## **Bias in Reanalysis**

Climate Change

**Bill Bell<sup>1</sup>**, Hans Hersbach<sup>1</sup>, Paul Berrisford<sup>1</sup>, András Horányi<sup>1</sup>, Joaquin Muñoz Sabater<sup>1</sup>, Julien Nicolas<sup>1</sup>, Paul Poli<sup>1</sup>, Raluca Radu<sup>1</sup>, Dinand Schepers<sup>1</sup>, Adrian Simmons<sup>1</sup>, Cornel Soci<sup>1</sup>,

Patrick Laloyaux<sup>1</sup>, Inna Polichtchouk<sup>1</sup>, Robin Hogan<sup>1</sup>,

Andrzej Klonecki<sup>2</sup>, Jon Mittaz<sup>3</sup>, Emma Turner<sup>4</sup>

[1] ECMWF, Reading, UK[2] Spascia, Toulouse, France[3] University of Reading, Reading, UK[4] Met Office, Exeter, UK











In reanalyses, we aim for :

- continuity
- accuracy in synoptics and climate
- well characterised uncertainties

**Biases** present challenges for all three aspects

We mitigate biases through:

## 'Pragmatic' approaches

- **observation domain** *e.g.* VarBC
- model domain e.g. weak constraint 4D-Var

'Ideal' approaches

 improved physical modelling of root causes



### Overview

Biases: How are they manifested in reanalyses? What do we do about them? ... and what next?

## **Observational biases**

- Radiances in ERA-Interim / ERA5 & future prospects
- Other innovations

## Model biases

- ERA5 stratospheric temperatures
- Weak constraint 4D-Var & model error forcing
- Using early sounding data (IRIS in 1970) to evaluate model error correction strategies

Mean state uncertainties - The systematic component

**Summary / Future Perspectives** 





## latitudinal bias (instrumental & forecast model contributions)

## Uncorrected First Guess Departures NPP ATMS-7 / K



### background errors

- bias ~ background error
  noise: **O**(0.1K)
- For early sounding obs, biases are larger
- DA methods assume observations unbiased wrt background/analysis
- Biases corrected using variational bias correction (VarBC)

Dee, QJRMS, 2005 Auligne et al, QJRMS, 2007



0.4

0.2

0

-0.2

-0.4

-0.6

-0.8





## Variational Bias Correction in ERA5 – bias model & typical corrections (example : ATMS-7)



Accounts for:

- Instrument errors (spectral, radiometric, ...)
- Forecast model errors
- RT model errors

Expect that in time, as instruments & models improve & datasets are reprocessed:

- The (mean) amplitude of bias corrections reduce; and
- The variance of the bias corrections reduce
- <u>Eventually</u> the corrections are (i) small & (ii) bounded by the uncertainties



# Instrument biases in the temperature sounding channels of MSU, AMSU-A and ATMS in ERA5 and ERA-Interim



## MSU-3 / AMSU-A7 / ATMS-8 Bias Corrections

Clima 2  $\checkmark$ Chang correction bias ERA Interim Bias corrections (**bold**) ± STDEV of bias corrections (grey) 0 ERA5 envelope Difference in bias corrections BC STDEV of ERA5 bias corrections Ь 1980 1985 1990 1995 2000 2005 2010 2015 ERA5 – ERA Interim ERA5.1 – ERA Interim

MSU-3 / AMSU-A7 / ATMS-8 (54.96 / 54.94 / 54.94 GHz) T –sounding, w.fn. peak at 270 hPa

- Improvements MSU > AMSU-A > ATMS (FY-3 MWTS / MWTS-2, Metop-SG MWS ?)
- Little change from ERA-Interim to ERA5
- Suggests model bias and RT related biases are less significant than instrument biases
- MSU & AMSU-A possible mechanisms identified:
- Radiometer non-linearities. Zou *et al* (JTECH, 2010)
- Spectral shifts. Zou *et al* (JGR, 2011), Lu and Bell (JTECH, 2013)

pernici

 But disappointing results in NWP testing so far (for AMSU-A, Lupu *et al*, ECMWF TM 770, 2016)



