Forecast-error-sensitivity to observations in the UM

Richard Marriott and Andrew Lorenc

9th Adjoint Workshop, Cefalu, Sicily (Oct 2011)

Met Office

Contents:

- The adjoint technique vs. nonlinearity
- The Met Office finite observation sensitivities (OS) method
    - Dealing with a non-quadratic inner-loop
- Why are only 51% of observation beneficial?
- A few results

The adjoint technique vs. nonlinearity

- The quadratic measure of forecast error ($J = \delta\mathbf{x}^{f\mathrm{T}}\mathbf{C}\delta\mathbf{x}^{f}$) is known to be the dominant nonlinearity of the OS problem
- Gradient of J (at point B) given by

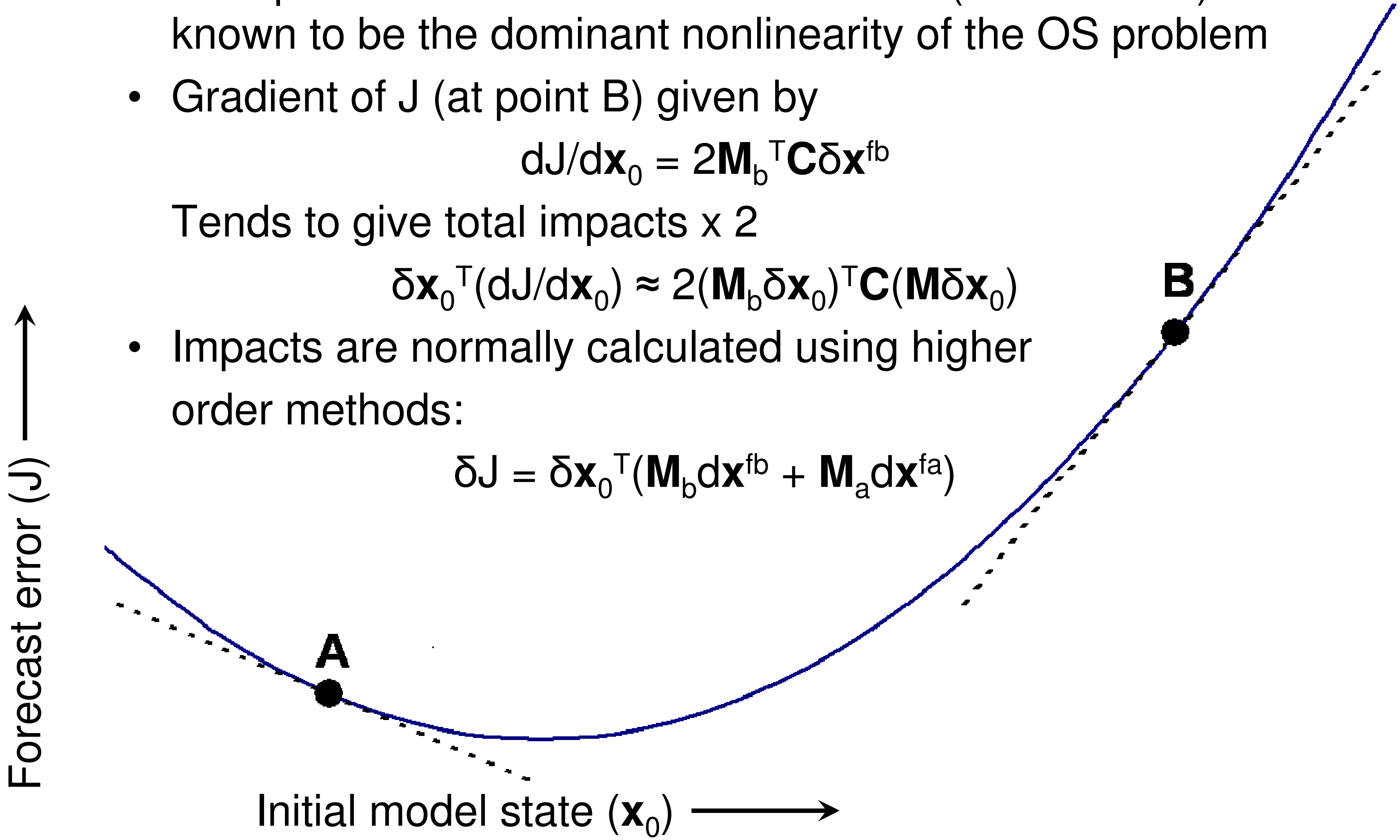
$$dJ/d\mathbf{x}_0 = 2\mathbf{M}_b^{\mathrm{T}}\mathbf{C}\delta\mathbf{x}^{fb}$$

  Tends to give total impacts x 2

$$\delta\mathbf{x}_0^{\mathrm{T}}(dJ/d\mathbf{x}_0) \approx 2(\mathbf{M}_b\delta\mathbf{x}_0)^{\mathrm{T}}\mathbf{C}(\mathbf{M}\delta\mathbf{x}_0)$$

- Impacts are normally calculated using higher order methods:

$$\delta J = \delta\mathbf{x}_0^{\mathrm{T}}(\mathbf{M}_b d\mathbf{x}^{fb} + \mathbf{M}_a d\mathbf{x}^{fa})$$



