Practical issues relating to the assimilation of GPS radio occultation measurements

Sean Healy

Acknowledgements
Jean-Noël Thépaut, John Eyre, Chris Marquardt

ECMWF

Outline

1) Importance of realistic error estimates.

2) Quality control of observations, prior to assimilation.

3) Choice of observation operator. Refractivity or bending angle. Why we chose different approaches at the Met Office and ECMWF.

4) Are 2D operators much better than 1D? (We might start arguing here!)

5) Information content of the measurement (Vertically thinning the data).

Importance of realistic error estimates

People from the NWP community appear to be obsessed with measurement errors, and we often seem to be very negative or argumentative in meetings like this one. We have to understand the limitations of the measurement type.

Historically, poorly understood measurement errors and correlations led to considerable difficulty in seeing any positive impact from satellite sounding data in the 1980’s. They were being treated as “*poor quality radiosondes*” (see Eyre, QJRMS, (1989), vol 115, 1001- 1026).

We need the need the error estimates in order to weight our data in the (4D-Var,3D-Var) assimilation system (or 1D-Var retrieval).

Recall the famous cost function…..

Background error cov. matrix

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x_b})^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x_b}) + (\mathbf{y_m} - \mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y_m} - \mathbf{H}(\mathbf{x}))$$

$$= J_b + J_o$$

Combined forward model/observation error cov. matrix.

$$\mathbf{R} = \mathbf{E} + \mathbf{F}$$

The errors tell us how close should we try to fit the data! Key point – we need realistic errors in order to get the balance between the background and forecast information correct. The value of cost function at convergence usually gives us some idea whether error estimates are realistic.

ECMWF 4D-Var Jo Tables

A useful diagnostic for checking errors. Initial fit to observations.

Variable	DataCount	Jo_Costfunction	JO/n	ObsErr	BgErr
RO	10901	23990.80144516	2.20	0.331E-03	0.000E+00
	----------	---------------------------	--------		
ObsType 10 Total:	10901	23990.80144516	2.20		

After screening + Var QC

4D-Var analysis fit

Variable	DataCount	Jo_Costfunction	JO/n	ObsErr	BgErr
RO	10894	9146.744474298	0.84	0.331E-03	0.000E+00
	----------	---------------------------	--------		
ObsType 10 Total:	10894	9146.744474298	0.84		

Quality control

All observations contain errors! That is why in assimilation we only try to fit observations to within their expected errors – why we need to know **R**!

However, occasionally observations contain unusual large errors (“gross errors”) and they do not belong to the population characterised by the covariance matrix **R**.

The purpose of QC is to identify the “gross errors” and remove them from the observation set (or reduced their weight!) before the 4D-Var minimization.

ECMWF approach

1) Screening step (ECMWF tech memo. 236, available at http://www.ecmwf.int/publications/library/do/references/list/14)

Reject an observed value if

$$|y_o - y_{nwp}| > \alpha(\sigma_o^2 + \sigma_b^2)^{1/2}$$

Observation error From **R**.

Typically $\alpha \approx 6$

Observed value

Simulated value from NWP forecast

NWP forecast error mapped to ob. space

VAR QC

Andersson and Jarvinen, QJRMS, (1999), 125,p 697-722

During the 4D-Var minimization, they try to estimate the "probabilty of gross error" (PGE) from the size of the **latest** observed minus simulated departure, normalised by the observation error.

$$Z = \left(\frac{y_o - y_{nwp}}{\sigma_o} \right)$$

They evaluate the PGE given the latest value of Z, using a Bayesion approach. What is the PGE given this value of Z? If the PGE is high, the observation gets little weight in the minimization.

KEY POINT: The variational assimilation (or retrieval) code will produce very strange answers if you pass it data with gross errors, but give it the usual weighting!

Assimilate bending angle or refractivity?

Met Office: refractivity with a 1D operator

The top of the Met office model was ~40km when we started to work on the assimilation trials. We used the GFZ refractivity values so they had performed the extrapolation and statistical optimization. Assimilating refractivity (up to 30km) was easier to code up, computationally inexpensive and it worked (Healy et al, GRL, 2005, v32, L03804,doi:10.1029/2004GL020806) Errors more complicated – eg, correlations introduced by statistical optimization, Abel Transform.

ECMWF: bending angles with a 1D operator

The ECMWF model top at ~65km – extrapolate it without introducing large errors. UCAR "Raw" bending angles assimilated up to 40 km. No "statistical optimization" required. Stepping stone towards implementing a 2D bending angle operator (2D operator required major coding changes in the 4D-Var assimilation system).

Extrapolation above the NWP model

Extrapolation above the NWP model is often considered to be problematic and is cited as a reason for assimilating refractivity N. In fact, its not a big problem.

Bending above the model top can be approximated with,

$$\Delta\alpha_{ext}(a) = 10^{-6}\sqrt{2\pi k_l a} N_l \exp(k_l(x_l - a)) \times \left[1 - erf(\sqrt{k_l(x_l - a)})\right]$$

Consider a bending angle with an impact parameter 40km above the surface and estimate the bending above the model top (0.1hPa, 65km).

