Observation bias correction schemes in data assimilation systems

John Eyre

ISDA, July 2016

→ The effects of model bias on bias correction of observations

- The bias correction problem in DA
- This study:
  - a very simple assimilation system
  - effects of model bias
  - some results
  - findings
- Implications, questions and conclusions

Acknowledgements

James Cameron

Bill Bell

Chris Burrows

Peter Weston

Masashi Ujiie

The bias correction problem in DA

- Standard DA theory assumes observations unbiased
- … or that they have bias-corrected ahead of the DA

- Bias correction is necessary for assimilation of radiances
- … for biases in the observations or their operators

- Two types of observation:
  - "Anchor" observations, assumed unbiased
    - may have been pre-corrected (e.g. sondes)
    - may still contain biases
  - Observations to be bias-corrected within the DA system

Types of bias correction scheme used within DA systems

Bias correction schemes can:

- attempt to remove biases:
  - relative to **background**, or
  - relative to **analysis**

- be "**static**" (one-off), or
- iterated to convergence (e.g. **VarBC**)

Bias correction literature

- “Static” bias correction (against background)
  - Eyre, ECMWF TM 176, 1992
  - Harris and Kelly, QJRMS, 2001
- VarBC (correction against analysis)
  - Derber and Wu, MWR, 1998
  - Dee, ECMWF Workshop, 2004; Dee, QJRMS, 2005
- “Off-line scheme” (like VarBC, but correcting v. background)
  - Auligné et al., QJRMS, 2007
- General papers on biases and DA and forecast model bias
  - Dee and da Silva, QJRMS, 1998; Dee, QJRMS, 2005

This study

- An attempt to understand scientific differences between Met Office old “static” scheme and new VarBC scheme
- Uses a very simple system (one variable)
- Explores the role of anchor observations
- Explores the role of model bias

For details see: Eyre, 2016. Observation bias correction schemes in data assimilation systems. Q.J.R.Meteorol.Soc.

This study – key result

- Bias correction of observations is not "passive"…

- In the presence of model bias,
  bias correcting a greater proportion of observations
  pulls the analysis away from the anchor observations
  and towards the model bias

→ A strategy of having a few, high-quality anchor observations is likely to lead to problems

A very simple assimilation system (1)

One scalar analysis variable

Scalar observations in same space as the analysis

Analysis, at n^th^ step:

$$\boxed{x_{a,n} = w_b x_{b,n} + w_1 y_{1,n} + w_2 y_{2,n}}$$

$x_{a,n}$ = analysis, $x_{b,n}$ = background

$y_{1,n}$ = <u>anchor observations</u>, $y_{2,n}$ = <u>observations to be bias-corrected</u>

w_j = analysis weights – general, not necessarily optimal, but …

$$w_b + w_1 + w_2 = 1$$

A very simple assimilation system (2)

Biases and random errors:

$$x_{j,n} = x_{t,n} + b_j + \varepsilon_{j,n}$$

truth bias random error

$$y_{j.n} = x_{t,n} + b_j + \varepsilon_{j,n}$$

→ $$\boxed{b_a = w_b b_b + w_1 b_1 + w_2 b_2}$$

A very simple assimilation system (3)

Forecast model:

$$\boxed{x_{b,n+1} = x_{f,n} = x_{a,n} + \delta x_{m,n}}$$

$$\delta x_{m,n} = \delta x_{t,n} + \delta b_{m,n} + \varepsilon_{m,n}$$

forecast increment   truth   bias   random error

Forecast model bias:

$$\boxed{\delta b_{m,n} = \alpha\,(x_{m,n} - x_{a,n})}$$

- a relaxation towards state $x_{m,n}$, which has bias b_m

- where the relaxation rate is α

A very simple assimilation system (4)

Combining these equations →

$$b_b = b_a + \alpha(b_m - b_a) = (1-\alpha)b_a + \alpha b_m$$

→

$$\boxed{b_b = \frac{\gamma b_m + w_1 b_1 + w_2 b_2}{\gamma + w_1 + w_2}}$$

$$\boxed{b_a = \frac{\gamma(1 - w_1 - w_2)b_m + (1+\gamma)(w_1 b_1 + w_2 b_2)}{\gamma + w_1 + w_2}}$$

weighted averages

where $\gamma = \alpha/(1-\alpha)$

Special case – no model bias

No model bias: $\alpha = \gamma = 0$:

→ $$\boxed{b_a = b_b = \frac{w_1 b_1 + w_2 b_2}{w_1 + w_2}}$$

If also, $w_2 = 0$

→ $$\boxed{b_a = b_b = b_1}$$

Bias correction strategy:

- introduce observations y_2 into DA system passively: $w_2 = 0$
- monitor bias in y_2 against background: $c_2 = b_2 - b_b$
- bias-correct y_2 : $y_2{}^* = y_2 - c_2$

These bias-corrected observations will now have bias:

$$\boxed{b_2{}^* = b_2 - c_2 = b_b = b_a = b_1}$$

PERFECT!!!

Effects of model bias (1)

With a **static bias correction scheme**, after 1st application:

using (O-B) statistics →
$$\boxed{c_2 = b_2 - b_b = b_2 - \frac{\gamma b_m + w_1 b_1 + w_2 b_2}{\gamma + w_1 + w_2}}$$

In principle, you can stop here.

*** But we tend to repeat the process in an ad hoc manner ***

If you repeat the process to convergence:

→
$$b_b = \frac{\gamma b_m + w_1 b_1 + w_2 b_b}{\gamma + w_1 + w_2}$$

If $b_1 = 0$, →
$$\boxed{\frac{b_b}{b_m} = \frac{\gamma}{\gamma + w_1}} \quad \boxed{\frac{b_a}{b_m} = \frac{\gamma(1 - w_1)}{\gamma + w_1}}$$

Effects of model bias (2)

With **VarBC**:

- we bias-correct against analysis, rather than against background,

- and we **do** iterate the process to convergence

→ $$\boxed{\frac{b_b}{b_m} = \frac{\gamma(1 - w_2)}{\gamma(1 - w_2) + w_1}}$$ $$\boxed{\frac{b_a}{b_m} = \frac{\gamma(1 - w_1 - w_2)}{\gamma(1 - w_2) + w_1}}$$

***

So we now have **4 equations** for **bias as a fraction of model bias**:

- for **background bias**, and for **analysis bias**

- correcting **against background**, and correcting **against analysis** (VarBC)

This study – parameters used

Baseline values – to mimic Met Office global NWP system

- Total observation weight, $\mathrm{Tr}(\mathbf{W})/p \approx 1-\{\mathrm{E}(J_{of})/\mathrm{E}(J_{oi})\}^{0.5}$

  where $\mathbf{W}$ is matrix of obs weights, dimension p,

  $\mathrm{Tr}(\ldots)$ = trace, $\mathrm{E}(\ldots)$ = expected value,

  J_{if} = VAR initial observation cost,

  J_{of} = VAR final observation cost.