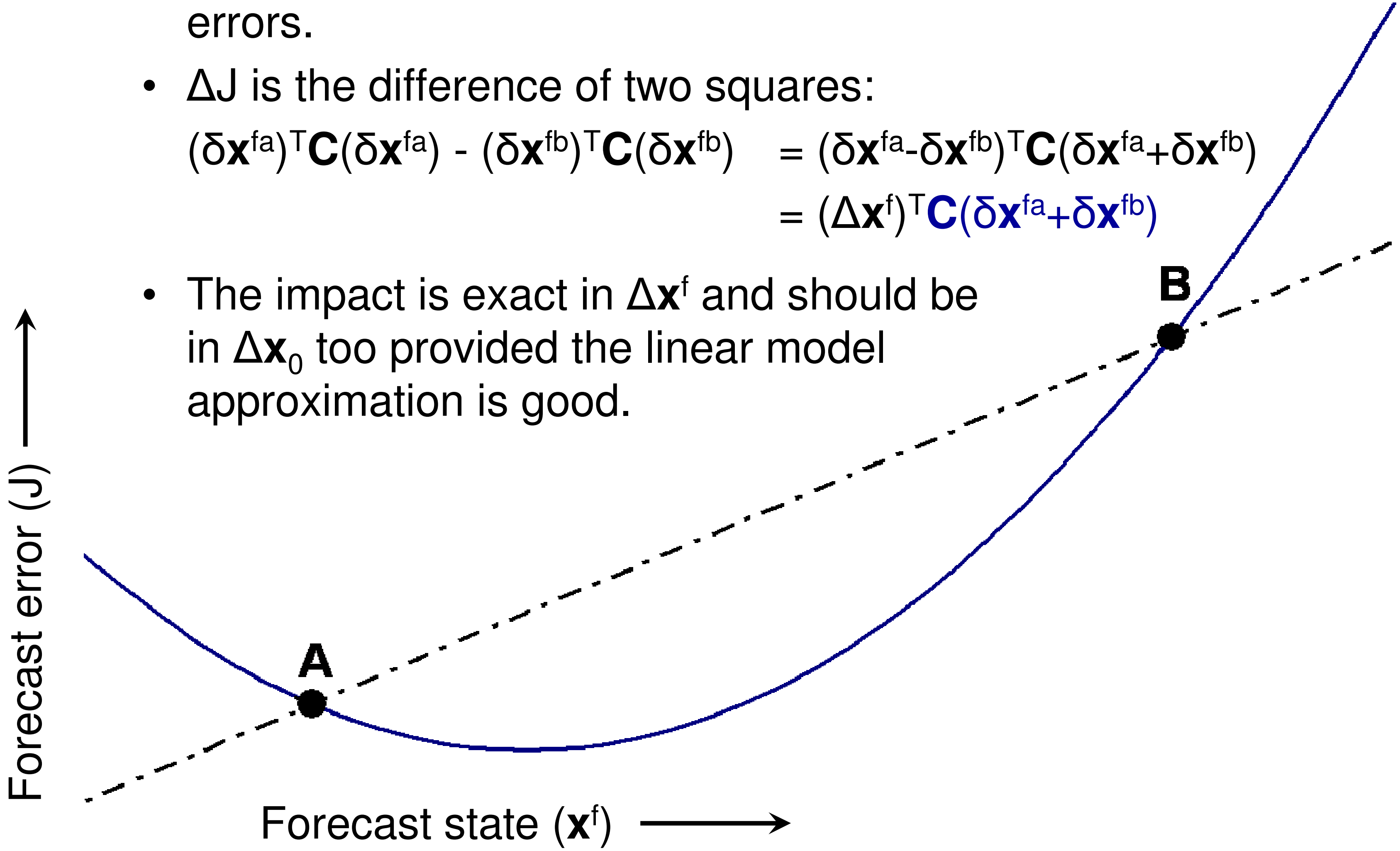
The Met Office finite OS method

- We wish to find the impact of finite increments.
- Using the finite gradient, $\Delta J/\Delta \mathbf{x}^f$, will avoid linearisation errors.
- ΔJ is the difference of two squares:

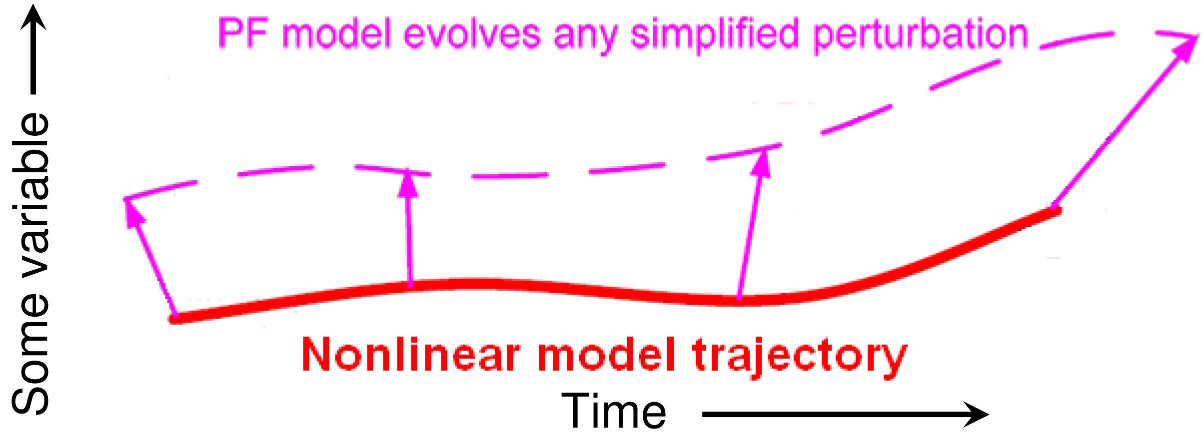
$$(\delta \mathbf{x}^{fa})^T \mathbf{C} (\delta \mathbf{x}^{fa}) - (\delta \mathbf{x}^{fb})^T \mathbf{C} (\delta \mathbf{x}^{fb}) = (\delta \mathbf{x}^{fa} - \delta \mathbf{x}^{fb})^T \mathbf{C} (\delta \mathbf{x}^{fa} + \delta \mathbf{x}^{fb})$$

