Climate model evaluation using GPS-RO data

Mark Ringer, Met Office Hadley Centre

ROM-SAF workshop, ECMWF, 16-18 June 2014

Outline

- Intro – using satellite data for model evaluation
- Evaluation of the new Hadley Centre model, including comparisons with ROM-SAF products
- Forward modelling and satellite simulators
- Use of uncertainties
- Conclusions, questions for climate working group

Why do we use satellite data?

- Evaluate physical processes relevant to reducing uncertainty in climate predictions
- Inform & prioritise key areas for developing and improving climate models
- Constrain climate change predictions – or at least try and determine if this is possible
- Detection & attribution of observed variations to natural and anthropogenic climate forcing
- Initialisation of models used for seasonal-to-decadal prediction

How do we use the data?

• "Traditional" method: comparison of high-level products – temperature, humidity, cloud amounts, etc – with their model equivalents

• "Model-to-satellite" approach: simulate what is actually measured, IR or microwave radiances, radar reflectivities, RO bending angles

• Development of process-based evaluation techniques including combining with other information such as reanalyses: e.g. compositing in terms of dynamical regimes

• Apply similar techniques to analysis of climate change simulations and feedbacks

• For "fast" physical processes – e.g. clouds, precipitation – we can also use comparisons with the global NWP model

• In combination with *in situ* data from the global observing network: e.g. from aircraft and other field campaigns

Comments

- We are not just interested in long-term climatologies and simple comparisons with models…although we still of course do lots these!

- The focus on processes also includes variability on timescales from diurnal to interannual…and longer if possible

- Sometimes relatively short, e.g. ~2-3 years, of high-quality data are potentially useful

- Key aim is to provide information which can be used to improve the models, e.g. to develop better physical parameterizations

- We often invest much time & effort to use new, high quality data, e.g. development of satellite simulators…but there needs to be a clear demonstration of its novelty and utility

The wider context: analysis of multi-model ensembles (CMIP)

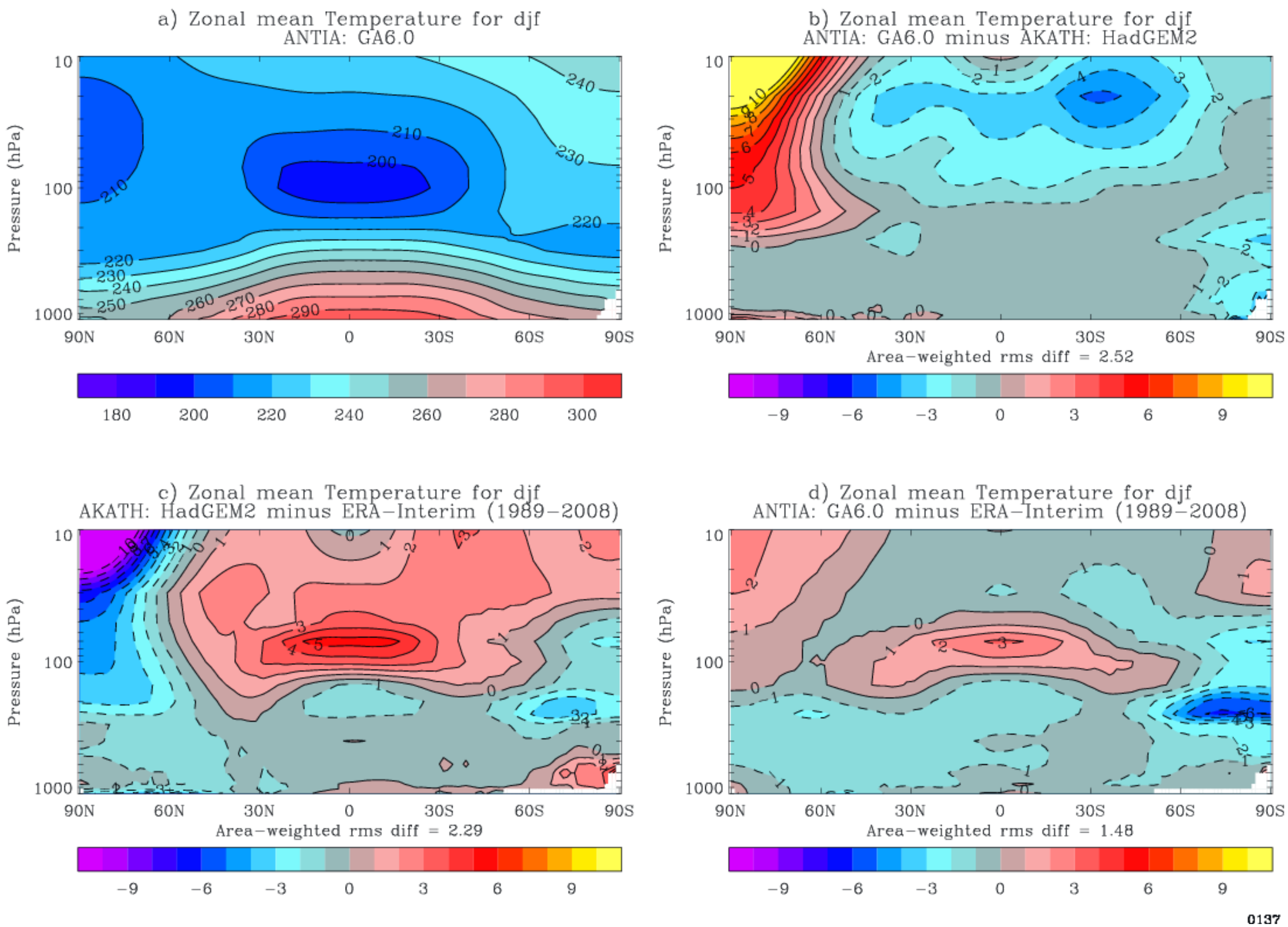
- Are climate models improving?
  - e.g. from CMIP3 → CMIP5
- Are some models demonstrably better than others?
- Are some quantities more robustly simulated than others?
- Do we see consistent strengths or weaknesses across the multi-model ensemble?
- Better comparison with obs → more reliable projections?
- Can we weight models based on comparisons with obs?
- Development of metrics, WGNE/WGCM metrics panel

http://www-metrics-panel.llnl.gov/wiki

HadGEM3 vs. HadGEM2

- A model for application across NWP, Seasonal, Decadal & Centennial timescales, and for regional prediction
- Includes new ocean (NEMO), sea-ice (CICE) and cloud (PC2) schemes, improved soil treatment, better representation of coastal regions, etc
- Hierarchy of models – at a range of horizontal resolutions – suitable for the different applications
- Is the basis of the next generation Earth System model, UKESM1, being developed with external partners
- Increased vertical resolution: 85 levels (vs. 38); lid at 85 km (vs. 39 km); 35 levels above 18 km

