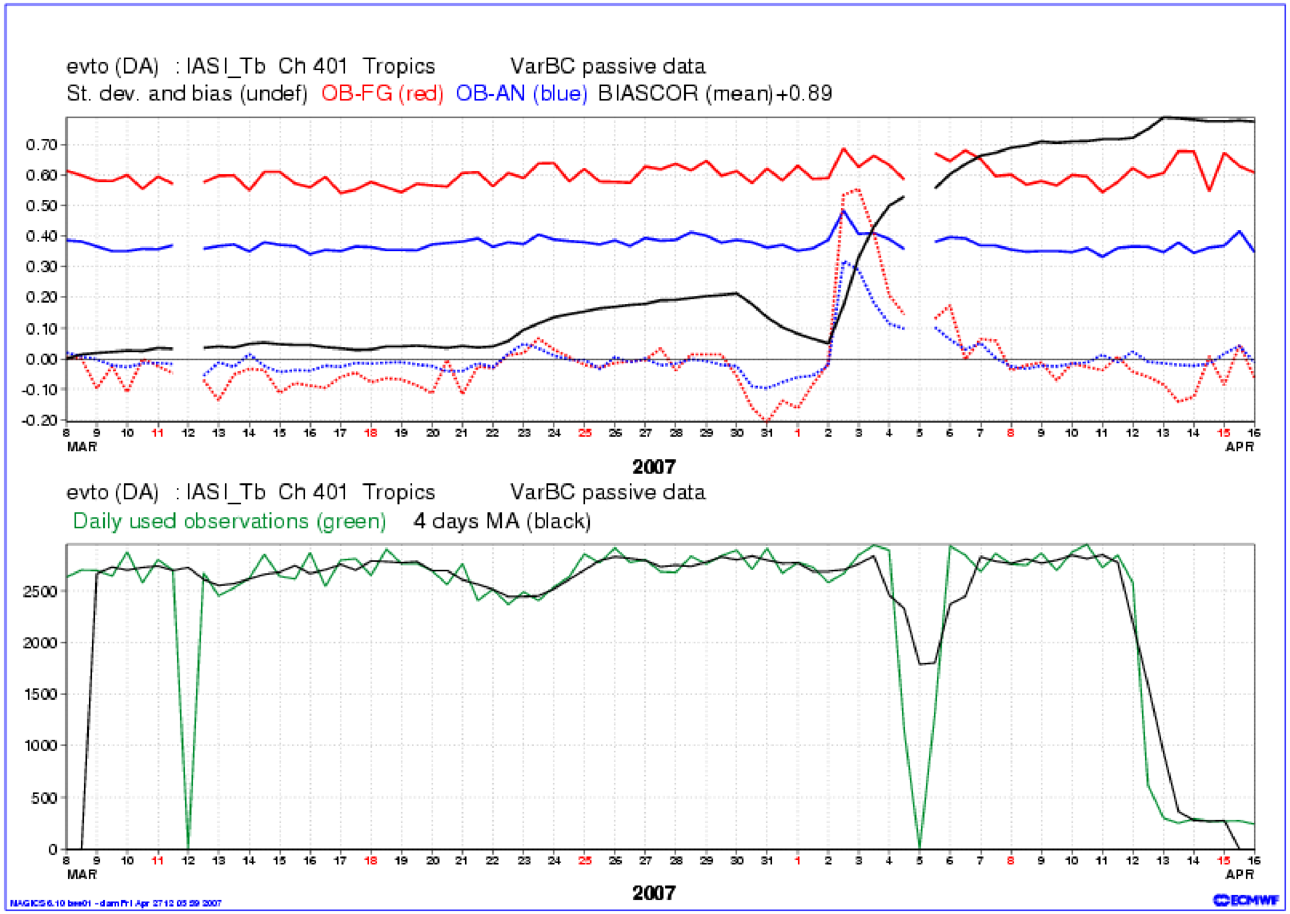
Performance:
Adaptive bias correction of NOAA-17 HIRS Ch12



p(0): global constant
p(1): 1000-300hPa thickness
p(2): 200-50hPa thickness
p(3): surface temperature
p(4): total column water