Plug in these values

$$N_l \approx 0.031 Nunits$$

$$k_l \approx 1/(7km)$$

$$a \approx (6371 + 40)km$$

$$x_l - a \approx 25km$$

And use

$$\exp(u)\left[1 - erf(u)\right] \leq \frac{2}{\sqrt{\pi}} \times \frac{1}{u\sqrt{u^2 + 4/\pi}}$$

Find the bending above the model top is **~0.7 microradians**. The observed bending angle error is typically **~6 microradians**. We don't need a very complicated method ensure the extrapolation errors are small compared with the observation errors.

2D GPS RO observation operators – are they worth it?

There is a significant amount of research into the development of fast 2D GPSRO observation operators. The hope is that they will reduce the errors (forward model/observation) associated with horizontal gradients.

Solution (1D-Var or 4D-Var) error covariance matrix for 2d operator:

observation errors

forward model error

$$\mathbf{A}_{\mathbf{2d}} = (\mathbf{B}^{-1} + \mathbf{H}^{T}(\mathbf{E} + \mathbf{F}_{\mathbf{2d}})^{-1}\mathbf{H})^{-1}$$

Background errors

Linearized 2d operator

Although they should be more accurate, **I'm** not sure they will produce a major improvement in the NWP forecasts *relative to the 1D operators*. WHY?

1) Often the 1D and 2D bending angle operators give almost the same results.

2) Some horizontal gradient errors are still buried in the 2D operators.

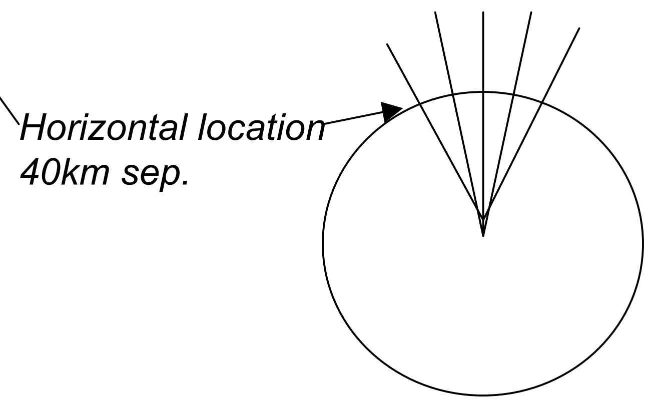
3) The horizontal width of the GPSRO weighting functions

4) Incremental 4D-Var and the role of the B-matrix in assimilation

Simple 1d vs 2d operator example

Set up a spherically symmetric, exponentially decaying atmosphere.

refrac(i,j) = 350.0*EXP(-i*0.25/6km)

