Forecast-error-sensitivity to observations in the UM

Richard Marriott and Andrew Lorenc
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 ($J = \delta x^T C \delta x^f$) is known to be the dominant nonlinearity of the OS problem.
- Gradient of $J$ (at point B) given by
  \[
  \frac{dJ}{dx_0} = 2M_b^T C \delta x^{fb}
  \]
  Tends to give total impacts $\times 2$
  \[
  \delta x_0^T \left( \frac{dJ}{dx_0} \right) \approx 2(M_b \delta x_0)^T C (M \delta x_0)
  \]
- Impacts are normally calculated using higher order methods:
  \[
  \delta J = \delta x_0^T (M_b dx^{fb} + M_a dx^{fa})
  \]
The Met Office finite OS method

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

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

- The impact is exact in $\Delta x^f$ and should be in $\Delta x_0$ too provided the linear model approximation is good.
The Perturbation Forecast (PF) Model

Met Office

• 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. $M_{PF} \approx \Delta x^f/\Delta x_0$
• Our observation sensitivity equation is then:

$$\frac{\Delta J}{\Delta y^o} = K^T \left( \frac{\Delta x^f}{\Delta x_0} \right)^T \left( \frac{\Delta J}{\Delta x^f} \right) = K^T M_{PF}^T C (\delta x^{fb} + \delta x^{fa})$$
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 $\rightarrow$ 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.
Correlations: LinHbckd=0.936; LinHanal=0.997
Improved impacts due to the correct observation set being used.
Other reasons for improvement?
• 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

<table>
<thead>
<tr>
<th>Assimilated observations</th>
<th>Ob/Anal-impact correlation (LinHbckd)</th>
<th>Ob/Anal-impact correlation (LinHanal)</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATOVS</td>
<td>0.810</td>
<td>0.988</td>
<td>0.178</td>
</tr>
<tr>
<td>Scatwind</td>
<td>0.984</td>
<td>1.000</td>
<td>0.016</td>
</tr>
<tr>
<td>GPSRO</td>
<td>0.978</td>
<td>0.988</td>
<td>0.010</td>
</tr>
<tr>
<td>AIRS</td>
<td>0.995</td>
<td>0.999</td>
<td>0.004</td>
</tr>
</tbody>
</table>

- 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

\[ J(\hat{a}) = \frac{1}{2}(\hat{a} - \hat{v})^T(\hat{a} - \hat{v}) + \frac{1}{2}\hat{a}^T U^T G^T R^{-1} G U \hat{a} \]

where \( G = H M \) 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

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

- 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 $\rightarrow$ 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 $\sim$49% of obs are measured as having detrimental impacts.
- Possible problems with verification against own analyses.

Questions and answers