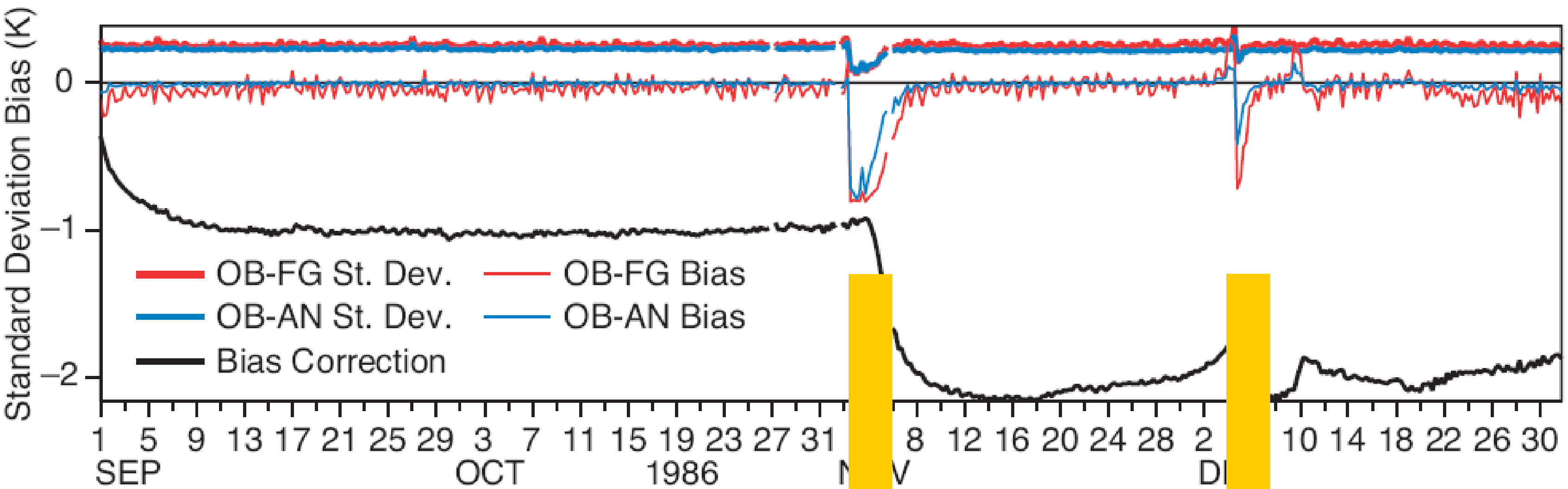
Performance:
Spinning up new instruments – IASI on MetOp

- IASI is a high-resolution interferometer with 8461 channels
- Initially unstable – data gaps, preprocessing changes