$$= (\Delta \mathbf{x}^f)^T \mathbf{C} (\delta \mathbf{x}^{fa} + \delta \mathbf{x}^{fb})$$

- The impact is exact in $\Delta \mathbf{x}^f$ and should be in $\Delta \mathbf{x}_0$ too provided the linear model approximation is good.



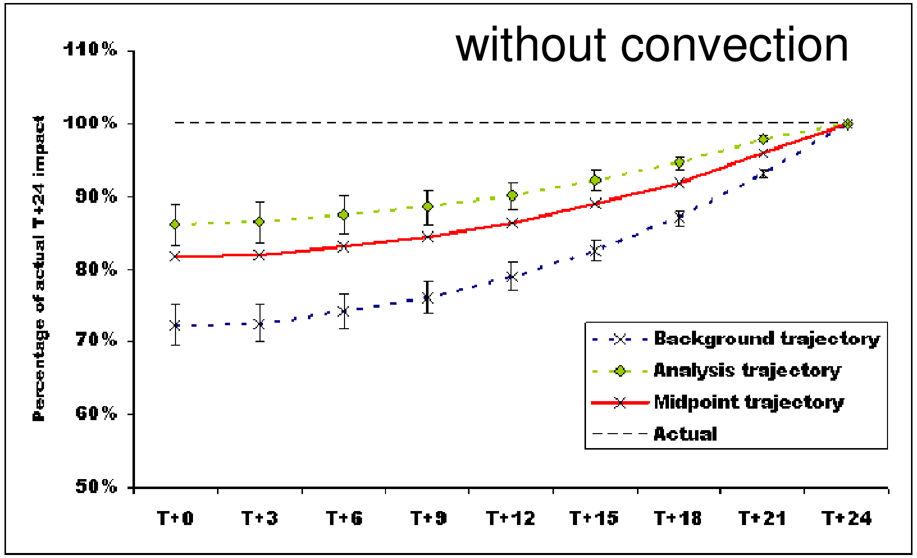
The Perturbation Forecast (PF) Model

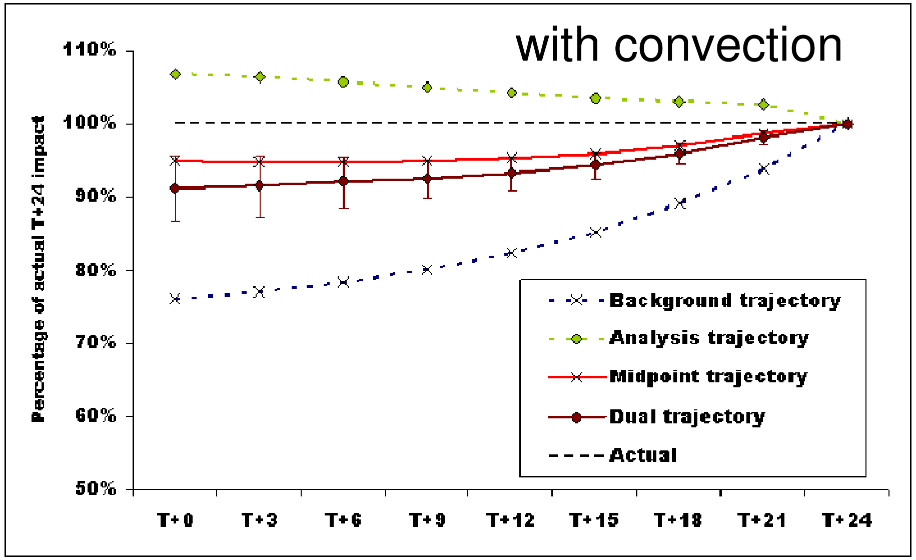


- Met Office 4D-Var does not use a tangent-linear model.
- Small-scale features should not be allowed to continually grow at the rate of infinitessimal perturbations such that they obscure large-scale features.
- Instead we use a regularised "PF" model which is designed to be a good approximation to the growth of a finite perturbation in the nonlinear model. I.e. $\mathbf{M}_{PF} \approx \Delta\mathbf{x}^f / \Delta\mathbf{x}_0$
- Our observation sensitivity equation is then:

$$\frac{\Delta J}{\Delta \mathbf{y}^{\circ}} = \mathbf{K}^T \left(\frac{\Delta \mathbf{x}^f}{\Delta \mathbf{x}_0} \right)^T \left(\frac{\Delta J}{\Delta \mathbf{x}^f} \right) = \mathbf{K}^T \mathbf{M}_{PF}{}^T \mathbf{C}(\delta \mathbf{x}^{fb} + \delta \mathbf{x}^{fa})$$