HadGEM3 vs. HadGEM2 Temperature profiles: DJF



HadGEM3
- HadGEM2

HadGEM2/3
- ERA

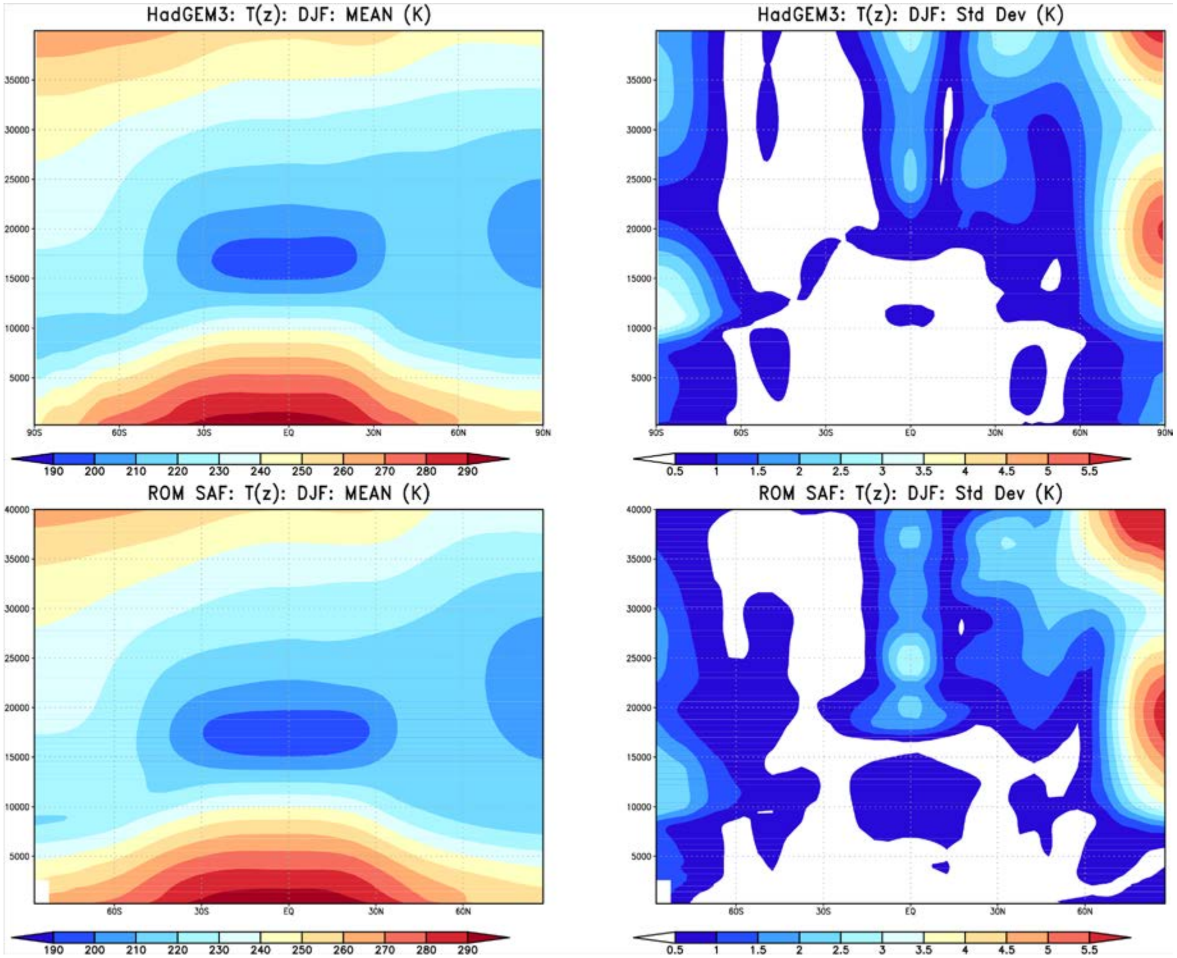
Met Office
Hadley Centre

Coupled model
Temperature profiles: DJF

Mean Std Dev

Model

Obs



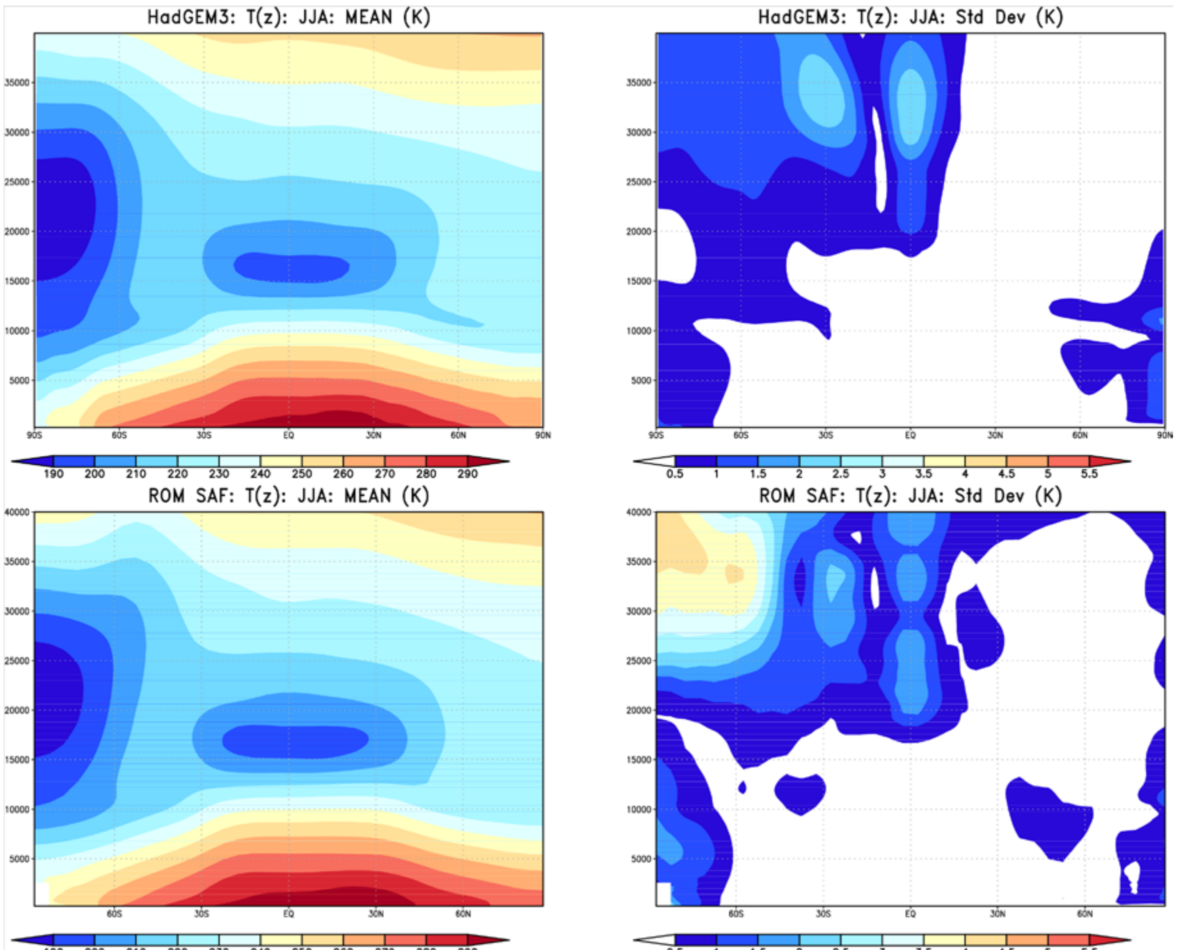
Coupled model Temperature profiles: JJA

Mean

Std Dev

Model