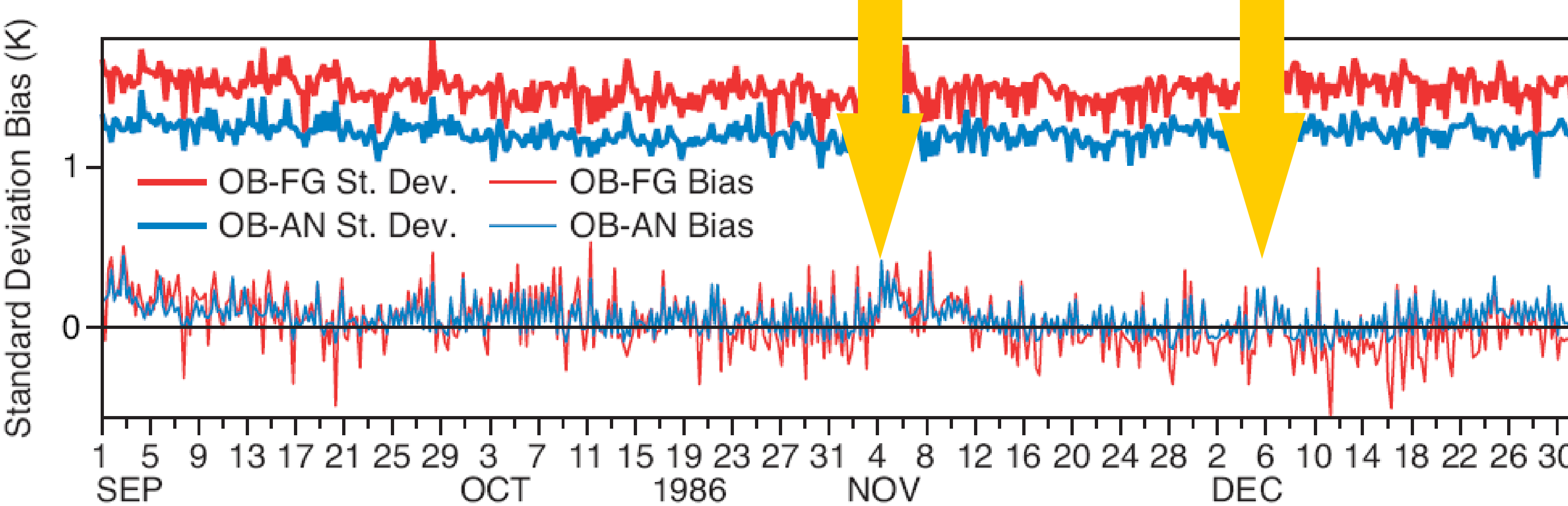


Performance:

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

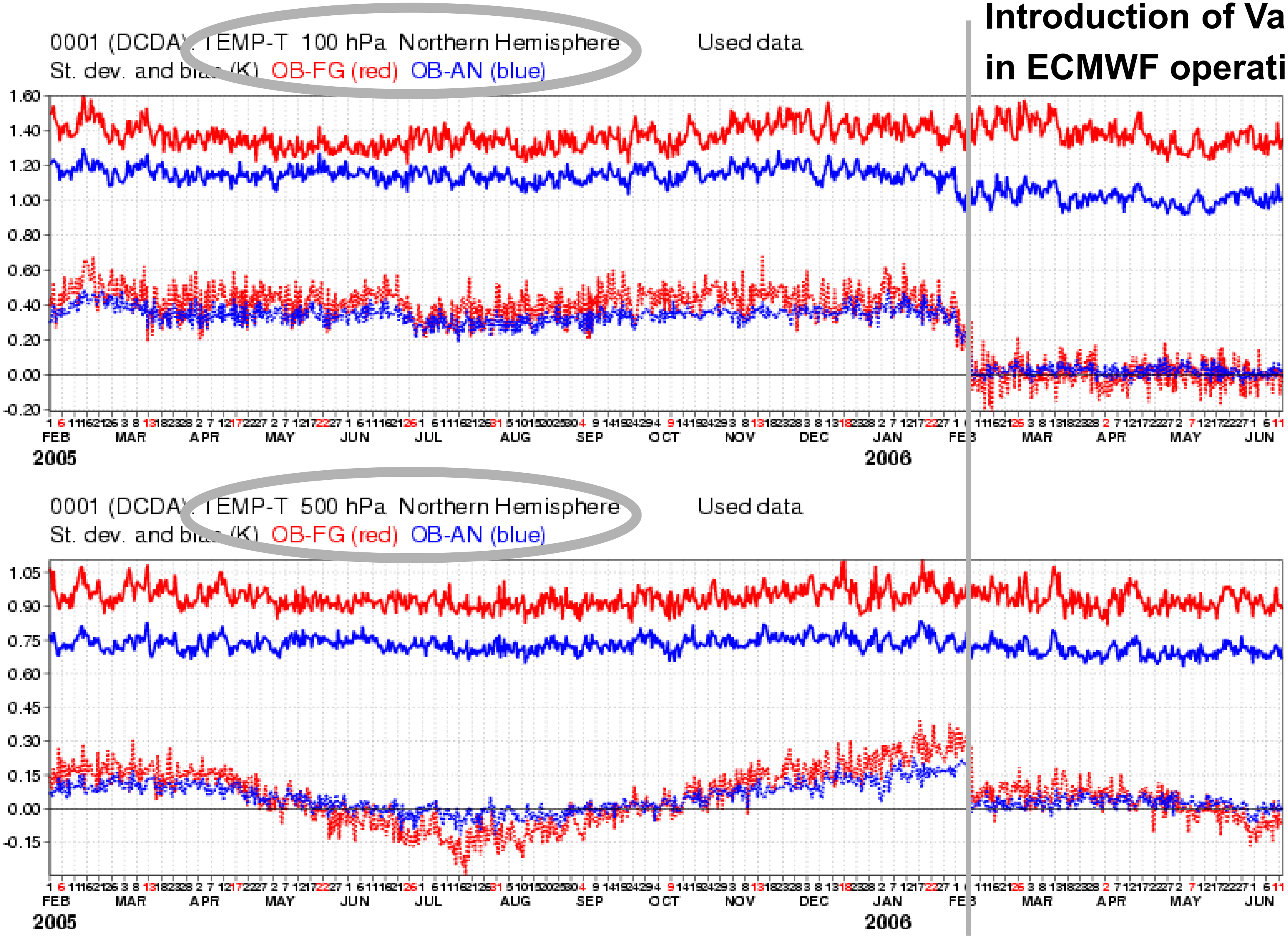


200 hPa temperature departures from radiosonde observations



Performance:
Fit to conventional data

Introduction of VarBC in ECMWF operations



Outline

- Introduction
  - Biases in models, observations, and observation operators
  - Implications for data assimilation
- Variational analysis and correction of observation bias
  - The need for an adaptive system
  - Variational bias correction (VarBC)
- Limitations of VarBC
  - Interaction with model bias
  - Assimilation in the upper stratosphere
- Summary

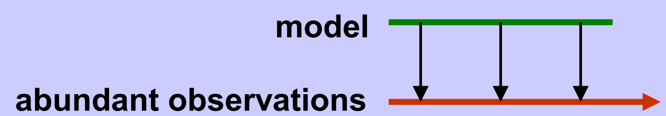
Limitations of VarBC:
Interaction with model bias

VarBC introduces some extra degrees of freedom in the analysis, to help improve the fit to the (bias-corrected) observations:

$$J(\mathbf{x},\beta) = (\mathbf{x}_b - \mathbf{x})^T \mathbf{B}_x^{-1} (\mathbf{x}_b - \mathbf{x}) + (\beta_b - \beta)^T \mathbf{B}_\beta^{-1} (\beta_b - \beta) + [\mathbf{y} - b(\mathbf{x},\beta) - h(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - b(\mathbf{x},\beta) - h(\mathbf{x})]$$

This works well where

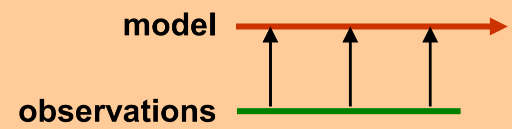
- the analysis is well-constrained by observations, and
- "anchoring" observations are available (e.g., radiosondes, GPSRO data).



VarBC will correct any biased observations and produce a consistent consensus analysis.

This may lead to undesired effects where

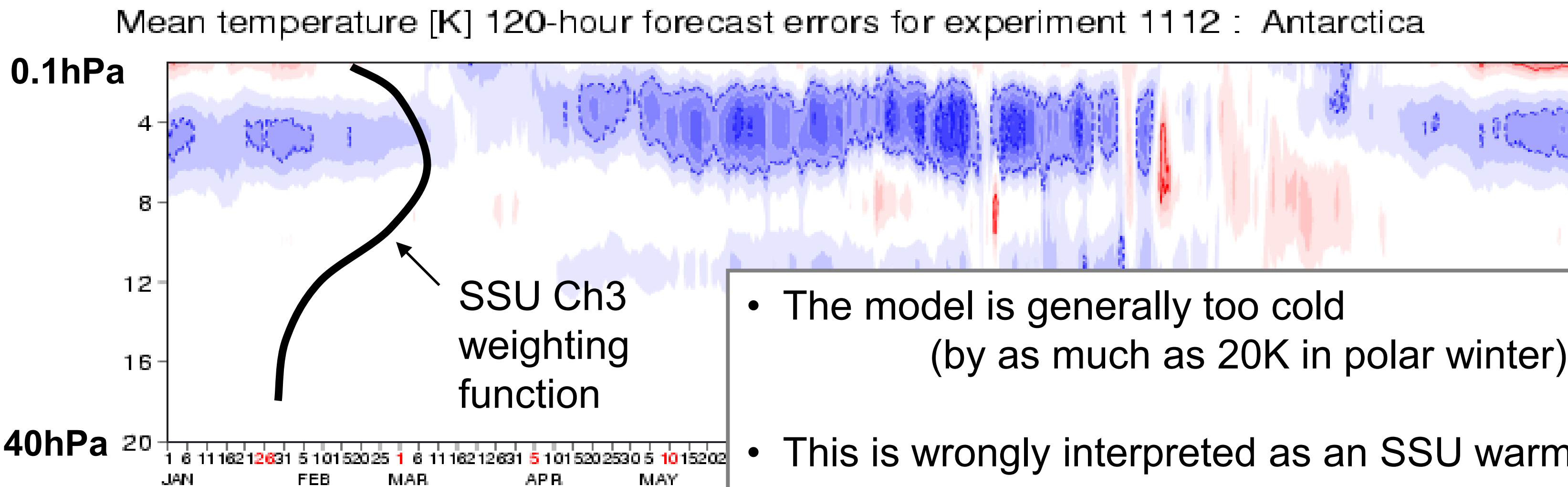
- model bias is present, and
- few observations are available, or
- only observations with VarBC are present.



