Atmos. Meas. Tech. Discuss., 7, C625–C631, 2014
www.atmos-meas-tech-discuss.net/7/C625/2014/

Atmospheric Measurement Techniques Discussions
Open Access

AMTD
7, C625–C631, 2014

Interactive Comment

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

Anonymous Referee #5

Received and published: 22 April 2014

Review of "Smoothing error pitfalls" by T. von Clarmann

General Comments

The paper "Smoothing error pitfalls" by T. von Clarmann discusses implications and limitations of the smoothing error concept in retrieval theory. The "smoothing error" is usually considered being an important diagnostic quantity of an optimal estimation retrieval. It measures the difference between the optimal estimate and the unknown true atmospheric state due to the influence of explicit a priori information. A major finding of the paper is that the smoothing errors for a coarse retrieval grid are not compliant with smoothing errors estimated for a finer grid, if the coarse grid data was simply interpolated to the fine grid and the corresponding smoothing error was determined by means of Gaussian error propagation. Solutions to this problems are presented.

In general, I found the paper well-written and interesting to read. I would also think it is of interest for many readers of AMT. However, right now I am not convinced that all key statements hold and I would dispute the conclusion that the smoothing error concept is "questionable", "untenable", or "failed". Clarification is also needed that the topic has been addressed in the literature before, in particular by Rodgers (2000). I will try to express my concerns in more detail in the following general comments. These concerns need to be carefully addressed and the manuscripts needs to be revised before it can be published.

In section 2 the paper presents a misleading definition and discussion of the smoothing error. In particular, Eq. (5) defines the smoothing error as

S_smooth = (I - A) S_a (I - A)^T.

However, Rodgers (2000, Eq. 3.17) introduced this as

S_smooth = (I - A) S_e (I - A)^T,

with S_e being the covariance of a real ensemble of states rather than the a priori covariance S_a. It seems S_e rather than S_a was introduced intentionally by Rodgers (2000), with the following argument (p. 49): "Many remote observing systems cannot see spatial fine structure, the loss of which contributes to the smoothing error. To estimate it correctly, the actual statistics of the fine structure must be known. It is not enough to simply use some ad hoc matrix that has been constructed as a reasonable a priori constraint in the retrieval." I think this is at the heart of the problem raised by the new paper of T. von Clarmann, i.e., that a coarse grid a priori covariance S_a cannot be simply interpolated to obtain an estimate of S_e on a fine grid. This would not be a completely new finding than.

Rodgers (2000, p. 49) also provides a simple solution to this problem: "If the real covariance is not available, it may be better to abandon the estimation of the smoothing error, and consider the retrieval as an estimate of a smoothed version of the state,

AMTD
7, C625–C631, 2014

rather than an estimate of the complete state." T. von Clarmann came to a similar conclusion in his paper, i.e., to better exclude the smoothing error from the error budget and to supply averaging kernels to the data user instead. However, this would also not be a completely new finding than?

A practical problem that I see with the solution of Rodgers (2000) and the one presented in the paper is that it can become very inconvenient to distribute full averaging kernel matrices along with every retrieval, simply in terms of data size. (I am thinking of current nadir sounders, which provide measurements for millions of footprints per day, or new limb sounders, which may require tomographic retrievals that are associated with large and complex averaging kernel matrices.) Another danger I see is that error budgets excluding the smoothing error could easily be mistaken as "total" errors by inexperienced data users? One likely can find good arguments for both, to include or exclude the smoothing error in the error budget. I would say it is a judgement call and suggest to discuss this more balanced in the paper.

Another major concern is related to the discussion of the "nature of the retrieved quantities": Section 4 first presents a theoretical solution to the error propagation problem, i.e., to simply evaluate the smoothing error on an infinitesimally fine grid. In principle, this would allow for the atmospheric variability on all spatial scales to be taken into account and the problems related to coarse grid representation would not arise. This solution is then discarded with the argument that the representation of the atmospheric state based on infinitesimal volumes of air itself is meaningless. This is illustrated in the paper in the following way: "For a single extensionless point in the atmosphere, the mixing ratio of a species is not a meaningful quantity: either, at the given point, there is a target molecule; then the mixing ratio is one. Or there is a molecule of another species; then the mixing ratio is zero. [...] For number densities and temperature, there are similar problems to define these quantities in any meaningful manner for an infinitesimal point." Based on this proof by contradiction is is argued that the smoothing error concept itself failed. In the conclusion (section 7) the paper reads: "It has further been shown that this problem cannot be solved by representing the atmospheric state on a ’sufficiently fine’ grid, because the estimate of the atmospheric state does not converge to a useful value when the grid approaches an infinitesimally find grid."