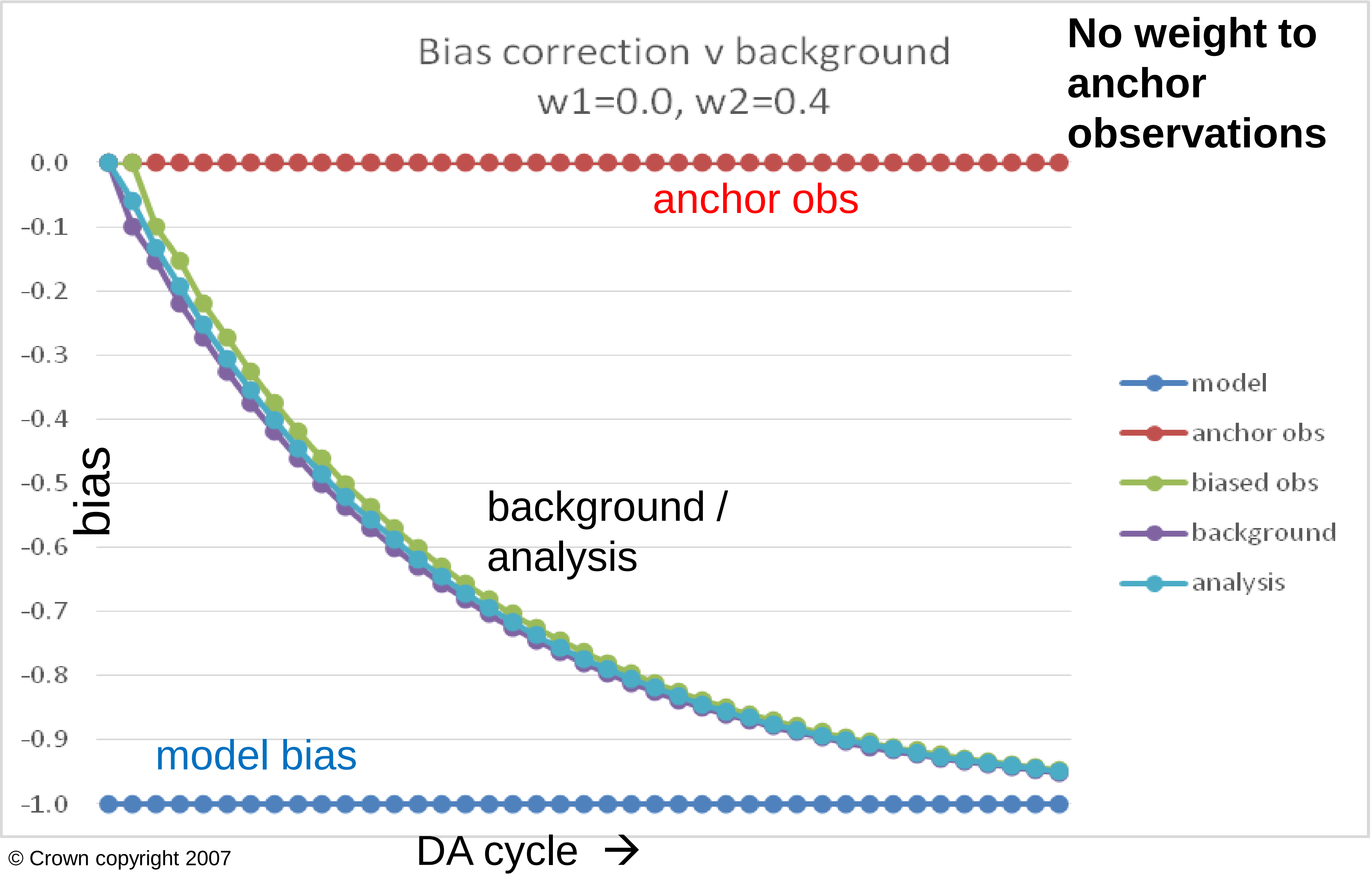
  For Met Office global 4D-Var, $J_{if}/J_{of} \approx$ 0.6-0.7,