## MSU-2 / -3 / -4 Bias Corrections



- Similar picture (to MSU-3) for MSU-2 and MSU-4.
- Changes in bias correction *wrt* ERA-Interim are generally small, with the exception of:

pernicus

European Commission

- Aqua AMSU-A 2003-2016
- The period from 2000 2006 (fixed in ERA5.1) see next few slides
- Largest discrepancies AMSU-9 (0.5K), but still detectable in AMSU-7 and AMSU-5

## Forward look: Re-characterisation of the MSU instruments NOAA-6 to NOAA-14

Т

(Emma Turner, Met Office)

Climate Change



- Largest sensitivity of bias to spectral shifts expected in MSU channel 3
- Simultaneous estimation of non-linearity and spectral shifts carried out – with several calibration models
- Plan to evaluate the impact of the new data / new . RT modelling in advance of ERA6 back extension.



Commission



Radiative Transfer Model biases in the IR sounders (HIRS, AIRS, IASI, CrIS and SSU)



## Improvements in RT modelling: HIRS Temperature Sounding Channels 2 - 7





## Improvements in RT modelling: bias corrections for Adv. IR Sounders in ERA5

Change



- AIRS, IASI and CrIS channels shown at ~14µm (710 - 717 cm<sup>-1</sup>) & peak in the range 430 - 480 hPa
- AIRS & IASI: assume [CO<sub>2</sub>] = 376 ppm CrIS assumes [CO<sub>2</sub>] = 396 ppm
- HIRS (& SSU & VTPR): assume time varying [CO<sub>2</sub>]



European

Commission

**CECMWE** 

#### Improved RT modelling : SSU Bias Corrections Climate SSU-2 SSU-1 SSU-3 Change peak 29km peak 38km Peak 44 km correction bias ERA-Interim uses SSU-3 as an -2 anchor until ATOVS (AMSU-14) -2 ERA5 uses SSU-3 as an anchor throughout × Minum -5 1985 1990 1995 2000 2005 1985 1990 1995 2000 2005 1980 1980 1985 1990 1995 2000 2005

Improved treatment of RT (cell pressure leaks) in ERA5 (Kobayashi, QJRMS, 2009):

- Reduced inter-satellite biases
- Reduced variance in bias corrections
- Reduced drift in biases (NOAA-7 during 1982-1985)





## Other innovations

- Reprocessing and data resue efforts for ERA6 (see Paul's presentation on Wednesday)
  - Conventional and satellite data
  - Copernicus Climate Change Service work & other work by agencies (EUMETSAT, ESA, NOAA/NASA, CMA)
- Improved bias models for ascending-descending / orbital / harmonic biases
  - for sounding radiances (Booton et al [2013], Bormann et al [2023])
  - also applied to microwave imager radiances



- Correction of biases related to thermal gradients in main mirror for Aeolus winds (Rennie at al, 2021)
- Constrained variational bias correction (Han and Bormann, ECMWF Tech Memo 782, 2016)





## Model biases in the upper troposphere and stratosphere



## The improved mean state for stratospheric temperature in ERA5.1

Climate Change



Monthly average observation-background differences from 1979 onwards for all assimilated bias-adjusted radiosonde temperature data (K) between 40 and 60 hPa, for ERA-Interim, ERA5 (based on 1979- $B_{cli}$  before 2000 and 41r2-B <sub>cli</sub> afterwards) and ERA5.1 (using 1979- $B_{cli}$  from 2000-2006).

Hersbach, H. et al., 2020, doi:10.1002/qj.3803

- ERA5.1 provides an improved mean state for stratospheric temperature.
- In the troposphere the difference between ERA5 and ERA5.1 is very small.

