Biogeosciences Discuss., 9, C4701–C4706, 2012
www.biogeosciences-discuss.net/9/C4701/2012/


Biogeosciences Discussions

# *Interactive comment on* "Reducing the model-data misfit in a marine ecosystem model using periodic parameters and Linear Quadratic Optimal Control" *by* M. El Jarbi et al.

**M. El Jarbi et al.**

mej@informatik.uni-kiel.de

Received and published: 14 October 2012

[11pt,a4paper]article amsfonts amsmath graphicx wrapfig graphicx amsmath graphicx epstopdf pdfpages tabularx

algorithmic longtable lscape amsthm caption epstopdf multirow multicol url tabularx font=small,labelfont=bf,singlelinecheck=false float subfig color ruled plain

equationsection

Comments to referee $3$: We thank you for your comments. We give answers here, referring to your numbering.



Answers to general comments:

(We will repeat the comments here, followed by our answers thus marked by *Answer:*)

$(1)$. There is a strong prior knowledge imposed on this problem, which plays a large part in explaining the success of the approach. The cost function, eqn (12), measures the deviations of the state and parameters from a reference values, and minimizes those by adjusting parameters in time. The reference used for the state are the observations, and so this part measures weighted squared discrepancy between observations and model predictions. The parameter part measures the weighted squared discrepancy between the parameters and the baseline values in Table 1. Parameters are estimated to minimize these (subject to satisfying the linearized model and satisfying the periodicity "constraint" on parameters). This setup is a powerful mechanism to preventing parameter drift since you effectively impose the annual mean value for each parameter, and estimate only seasonal anomalies. This also gets around the well-known parameter identifiability problem for ecological models by fixing the parameters in relatively narrow ranges, so parameters don't start trading off against one another to improve fit. The paper should be very clear and up front about this aspect – it is what makes the approach work and why the results look so good. It seems like, in practice, the procedure requires good static parameters to start, and then it can get deviations (seasonal cycles) that better fit the data. These things should be made clear.

*Answer*: We will discuss this point and give some possible reasons in the final version. Possible reasons for the good quality of the fit using this method here are

- the use of observational data for the reference trajectory in the linearization
- the application of the method separately on each time interval between two succeeding observation points (in time)

- the usage of a very short temporal discretization on each of these intervals
- the usage of exact derivative information for the system matrices in linearization (using Automatic Differentiation, a point we did not mention in the first version).

Still the result in this application is impressive, and might not be obtained in other examples.

2. The linear optimal control method is not a state-of-the-art methodology. Such approaches were experimented with quite extensively in the early days of data assimilation. For example, the extended Kalman filter relies on such an assumption. Linearization approaches have been superseded by variational and ensemble methods that incorporate fully the nonlinearity in the forward model (the main use of linearization now – or the so-called tangent linear model – is to produce a gradient to aid in minimizing the cost function). The linear optimal control approach does works, but it is awkward to implement and I doubt many data assimilation practitioners would use it. There is nothing wrong with it, however, and I think it is suitable for illustrating the value of time dependent parameters, and flexible seasonal cycles for. I can see how the essence of the approach could be implemented using more modern data assimilation methods, and this is what I find interesting. Maybe some discussion on this is warranted.

*Answer*: This is a good point, we will include some discussion in the final version.

3. The sensitivity of the results to changes in weighting "R" is interesting. This measures the extent to which parameters are allowed to vary, traded off against the fit of the model to the data (since Q is being fixed). This is, of course, is yet another parameter that can be tuned to help better fit the data. In Figure 5, it is shown that it does not make much difference in the state estimates. But it yields quite different values for the parameter evolution in subsequent Figures– their cycles are often similar in shape but their magnitudes and offsets are different. I think, ultimately, you want



a more quantitative criteria for estimating R. There are two ways I have approached this in similar problems: through cross-validation, and by estimating the ratio of R to Q explicitly. Maybe mention?

*Answer*: This is a interesting point, we will try the two ways to mention where possible.

4. Is there an issue of over-fitting here? When I see model predictions (for a model with a lot of parameters) fitting data exactly (as in Figure 2), over-fitting is the first thing that would come to mind for any Statistician. You are estimating parameter anomalies at every time step. The matrix Ricatti equations can be cast as a sequence of linear regressions, where here the number of parameters likely exceeds number of observations. But at the same time things are tied together by the dependence structure in time so reducing the effective degrees of freedom. There is a lot of flexibility in the time varying parameters, and in general you are trading temporal smoothness off against fit via R. This is more of a discussion point, as an obvious issue when things (state estimates) look too good to be true.

*Answer*: We will follow this discussion point in the final version.

5. Section 4.1 A better metric to assess the effectiveness of the seasonally varying parameters might be predictive skill, rather than just fit of the model state to the observations. The flexibility in the parameter values guarantees a good fit, but a more true test is: can you predict future states, or even retrospectively fill in withheld observations (i.e. cross-validation).

*Answer*: We will include such kind of example in the final version.

$6$. Section 2.1. I am a little unclear on the model – it states that is uses output from a basin scale model of the North Atlantic but then states it simulates the water column at one place (the BATS site, I presume). I think it is just a 1D location specific model. This should be clarified.

*Answer*: It is one-dimensional marine ecosystem model that simulates one water column at a given horizontal position. We will make this clearer in the final version.

$7$. Theorem 1. Note sure if you should state this as a theorem here. It is a well-established result and should likely be presented as an clear step by step algorithm, to aid in implementation by interested readers.

*Answer*: We will it present as an clear step by step algorithm.

$8$. Figure 3: How did the optimized static parameters compare with the Table 1 values? I presume the fit is better, since you likely used Table 1 parameters as starting values for the optimizer.

*Answer*: We use the table 1 values during the whole year. In the first year the choice of the cost function will indeed somehow force "periodicity" to the constant reference parameters. This effect can be reduced by choosing appropriate small values in the matrices $R_k$ in the first year. In the following years the difference of the current parameter (vector) $u_k$ to its counterpart from the year before is minimized. This in forces periodicity. The crucial point in adjusting the matrices $R_k$ throughout an optimization run is to both allow

- for sufficiently large deviation from the constant reference parameters in the first



year to enable their temporal variation

- and for smaller deviation and thus more or less strict periodicity in the following years.

A figure with the optimized parameters compared with the table 1 parameters will be added.

$9$. Figure 4: The model-data misfit looks strange and blocky. Why is this?

*Answer*: Will be changed using a more appropriate colormap.