Working Group 1: Treatment of biases

Chaired by Paul Poli (EUMETSAT) and Patrick Laloyaux (ECMWF)

Blocks of questions
Diagnosing biases
Bias sources
Reference data
Bias correction methods
NWP vs climate applications of data assimilation

Participants	
Mark Buehner	Dinand Schepers
James While	Bill Bell
Sergey Frolov	Roger Randriamampianina
Dick Dee	Devon Francis
Amal El Akkraoui	Xavier Calbet
Dave Tobin	Paola Corrales
Qiwen Sun	Christopher Merchant
Xavier Calbet	Christina Köpken-Watts
Massimo Bonavita	Timo Hanschmann
Mazlar Bani Shahabadi	Viju John
David Tobin	Rob Roebeling

Diagnosing biases

What is the current chain of events to establish whether biases in a new instrument/observing system are acceptable (or not), for assimilation?

- Monitor data against operational model, statistics of observation departures:
  - check (ideally) 'zero bias', gaussian distribution, bias smaller than standard deviation
  - Impact of forecast scores. No urgency to tackle biases, unless bias corr. isn't working.
- Exchanges between NWP centers and data providers vary. Some frameworks do exist (e.g. International TOVS working group, Cal/Val communities), but are the NWP centers asking the right/specific questions: e.g., 'can you tell me what, in your instrument design, could lead to potential biases?' ➔ Target new discussions on biases in these fora?

What tools and data are used for bias diagnostics? What's missing?

- Loads of statistics, but they typically miss the underlying physical context (cf. comment above), to bring more understanding (in observation and model biases)
  ➔ Open-access databases of statistics, and observation vs model underlying data?
  (…possible link between NWP+GSICS)
- Knowing what could lead to biases (nonlinearity etc…) bias predictors must be revisited
  ➔ Internationally-coordinated study to revisit bias predictors? (different NWP models etc…)

Are there diagnostic methods to separate sources of biases?

- Offline approach could help to better disentangle sources
- Distinguish the sources feasible only if (spatio-temporal) patterns differ
- Importance of anchor observations and reference observations (e.g. GRUAN)

Bias sources

*How may we progress with gathering and exchanging knowledge on **instrument biases**?*

- Instead of 'physical intuition', we need instrument bias simulations, constructed from the measurement equations and tables of effects (e.g. FIDUCEO)
- Easier to do with new sensors; case studies relevant for historical climate sensors (e.g. MSU)
  ➔ Request such simulations be carried out/shared for new missions (e.g. IASI-NG)?
  ➔ Revisit selected climate missions
- Metadata exchange, ideally via RTTOV, e.g. for Spectral Response Function uncertainties
  ➔ Develop RTTOV to propagate some of the instrumental design uncertainties (e.g. SRF or central frequency shift…)

*How may we progress with gathering and exchanging knowledge on **representation biases**, e.g. including observation operators, radiative transfer models, turbulence…?*

- Using high-resolution measurements (e.g. Lidar) and very-high-resolution models to simulate/quantify representativeness impact on biases
- Using LBLRTM to quantify spectroscopy/fast RT impact on biases
- Use then OSSE, with all sources included, to check error budget closure
  ➔ Organize & fund 3 classes of studies above

Reference data

What data are available to serve as references in a data assimilation system?

- GNSSRO, Radiosondes (even though bias-corrected at ECMWF, even GRUAN?), Argo,…
- Forcings (generally uncorrected… but based on observations also, e.g. SST/sea-ice)
- Hyperspectral IR, at least for the temperature channels (humidity channels more challenging because of spectroscopy and representation) ➔ Try out cVARBC for these channels

Would we need more / others?

- Temporal and spatial resolution is limited for NRT 'reference' data
- Observation campaigns may be useful sources, unsufficiently used?
  ➔ Review what obs. campaigns would support revisiting biases in near-recent instruments
- Sparsity is an issue to anchor a system
- We would especially need extra reference data for tropospheric humidity
- For high-altitude (atm.) and higher-depth (ocean), we need more anchor data (mesosphere and deep Argo); Increases in high-resolution, coupling call for more surface flux data
  ➔ Is this all captured in the WMO/GCOS requirements?

Are their present error characterizations sufficient?

- No, we do not know from the data a priori if they need correction or not until testing, in a data-driven world we'd need to know what is the uncertainty beforehand
- In practice, traceable uncertainties are needed for all sources which pretend to be treated as 'references' ➔ Traceable uncertainties needed for GNSSRO and Hyperspectral IR
- One issue is that we've accepted to take on biased observations for too long.

Bias correction methods

How do we specify that some data are 'references' in a data assimilation system?

- By default any observation dataset is considered as a reference
- cVARBC is one algorithm possible to choose how much we want to control the bias (from fully-controlled bias to completely unknown bias)
- Using the uncertainty estimates (ideally) produced (earlier), for ref. data, we could use this information in cVARBC. ➔ Try out cVARBC for instruments with known uncertainties

*How do we account for **observation biases** in data assimilation (e.g. VARBC), and what improvements do we need? [a) constraints, b) non-convergence cases, c) background error term, d) bias predictor choices, other...?]*

- Sharing varbc correction and model bias definition between centres ➔ See open-access databases
- Realizing an ensemble of corrections could be one way to advance. (ensVARBC)
- Offline approach (anchor-only DA system) provides a way to get more control
- Background error term in VarBC, revisit with recent work at NCEP

*How do we account for **model biases** in data assimilation, and what improvements do we need?*

- There are two approaches at the moment: weak-constraint 4D-Var (ECMWF) and 3D-Var anchor-only analysis (ECCC)
- Weak-constraint is not necessarily limited to 4DVAR but could be used in 3DFGAT
- Departures for anchor observations could be used to identify better the model error

NWP vs climate applications of data assimilation

Do we have a baseline for today's biases in data assimilation systems? (NWP & reanalysis) Should we attempt to do better?

- No – and biases can be large (cf. user feedback). A reanalysis of the present times based on best observations would provide a benchmark/starting point for the future & the past. This benchmark should be uncertainty-characterized. Two lines of work needed: model stochastics with perturbations of key parameters, and analysis with varying amounts of obs.

How can we improve representation of climate signals in reanalysis?

- Link with CMIP protocols (for the runs and output) would be desirable in the medium-term. Model-only runs are absolutely essential in reanalysis, to guide users in the origin of possible biases. Low-resolution runs with varying model parametrizations would help explore, gradually, the space of model biases.
- Minimal international coordination could be done on the baseline of agreed years for such short productions.

Is there a case for different approaches in NWP and reanalysis regarding biases?

- There's a tension between reanalysis striving to deliver unbiased solutions for the past, and initial conditions for reforecasts (for seasonal forecasting). The tie of reanalysis to NWP makes it impossible to realize both perfectly at the same time.
- One trade-off would be to deny some key datasets retained for validation. Generally, reanalyses allow to withdraw whole components of the observing system (or fractions of). This feasibility is not explored enough today, in a systematic fashion. The forcings should also be perturbed.

What lessons have we learnt from reanalysis for NWP regarding biases, and vice-versa?

- Better understanding of model biases (e.g. ERA5 stratosphere).
- Help specify new missions, addressing identified past flaws. Reanalysis community should review the GCOS/WMO requirements and ensure their needs are covered. E.g. protection of frequencies for long-flying sensors.
- Observation feedback archive from reanalysis is still an unexplored resource.
- Regardless of other developments, in-orbit calibration missions (TRUTH/CLARREO) remain needed.