Obs

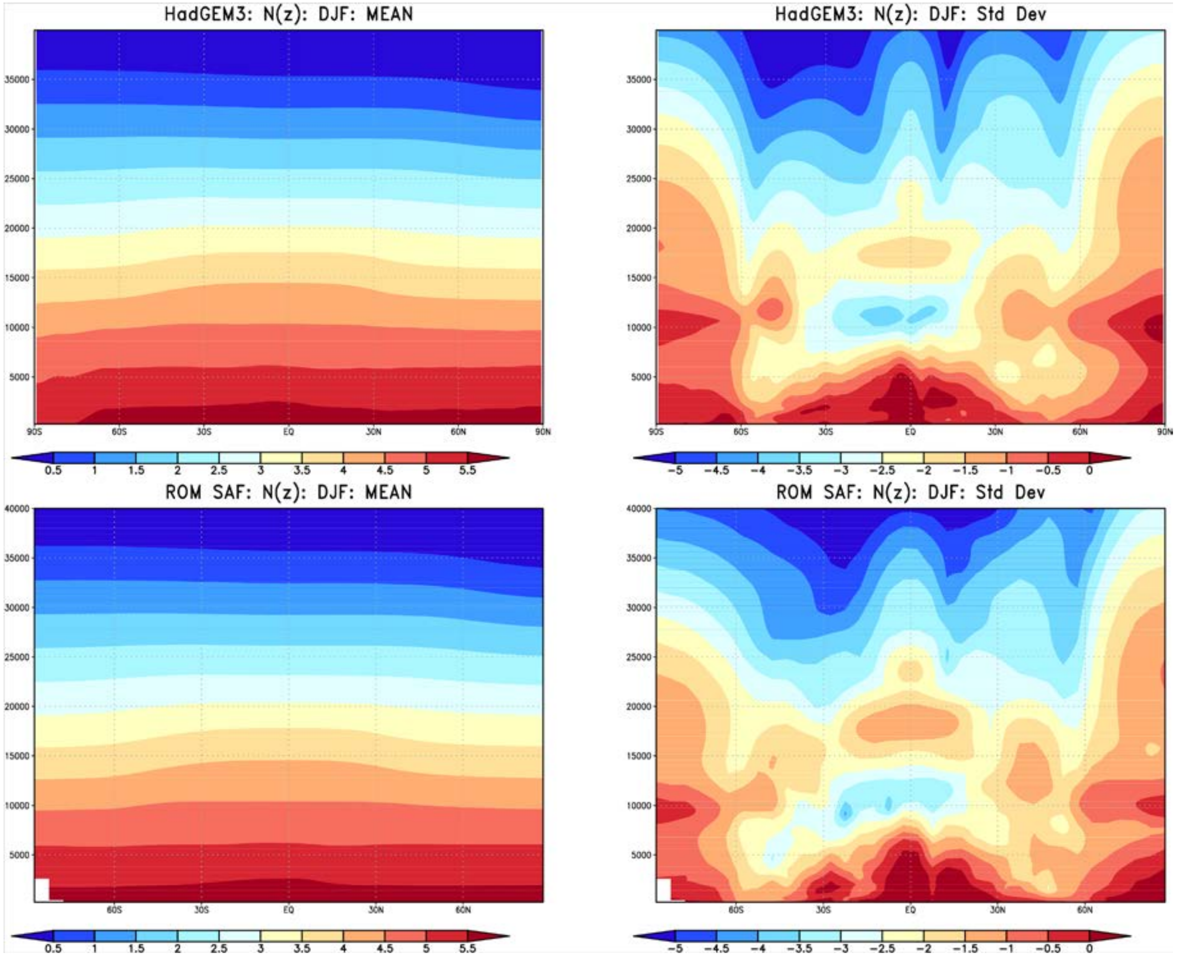


Coupled model
Refractivity profiles: DJF

Mean Std Dev

Model

Obs

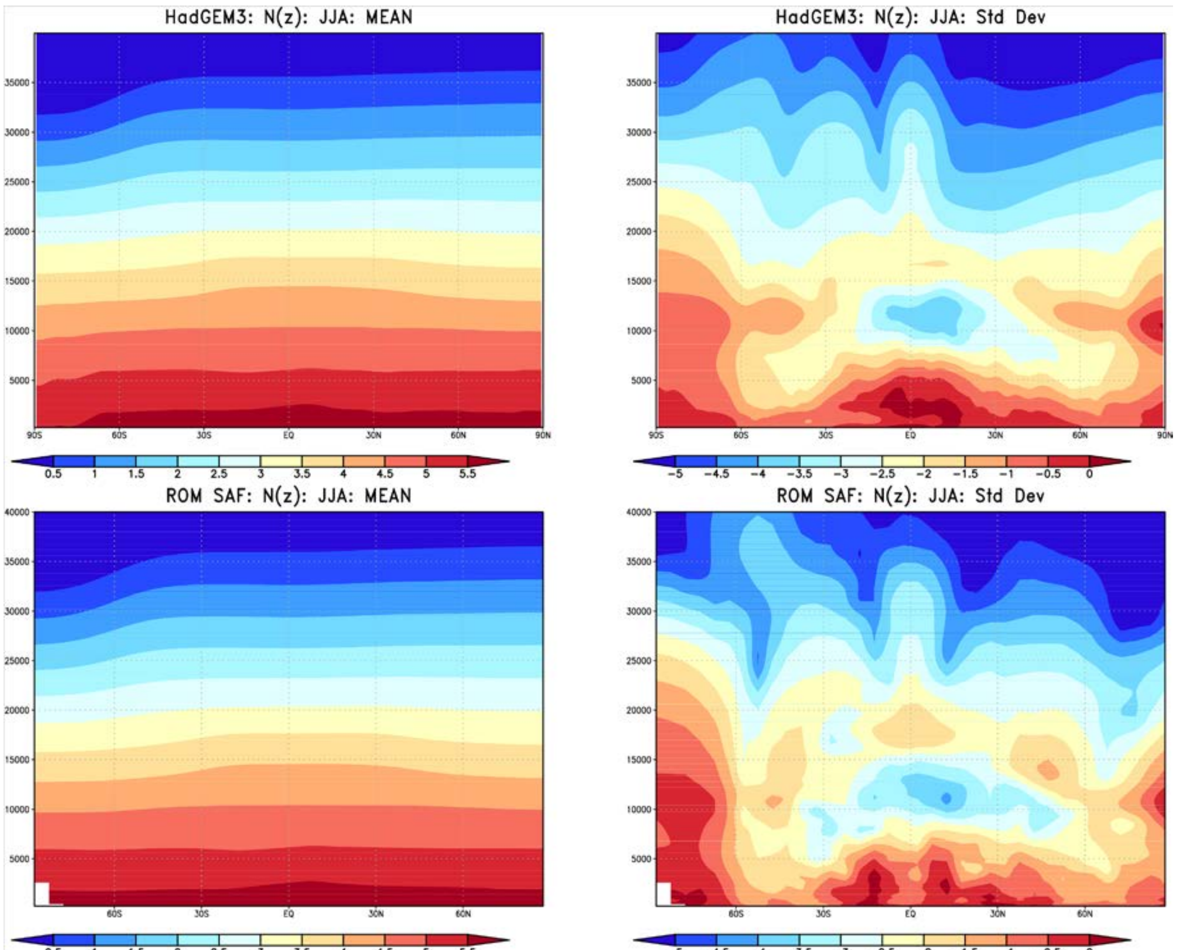


Coupled model
Refractivity profiles: JJA

Mean | Std Dev

Model

Obs

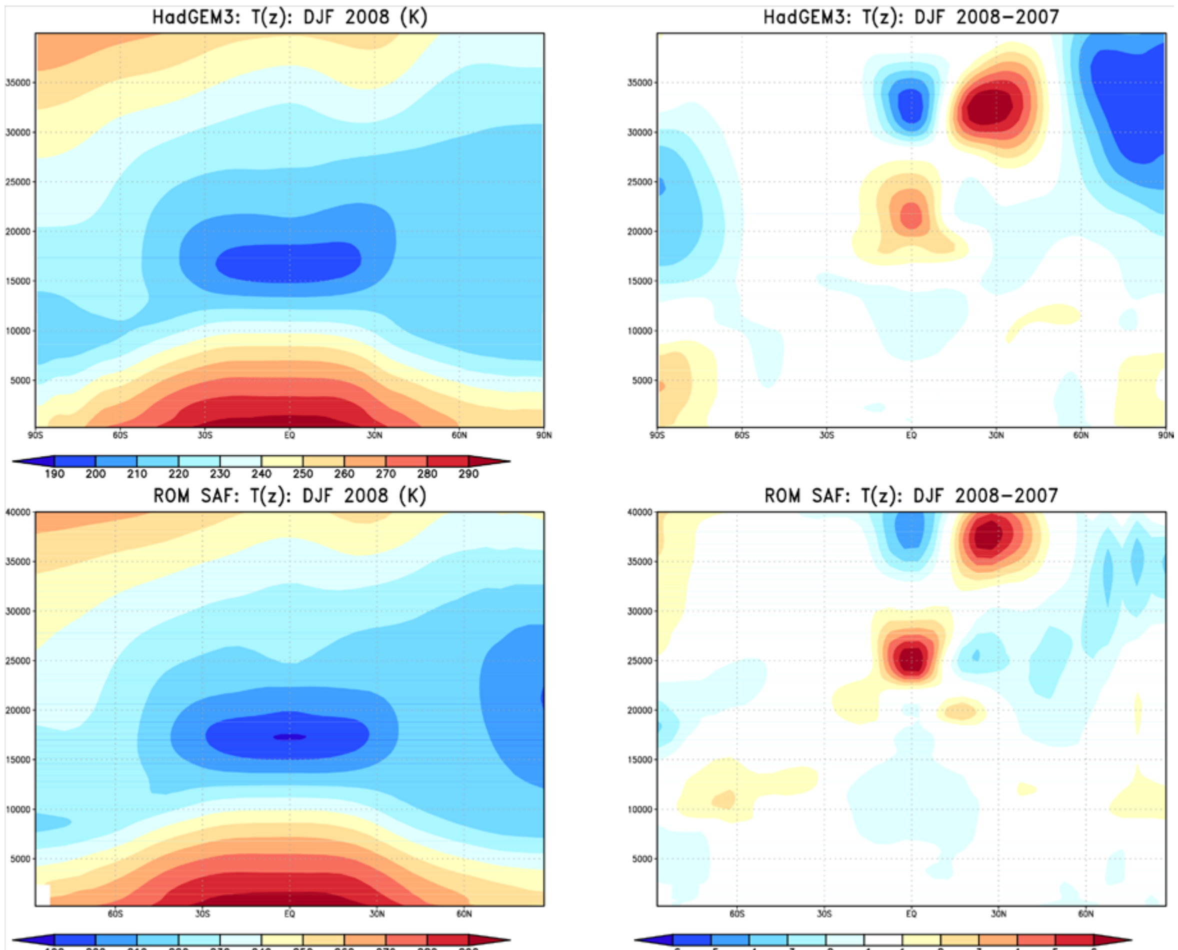


SST-forced (AMIP) simulations
Temperature