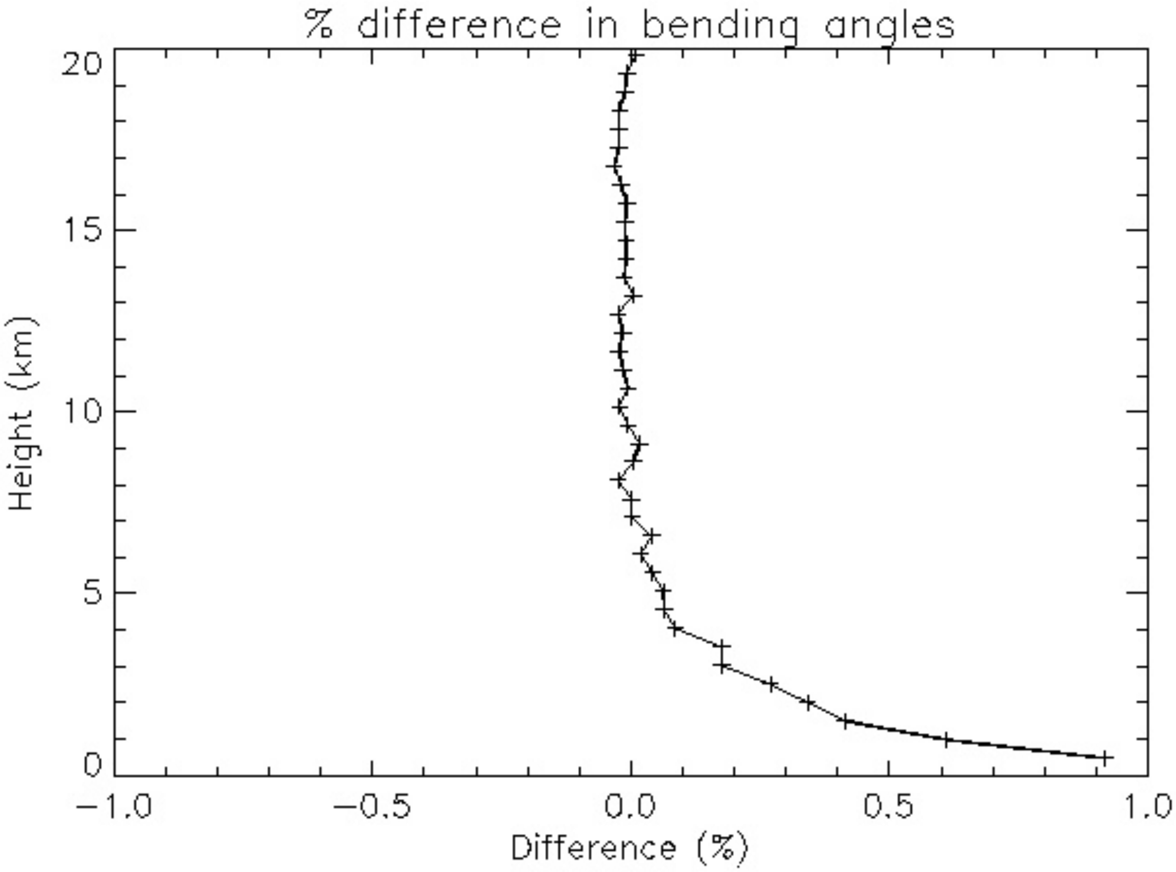


Add on the **STRONG** linear perturbation

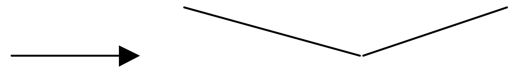
refrac(i,j) = refrac(i,j) +150.0*(kcen-j)/(kcen-1)

*exp(-i*0.25/4km); j=1,21, kcen = 11 (central profile)

Compare the bending angle values calculated with the 1D operator at kcen, with a 2D Runge-Kutta ray-tracer.
RK starts ray at tangent point, derived from the impact parameter!

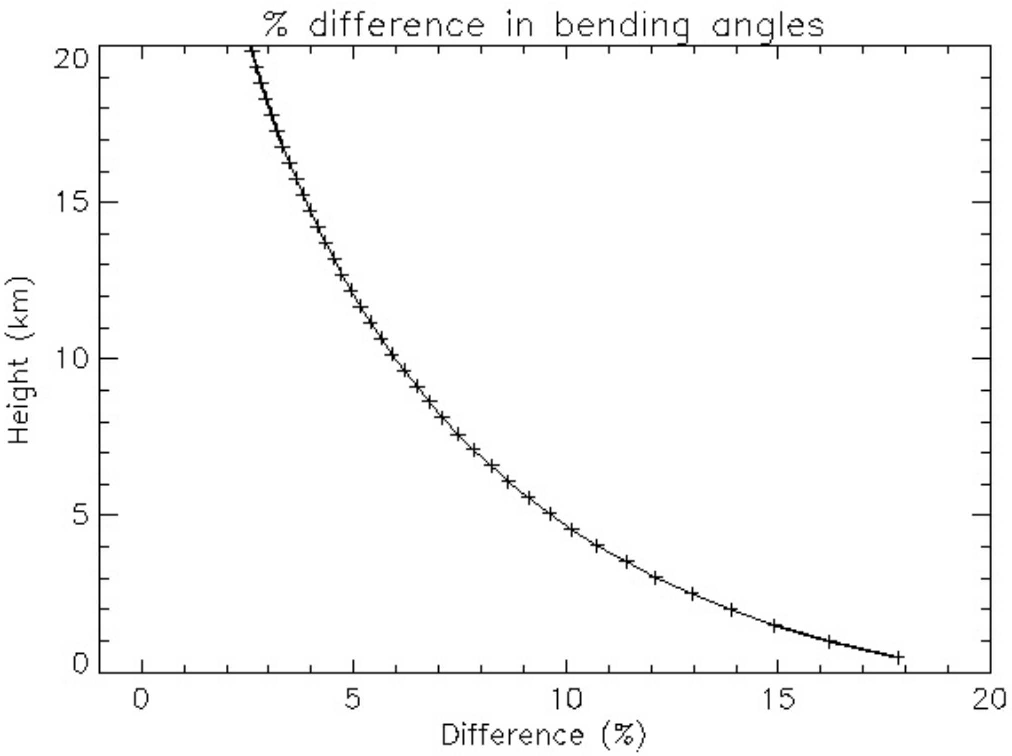


Consider even perturbation



*refrac(i,j) = refrac(i,j) +150.0****ABS****(kcen-j)/(kcen-1)*

**exp(-i*0.25/4km)*



So the 1D operator is fine for odd perturbations, but behaves poorly for even (symmetric) perturbations.

To first order, errors caused by odd perturbations cancel (Melbourne et al, 1994).

I think that’s a bit too simplistic!

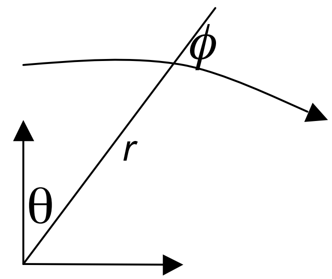
In fact, the 2D Runge-Kutta results cannot be considered “truth” because the **derived impact parameter** is used to determine the ray tangent height!

