Atmos. Meas. Tech. Discuss., 7, C967–C972, 2014
www.atmos-meas-tech-discuss.net/7/C967/2014/

Interactive comment on “Smoothing error pitfalls” *by* T. von Clarmann

C. Rodgers (Referee)

c.rodgers@physics.ox.ac.uk

Received and published: 21 May 2014

Review of ‘Smoothing Error Pitfalls’ by Thomas von Clarmann

General Comments

The aim of this paper is to show that the concept of smoothing error leads to problems of interpretation that make it not a useful quantity in characterising retrieval errors. I agree that smoothing error is a tricky concept, and should not form part of the error analysis supplied to data users. However it is still a useful concept if handled correctly. The error analysis is best described in terms of the retrieval noise covariance (uncorrelated between separate retrievals), the systematic error covariance (may be correlated

between retrievals), and the averaging kernels on clearly specified representation of the state. (See Rodgers (2000) 11.2.6.) Smoothing error is useful in certain circumstances, but is for the data users to compute based on their own assessment of the state ensemble covariance, and the spatial resolution needed for their particular applications.

Detailed Comments

3303 L1: An optimal estimation scheme is one which optimises something. It covers all sorts of possibilities, of which MAP is just one. I did not retitle the MAP approach, I simply used the name which had been normally used in the literature for many years.

3303 L5: I don’t think the use of the term ‘a priori’ in analytic philosophy is relevant.

3303 L12: “indirect prior knowledge" does not sound like “knowledge" at all. It is simply a representation, used as a practical neccessity (we can only deal with a finite number of rational numbers), which has the effect of constraining the solution to a particular region of state space.

3306 L4: As other reviewers have pointed out, S_a should be S_e. Incidentally the error analysis in Rodgers (2000) chapter 3 applies to any retrieval method. That analysis only uses S_a as a linearisation point.

3306 L18–20: The choice of grid is important. It needs to be chosen so that it can represent all scales of variability that are important to the problem in hand, and so that information present in the raw measurement is not lost. S_e matrices from real data should be constructed in such a way that the are **not** singular. And, of course, due allowance should be made for the errors in the ‘real’ data. See Rodgers (2000) section 10.3.3.

3306 L24–8: If you can think of a better term, please suggest it! If you think of eqn (4) in terms of a fourier expansion or in terms of the eigenvectors of A, you will generally find that smaller scale components are suppressed more than large scale components, hence 'smoothing'. If the jacobians are, as is commonly the case, more or less single peaked smooth functions of finite width, you would expect a smoothing effect.

3308 L5–7: The statement in parentheses could be omitted without loss.

3308 L9: You should remind the reader here that $VW = I_{coarse}$ but $WV \neq I_{fine}$.

3308 L19: What makes you think it has remained unnoticed? Isn't it obviously the case? The same arguments apply as for S_e. Smoothing error should not in principle be calculated on a grid which is too coarse to capture all relevant atmospheric variability. Or, if you do so, you must be aware that you have introduced another source of error âĂŤ- representation error, describing variability in the null space of the representation you are using.

3309 L1: What precisely do you mean when you say "Gaussian error propagation is not compatible with the smoothing error concept". What goes wrong? If I understand you properly, then what you mean is that the smoothing error calculated correctly on a fine grid does not equal the smoothing error calculated on a coarse grid, and then interpolated to a fine grid. Why would anyone expect otherwise? It is not that the smoothing error concept or Gaussian error propagation has gone wrong, it is that you are not comparing like with like.

Gaussian error propagation applies, but it gives you the error of the coarse smoothing error when interpolated to the fine grid, not the smoothing error calculated on the fine grid. That is given by the corrected version of eqn (16).

Incidentally, it can be shown (see appendix) that eqn (15) is related to corrected eqn (16) by replacing $S_{e,fine}$ by $WVS_{e,fine}(WV)^T$, i.e. $S_{e,fine}$ with the finer scales removed.

3310 L22–24: Gaussian error propagation does not have to be abandoned, it just has to be used appropriately. Smoothing error should be calculated on the fine grid, and

then transformed to the coarse grid, if required.

3311 section 4: There is no doubt variability on all scales, and we may not be able to estimate it from independent measurement or from theory (though turbulence theory may be helpful). However there will be a scale below which we can be reasonably sure that our instrument has effectively a zero sensitivity. We should choose a representation such that scales larger than this can be represented. (See Rodgers (2000) 10.3.1.2 and 10.3.1.2.) A data user who wishes to extend the smoothing error to smaller scales (if he has a suitable $S_{e,fine}$!) can then assume that *all* variation at smaller scales is in the smoothing error.

3312 L5–7: It is simply wrong to attempt to do this, as discussed above.

3312 L8–13: We do not attempt to retrieve at dimensionless points. The state vector is always a set of numbers which define a continuous function. It is the continuous function that we retrieve, and which is required to evaluate the forward model.

3312 L13–20: As mentioned above, this is exactly what I recommend, e.g. in Rodgers (2000) 11.2.6. (Though very briefly – I was hurrying to finish the book!)

3312 L20-24, I don't understand this sentence.

3313: Much of this will need a different emphasis in the light of the difference between S_e and S_a.

3314 L6: useful

3315 L17-19: The formalism can be applied ML retrievals. All you need is the representation function (which is the same as W), and an $S_{e,fine}$ matrix. The smoothing error can be evaluated on the fine grid.

3315 L25: for "implicit smoothing error" I would use "representation error" as a more descriptive term.

3315 next sentence: I don't think it has been shown adequately in a practical sense.

Extensionless points are not relevant. You could e.g. use a fourier series.

3316 L4 onward: I do think the problem is purely philosophical and practically irrelevant! Smoothing error does comply with Gaussian error propagation if done correctly.

3316 L15: I agree strongly that the averaging kernel should be supplied. I also feel that the smoothing error should *not* be supplied. The users can evaluate that themselves if they want it.

Comments on Thomas's reply:

C797 Implications for sect 5, L1-6: These are exactly the reasons that TES chose to provide averaging kernels with their data set.

C798 2. Possibility...: But there will always be a scale beyond which it doesn't matter for practical purposes.

Appendix: Eqns 15 and 16

The correct version of eqn 16 requires $A_{fine} = WGK_{fine}$:

$S_{sm,fine} = (I_{fine} - WGK_{fine})S_{e,fine}(I_{fine} - WGK_{fine})^T$

Equation 15 uses $A_{coarse} = GK_{coarse} = GK_{fine}W$.

$S_{sm,fine} = W(I_{coarse} - GK_{fine}W)S_{e,coarse}(I_{coarse} - GK_{fine}W)^T W^T$

Substitute $S_{e,coarse} = VS_{e,fine}V^T$ and bring the W's inside the brackets:

$S_{sm,fine} = (W - WA_{coarse})VS_{e,fine}V^T(W - WA_{coarse})^T$

Substitute $A_{coarse} = GK_{fine}W$ and move the right hand **W**'s:

$S_{sm,fine} = (I_{fine} - WGK_{fine})WVS_{e,fine}(WV)^T(I_{fine} - WGK_{fine})^T$

This is the same as eqn 16, but with $S_{e,fine}$ replaced by $WVS_{e,fine}(WV)^T$, which is

$S_{e,fine}$ projected onto the coarse grid, and interpolated back on to the fine grid, thus removing fine scale information.