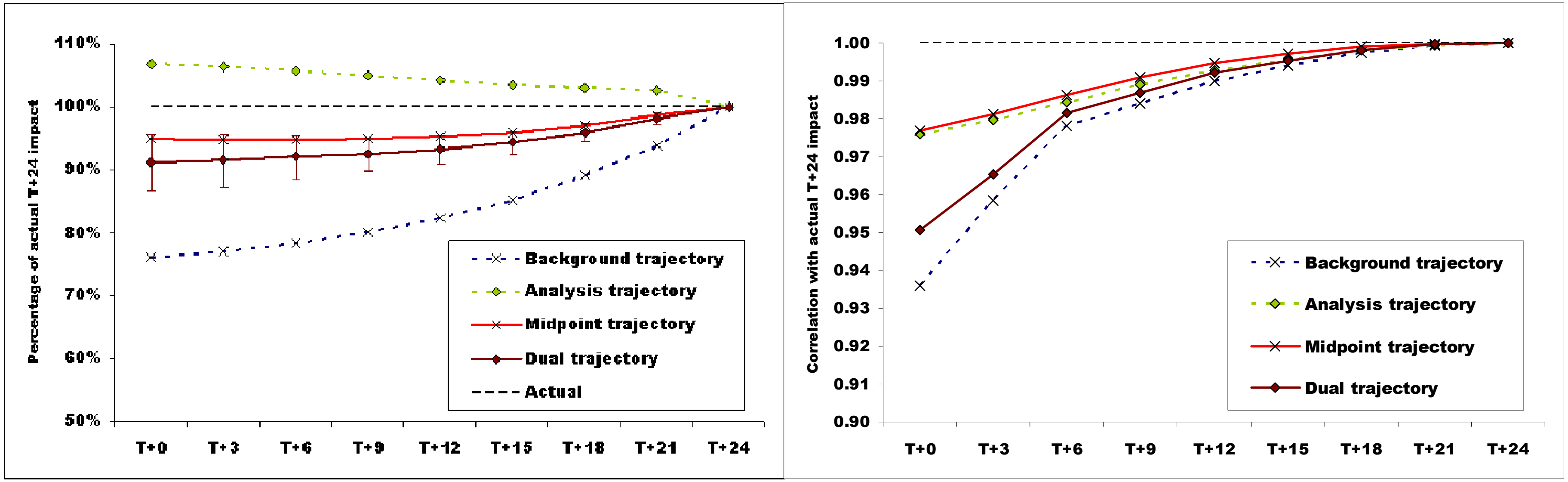
Linearisation of the PF model





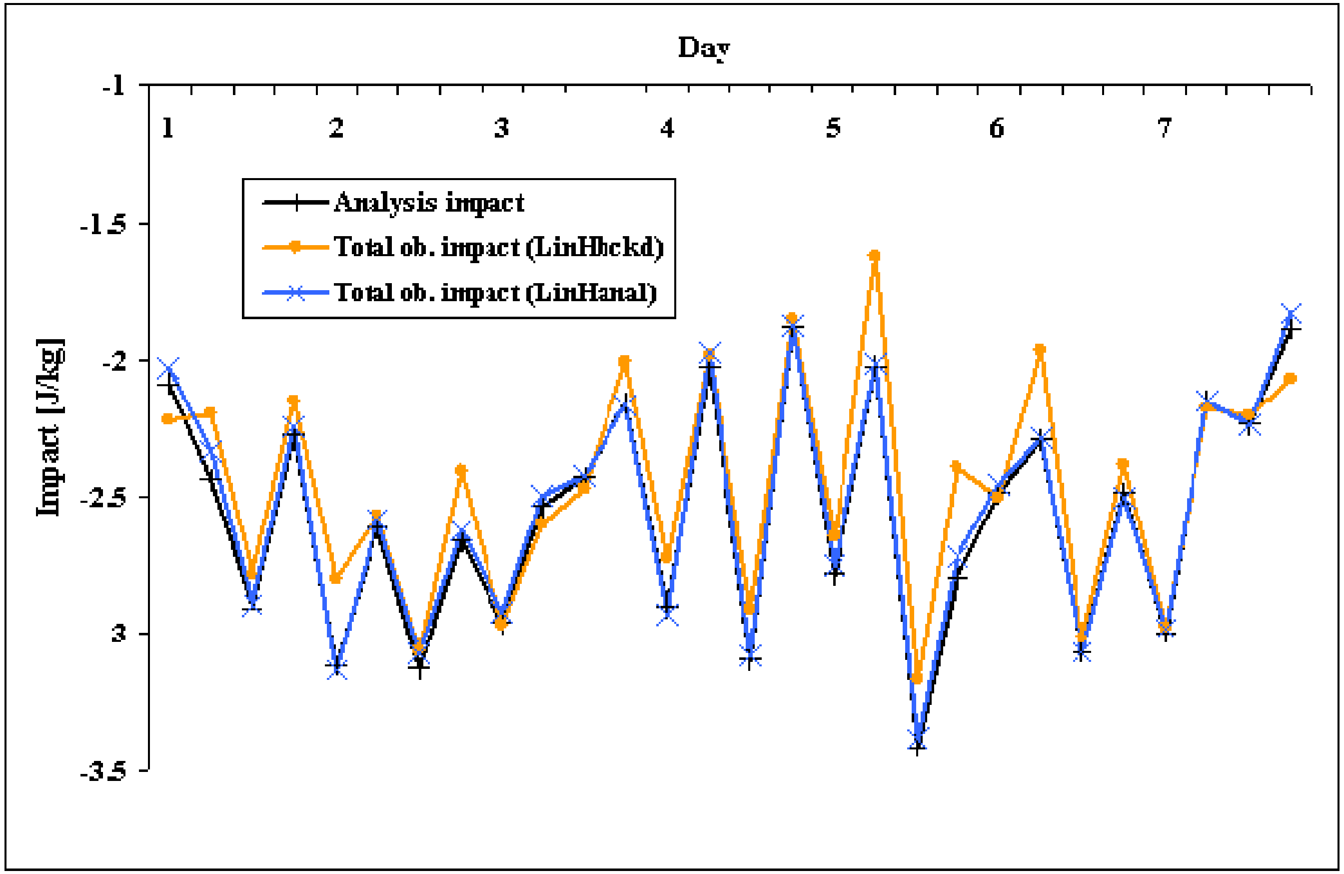
- “Trapezoidal quadrature methods”, i.e. dual-trajectory methods, recover ~78% of forecast impact (no moist physics). Due to our finite forecast-sensitivity we get similar results with only a single adjoint model run, no matter which trajectory we linearise about. (Midpoint trajectory is more likely to be more accurate.)
- Enabling moist physics allows recovery of ~95% of the impact.
- Moist physics improves correlation with forecast impacts at T+0 for midpoint method (0.96 → 0.98).