"Buried" approximations

We do not measure bending angle and impact parameters directly! **They are derived using assuming spherical symmetry!**

The impact parameter is assumed to be constant along the ray-path. It is not constant when there are horizontal gradients. The actual variation is given by

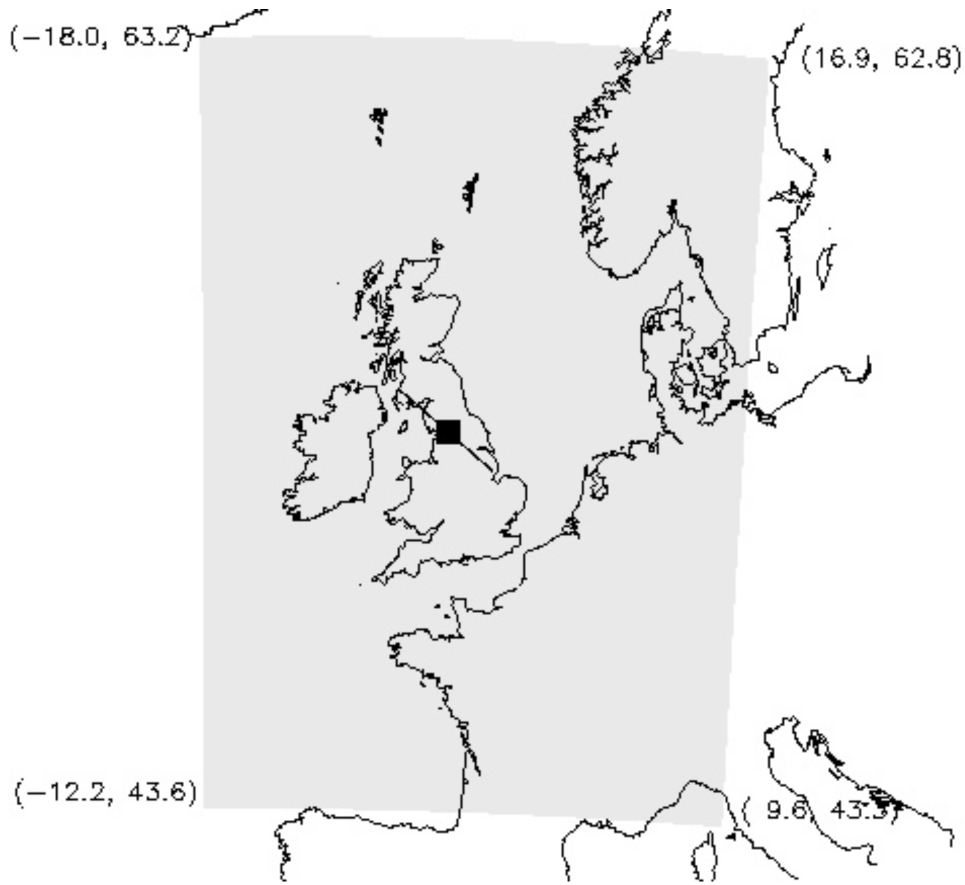
$$\frac{da}{ds} = \frac{d(nr\sin\phi)}{ds} = \left(\frac{\partial n}{\partial \theta}\right)_r$$



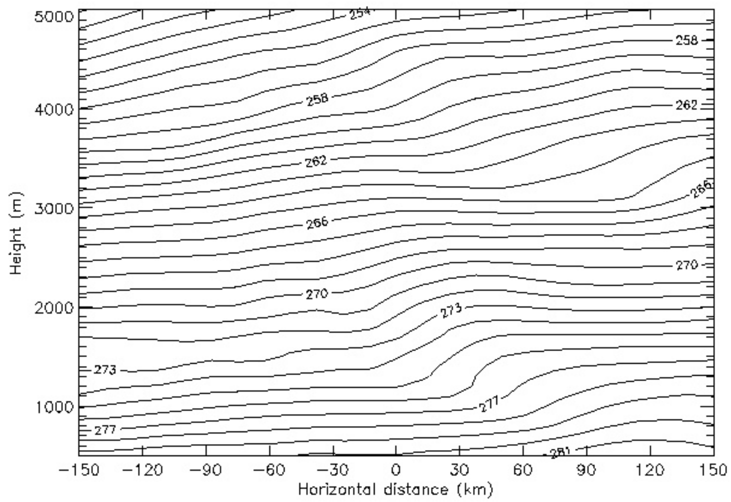
In the 2D bending angle operators we assume the derived impact parameter gives us the tangent point height! We start the ray of from the assumed tangent height. That is a source of horizontal gradient induced forward model error!

