Ocean Sci. Discuss., 6, C69–C74, 2009
www.ocean-sci-discuss.net/6/C69/2009/

Ocean Science
Discussions

Interactive comment on "Application of the Gaussian anamorphosis to assimilation in a 3-D coupled physical-ecosystem model of the North Atlantic with the EnKF: a twin experiment" *by* E. Simon and L. Bertino

M. Bocquet (Referee)

bocquet@cerea.enpc.fr

Received and published: 29 April 2009

In this paper, the authors make use of non-Gaussian techniques in conjunction with an ensemble Kalman filter (EnKF) applied on a biological model of phytoplankton coupled to an ocean model. Specifically they implement a Gaussian anamorphosis of their biological variables before the Gaussian analysis of EnKF, and the reverse transformation after the analysis. They concentrate on synthetic experiments, but the context and the models are realistic and complex. In the context of these twin experiments, the anamorphosis improves the tracking of the biological system very significantly.

1 General comments

To my knowledge, this is the first time Gaussian anamorphosis is applied to the data assimilation analysis step in the context of a large scale biogeochemical and/or ocean model. This is why this paper should be considered as useful and important.

They are quite a few typos in the text, and I list some of them below. The English is not perfect, and needs to be corrected. I suggest that the authors have their paper read by a native English speaker (two French authors and at least one French reviewer might be misleading). The sentences are sometimes too long. Also, there is an excessive use of logical articulation words.

Yet the exposition of the ideas are very clear, direct and easy to read, even though the methodology is not simple.

I have a few comments/suggestions/questions on the paper:

1. What makes the anamorphosis so successful ? Is it essentially because it constraints the variables to be positive, or is it more ? In atmospheric pollutant dispersion, one can show that the positiveness constraint is the major contribution to the non-Gaussian nature of the analysis. If so, would a truncated Kalman filter (Lauvernet et al., 2009, cited in the manuscript) just as useful as the anamorphosis ?

2. The localisation of the anamorphosis in time and space seems an important source of improvement. The idea put forward in the conclusion could be elaborated slightly more.

3. The observations are perturbed by a log-normal law, which presumably (?) lend itself well to anamorphosis. How do you expect the anamorphosis outcome to degrade with real errors, possibly multimodal ?

4. A short appendix on the mathematical details of the anamorphosis would be helpful to most readers (see remarks below).

I believe that, provided the questions are adequately answered and the minor corrections taken into account, this paper should be published in Ocean Science, and could become an important contribution to the issue of non-Gaussian statistical modelling in geophysical data assimilation.

Reviewer: Marc Bocquet

2 Minor questions and comments

- p.619, l.29: "stochastic" → "probabilist"
- p.621, l.2: Reference for EnKF: Evensen could have been cited earlier at the first occurrence of EnKF
- p.621, l.9-l20: Other methods have been tested in geophysics in the context of data assimilation. For instance Hólm ("Assimilation and modelling of the hydrological cycle: ECMWF' status and plans", E. Hólm et al., ECMWF Technical Memorandum 383, September 2002, see specifically p.32-36) used a Gaussian anamorphosis for humidity in meteorological models. Bocquet uses a fully non-Gaussian data assimilation framework in air pollution modelling (e.g. "Inverse modelling of atmospheric tracers: Non-Gaussian methods and second-order sensitivity analysis" M. Bocquet, Nonlin. Processes Geophys., 15, 127-143, 2008.)
- p.621, l.26: "two last" → "last two"
- p.621, l8-9: The paragraph break should be avoided here.

- p.619, l.13: "An other" → "Another"
- p.619, l.17: "dependent of the" → "dependent on the"
- p.619, l.8: "in comparison of" → "in comparison to"
- p.618, l.12: What is a Gaussian space ? This beginning of an explanation is only given on p.622, l.11 and further on.
- p.618, l.7: "physical/biological limitations": the ellipsis is awkward.
- There is a lack of general references in the beginning of the introduction, before the data assimilation concepts are presented.
- p.622, l.21: "is based of" → "is based on"
- p.622, l.25: $\mathbb{N}_N$ is not a standard definition for $1, 2, .., N$. Please define it.
- p.623, l.18: "By assuming that H is linear" is surprising but later discussed. Maybe you should tell the reader that a discussion follows.
- p.623, l.22: "a Normal law" → "a normal law"
- p.623, l.23: "The return" → "The pull-back" (dedicated mathematical terminology)
- p.624, l.3, p.626. l.1, and p.630, l.20: "Remarks": please make sentences.
- p.624, l.13: "In the case when H is extracting measurements from the state vector, this is not an issue." Why so ?
- p.624, l.3: It is certain that marginal distributions will not capture the full non-Gaussianity, especially from cross-correlations. I have personally checked that on a very simple but highly non-linear (and therefore non-Gaussian, as far as pdf are concerned) Lorenz-63 model. But dealing with the problem with marginals is

perfectly acceptable for high-dimensional problem, and can hardly been avoided. A way out could be independent component analysis but could be numerically too expensive.

- p.625, l.11-15: It would be good to write some of the details in a short appendix. Especially you could tell the readers that do not know anamorphosis what it implies in terms of mathematics.
- p.625. l28: "summarized Fig.1" → "summarized in Fig.1"
- p.627, l.7: "with a minimum thickness of the top layer of 3 m" → "with a minimum thickness of 3m at the top layer"
- p.627, l.21: Drop the "So", because the reader cannot infer that the exact number of variables is 7 from what you have just explained
- p.627, l.22: "is illustrated Fig.2" → "is illustrated in Fig.2"
- p.628, l.1: Drop the "it means"
- p.629, l.9-14: What is the bias reduction for exactly ? I suppose it is to debias the observation error (E$[y_n - y_n^t] = 0$) ? But would it be the same in *Gaussian space* ? Without the correction term, you would debias the median, so that there would be no bias in Gaussian space (?). The objective is in any case, to match the BLUE requirements for the EnKF analysis.
- p.629, l.15-24: Couldn't you write in one sentence, that you generate the biological fields with this dynamical ensemble ?
- p.629, l.25: Mitchell and Houtekammer were the first to introduce localisation in EnKF. You could cite them too.

- p.631, l.13: "histogram of the transformed" is making reference to the anamorphosis procedure, which could be confusing for a reader unaware of its details. That is why again a short appendix on the maths of the method would be useful.
- p.631, l.14: "gaussians" → "Gaussians"
- p.632, l.3, Eq.(9): It seems that you do not make use of STD later in the text. If so, do not introduce it.
- p.634, l.8: "as said previously" → "as stated previously"
- p.634, l.13: "and the data assimilation" → "and data assimilation"
- p.634, l.14: "can not" → "cannot"
- p.634, l.15: "the solution issued from" → "the solution obtained from"

Interactive comment on Ocean Sci. Discuss., 6, 617, 2009.