Linearisation of the PF model



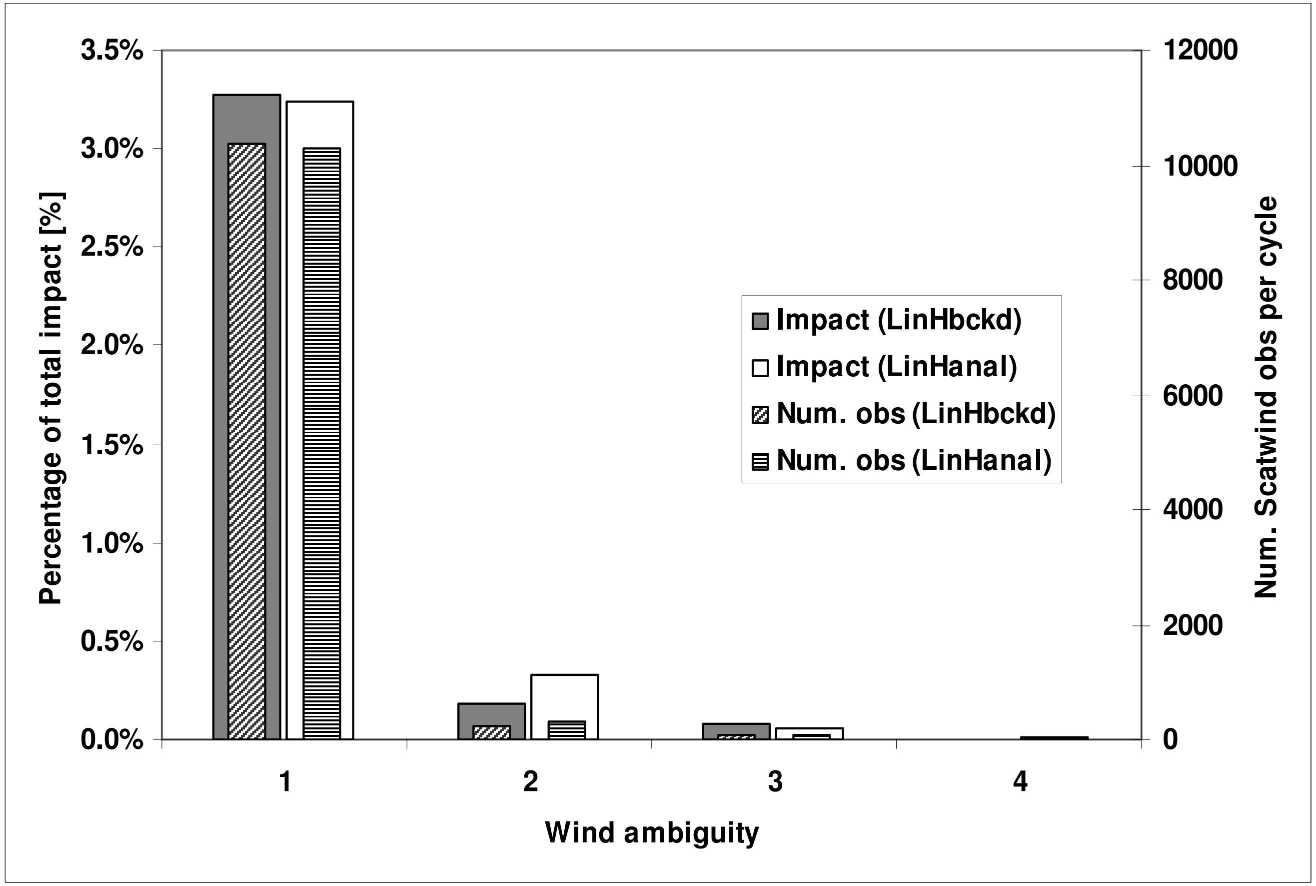
- Averaged (“midpoint”) trajectory best by both measures.
- No benefit seen from running a second adjoint forecast.
- Analysis trajectory impacts are strongly correlated but biased.