However, the complete section 4 does not convince me at all, I am afraid. Starting with the illustrating example of the mixing ratio at a single point: a point itself is by definition infinitesimal (an "infinitesimal point" or "extensionless point" as referred to in the paper is a tautology). In contrast, a molecule has finite extent. I would agree that it does not make sense to try to define mixing ratios in this context. One would need a finite volume to do this. Likewise, a definition of thermodynamic quantities such as temperature does not make sense on the basis of individual points in space. Temperature is a macroscopic quantity that requires a canonical ensemble with a certain finite volume or system size to be defined. However, I think one cannot turn this around and argue that as we cannot define temperature for an infinitesimal volume (a point) the concept of temperature itself is meaningless. This contradicts basic principles of statistical physics. The same would also apply for the smoothing error problem addressed in the paper, I think?

Instead, I would think that this is a scale problem. It is certainly true that atmospheric variability occurs on all scales. However, one would really need to know how variability varies from larger to smaller scales, to judge if the smoothing error problem will be relevant on all scales. For example, I would think that "true data" or "real ensemble data" (Rodgers) with a spatial resolution on the order of a few meters would be perfectly fine to estimate smoothing errors for today’s satellite remote sensing measurements? There would be no need to go for the infinitesimal limit?

I think the problem discussed here arises in practice, because such perfect, high-resolution data are usually not available. However, this is not a problem of the theory itself, but an issue that needs to be specifically addressed for individual problems. Based on that I would also support the concern raised by referee John Worden that the paper needs more illustrative and real-world examples to demonstrate the relevance of the problem. I learned from the online discussion that such practical problems came up in a comparison of MIPAS and TES data. I was also wondering, if the problem would not be of even larger relevance for nadir measurements? Limb measurements, as discussed in the paper, are often not affected by a priori and regularization as much? In its current form the paper is focused on limb applications.

Finally, I noticed that Rodgers (2000, chap. 10.3.1) suggests another practical solution for scaling an a priori covariance from coarser to finer grid: "If we only have an estimate of S_za [coarse grid] we cannot assume that S_a = W S_za WˆT [fine grid], as this will be singular [...] and will carry the implication that any structure not represented by the interpolation has zero variance. If we really have no other information, we should make and educated guess, for example by changing the zero eigenvalues of the eigenvector decomposition of W S_za WˆT to values conservatively extrapolated form the non-zero eigenvalues." This sounds like a very reasonable approach to me. Its applicability could be directly tested with the idealized example presented in section 3 of the paper, for instance. This possible solution should be mentioned in the paper at least.

Specific Comments

p3303,l5-7: I think it is fine to explain what we mean by "a priori" in the context of retrieval theory. I think there is no need to explain its meaning in analytical philosophy in AMT.

p3306,l24-28: It is okay to point out that the term "smoothing error" may be misleading in certain situations, as the retrieval may not always provide a smoothed version of the true profile. However, the standard case is that a smooth a priori profile is combined with a more fluctuating true profile, yielding a smooth optimal estimate. So, I think the term "smoothing error" is okay and I do not see a need to refer to it as "so-called" smoothing error (e.g. p3304,l3 and in other places). Mentioning this "standard case" in the paper would provide a more balanced view.

p3307,l17-18: Rather than saying that the error propagation rules are "generally ac-

AMTD

7, C625–C631, 2014

cepted" for all cases except grossly non-linear functions one might say that this is actually the way how the "moderately non-linear" and "grossly non-linear" case are discerned? I thought the difference of the two is that "moderately non-linear" means "linear within the error bars" whereas "grossly non-linear" means that a linear approximation is not applicable for error estimation? Also, I guess you actually mean "moderately non-linear case" rather than "moderately linear case" in the heading of section 3.1?

p3308,l5-14: It would be nice to adopt the notation of Rodgers (2000, chap. 10.3.1.1). This is basically the same discussion and it would be easier to relate your paper and the discussion of Rodgers (2000) to each other.

p3308,l15: Maybe briefly introduce what is meant by "parameter error" here. The term was not introduced in the paper before.

p3309,l10-15: I was wondering if it is possible to rewrite Eq. (15) in a way that it corresponds to Eq. (16) plus an additional term, which represents the difference of the two equations? This could be helpful to further understand the differences between Gaussian error propagation and direct linear analysis?

p3310,l20-21: The vertical correlation length of 1 km in this illustrative example seems to be rather short. I am sure it helps to demonstrate the point as it implies significant variability on smaller scales? Real vertical correlation lengths may be larger, though? (Larger values are often used in regularization matrices, I think.)

p3312, l8-13: It would be good to repeat here or in another place that the representation of the continuously varying atmospheric state does not only depend on a fixed number of state variables on a fixed set of grid points, but also on a set of interpolation rules that define the state anywhere in-between. (Isn't that also a reasonable way out of the dilemma that the atmospheric state cannot be defined for an individual point?)

p3315, l1-10: I think that these statements require their mathematical derivation to be shown. I could not directly see that these are true.

AMTD

7, C625–C631, 2014

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Technical Corrections

p3309,l19-p3310,l21: This paragraph is rather long. Better split in 2-3 paragraphs.

p3314, l5: Suggest to remove "in their paper".

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 3301, 2014.

AMTD

7, C625–C631, 2014

Interactive Comment