DJF 2008 | 2008-2007

Model

Obs

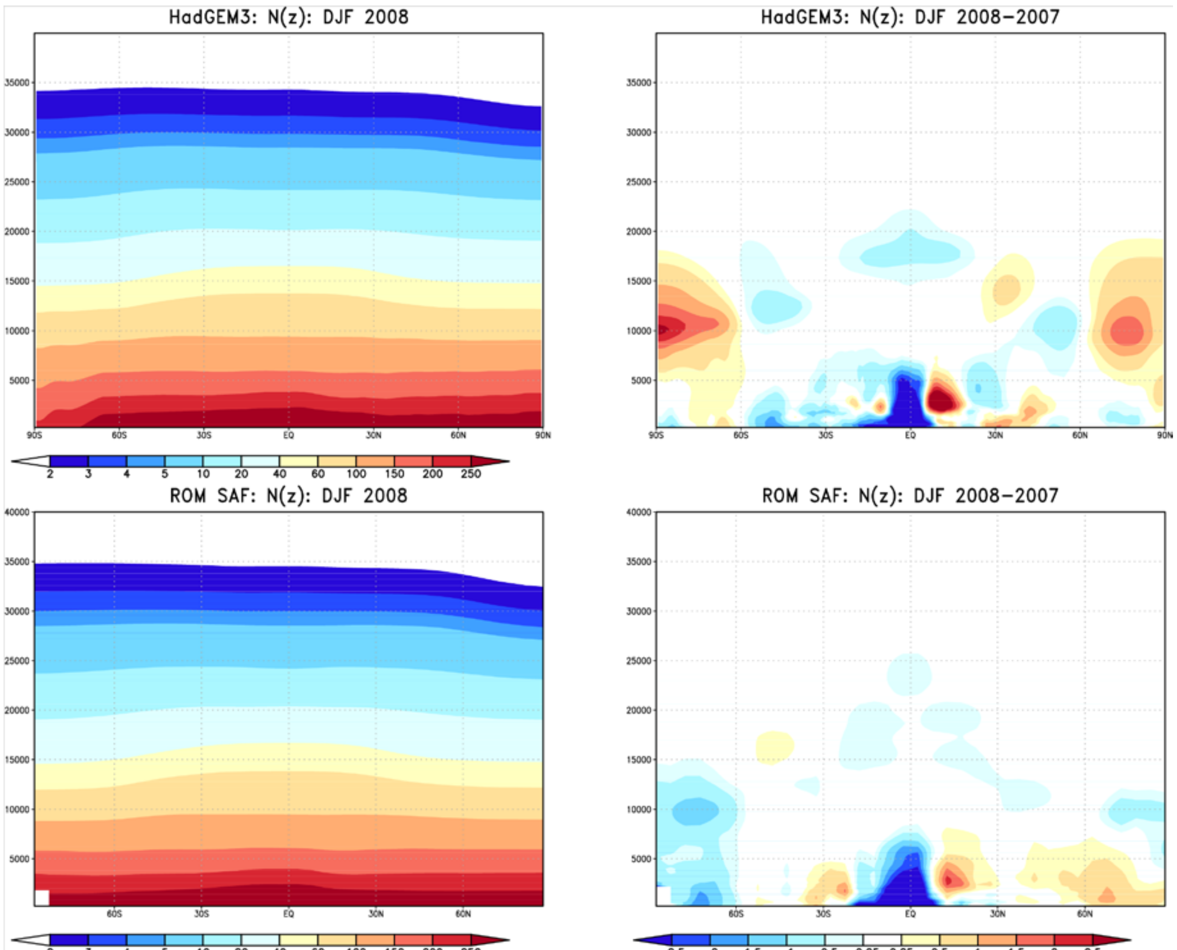


SST-forced (AMIP) simulations
Refractivity

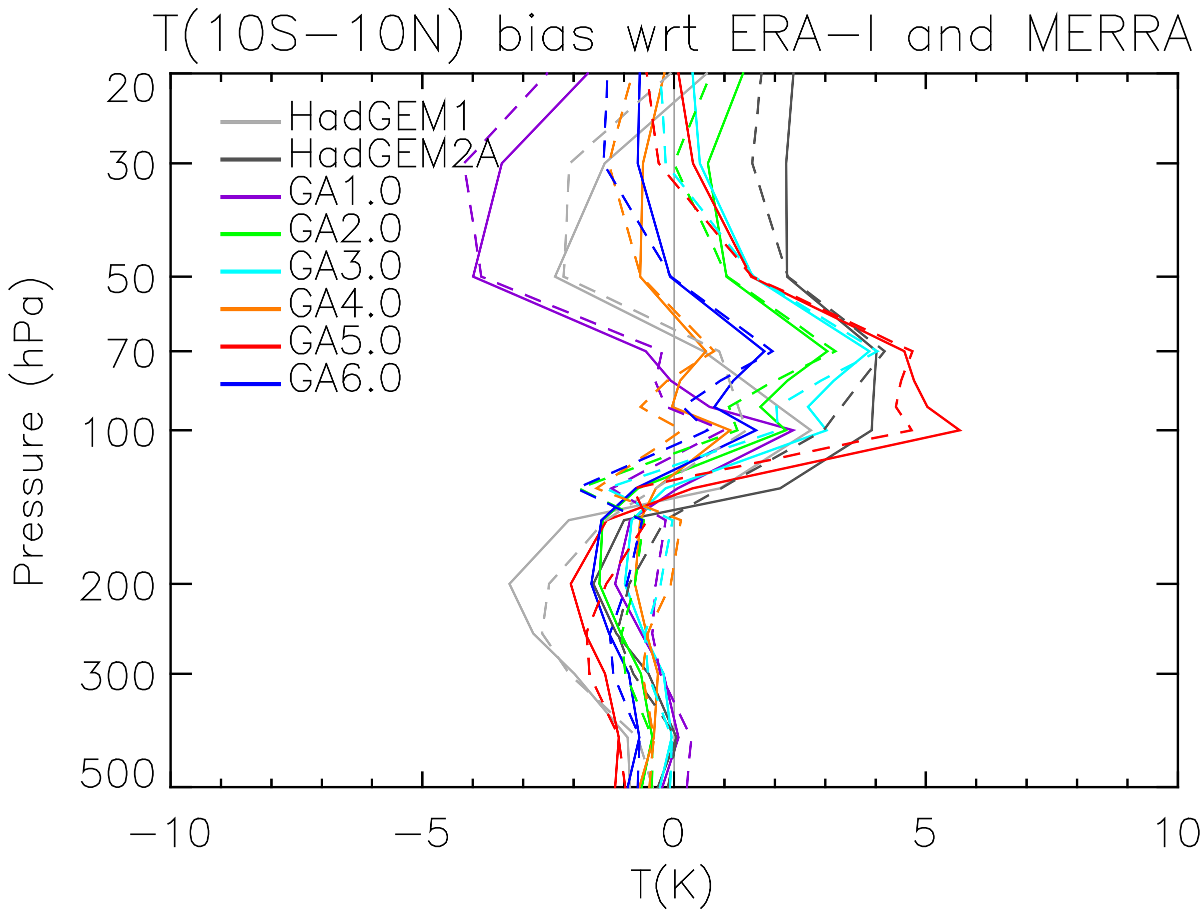
DJF 2008 | 2008-2007

Model

Obs

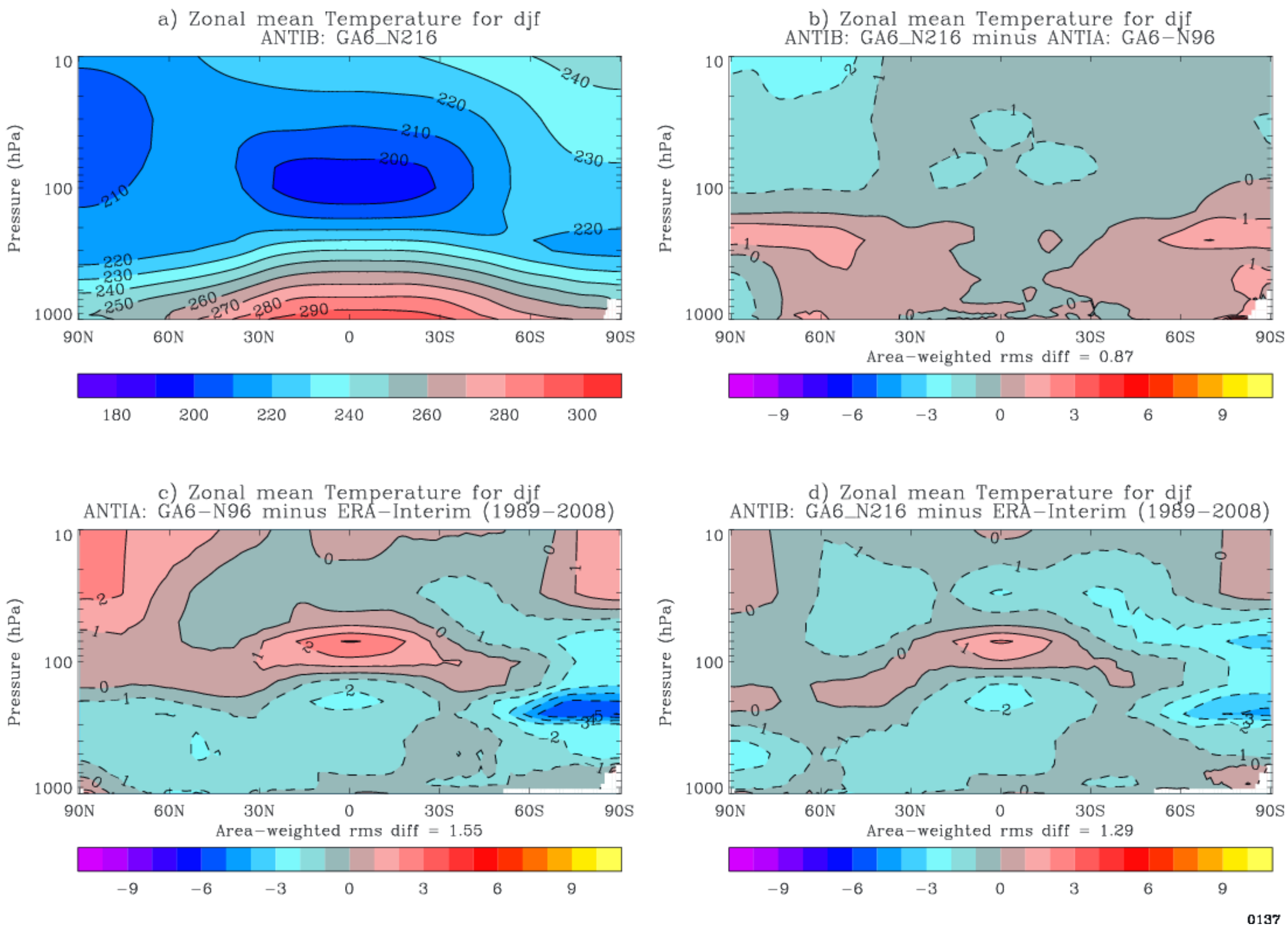


Continuous model development: Equatorial temps vs. reanalyses



Credit: S. Hardiman