An example – ray tracing through a mesoscale model front, 12km by 12 km resolution

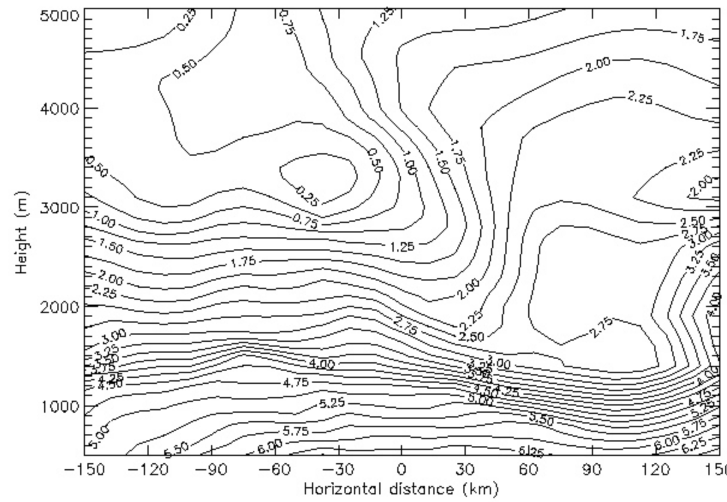
(Healy, JGR, 2001, v106, 11875-11889)



