**Atmospheric
Chemistry
and Physics
Discussions**

Interactive
Comment

Interactive comment on "Technical Note: Trend estimation from irregularly sampled, correlated data" *by* T. von Clarmann et al.

Anonymous Referee #1

Received and published: 10 February 2010

The trend of a selected atmospheric state variable is often derived by fitting a line to measurements of this variable performed at different times. The regression line is often determined regardless of the correlations existing among the measured data. The authors suggest that neglecting correlations among measurements is a rough approximation, as the correlations themselves have an impact on the estimated trend. Therefore, correlations among measurements should be properly taken into account by representing the errors of the measurements with their full covariance matrix. The authors also suggest that periodic variations of the state variable under study can be jointly fitted with the trend in a straight forward manner, even if the shape of the periodic function is not known.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

General

The subject of this Technical Note (TN) is very important for the atmospheric science community and is certainly worth of publication in an ACP Technical Note. Correlations among measurements are too often disregarded, unnecessarily growing the error budgets of the derived parameters / trends.

Despite the good scientific significance of the subject treated, I have, however, serious concerns regarding the scientific quality of the approach used. In particular, in the TN the discussion is always very theoretical and there is no occasion in which the presented theory is applied to an even small test dataset. If the authors plan to postpone the presentation of examples of application of their theory to a second forthcoming paper I do not agree with them for several reasons.

- I have found some (too many !) errors / imperfections in the presented equations, and I cannot guarantee that I found "all" of them. I have seen that in some occasions the main author of this paper delegated the reviewers to do an extensive debugging of his work. I do not agree with this approach. Instead, together with the theory, the authors should present a set of convincing examples in which their theory is successfully applied to test data. On their own, numerical examples are a kind of validation, increasing the trust of the reader on the presented theory.

- The presented theory could easily embed unpredictable bottlenecks or problems that show-up only when the algorithm is implemented into a computer program and applied to data (e.g. convergence, stability problems, etc., see the specific comments below).

- The authors claim that neglecting correlations has an impact on the derived trend. I agree, however, how large is the effect ? If the effect is small one could decide that the complication of using the full error covariance is not worthwhile. The

authors should quantify the roughness of the approximation of neglecting corre-
lations with an example.

To conclude, I do not recommend publication of this TN in ACP due to the poor scientific
quality of the approach used. Of course I am ready to change opinion if the authors
decide for a major revision of their paper, with inclusion of several numerical examples
addressing the above specified concerns and the specific comments outlined below.

Specific comments

1. p.27679, l.1-3: it is not clear how the vector $\mathbf{x}$ is set-up.

2. p.27679, l.15: .... =0

3. p.27679, l.16: the first "$=$" does not hold.

4. p.27680, l.12: this equation can be simplified to get the same form as the equa-
   tion for a.

5. p.27681, Eq.(13): there are two serious errors showing that this formula was
   never tested in a practical case. Please check matrix products, some of them
   are not possible due to dimension mismatch. The transpose sign is in the wrong
   place.

6. p.27681, l.13-18: This sentence is too long, I was not able to follow it.

7. p.2782: I was not able to follow the discussion here. Long sentences and lack of
   practical examples do not help the reader. If $\mathbf{u}$ is a global or multi-site field, the
   covariance matrix of Eq.(14) characterizes the spatial variability of $\mathbf{u}$ within the
   considered sample of satellite measurements. Wouldn't it be reasonable to scale

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

this matrix according to the actual distance between the two stations mentioned at l.2 of the same page ?

8. p.27683, ll.1-18: This reasoning is difficult to follow and rather speculative if not supported by a practical example.

9. p.27684, ll.4-14: I do not see the reason why these derivatives are introduced here. At the end, the final analytical solution is not given and the *lazy* reader is invited to solve numerically the minimization problem. I would either remove equations from 18 to 21 or present also the final analytical solution (seems not too difficult to get). The user could also be interested to know if secondary minima of this χ^2 are expected and if the inversion is well-posed / conditioned. This could help in the choice of the minimization algorithm. Here I got, once again, the impression that the authors themselves never tried the approach they are proposing.

10. p.27685, l.1-5: Again, this reasoning seems rather speculative if an example is not provided.

11. p.27685, Eq.(22): Is this a scalar or vector equation ? Please clarify and use consistent symbols, see also below.

12. p.27685, Eq.(23): in this equation a scalar is added to a vector ... please correct as appropriate. How do you define exactly the vector **c** of this equation ? It should be linked with $c_{month}(x)$ of Eq.(22), however the symbol used is different.

13. p.27685, l.17: Binning of what ? In which domain ? Under which circumstances binning or averaging is to be avoided ? Please provide an example. Here c became a scalar, in Eq.(23) it was a vector, please make a decision and then use consistent notations.

14. p.27685, l.24: $month = month + 12$ ... therefore I conclude $0 = 12$?? Please correct.

15. p.27685, Eq.(25): same problem of Eq.(23).

16. p.27686, ll.7,8: In my view there is no guarantee of success here. There could be multiple minima of the cost function or the problem could be ill-posed. The authors should illustrate an example of successful application of this method to real data. This would give some confidence that the proposed approach is feasible, at least in some cases.

17. p.27686, ll.12,13: is there some physical justification for smoothing the (scalar ?) function c ? It will depend on what is c. Please illustrate a practical test case.

18. p.27686, Eq.(28): same problem of Eq.(23).

19. p.27686, ll.20,21: How to choose γ ? The second order cyclic differences matrix would seem more appropriate here. Why do you suggest the first order differences matrix ?

20. p.27687, Eq.(29): same problem of Eq.(23). How would you determine $\mathbf{D}$? Have you ever tried this implementation or is just a speculation ?

21. p.27687, Eqs(30-31): These equations are correct only if the derivatives appearing therein (vectors) are defined in a very unusual way. Please provide an appropriate definition of the derivative vectors and correct the equations if necessary. Note also that the usefulness of these equations depends on the method used for the minimization of the cost function. If a stochastic method is used (e.g. to avoid secondary minima) these expressions are useless.

22. p.27687, l.18: *surprising small uncertainties*... please show the example you have in mind here.

23. p.27688, l.8: is the hypothesis of *normally distributed errors* really necessary ?

24. You state that if $\mathbf{L}$ is a cyclic first order differences smoothness constraint:

    - p.27688, l.14: the number of the degrees of freedom (of the χ^2, I guess) is the rank of the regularization matrix $\mathbf{L}^T\mathbf{L}$;
    - p.27688, l.15: $rank[\mathbf{L}^T\mathbf{L}] = n - i_u$;

    I think none of these statements is correct. Please include an analytical proof or provide numerical evidence of their validity. In alternative, if references exist where these statements are demonstrated, please cite them. I am sorry that I can not suggest what are the "correct" expressions here because Eqs.(28-29) need first to be clarified (as already mentioned above).

25. p.27688-27689, Sect.6, Application areas. Again: this section looks purely speculative if not supported by at least one pertinent example.

Interactive comment on Atmos. Chem. Phys. Discuss., 9, 27675, 2009.