Linearisation of VarAdjoint



- Correlations: LinHbckd=0.936; LinHanal=0.997
- Improved impacts due to the correct observation set being used.
- Other reasons for improvement?

Scatwind dealiasing



- Category 1: Obs -1%, Impact -1%;   Category 2: Obs +28%, Impact +83%
- The total Scatwind impact share increased from an average of 3.5% to 3.6%, so a 3% improvement.

Other effects of VarAdjoint linearisation

Assimilated observations	Ob/Anal-impact correlation (LinHbckd)	Ob/Anal-impact correlation (LinHanal)	Diff.
ATOVS	0.810	0.988	0.178
Scatwind	0.984	1.000	0.016
GPSRO	0.978	0.988	0.010
AIRS	0.995	0.999	0.004

- Single ob-type assimilations performed.
- No VarQC for ATOVS, GPSRO or AIRS.
- Improvement suggests that the gradient at the analysis point better represents nonlinear K.

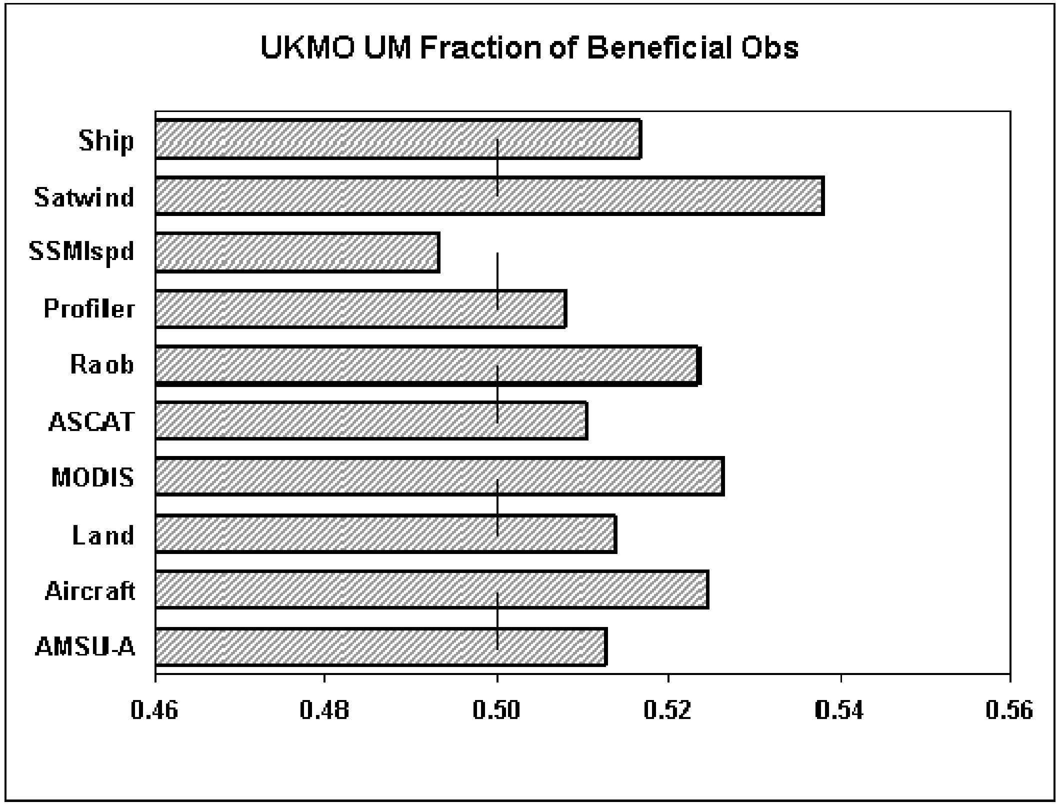
Met Office OS System setup

- Implemented in global model
- Impact on 24-hour forecasts
- Moist energy-norm (u, v, theta, p, q) using latent heat of condensation
- Penalty calculations and adjoint steps performed at Var-resolution on simplified states
- Finite forecast sensitivity calculated
- Single adjoint model integration (linearised about averaged trajectory) with moist physics enabled
- Use Var descent algorithm to minimise Observation Sensitivity cost function

$$J(\hat{\mathbf{a}}) = \tfrac{1}{2}(\hat{\mathbf{a}} - \hat{\mathbf{v}})^T (\hat{\mathbf{a}} - \hat{\mathbf{v}}) + \tfrac{1}{2}\hat{\mathbf{a}}^T \mathbf{U}^T \mathbf{G}^T \mathbf{R}^{-1} \mathbf{G} \mathbf{U} \hat{\mathbf{a}}$$

where **G** = **HM** and **H** is linearised about the analysis state

Why are only ~51% obs beneficial?