(see A. Simmons et al, ECMWF Tech Memo 859, Jan 2020)





## Model error manifested in biased first guess departures

## NOAA-18 AMSU-A8





## Climate Change

## Model Error / AMSU-A Mean First Guess Departures in ERA5



ERA5 mean first guess departures shown for AMSU-A

Error bars represent  $(\pm 1\sigma)$  spread over the lifetime of each sensor

Consistent picture of :

- a cold model bias mid-trop to mid-strat
- a (larger) warm model bias above 10 hPa

Broadly consistent with analysis increments in ERA5 (below, from Fig 16, Hersbach et al, 2020)



## Model Error / AMSU-A Mean First Guess Departures in ERA-Interim

Climate

Change





## Model Error / AMSU-A Mean First Guess Departures in proto-ERA6 testing (CY48R1)



- Based on JJA 2020 CY48R1 experiment
- Significant changes since ERA5 cycle (CY41R2, 2016):
  - Weak Constraint 4D-Var above 100 hPa
  - Improved dynamics: quintic interpolation
  - Clear sky -> all sky scheme for tropospheric channels
- Overall model in better agreement with observations







• General problems in reanalysis temperatures above 10 hPa well documented (see SPARC-RIP report 2021).



Climate Change

#### (2) Cold model bias in UTLS - exposed in early period



- Very few observational constraints on stratospheric temperature analysis in the early 1940s so UTLS cold bias is exposed.
- Analysis increments in 10-200 hPa layer very small 1940 (< 20mK above 100 hPa as a global mean)</li>



100-1 hPa







- General problem foreseen & analysed in Eyre (QJ, 2017): with VarBC, if radiances are dominant (cf anchors) model bias is reinforced
- VTPR channels 1 & 2 bias corrected using VarBC reinforcing model cold bias
- Despite clear benefits (from assimilating VTPR) in improving synoptic analysis mean state exhibits a discontinuity.
- VTPR exhibits significant radiometric and spectral errors ⇒ we need VarBC





#### (4) Impact of model cold biases 2000-2006





• ERA5 and ERA5.1: See previous slides





## Upper stratospheric biases in ERA5: Temperature anomalies relative to ERA5 climate



**IRIS** experiments

 Generally, ERA5 temperature analyses above 10 hPa exhibit biases and discontinuities

 Particularly large biases evident in southern polar winter (>> 6K in the plot shown)

100-1 hPa

3.6

2.4

1.2

0.0

-1.2

-2.4

- -3.6 - -4.8 -6.0

- Repeatable from year-to-year (before 1972)
- Reduced following the assimilation of VTPR data (Nov 1972 - Jan 1979)





## Investigating biases using ealry hyperspectral sounding data (Nimbus-4 IRIS, 1970)



- IRIS data has been shown to be valuable in improving SH analysis quality (April 1970 January 1971)
- Valuable for assessing biases in ERA5 in previously unobserved regions (eg S. Polar upper stratophere)
- Highest peaking channel is particularly valuable







Climate Change



• During the GNSS-RO era (2006 - ) the stratospheric temperature analysis is realistic





#### Climate Change





- During the GNSS-RO era (2006 ) the stratospheric temperature analysis is realistic
- In the early period (1940-75) of the reanalysis, few observations constrain the analysis ⇒ model biases are exposed. At 5hPa, temperatures are 10 – 25 K warmer in mid-winter, relative to 2006-2022





#### During the GNSS-RO era (2006 - ) the • stratospheric temperature analysis is realistic

- In the early period (1940-75) of the reanalysis, few observations constrain the analysis  $\Rightarrow$ model biases are exposed. At 5hPa, temperatures are 10 – 25 K warmer in mid-winter, relative to 2006-2022
- ERA5 (41R2, 2016) in 1970 is at the top end of this range, with temperatures of 230K in mid-winter