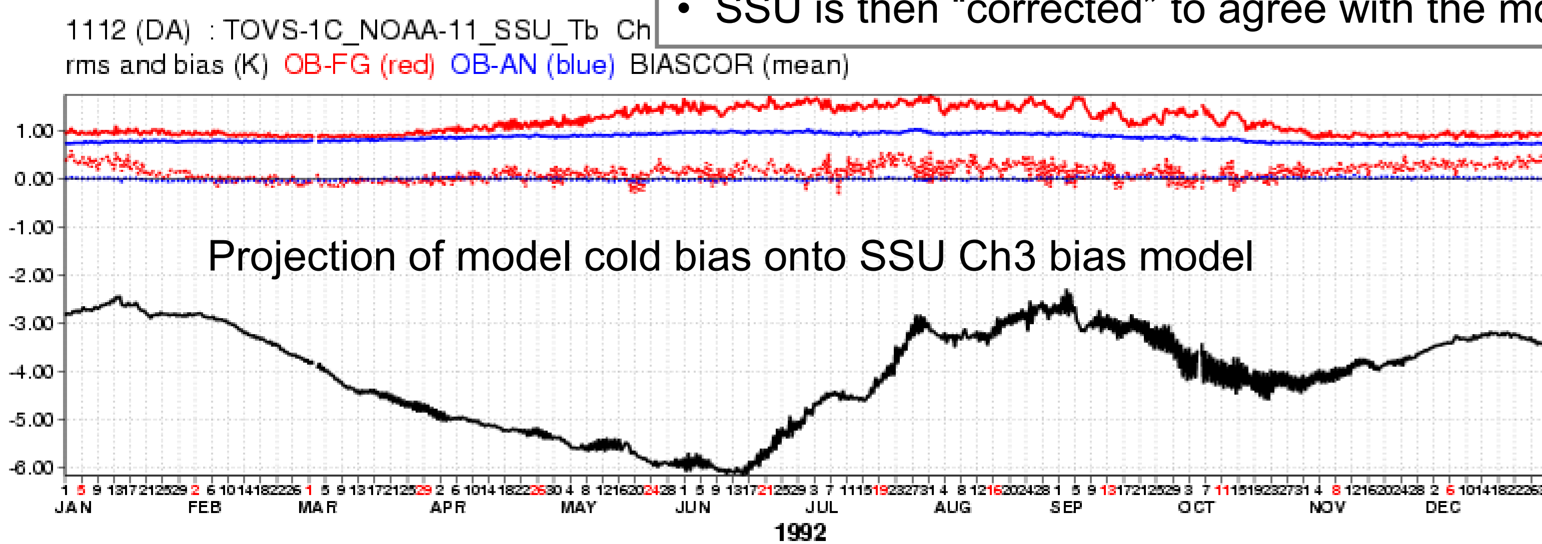
VarBC will, over time, force agreement with the model background.

VarBC may wrongly attribute model bias to the observations

Limitation of VarBC: Interaction with model bias



- The model is generally too cold (by as much as 20K in polar winter)
- This is wrongly interpreted as an SSU warm bias
- SSU is then “corrected” to agree with the model

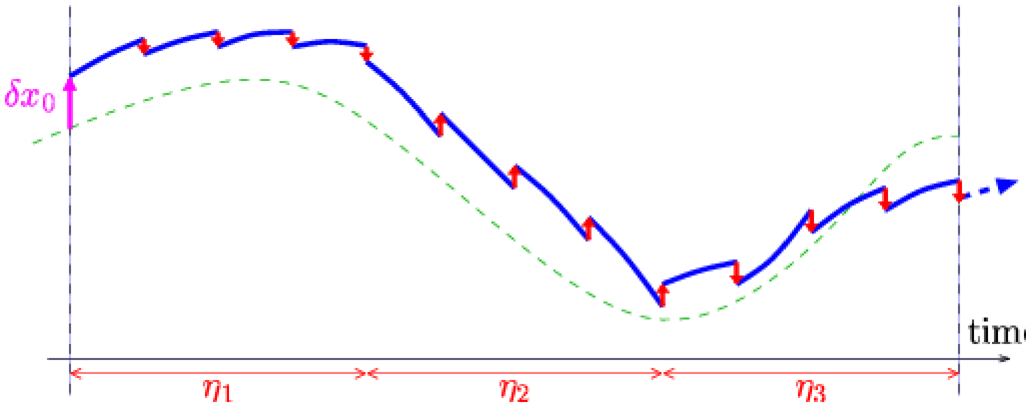


Weak-constraint 4D-Var (Y. Trémolet)