Increasing horizontal resolution



N216-N96

N96/N216 - ERA

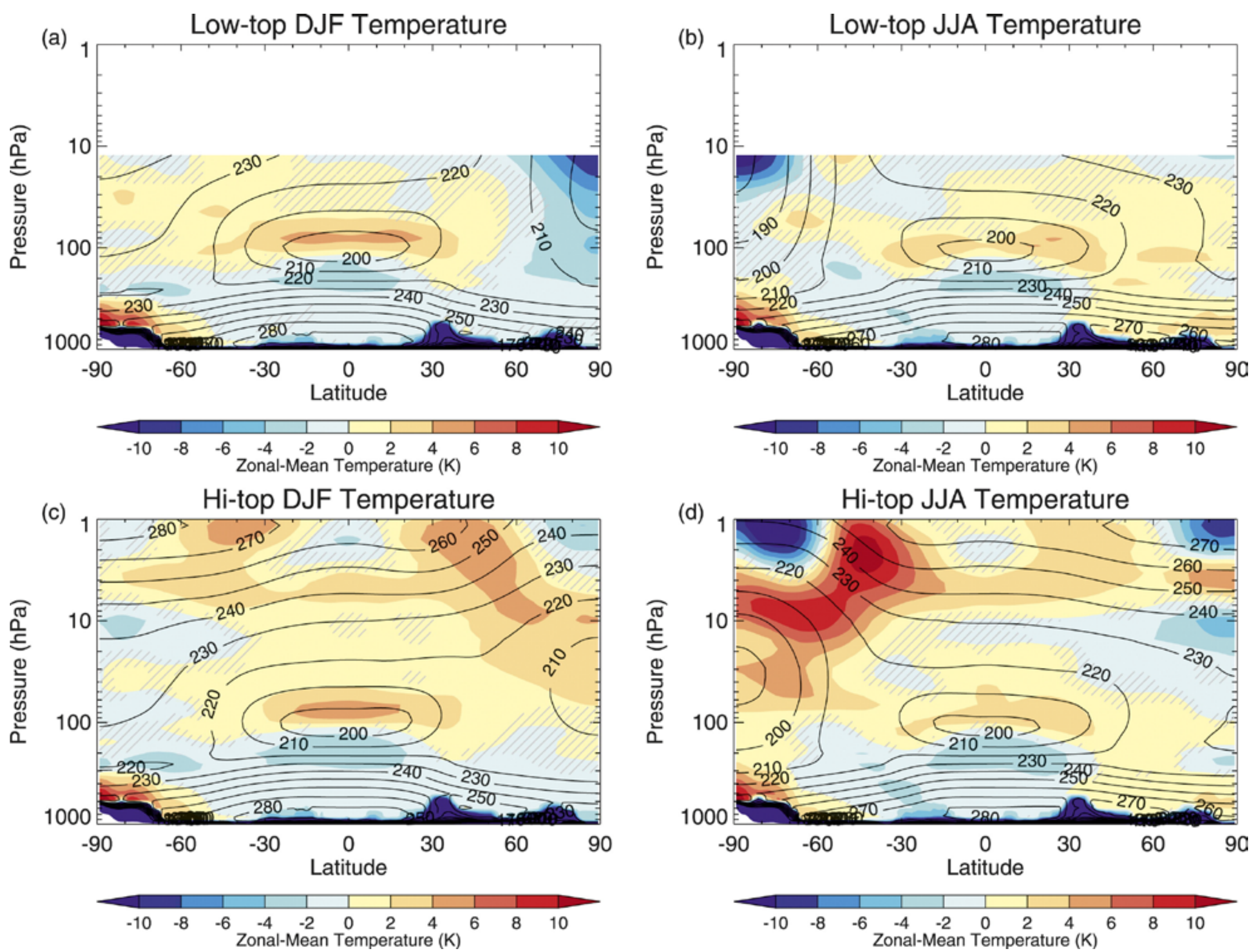
Vertical resolution and model lid

S. Osprey et al., J. Clim., 2013

DJF JJA

Low-top

High-top



Forward modelling and satellite simulators: the "model-to-satellite" approach

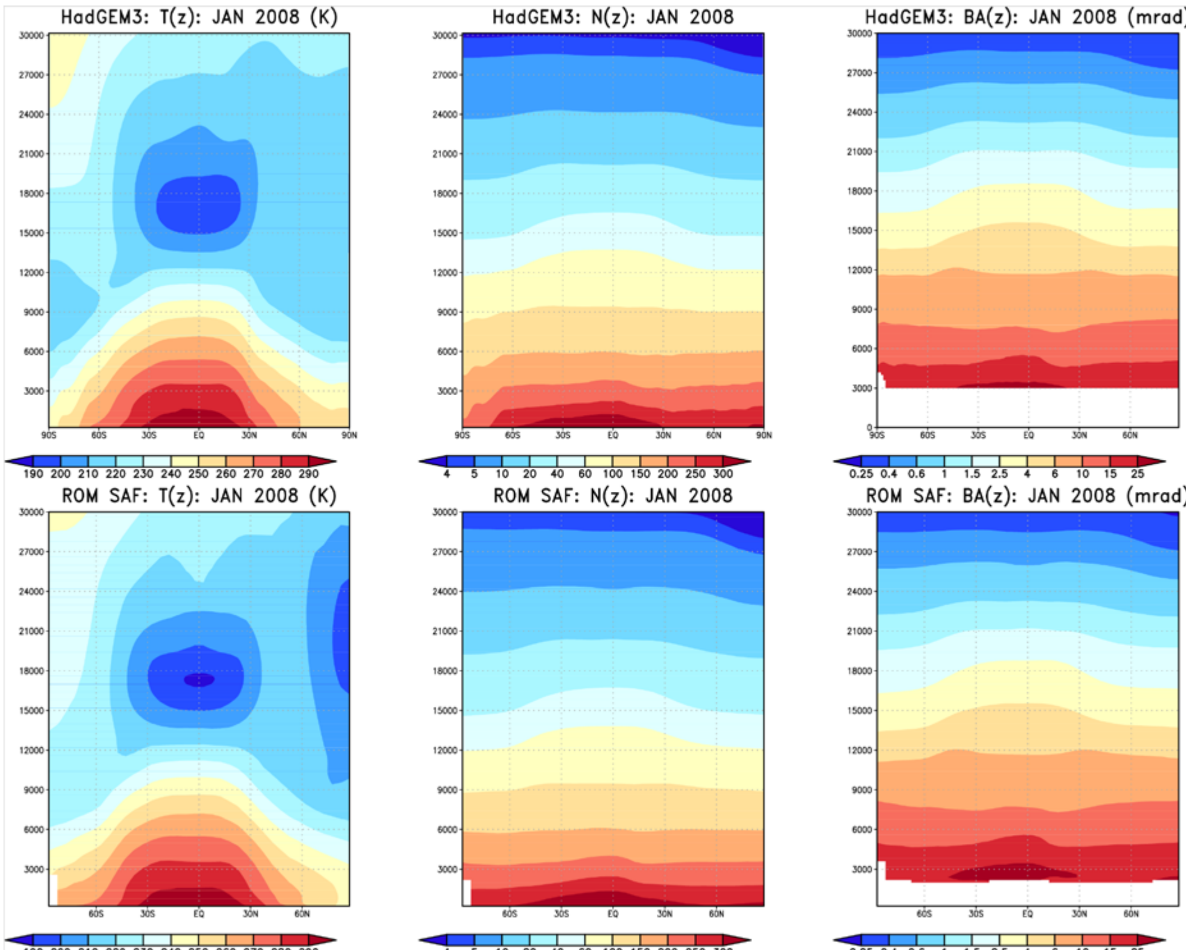
- Standard approach assumes model and retrieved quantities are equivalent
- Forward modelling avoids ambiguities between model and satellite-retrieved parameters
- Allows us to make full use of information content of measurements
- Comparison, including uncertainties, now in radiance/reflectivity/bending angle space
- …as in NWP!