  and so $\boxed{\mathrm{Tr}(\mathbf{W})/p \approx 0.2}$

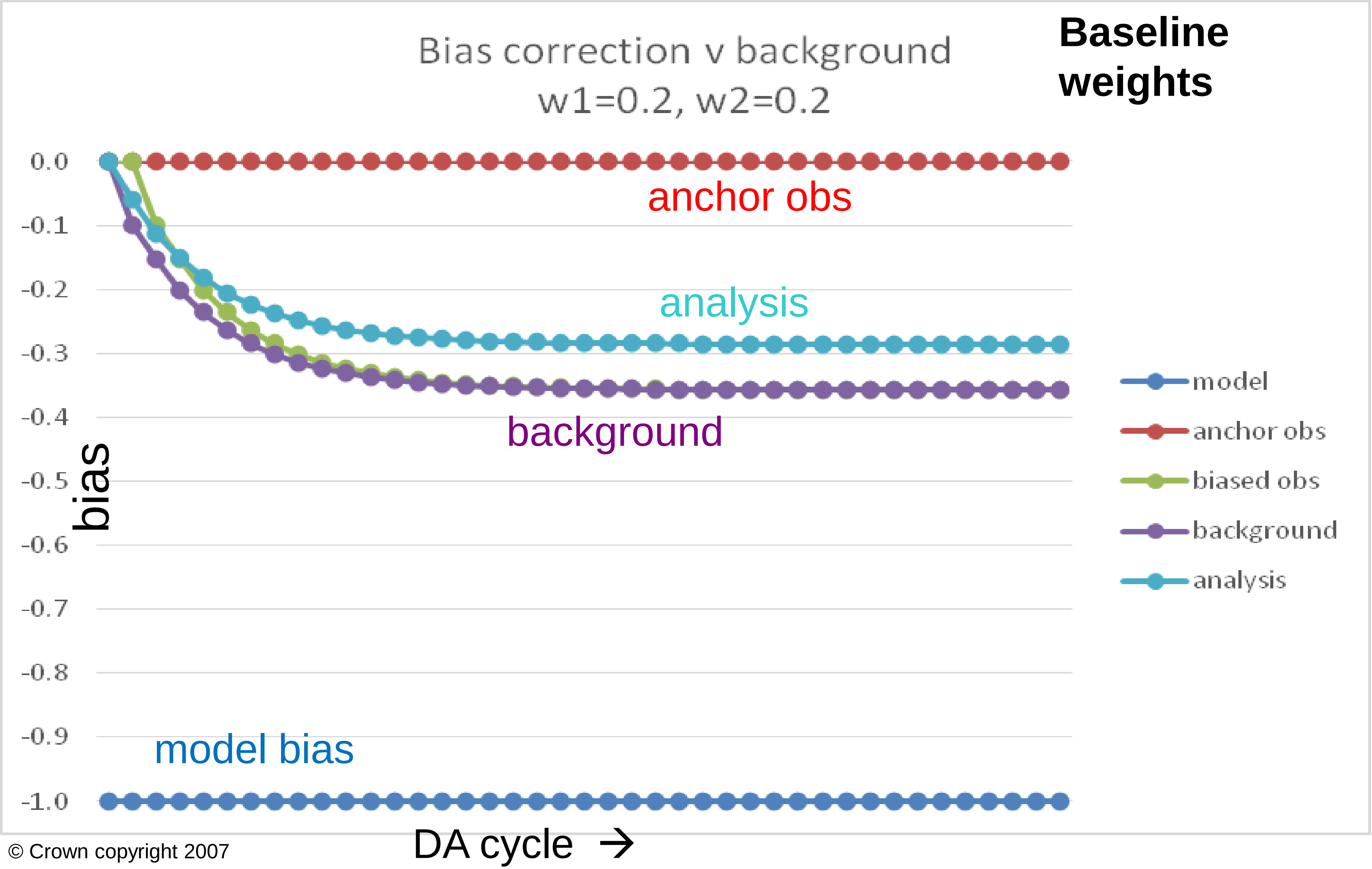
- FSOI results → $w_1 \approx w_2$ → $\boxed{w_1 = w_2 = 0.2}$

- Model relaxation time ≈3 days → $\boxed{\gamma = 0.1}$ (per DA cycle)