Include model error in the control vector:

$$J(x_0,\eta) = \frac{1}{2}\sum_{i=0}^{n}[\mathcal{H}(x_i)-y_i]^T R_i^{-1}[\mathcal{H}(x_i)-y_i] + \frac{1}{2}(x_0-x_b)^T B^{-1}(x_0-x_b) + \eta^T Q^{-1}\eta$$

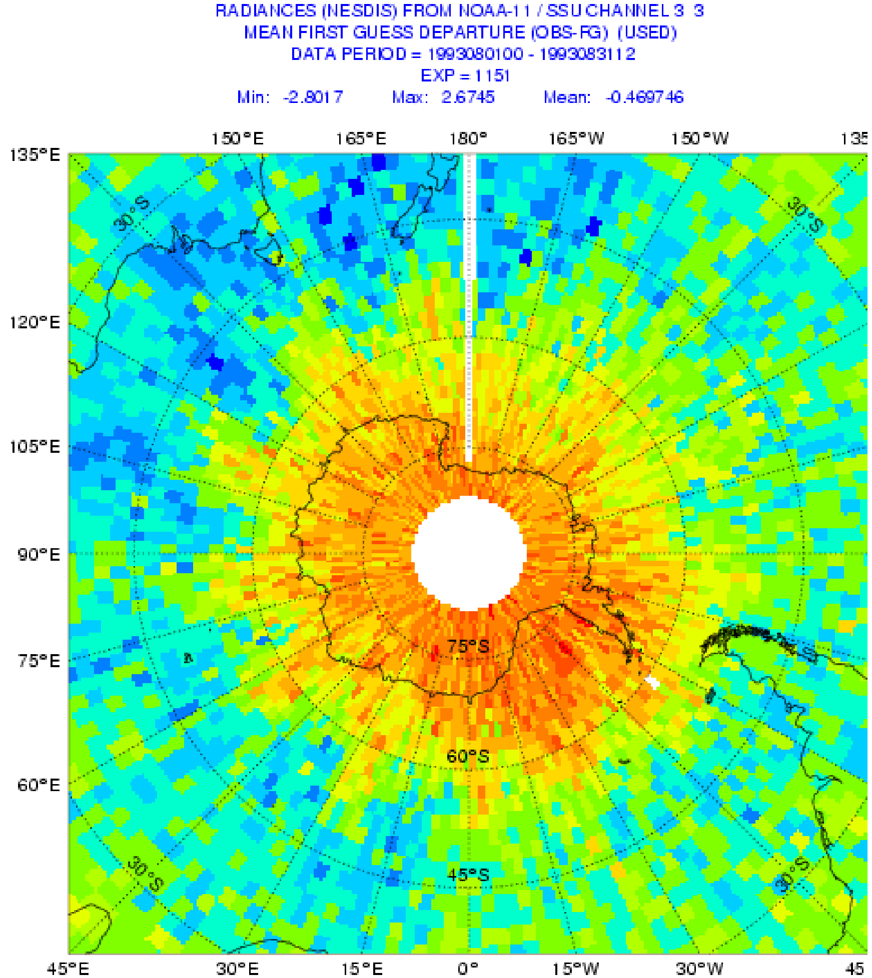
with $x_i = \mathcal{M}_i(x_{i-1}) + \eta_i$.