14130

14120

AU9 10

1970





280.0

During the GNSS-RO era (2006 - ) the • stratospheric temperature analysis is realistic

Impact of assimilating IRIS on S. polar stratospheric biases

- In the early period (1940-75) of the reanalysis, few observations constrain the analysis  $\Rightarrow$ model biases are exposed. At 5hPa, temperatures are 10 – 25 K warmer in mid-winter, relative to 2006-2022
- ERA5 (41R2, 2016) in 1970 is at the top end of this range, with temperatures of 230K in mid-winter
- The CONTROL (48R1, 2022) exhibits the same warm bias







280.0

During the GNSS-RO era (2006 - ) the stratospheric temperature analysis is realistic

Impact of assimilating IRIS on S. polar stratospheric biases

- In the early period (1940-75) of the reanalysis, few observations constrain the analysis  $\Rightarrow$ model biases are exposed. At 5hPa, temperatures are 10 – 25 K warmer in mid-winter, relative to 2006-2022
- ERA5 (41R2, 2016) in 1970 is at the top end of this range, with temperatures of 230K in mid-winter
- The CONTROL (48R1, 2022) exhibits the same warm bias
- Assimilating IRIS gradually brings temperatures to more realistic values. Note: increase (*I*) from 16<sup>th</sup>-24<sup>th</sup> July is associated with an outage of IRIS observations







280.0

During the GNSS-RO era (2006 - ) the stratospheric temperature analysis is realistic

Impact of assimilating IRIS on S. polar stratospheric biases

- In the early period (1940-75) of the reanalysis, few observations constrain the analysis  $\Rightarrow$ model biases are exposed. At 5hPa, temperatures are 10 – 25 K warmer in mid-winter, relative to 2006-2022
- ERA5 (41R2, 2016) in 1970 is at the top end of this range, with temperatures of 230K in mid-winter
- The **CONTROL** (48R1, 2022) exhibits the same warm bias
- **Assimilating IRIS** gradually brings temperatures to more realistic values. Note: increase (*I*) from 16<sup>th</sup>-24<sup>th</sup> July is associated with an outage of IRIS observations
- Using **Constrained VarBC** (Han & Bormann) reduces the bias absorbed by VarBC, and accelerates cooling of the analysis towards more realistic values.



## Standard 4D-Var formulation

Climate Change

4D-Var is a common algorithm to find the optimal initial state by minimising the discrepancies with the prior estimate and the observations



Model's equation

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

4D-Var cost function

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

- → Standard formulation assumes that the model is perfect
- → A model trajectory is entirely determined by its initial condition





## Weak-constraint 4D-Var formulation

We assume that the model is not perfect, adding an error term  $\eta$  in the model equation

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

The model error estimate η contains 3 physical fields (temperature, vorticity and divergence)



→ Introduce additional degrees of freedom to fit background and observations

→ A model trajectory is entirely determined by its initial condition and the model error forcing

➔ Concept of scale separation introduced between background and model errors

 $\rightarrow$  Constant model error forcing over the assimilation window

Laloyaux et al., Exploring the potential and limitations of weak-constraint 4D-Var, 2020





## Weak-constraint 4D-Var formulation

We assume that the model is not perfect, adding an error term  $\eta$  in the model equation

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

The model error estimate  $\eta$  contains 3 physical fields (temperature, vorticity and divergence)

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



→ Introduce additional degrees of freedom to fit background and observations

- → A model trajectory is entirely determined by its initial condition and the model error forcing
- ➔ Concept of scale separation introduced between background and model errors
- $\rightarrow$  Constant model error forcing over the assimilation window

Laloyaux et al., Exploring the potential and limitations of weak-constraint 4D-Var, 2020







## Model error climatology derived from weak constraint 4D-Var estimates of model error



## Model error forcing experiments in 1970 – impact on upper stratospheric temperatures

Climate Change



- For strong constraint & model error forcing experiments: increase in resolution (28km to 9km) helps lower minimum temperatures (230K->223K in June 1970)
- Model error forcing (both types) results in additional cooling of ~5K, with minimum temperatures of 217K
- ... but doesn't bring temperatures to the minimum temperatures expected (from IRIS assimilation experiments) of ~210K
- expect ERA6 (TCo799) will be closer to behavior of TCo1279 experiment shown here.