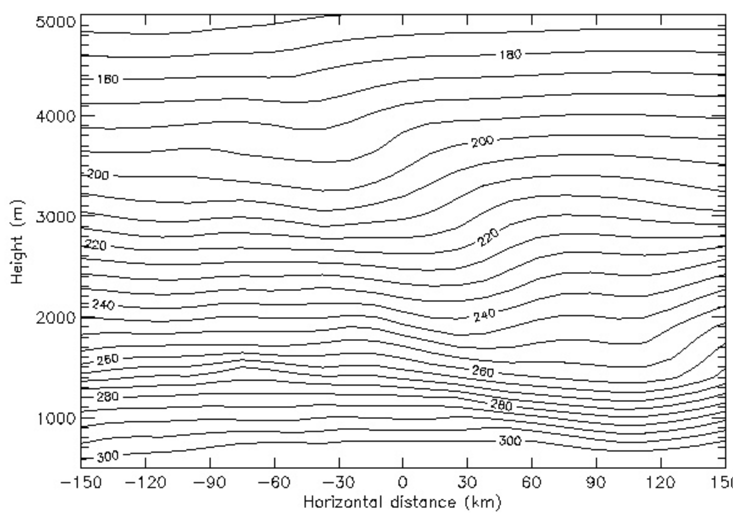
Temperature



Specific humidity

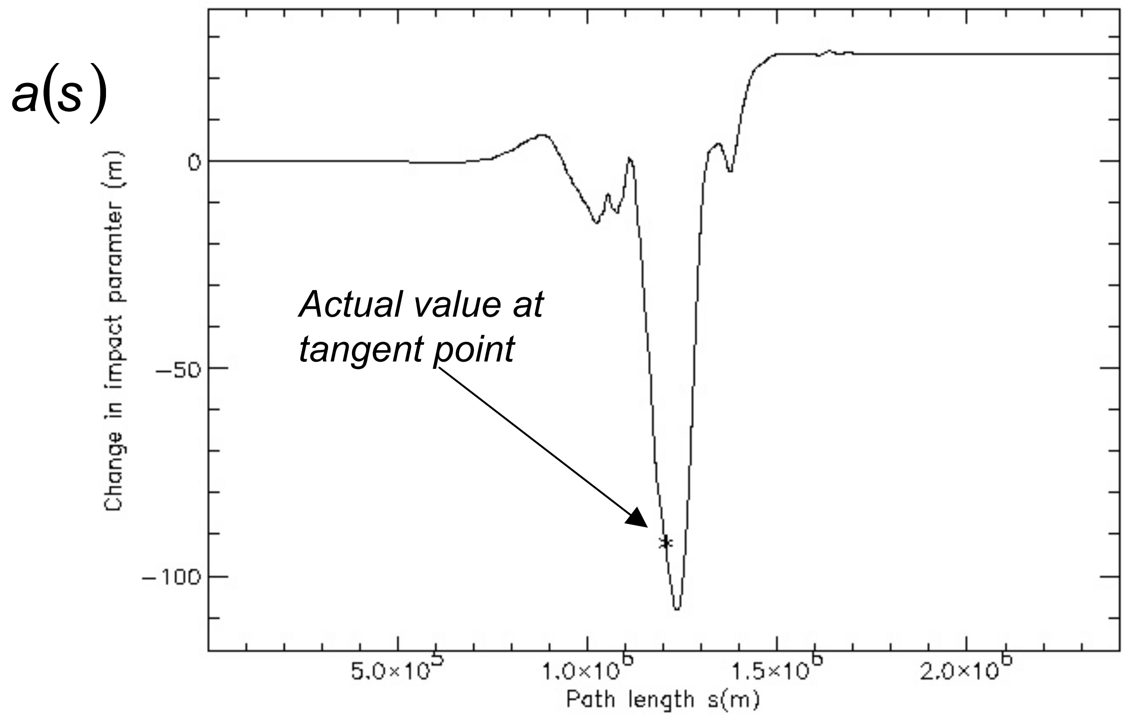


Refractivity



ECMWF

Variation of impact parameter along path

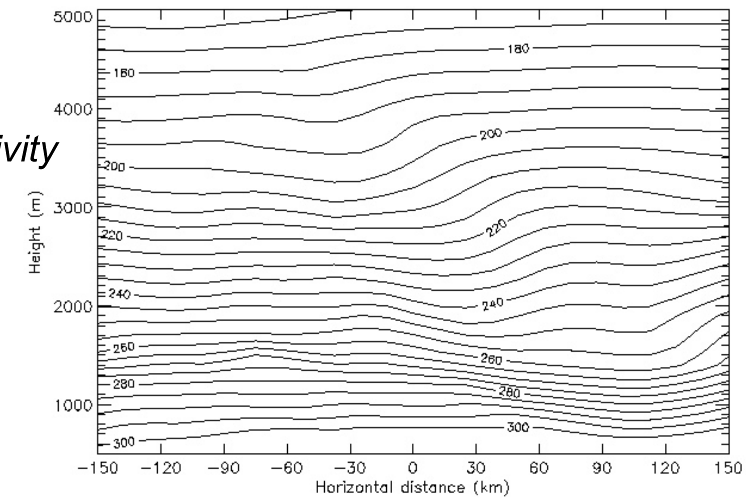


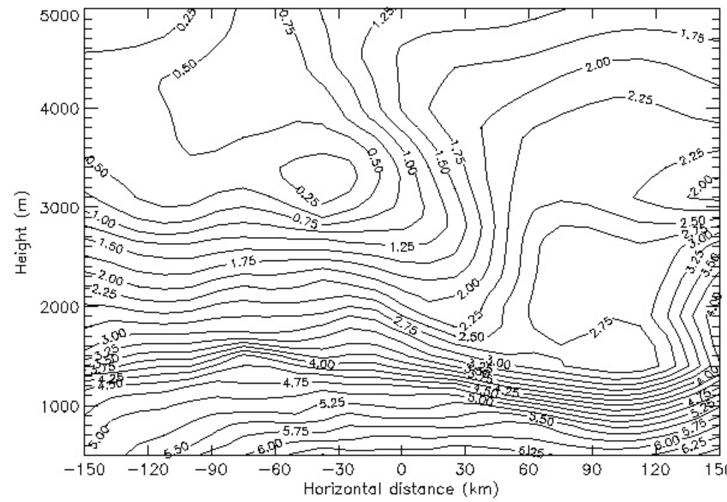
The impact parameter value at the tangent point is ~100m less than the derived value that we use in the fast forward models. This led to a 20% error in the simulated bending angle forward model that started the ray from the assumed tangent point!

But the NWP model can tell us about the horizontal gradients, so we can account for this error(??)

Perhaps, but does the NWP model have sufficient resolution in the horizontal to model the impact parameter variation accurately. EG. How would a global model with ~ 100 km resolution reproduce

Refractivity





Spec.hum.

The inability of a "coarse" NWP model to reproduce the **real** horizontal gradients leads to "**Representation error**", **<u>a form of forward model error</u>**. *Have we just ended up renaming the horizontal gradient error?*

Horizontal resolution is improving!

ECMWF currently runs at T519 resolution ~40km.

$$40km = \frac{\pi \times 6371km}{519 + 1}$$

But, in **incremental** 4D-Var the increments (*the corrections made to the forecast by the assimilation system*) are actually calculated at T159 (~125km).

This approximation is fine because the assimilation system is trying to correct large scale errors.

(ECMWF will move to T799 and T255 later this year)

Assume the forecast model has sufficient resolution…..

But the forecast gives a pretty poor representation of the horizontal gradient – ie, the forecast is wrong!

Will assimilating RO change the horizontal gradients?

1) Its accepted that the width of the horizontal weighting function of the RO are broad (~200-300km). Also the 4D-Var **B** matrix will **filter** scales of the increments you make. The averaging kernel will tell us something about the filtering/smoothing

true state

Linear theory, given perfect obs.

$$\mathbf{x_a} - \mathbf{x_b} = \mathbf{G}(\mathbf{x_t} - \mathbf{x_b})$$

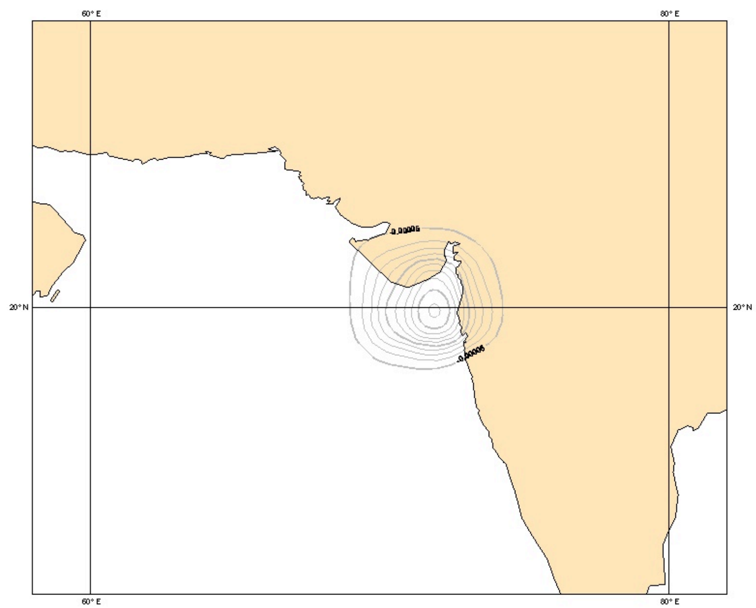
$$\mathbf{G} = (\mathbf{B}^{-1} + \mathbf{H}^{T}(\mathbf{E} + \mathbf{F_{2d}})^{-1}\mathbf{H})^{-1}\mathbf{H}^{T}(\mathbf{E} + \mathbf{F_{2d}})^{-1}\mathbf{H}$$