Simulation of BA using ROPP SST-forced run: JAN 2008

T(z) N(z) BA(z)

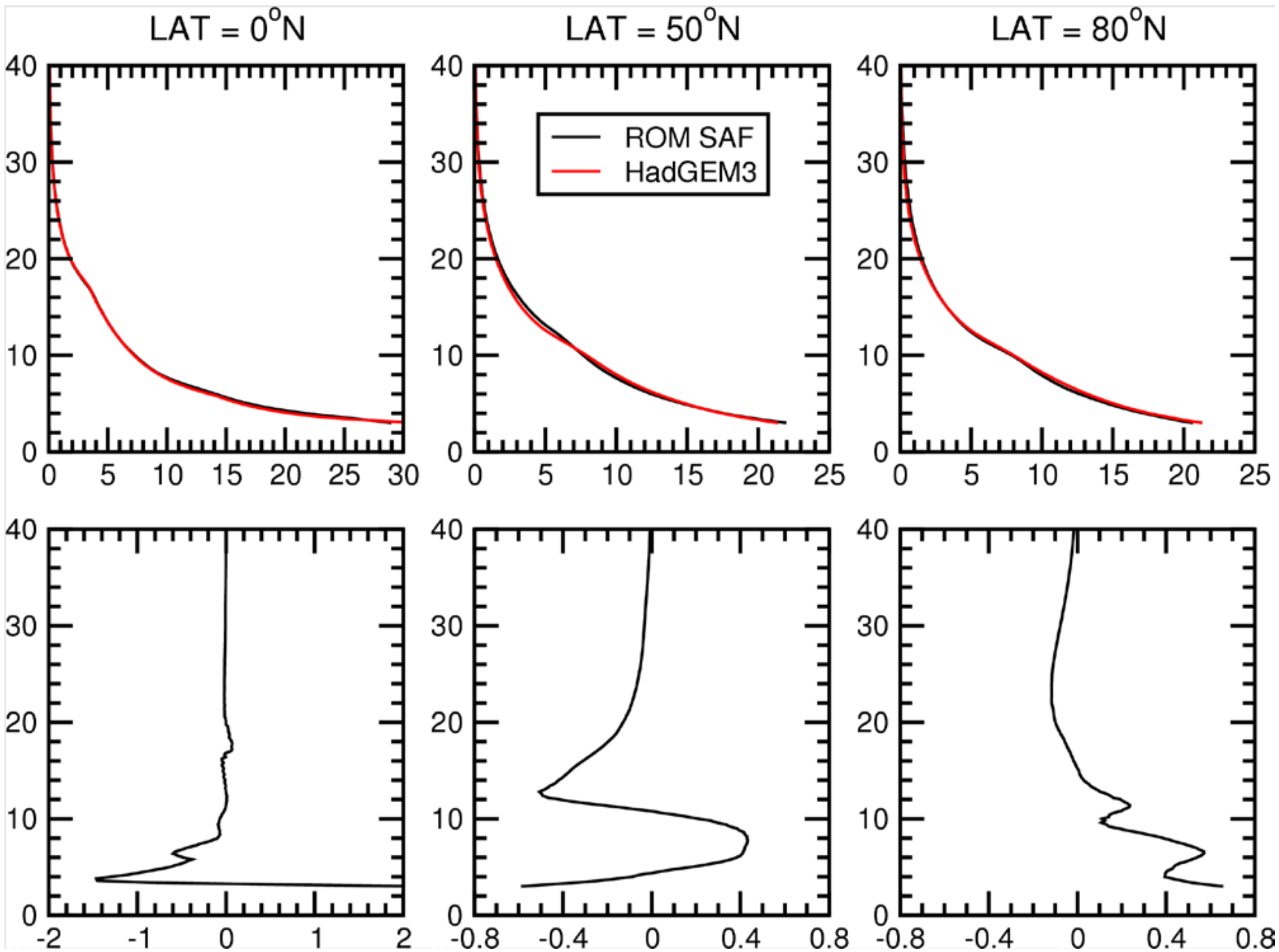
Model

Obs



BA profiles at selected latitudes

Eq | 50°N | 80°N



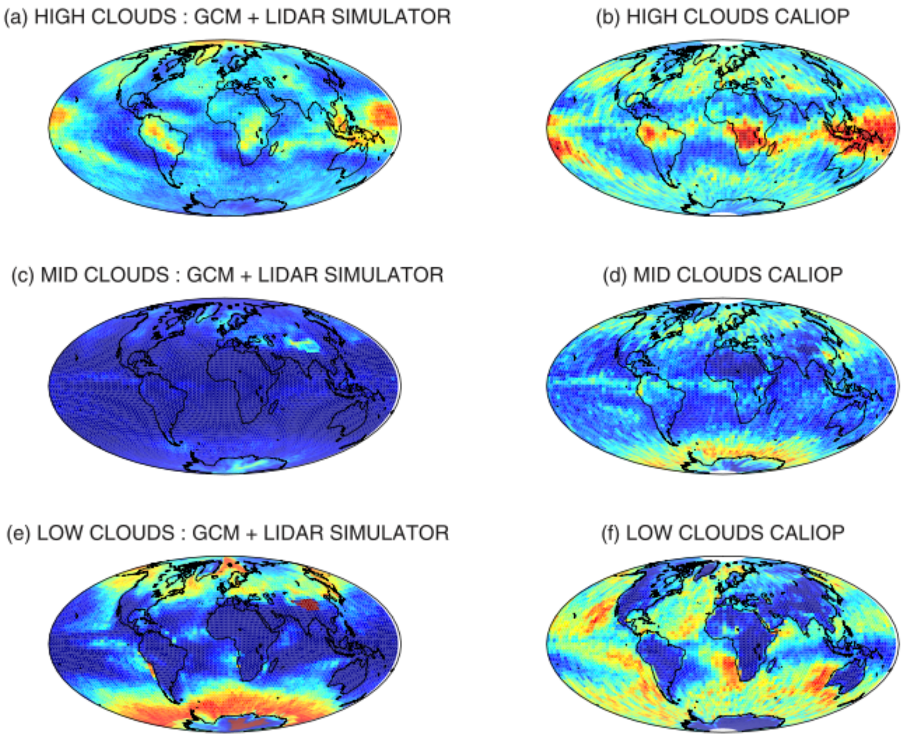
Model-OBS

The "satellite-to-model" approach

• Interpretation of lidar backscatter ratio in terms of cloud products (e.g. cloud fraction) requires set of criteria that depend on vertical resolution at which lidar scattering ratio is measured or computed.

• To make consistent comparisons between models and CALIPSO data, a GCM Oriented CALIPSO Cloud Product (GOCCP) data set has been derived from CALIPSO Level-1 data.

• This new data set is consistent with the CALIPSO simulator outputs derived from models.



Is this an issue for GPS-RO, e.g. high vertical resolution?

Using observational uncertainties

How good is our model?

Is it really improving?

In our model we wish to simulate a quantity: X_{MOD}

How close is this to reality? X_{OBS}

We wish to avoid both overconfidence: $X_{MOD} = X_{OBS}$

And rejecting the model unfairly: $X_{MOD} \neq X_{OBS}$

So we'd like to know: $X_{OBS} \pm \Delta X_{OBS}$

In the absence of observational uncertainties what can we do?