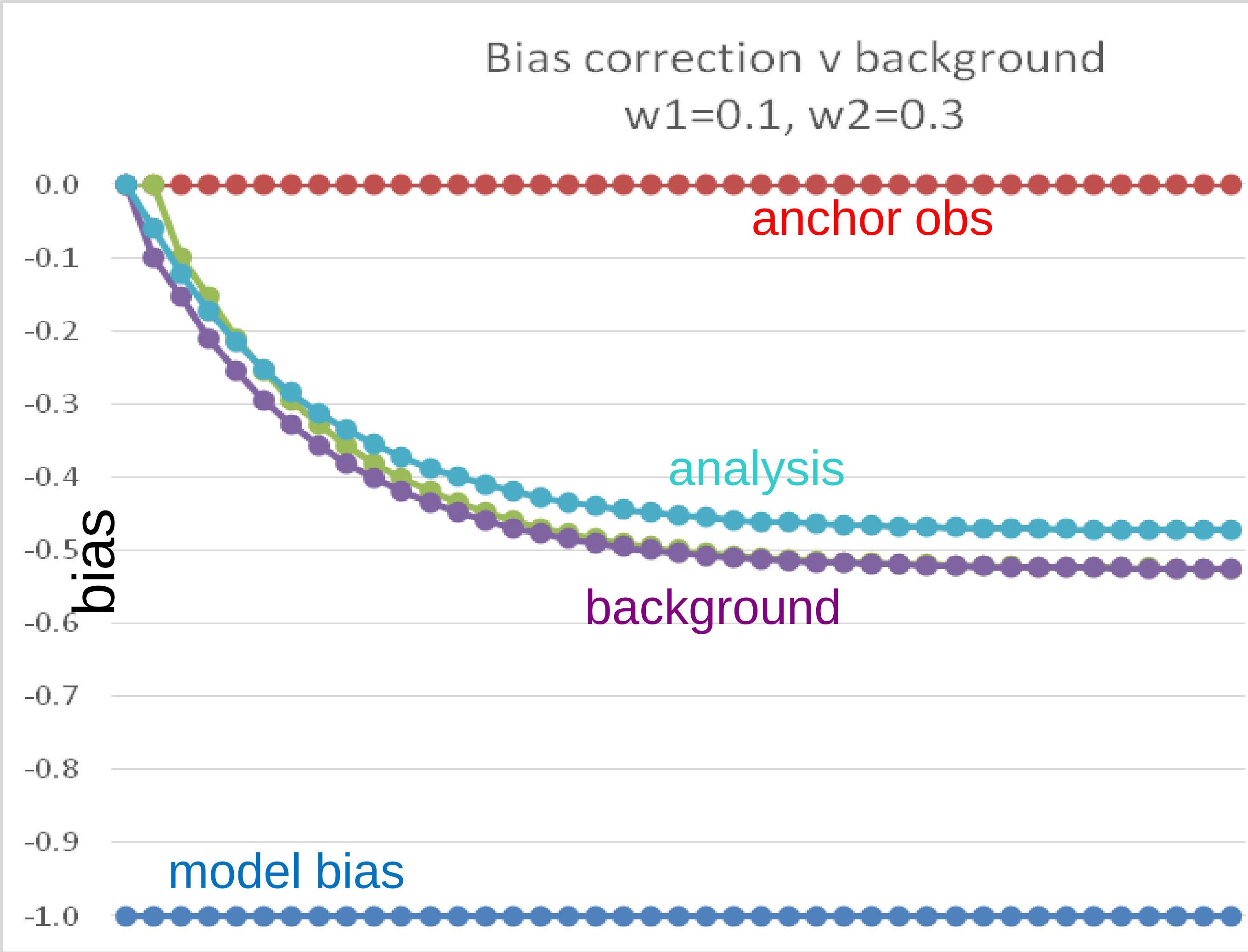
The convergence process (1)



The convergence process (2)