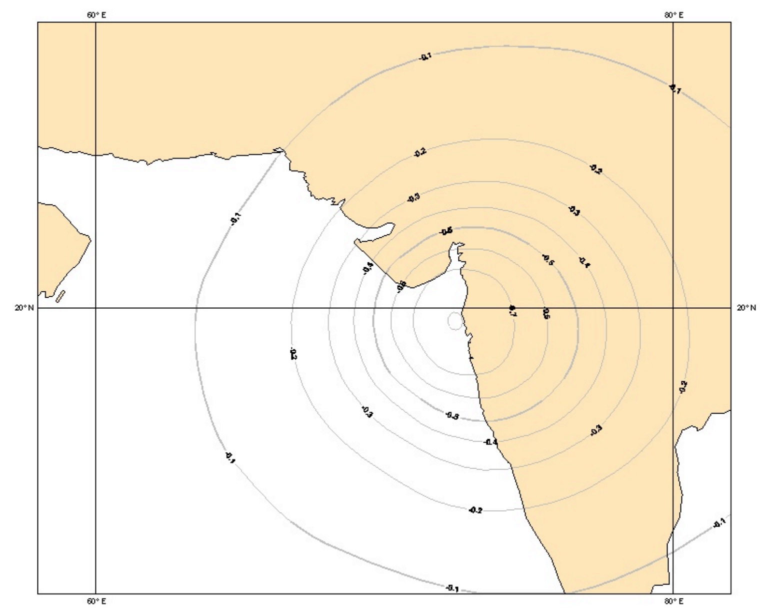
Smoothing error: If we didn't need a **B** matrix **G** would be the identity. **B** stops us from changing the fine-scale horizontal structure.

Purser et al. Mon.Wea.Rev., 2003, v131, p1524-1535

"The inherent smoothness of the background field errors e_b, and hence that of covariace **B** of these errors, is therefore imprinted on the analysis increments themselves".

EG. The single profile results shown in previous talk (1D op results).





ECMWF

Summary (1d vs 2d)

I would argue that simply moving to a 2D operator will not automatically solve all the problems associated with horizontal gradients.

Horizontal gradient errors are often buried in the 2D forward model.

We need to compare forecast impact experiments with 1D and 2D operators to assess the benefits of the 2D method. Hope to start these experiments at ECMWF soon.

Information content - vertical thinning

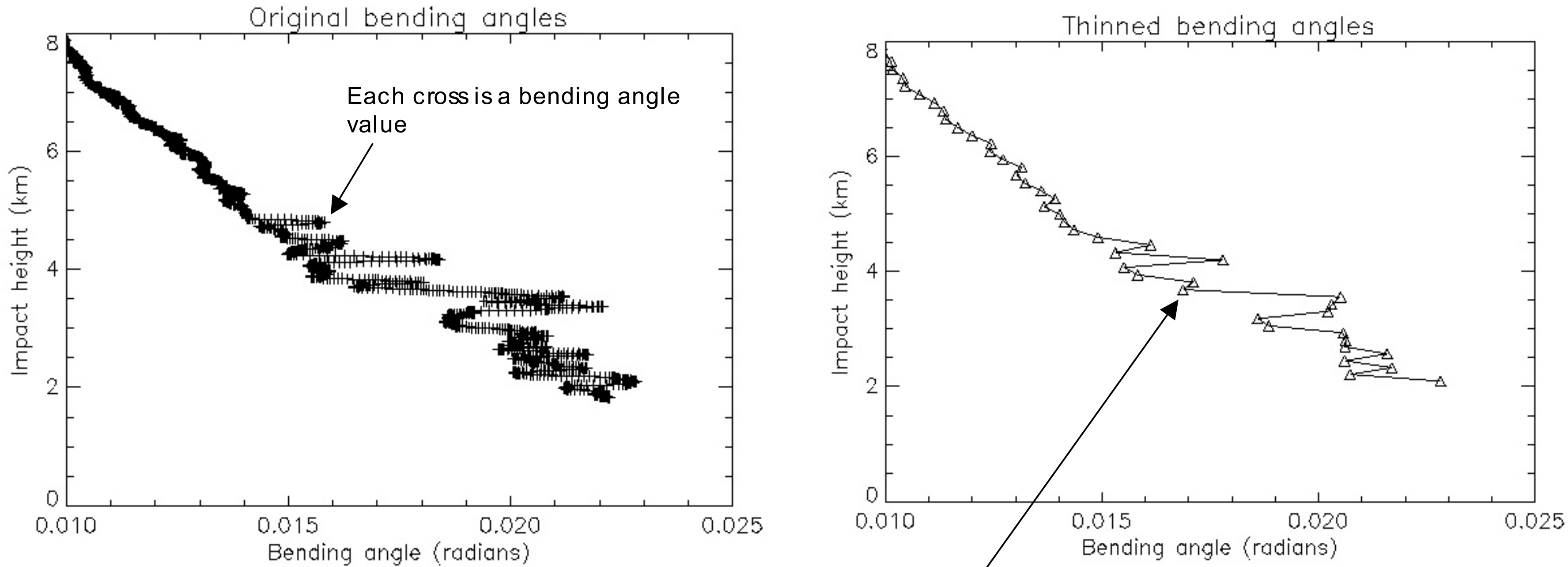
ECMWF use the UCAR bending angle profiles which have a very high sampling rate in the vertical. Typically a file might contain 3000-4000 bending angle values up to around 40km above the surface. These are only separated by a few meters in the vertical near the surface.