## Verification of impacts of MEF: background fits to IRIS and radiosondes





Background fits to radiosonde temperatures  $20^{th}$  April –  $26^{th}$  August 1970





## Verification of impacts of MEF: background fits to IRIS and radiosondes

0



Background fits to radiosonde temperatures 20<sup>th</sup> April – 26<sup>th</sup> August 1970



- NN MEF improves bias and synoptic performance
- IRIS provides unique insight into biases in otherwise observation sparse domains
- But significant biases remain (work in progress)

10 <sup>-2</sup>

10 -1

10<sup>0</sup>

10

10

/ hPa

Pressure

## Prospects for reducing model biases in the stratosphere

Climate



Improvements are anticipated from more accurate physical modelling, including:

[1] revised radiation scheme

[2] improved dynamical core

[3] reduction of H<sub>2</sub>O in lower stratosphere

[4] Improved representation of GWD







•

•



- **Uncertainties are derived from an inspection / understanding of the system** rather than ascribed after comparison with independent observations
- **Validation** should, ideally, involve a comparison of independent estimates, each associated with it's own independent uncertainty estimate





## Possible approaches to determining mean-state uncertainty

#### The observing system component

- Defined here as "uncertainty in mean state arising from uncorrected biases in the observing system & choice of observing system configuration"
- OSEs with different plausible configurations of observing system, for each epoch
- Simplest approach: withdraw 'redundant' components of observing system and evaluate change in the mean state (next slide)
- Other factors: choice of observational data, bias model, QC/thinning, observation errors, ...





## Possible approaches to determining mean-state uncertainty

#### The model component

- Defined here as "uncertainty in mean state arising from uncertain model parameters and forcings"
- Changes in time, due to the changing observing system
- OSEs with perturbed model parameters & alternative choices of forcings
- Key model parameters? draw upon experience of EPS and climate modelling communities
- Sample time dependence using paired down modern observing system, or run in past epochs



 Perturbed by magnitudes consistent with documented uncertainties and/or giving rise to no significant degradation in forecast skill in OSEs





## Validating the mean state uncertainties

Several components of the observing system could be considered 'reference' quality:

- **GNSS-RO** direct traceability chain to time standards
- **GRUAN radiosondes** available post-2010 in numbers
- CrIS well characterised uncertainties
- **GMI** reference MW imager mission

Use (a subset of) these observations passively (*i.e.* withhold from the analysis) to assess the uncertainty estimates from a **benchmark** period in the ERA6 reanalysis (~ 2010-2020, or 2015-2025)

*Benchmark\** - defined in this context as " associated with robustly defined uncertainties ideally validated through comparison with traceable independent measurements "

(borrowed from CLARREO, TRUTHS mission concepts)





### Summary

- In the treatment of biases many steps forward at ERA5 (RT model biases in IR), some sideways steps (MSU/AMSUA) and some backward steps (stratospheric biases).
- In the short/medium term the prospects are good for improved pragmatic correction (WC 4D-Var / WC 4D-Var MEF) as well as corrections at source (reprocessed data [Paul's talk], stratospheric model biases, improved RT modelling)
- Should we use the 'redundancy' of the very recent satellite era (~2010-2020) to withhold some (subset of) very high quality observations (GRUAN, RO, CrIS, ...) and use these to independently validate ERA6 during a **benchmark** period in the reanalysis (at the cost of a small degradation in analysis quality) as a first step towards methods for establishing the full uncertainty budget for reanalyses products ?





## Thanks for listening !