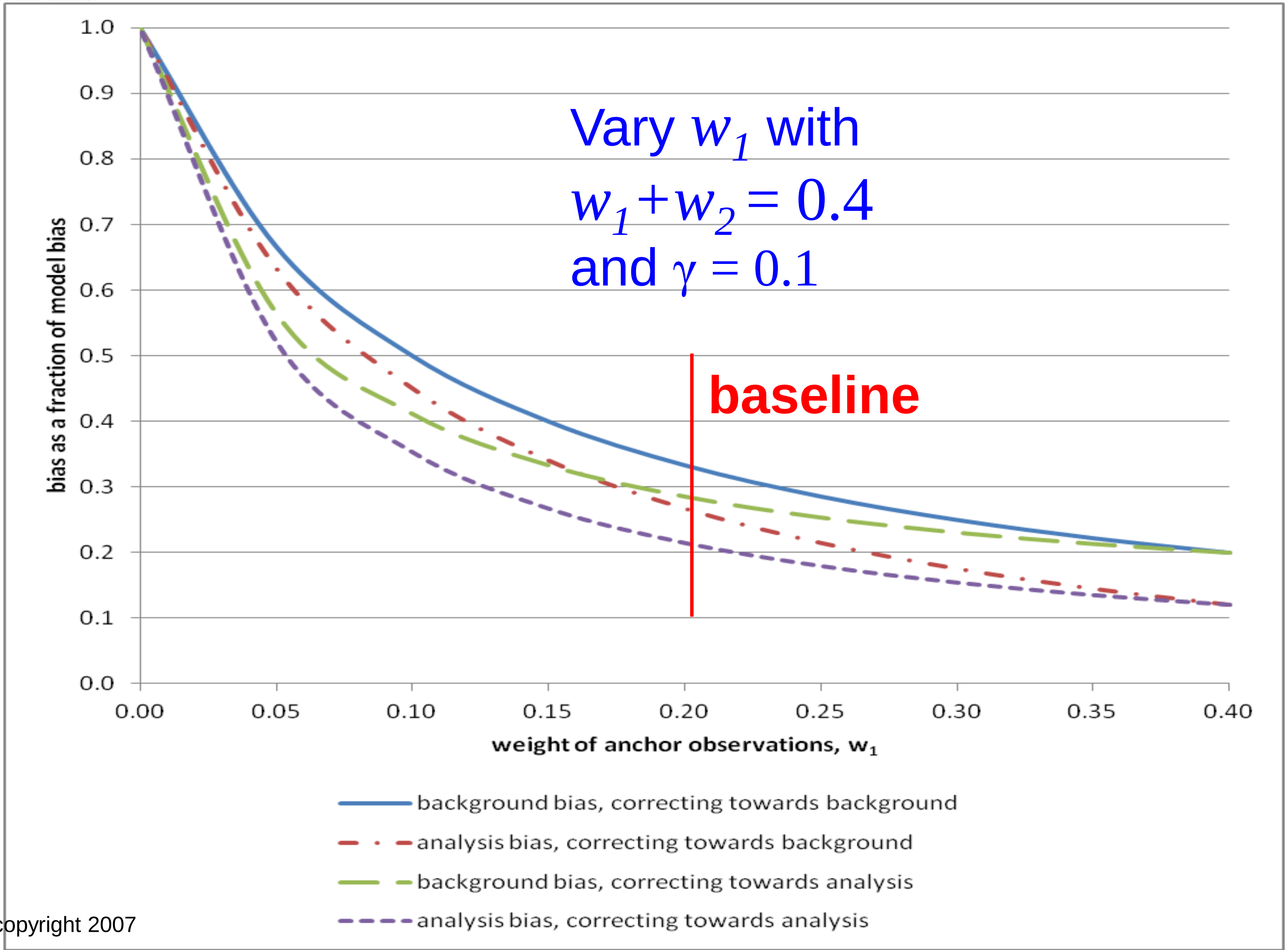
The convergence process (3)