We don’t want to assimilate 4000 bending angles per profile. Its computationally expensive do do this in the 4D-Var. In addition, we think we extract all the useful information from a much smaller subset of bending angles. **We don’t have 4000 independent pieces of information!**

We reduce the number of bending angles to ~180 up to 40km. The separation is around ~100m near the surface, to 300 m near 40km.

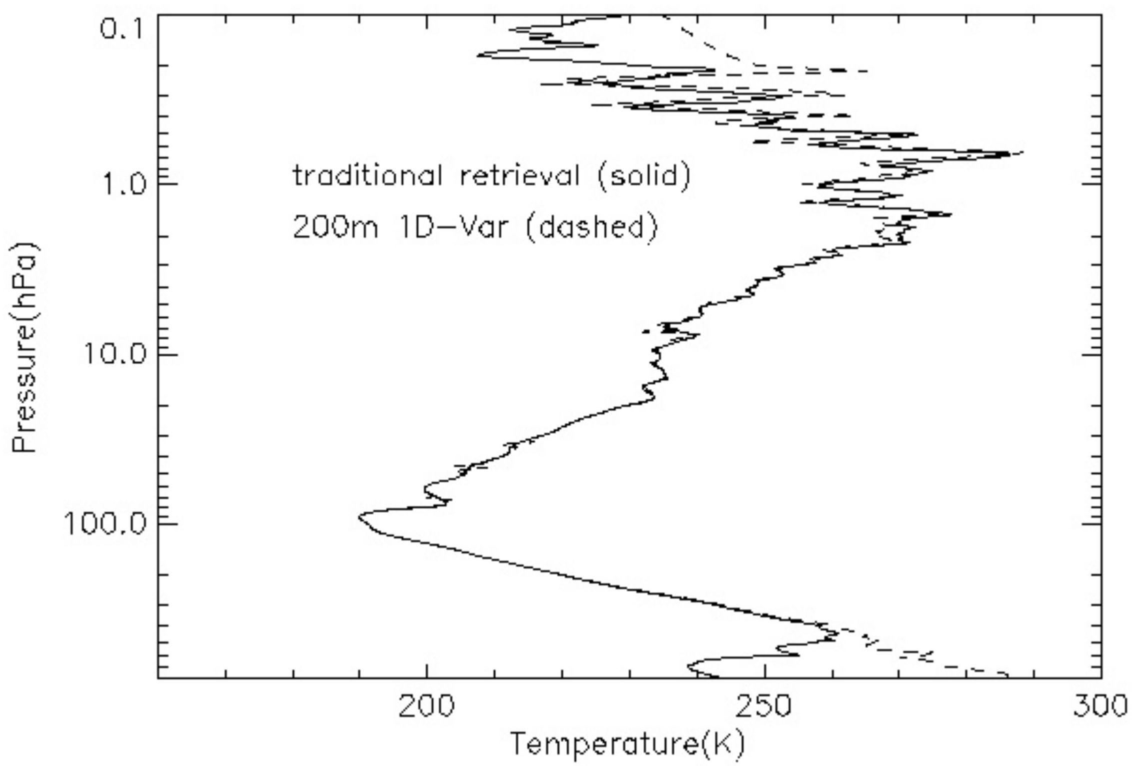
Example: atmPrf_CHAM.2003.213.23.19.G21_0028.0002_nc

(Contained 4016 bending angle values)



The bending angles we assimilate (triangles) are evaluated by interpolating the full resolution data to a set of fixed impact heights.

Example 2: Temperature profile over Nairobi (from Chris Marquardt, Met Office)



The dashed line is retrieval using data with the same thinning as we use in the assimilation.

Summary

1) Its important to provide realistic observation error estimates and QC the data to remove observations with gross errors.

2) Assimilating refractivity or bending angle is ok. Refractivity is easier for getting started. Bending angle requires more coding but extrapolating above the model top isn't a big problem.

3) No one has yet demonstrated 2D operators will provide significantly improvements over 1D operators. Just because you're using a 2D model doesn't necessarily mean the horizontal gradient errors have been solved.

4) We don't assimilate 3000 bending angles per profile. We thin by a factor of ~20. 150-200 bending angles will contain most of the information.

ECMWF