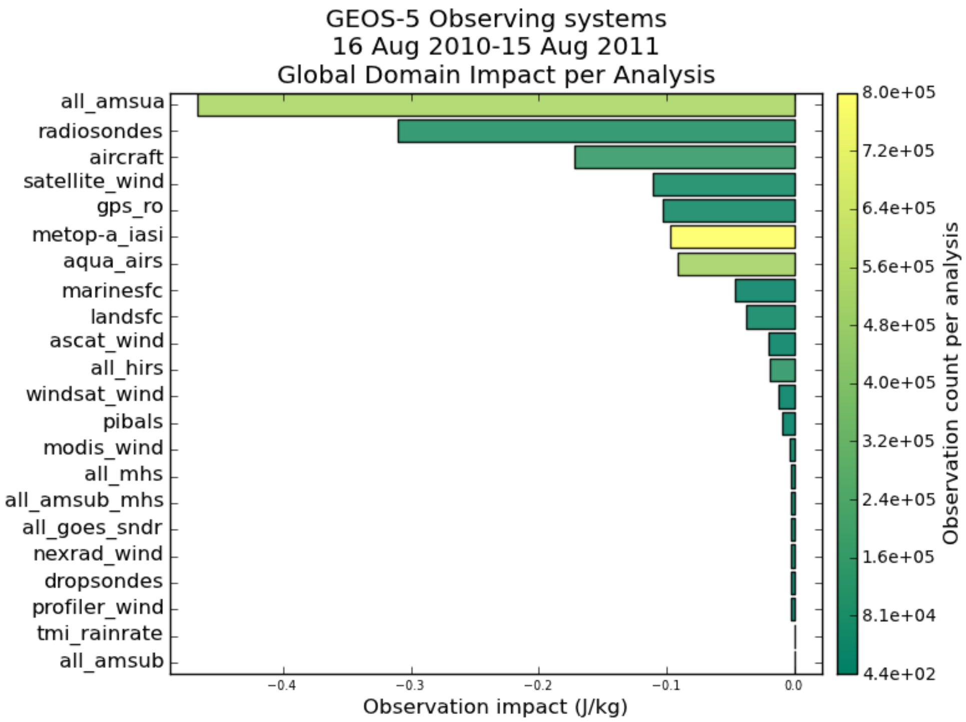
1) Random verification errors in analyses
2) Random observation errors
3) Error growth in the model

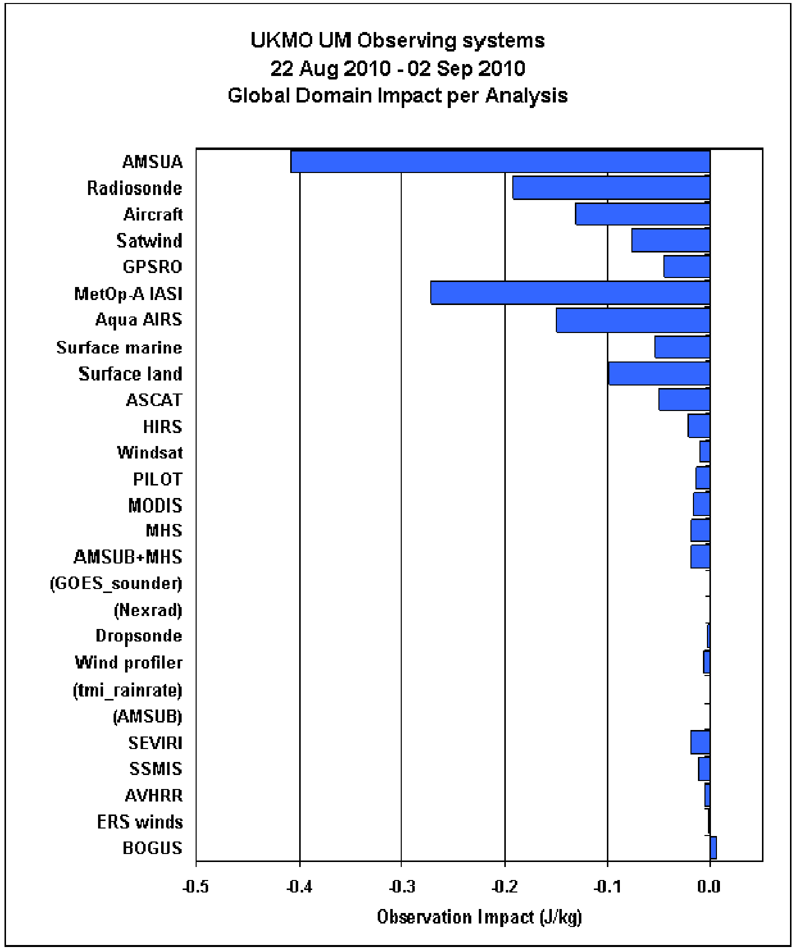
Toy model results

	Errors in assumed background covariance: **B**	Obs error variance: **R**	Error in verifying analysis: **A**/**B**	Error-growth per day: **M**	Mean relative impact	% useful
A	None	0	0	1	-12.0%	100%
B	None	0	**0.707**	1	-6.9%	67%
C	None	**1**	0	1	-6.0%	64%
D	None	0	0	**2, 0.5**	-11.7%	66%
E	None	**1**	**0.707**	**2, 0.5**	-4.3%	58%
F	**+50%, -50%**	1	0.707	2, 0.5	-3.0%	54%

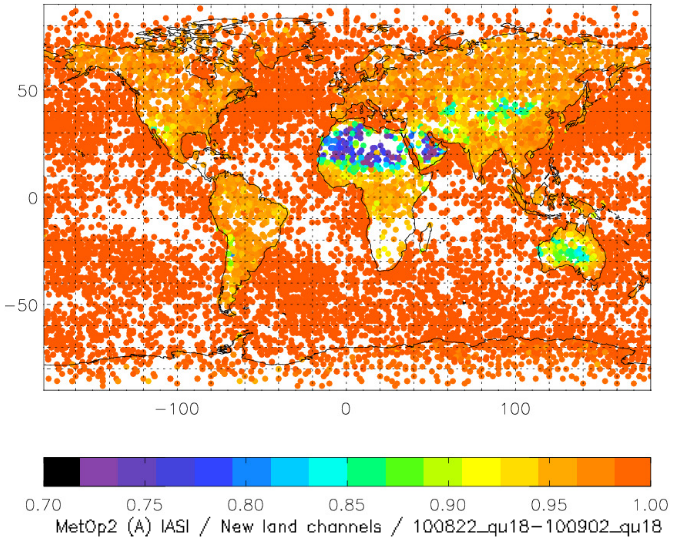
- Perfect obs with perfect **B** improve analyses but not necessarily forecasts.
- The effect of the incorrect partitioning of increments between error-modes is on a similar scale to that of random ob and verification errors.
- The fraction of beneficial obs could be improved by ~4% by improvements in **B**.