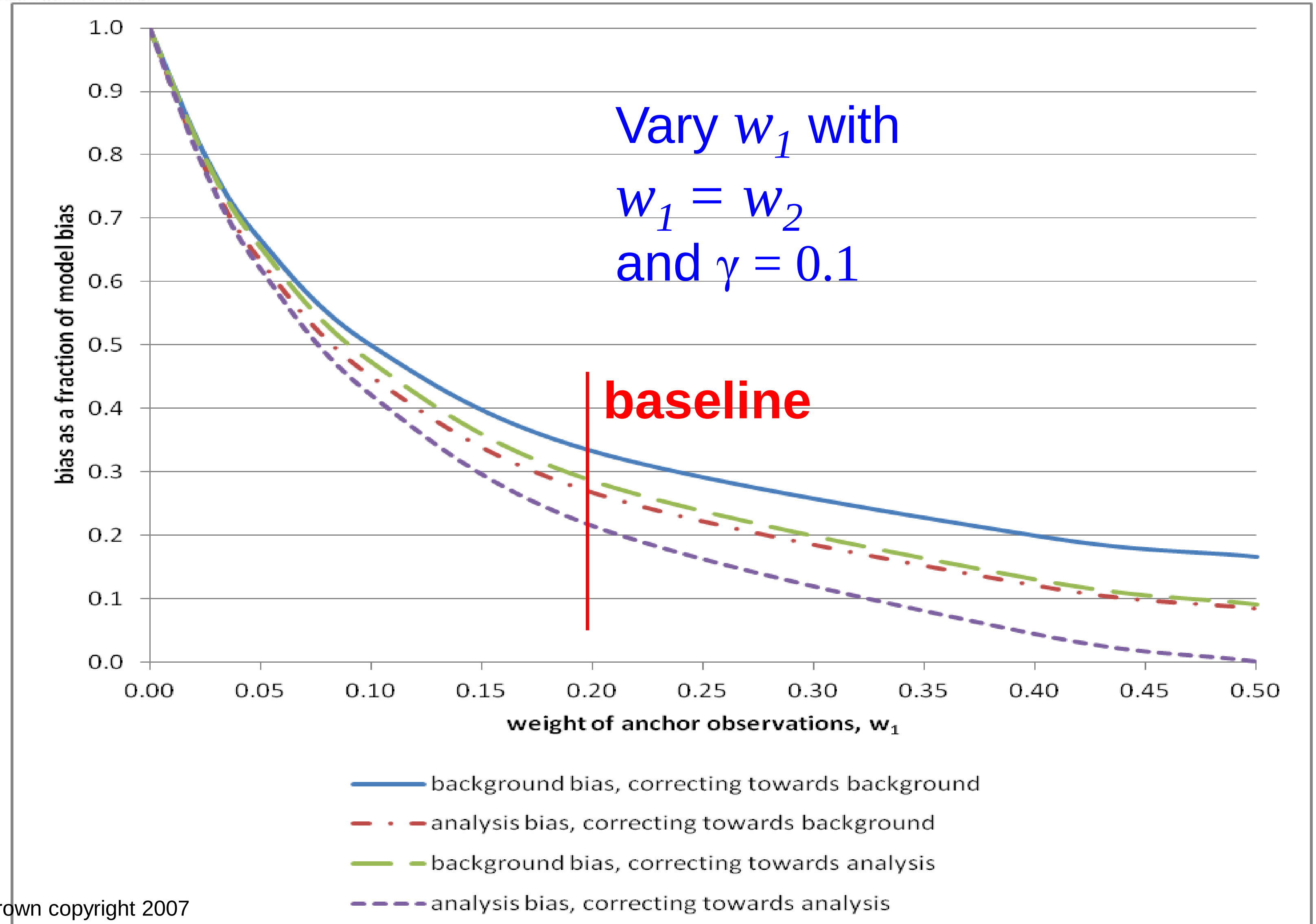
Reduced weight to anchor observations

DA cycle →

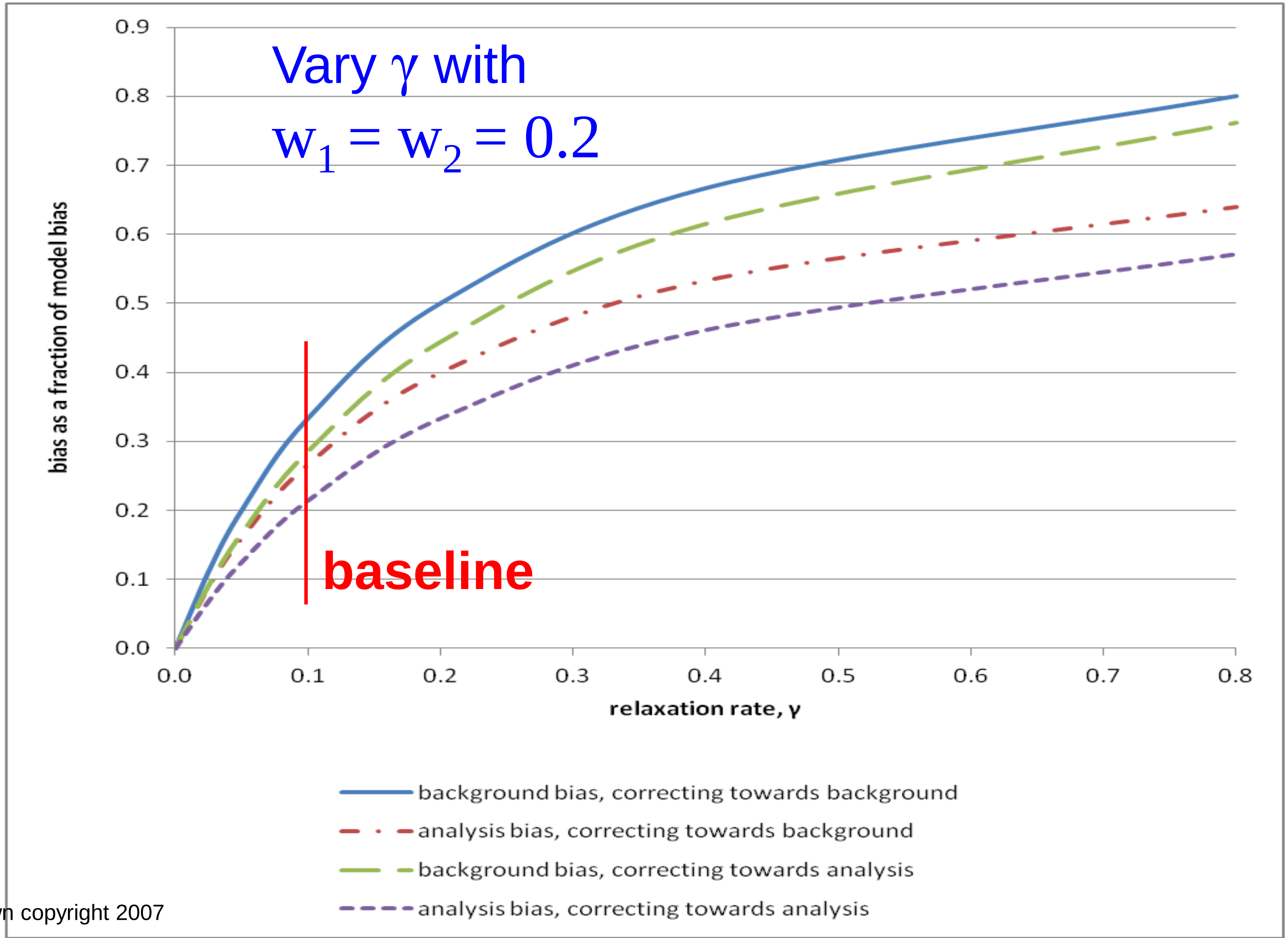
Results at convergence (1)



Results at convergence (2)



Results at convergence (3)



Some findings (1)

- In asymptotic limit, **biases in background and analysis** – when correcting v. background or v. analysis – **are weighted averages of model bias and bias in anchor observations**.

- When **more** observations are bias-corrected, less weight is given to anchor observations and **more** weight to model bias.
- This effect is less pronounced when correcting v. analysis (VarBC) than when correcting v. background … but difference is small.

- In VarBC, effect of model bias is realised quickly; …
- …in static scheme not fully realised, or only through repeated application of scheme.

Some findings (2)

- Baseline values used in this scheme are intended to be representative of global NWP system

  → **background/analysis bias ~0.21-0.33 of model bias** ! …

- … but much variation expected within model domain – according to observation density, fraction of anchor observations, height, model variable

Implications and questions

- Effect of more and more radiances?
- Role of RO?
- Bias correction of radiosondes?
- Role of GRUAN?
- Masking in VarBC?
- Choice of bias predictors?
- Need for other bias correction strategies?

Conclusions

- In the absence of model bias, bias correction of observations is relatively straightforward.
- VarBC is less affected by model bias that an equivalent scheme attempting to remove bias relative to the background,
- … but difference is small compared with model bias itself.
- With baseline values used here background / analysis biases are 0.21-0.33 of model bias – larger than expected.
- **The effect of model bias on both background and analysis biases will increase as the relative weight given to the anchor observations decreases → important implications for observation bias correction strategies.**

Met Office

Thank you! Questions?