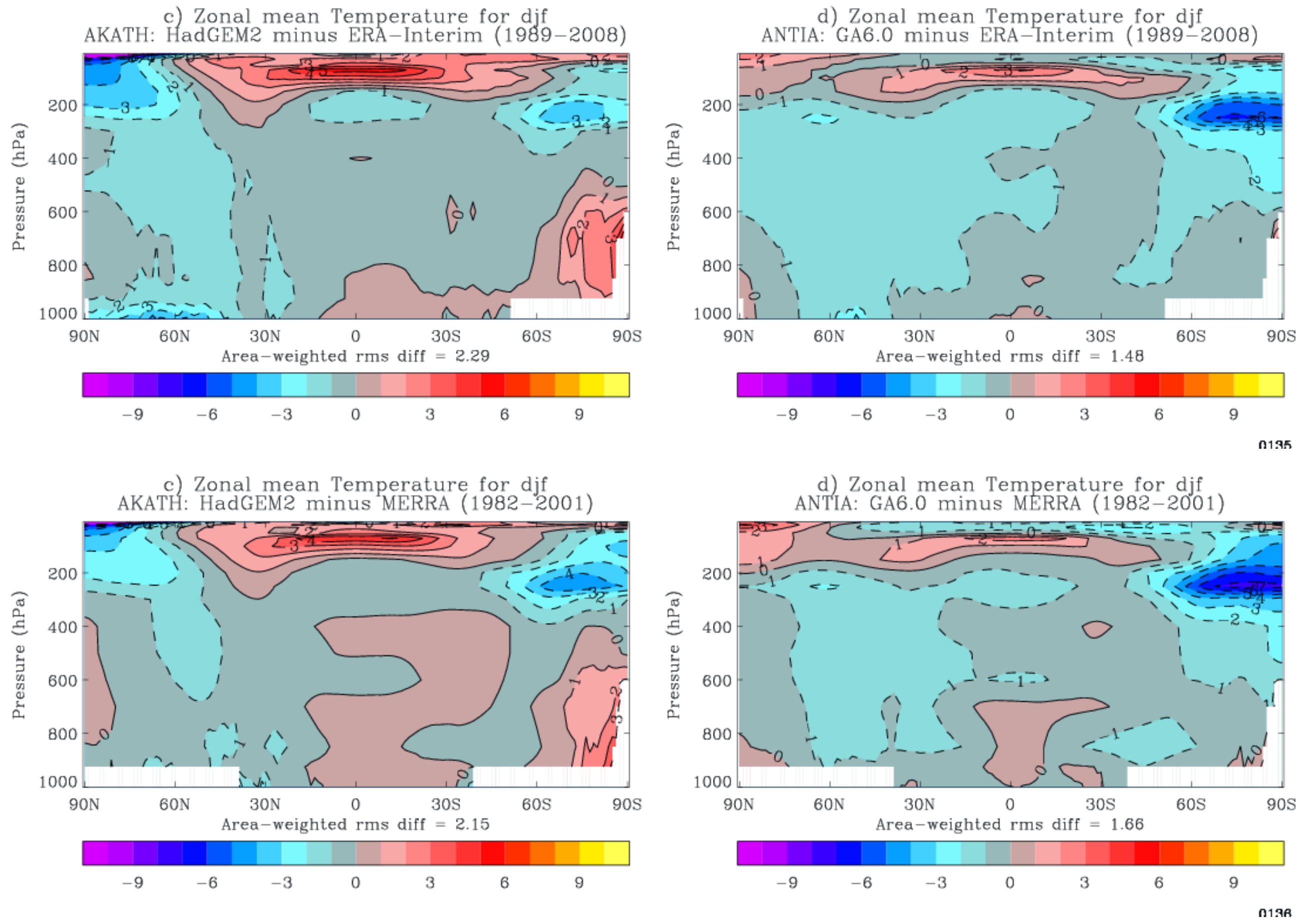
- Treat all data sets as equally plausible?
  - "principle of indifference"; no evidence to do otherwise.
  - e.g. uncertainty = range spanned by available data sets
- Improved information content = better data?
  - e.g. more channels, active vs. passive sensors, etc
- Improved sensors/technology = better data?
- Improved retrieval algorithms and methods = better data?
- Subjective assessment, based on expertise/experience?
- Rough guess ±10%, 50%, 100%,...?

E.g. using different reanalyses

HadGEM2 HadGEM3

vs. ERA

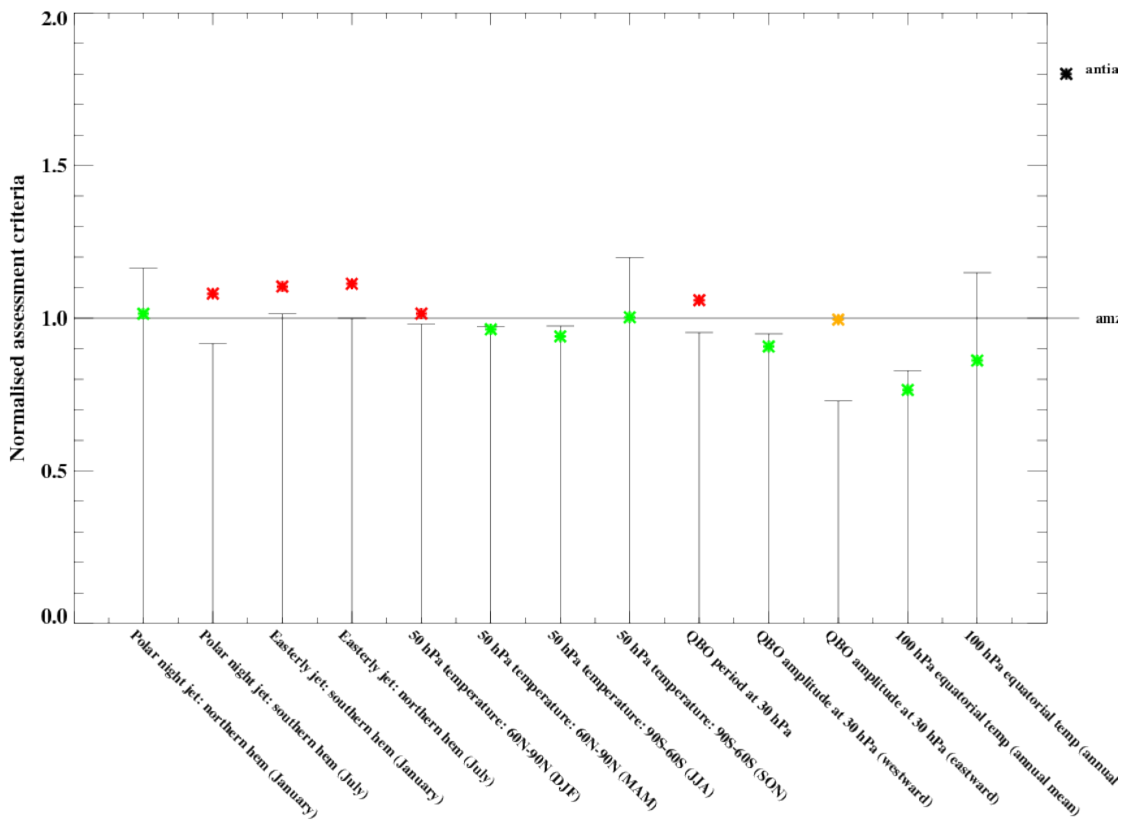
vs. MERRA



Metrics: Met Office auto-assess tool

- Attempt to use multiple obs data sets to assess improvement/deterioration
- Uses range of available obs as “uncertainty”
- Process-by-process: here uses stratospheric temps, QBO, jet strengths

Better ***Worse*** ***Neutral***



Climate model applications – recap

- Model development & evaluation, improvement of physical parameterizations
- Development of metrics for multi-model inter-comparisons
- Testing benefit of increasing horizontal & vertical resolution
- Seasonal to decadal prediction
- Detection & attribution of climate change
- Constraining climate projections, etc

Potentially need to consider all of these when constructing observational data sets for model evaluation and determining the associated uncertainties.

Concluding remarks

- GPS-RO is clearly an exciting new source of information for climate model evaluation, development and testing
- Primary interest is the UTLS, where information content is greatest and the benefit to NWP and reanalyses has already been demonstrated; tropopause height product also useful
- Currently tend to use reanalyses, so new observations very welcome (but ERA-I uses GPS-RO..?)
- Ensuring GPS-RO fulfils this potential requires collaboration between climate modelling centres such as the Hadley Centre and the GPS-RO community

Possible questions for the climate working group

- What products/parameters are required for climate model evaluation and development?
  - Temperatures? Refractivities? Bending angles? Tropopause height?
  - Temporal and spatial resolution?
- How can we ensure that the data are used to best effect?
  - Provision of forward models (e.g. ROPP) for inclusion in climate model satellite data simulators?
  - Characterization of uncertainties?
- Are different products required to test model simulations of trends and for detection & attribution studies?
- Engagement with WCRP metrics panel? Data sets for Obs4MIPs?
- How do we convince climate modellers of the benefits of using GPS-RO?
  - ROM-SAF pilot study?