Biogeosciences Discuss., 9, C4005–C4010, 2012
www.biogeosciences-discuss.net/9/C4005/2012/

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.

Anonymous Referee #2

Received and published: 14 September 2012

In this study, a one-dimension NPZD style ecosystem model is linearized in order to optimize 12 of its parameters. These parameters are assumed to be time-dependent and a term in the cost function that is used in the otimizations allows the authors to introduce annual periodicity into parameter values. The model with periodic parameters shows a much improved fit compared to the standard model with optimal constant parameters.

Despite the extra effort that is required to obtain time-varying, periodic parameters, the authors can clearly show their advantages. Time-varying parameters for marine ecosystem models have only been used in very few studies and the technique and the

results used here are certainly worthy of publication. However, in my opinion some major revision are necessary before this paper can be published.

general comments:

(1) The mathematical notation is lacking, often inconsistent and hard to follow. Big parts of Section 2 and 3 are especially hard to understand as some of the symbols are not properly introduced and others change from one subsection to the next. Refer to many of my minor comments.

(2) Nearly all of the results that are shown in this paper are taken from the surface of the model. According to Fig.4, the model fit is much better in the surface than below. It is important to quantify results in other layers as well. A table with the cost function values in each layer for the two optimization approaches would be useful. A result plot for one of the lower layers would also be a valuable addition. The less ideal fit might require some additional discussion.

(3) One of the stated goals of this study is "allowing for straightforward application to time periods outside the range of observations". Yet the authors do not follow up on this and use the same BATS data set for optimization and validation of their results. An application of the model with time-varying parameters at a time span with observations not used for optimization would really show the advantages of this technique.

(4) Marine ecosystem models are known to be highly nonlinear, so that an attempted linearization might not always be successful. Why does it work so well here? A brief discussion or at least a mention of the issues associated with linearization should be made.

specific comments:

abstract: Include "one-dimensional" in the model description. Mention upfront that time-varying parameters are introduced and also mention how many of the parameters are optimized.

p 10208, l 14: Please divide this large paragraph up, it is hard to read.

p 10209, l 5: Again, add "one-dimensional", e.g. "use one-dimensional models which simulate a single water column only"

p 10209, l 11: Define "NPZD" and other abbreviations right away.

p 10209, l 17: "This reflects the aim to obtain a model that is applicable for arbitrary time intervals." There are other practical reasons why parameters are assumed to be constant, it is easier to obtain them and the risk of overfitting the data is much lower. These points could be mentioned here.

p 10210, l 11: This sounds a bit misleading, the goal of "improved model parametrizations that can model these variations even with constant model parameters" is not achieved in this study, please rephrase.

p 10210, l 13: What is "optimal control", the term control is only defined later.

p 10210, l 15: "This avoids the process of parametrization in the sense that we do not have to know or assume how the above mentioned periodic functions look like." This sentence is hard to understand.

p 10210, l 17: Point out the two aims more explicitly in the text above.

p 10211, l 5: Break up this long sentence.

p 10211, l 25: Rephrase, \vec{u} is a vector containing the biological model parameters.

p 10212, l 2: What is q_n?

p 10212, eq (2): So x^1 = P. It would be more consistent to stick to one notation, e.g. x_P. The x_l notation could still be used by writing l \in {N,P,Z,D}. It would also avoid confusion with z (depth) and Z (zooplankton).

p 10212, eq (2): The sinking term for detritus is included here again.

p 10213, l 9: Now ZOO is introduced in addition to Z and x^3, please stick to one notation.

p 10213, l 11: Make it more obvious that x^3 (= x_Z) is dependent on z (depth).

p 10213, l 17: Why is J dependent on \mu and u which were not defined previously?

p 10214, l 4: "Note that these numbers may be quite different for the different years". Why are the F_{mj} in eq 4 then normalized by N_{mj}? This approach appears to emphasize observations in those years with little data, is there a reason for it?

p 10214, l 9: What motivates this choice measurement uncertainties?

p 10215: The use of "controls" and "control variables" is confusing. Maybe mention once that the control variables are the parameters, then just stick to the term parameters.

p 10216, l 9: Is is it a "time-step of length \tau"?

p 10216, eq 6: Is it an l or an I and what does it mean? Do the n signify time coordinates? Please introduce each new symbol. Without this information, the rest of the description remains guesswork.

p 10216, l 23: Mention the "4x4" and other properties right away in line 11.

p 10217, l 15: Now commas are used to separate the indices in the subscript, please use this throughout the paper.

p 10217, l 16: The first 3 bullet points are not consistent: if there are N_j observations, the steps cannot go from 0 to N_j + 1.

p 10217, l 21: Why is x suddenly used for observational data, or what is x^{ref}?

p 10217, eq 9: Is n suddenly defined as a real number, or is I_{j,i} (the interval) really {n_{j,i}, n_{j,i}+1, ... ,n_{j,i+1}}.

p 10217, eq 9: z_{n} is a poor choice notation-wise, as z was depth previously and Z stands for zooplankton.

p 10217: The notation is confusing: there is n which runs to M but there is also an $n_{j,i}$ while i runs to N_j (previously there was an N_{mj}). I like the fact that j runs to j_{max}, maybe this notation can be extended.

p 10219, l 4: How does the constraint enforce periodicity, when u_n^{ref} is not periodic (according to eq 11).

p 10219, l 10: The t_n notation does not seem to add anything, please drop it.

p 10219, l 13: Why the new notation? The state was x, the parameters u. What happened to the relationship from p 10218 l 5? Stick to one notation.

p 10220, l 1: Are the parameter values in u or in v now, where does the u come from?

p 10221, l 2: "Bb_n"

p 10222: Are the parameters periodic in Section 4.1 already? Mention this explicitly. It is not quite clear if the results from Fig.1 correspond to the parameters in Fig. 6-8.

Fig. 2,3: For a better comparison it would make sense to join these 2 figures together and have the information from the LQOC and SQP in one plot.

Fig. 4: The plots are dominated by a few single pixels with high costs, the rest consists of shades of blue. An adaptation of the colormap would make this figure more informative. Also, if there are pixels without observations, they should be in a color which does not appear in the colormap (e.g. white) to make it distinguishable from a perfect fit.

Fig. 6,7,8: These plots can probably be summarized in one, maybe two plots by showing only 2 years and mentioning the almost perfect periodicity. It would also be beneficial to add the optimized constant (SQP) parameter values and the parameter bounds from Table 1 (as straight lines) into the plot as well for comparison purposes.

Fig. 9: There appear to be 2 legends for each plot, summarize this information in one

legend.

Interactive comment on Biogeosciences Discuss., 9, 10207, 2012.