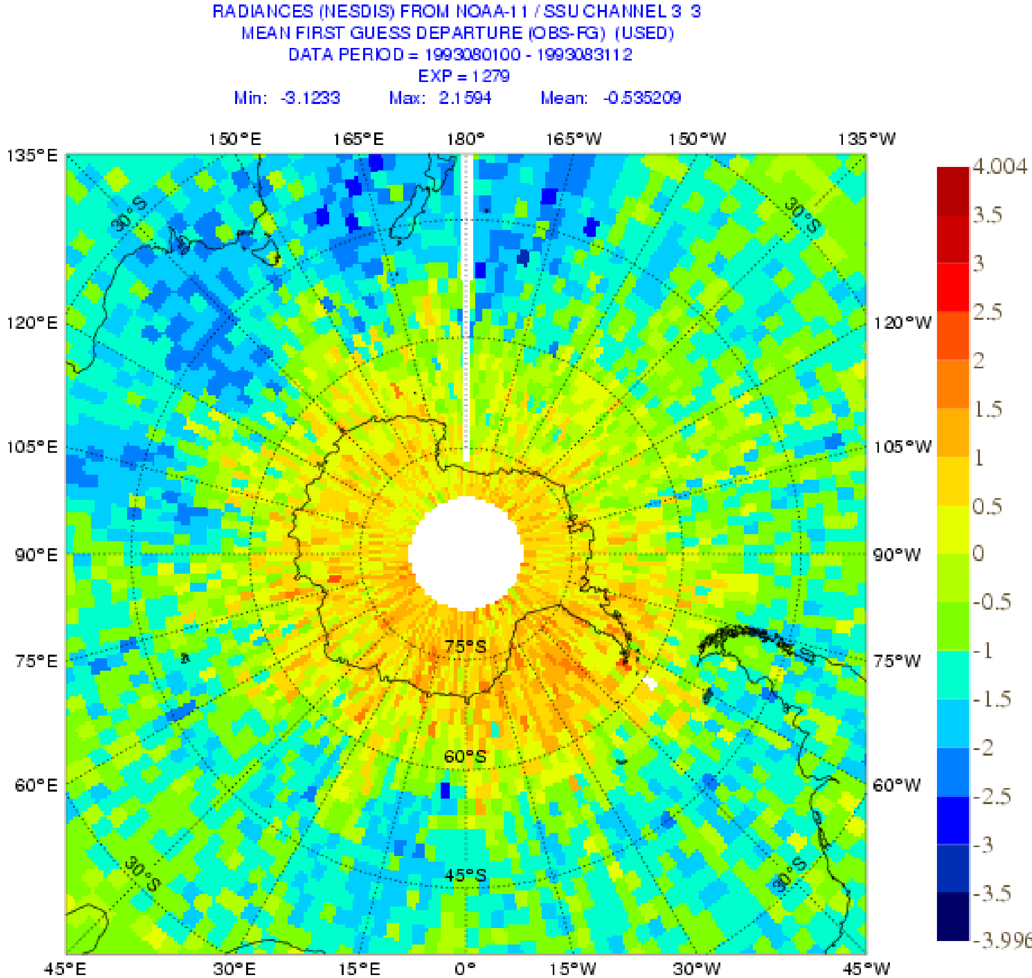
Constraint is determined by Q: for example, stratosphere only

SSU Ch3 mean radiance departures – Aug 1993

ERA-Interim



ERA-Interim + weak constraint



Summary

Biases are everywhere:

- Many observations cannot be usefully assimilated without bias corrections
- Manual bias correction for satellite data is no longer feasible
- Bias parameters can be estimated and adjusted during the assimilation
- VarBC works well in situations where
    - there is sufficient redundancy in the data; or
    - there are no large model biases

Some questions:

- How best to represent observation bias with a few parameters?
- Should VarBC be applied to non-radiance data as well?
- How much fixed (unbiased) information does the system need?
- How best to handle model bias in data assimilation?

Some references and additional information

Harris and Kelly, 2001: **A satellite radiance-bias correction scheme for data assimilation.** *Q. J. R. Meteorol. Soc., 127, 1453-1468*

Derber and Wu, 1998: **The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system.** *Mon. Wea. Rev., 126, 2287-2299*

Dee, 2004: **Variational bias correction of radiance data in the ECMWF system.** *Pp. 97-112 in Proceedings of the ECMWF workshop on assimilation of high spectral resolution sounders in NWP, 28 June-1 July 2004, Reading, UK*

Dee, 2005: **Bias and data assimilation.** *Q. J. R. Meteorol. Soc., 131, 3323-3343*

Dee and Uppala, 2009: **Variational bias correction of satellite radiance data in the ERA-Interim reanalysis.** *Q. J. R. Meteorol. Soc., 135, 1830-1841*

Feel free to contact me with questions:

Dick.Dee@ecmwf.int