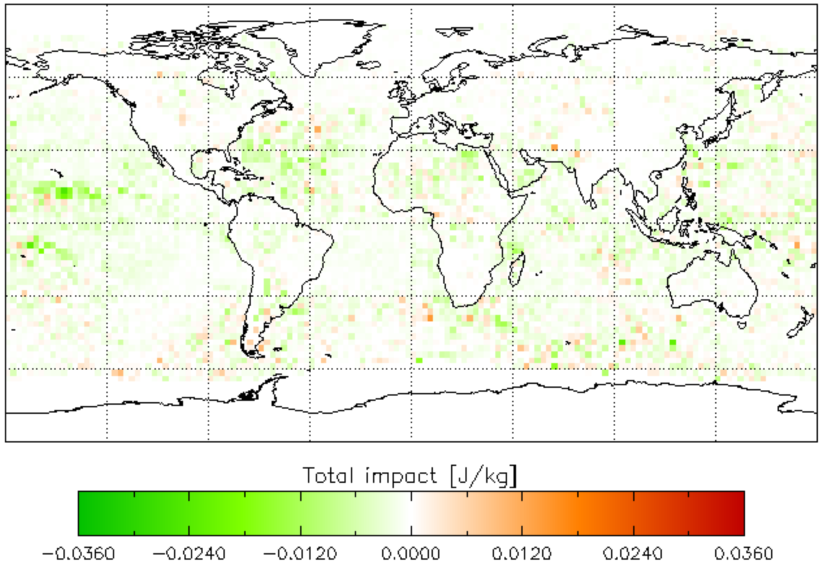
Comparison of impacts with GEOS-5



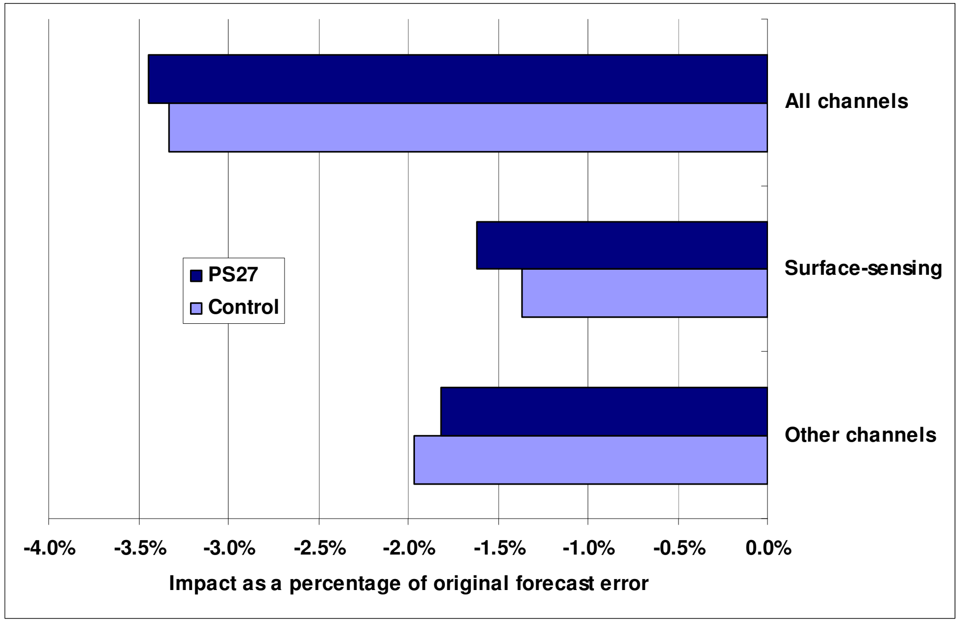


A few results: IASI Impacts





- New surface emissivity atlas (and hybrid) trialled in Parallel Suite (PS27) – now operational.
- Normalised impact for "other channels" decreased. (So too did overall impact – 25% down to 23% error reduction.)
- Thought to be problem with verifying DA changes against analyses.



Future work

Immediate plans:

- Technical change to utilise pre-conditioning
- Interface with ODB for more efficient analysis of impacts
- Investigate the effect on relative impacts of running at reduced resolutions

Longer term plans:

- Investigation of forecast error metrics, ideal forecast lengths, etc. for implementation in high-resolution models

Summary

- Finite forecast error gradient → Cheaper system; simpler to interpret impacts.
- No benefit seen from running two adjoint forecasts.
- Moist physics improves recovered impacts. (82% to 95%)
- H in VarAdjoint linearised about analyses gives better results (even with no VarQC).
- Error growth in model partly explains why ~49% of obs are measured as having detrimental impacts.
- Possible problems with verification against own analyses.

Questions and answers