Improvement of error covariance matrix computation in variational methods

Sibo Cheng
EDF R&D
LIMSI, CNRS, Univ.Paris-Sud, Université Paris-Saclay, F-91405 Orsay

Supervisor(s): Dr. Didier Lucor (LIMSI, CNRS), Dr. Jean-Philippe Argaud (EDF R&D), Dr. Bertrand IOOSS (EDF R&D) and Dr. Angélique Ponçot(EDF R&D)

Ph.D. expected duration: Dec. 2017 - Dec. 2020

Adress: 7 Boulevard Gaspard Monge, 91120 Palaiseau

Email: sibo@limsi.fr

Abstract:

The idea of variational data assimilation methods (e.g. *3D-VAR*, *4D-VAR*) consists of finding a compromise between background predictions and instrumental observations where the associated weights are provided by prior error covariance matrices. A key element is the improvement of background error covariance matrix (often denoted by B), giving the constraint of shortage of experimental information. The mis-specification of background error covariance matrix structure can be problematic in terms of reconstruction/prediction accuracy as well as output error covariance estimation. Continuous attention and effort has been given to this topic, several methods are developed in order to improve the B matrix computation. Considered as the main contributor of background error in a dynamical data assimilation chain, the model error covariance matrix computation is also carefully studied, as described in the overview [2]. These methods we find in literature are more appropriate in a successive data assimilation procedure while for our applications, we are also interested in short term prediction and statistic reconstruction.

A great effort has also been carried out to diagnose and improve the covariance matrices modelling *a posteriori*, in particular the diagnostic and iterative methods developed by Météo-France [1] (also known as Desroziers iterative method). This method adjusts sequentially the background-observation error covariance ratios based on posterior indicators. Recent efforts are also investigated to apply diagnostic methods in local sub-spaces, which could make the covariance rebuild more flexible especially for high dimensional or multivariate problems. However, we notice that the Desroziers iterative method only modifies a multiplicative coefficient of matrix B which means the assumed prior error correlation can not be corrected. The efficiency of this method could also be limited when lack of historical data due to sampling errors when evaluating posterior indicators.

Inspired by existing industrial practice, consisting in repeating several times the assimilation procedure with the same observations, we have developed two novel iterative methods: **CUTE** (Covariance Updating iTerative mEthod) and **PUB** (Partially Updating BLUE method) for building background error covariance matrices in order to improve the assimilation result under the assumption of a good knowledge of the observation error covariances.

Using a linear observation operator, we have compared *CUTE*, *PUB* with the Desroziers approach, starting by a mis-specified assumed background matrix B_{A} in a twin fluid mechanics experiment framework together. The improvement in terms of assimilation accuracy is similar for all three methods. However, experiments show that the two new methods own a significant advantage concerning output correlation recognition under the assumption that the background error is dominant over the one of the observations. We draw the evolution of reconstruction error against the number of iterations in Fig.1 [left].

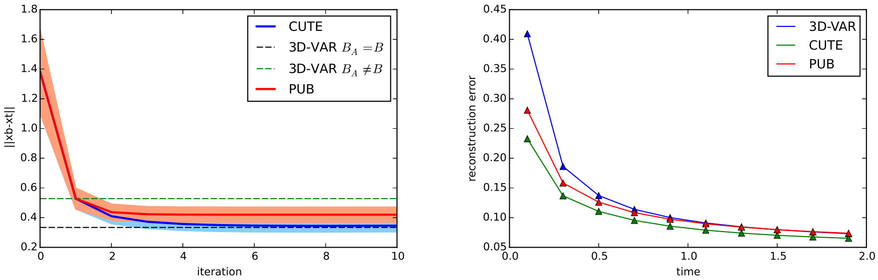

Figure 1: Comparison of standard *3D-VAR* method with *CUTE* and *PUB* in twin experiment framework in both a static reconstruction [left] and a successive data assimilation chain [right], starting with a mis-specified matrix B_{A}. For the static reconstruction, we draw the evolution of assimilation error and the standard deviation associated(transparent zones) throughout *CUTE*, *PUB* iterations, comparing to the one-shot *3D-VAR* algorithm when the background matrix is provided ($B = B_{\mathrm{A}}$, considered as the optimal target) or not ($B \neq B_{\mathrm{A}}$) . In the successive process, *CUTE* and *PUB* are only applied at the beginning (first reconstruction) of the process, following by standard *3D-VAR* algorithm latter. The initial background error is set to be 100 times higher than the observation error.

In a successive data assimilation chain, significant improvement provided by these two novel methods has also been identified compare to the flow-independent *3D-VAR* method, especially for short-term prediction. This advantage can be kept longer in the dynamical process as shown in Fig.1 [right] if the assumption of high level background/model noise is well fulfilled.

In order to get a more careful diagnostic, our effort has also been given to separate the state space into several well chosen sub-spaces where posterior diagnostic could be carried out independently. Instead of spatial distance based segmentation (as in [3]), we make the space separation by an observation based connection network. Unsupervised graph based community detection algorithms are therefore considered helpful for state space separation. For instance, reasonable positive results have been found in twin experiments by applying Desroziers iterative method in spatial distance independent sub-spaces using artificially simulated transformation operators.

Future focus will be given on the performance of our newly developed methods (*CUTE*, *PUB* as well as observation based connection networks) in more realistic/sophisticated industrial models.

References

[1] G. Desroziers and S. Ivanov. Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, 127:1433 – 1452, 04 2001.

[2] P. Tandeo, P. Ailliot, M. Bocquet, A. Carrassi, T. Miyoshi, M. Pulido, and Y. Zhen. Joint estimation of model and observation error covariance matrices in data assimilation: a review. *Arxiv, accepted for submission to Monthly Weather Review*, 07 2018.

[3] J. A. Waller, S. L. Dance, and N. K. Nichols. On diagnosing observation-error statistics with local ensemble data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143(708):2677–2686, 2017.

Short biography – Graduated from a master in applied mathematics, I started a PhD in 2017 at UPSud under a CIFRE convention with EDF R&D, where I completed an internship previously. The present subject is motivated by industrial problems in data assimilation reconstruction confronted by EDF.