Ocean Sci. Discuss., 3, S415–S423, 2006
www.ocean-sci-discuss.net/3/S415/2006/

Ocean Science
Discussions

Interactive comment on “Reduced-order optimal interpolation for biomass assimilation” *by* G. Crispi et al.

G. Crispi et al.

Received and published: 8 September 2006

Authors’ comments to Referee #2 for the manuscript “Reduced-order optimal interpolation for biomass data assimilation” by Crispi, Pacciaroni and Viezzoli for the Ocean Science, MFSTEP Special Issue. September 2006. The new title is: “Simulating biomass assimilation in a Mediterranean ecosystem model using SOFA: setup and identical twin experiments”.

General comments

The work is revised along three main routes. The links between the Observing System Simulation Experiment in terms of identical twin experiments and the general objectives of the work are made more precise. The methods and the parameter in Table 2 of SOFA are introduced. The root mean squared differences in the new Fig. 8 are used for investigating the convergence of the free and assimilated run in terms of the phytoplankton winter bloom. The free run does not converge in rms to the control (previous version reference) values of the phytoplankton, Fig. 8a, but only in its basin average, Fig 6a (previous Fig. 7a). Assimilated run has lower rms values, about half the initial rms originated by the phytoplankton summer initial conditions. Inorganic nitrogen reaches significant rms (Fig. 8c) not corrected by the assimilation method. We have cut Fig. 6 and considered strong and weak point of this methodology in the conclusions. Our comments and adds are introduced beginning with -C- , after each specific comment and technical correction by Referee #2. We quote the revised text between asterisks.

Specific comments

1) The link between the general objective of the paper and the experiment that is performed is not explained. Why do the author apply a perturbation on the phytoplankton initial condition? Is it the dominant source of error in the system? Why controlling initial condition errors and not modelling or forcing errors? Why applying a perturbation on the phytoplankton and not on the other state variables? To what extent will the conclusions hold in realistic experiments, in the presence of other sources of error? Why is it an appropriate preliminary step? One can doubt that the answers to these questions will always support the choice of the experimental protocol. It is thus necessary that the authors explain their choice and state to what partial conclusions that kind of experiment can lead (and also on what matters no conclusions can be drawn). Without such explanations, the experimental protocol appears oversimplistic to investigate the general objectives stated in the introduction, and the conclusions of the paper remain almost meaningless.

-C- Shifting the phytoplankton initial conditions from winter to summer values determines a reduction of the energy of the ecological variables as a system. Moreover reducing the energy to the biochemical cycle changes equilibria of the ecosystem. With this exercise we want to study the times and paths of the recovery of the initial states

OSD

3, S415–S423, 2006

under assimilation in this more oligotrophic ecosystem after the low phytoplankton conditions. Some recovery happens for each variable, while for the integral quantity, i. e. total nitrogen, which is significantly modified at the beginning of the run some problems appear. Both phytoplankton and inorganic nitrogen are the most important error sources in the system (Fig. 8 a; c). At the beginning phytoplankton is dominant, after two weeks N becomes important and greater (nearly double in this ITE) than P. This is for us a simple way of doing a preliminary work and significant numerical simulation of the biomass biogeochemical assimilation at the condition of fixing ecological parameters and forcings. Moreover the variable to be assimilated is the biomass and to modify the initial conditions of this variable gives advantages and limits of using this univariate methodology. In our opinion we cannot extend these results in presence of other sources of error. Their influence may require other methods with higher consuming times because a clear integral quantity cannot be at hand.

2) Without a minimal description of the assimilation scheme, it is impossible for the reader to understand, to interpret, and even less to reproduce the results. Maybe adding a few lines (in addition to the reference that is given) would be sufficient to remind the general features of the assimilation method. But the statistical parameterization is specific to this work and must be fully documented in the paper. How is the observation error covariance parameterized? How is the background error covariance parameterized? I assume that the statement that the scheme is univariate means that surface biomass data are used to update only the phytoplancton variable, but over the whole water column. In order to do that you need to provide vertical error covariance for the phytoplankton? How is it parameterized? What is the phytoplankton background error horizontal correlation structure? On what grounds did you set the values of these parameters? In order to answer these questions, a new section describing the assimilation scheme is needed.

-C- We have introduced these information in the methods paragraph. The new text is: *This reduced-order optimal interpolation system operates under the condition that the

OSD

3, S415–S423, 2006

background analysis error matrix is calculated from the background error variances, calculated from the previous analysis error variances, and from the correlations, estimated on the observational data and taken fixed during the simulation. Thus SOFA evaluates directly the horizontal correlations at observation locations, assuming them vertically uncorrelated. The observational error covariance matrix is diagonal and the parameters for these identical twin experiments are given in Table 2.*

3) Only relative errors of the assimilation run with respect to the free run are presented. I think it would be important for the reader to have an idea of the amplitude of the error for each variable. It would also help a lot the interpretation of the results. For instance, since the perturbation is only on the initial condition, it may be that, for some variables, the system is relaxing by itself close to the true trajectory. (Fig 7a shows that such relaxation to the truth without assimilation occurs at least for the basin average of the phytoplankton concentration.) In such a situation, it may be not very useful to know that the relative errir is asymptotically becoming 30% lower in the assimilation run, if the amplitude of the error in the free run is already becoming very small. Generally speaking, I believe that the presentation and interpretation of the results should be improved and clarified.

-C- There are two points for the effectiveness of all the data network adopted here: firstly the amplitude of the phytoplankton along all the simulation run; secondly the weight of the phytoplankton error with respect to the other variable, also in the specific comment 1). For giving an insight in the results of this application we plotted in a new figure, Fig. 8, the root mean squared differences of the four variables, in sequence P, Z, N, D. In these four plots we present the root mean squared differences of the free run (squares) and assimilation run (triangles) with respect to the control run (previous reference run). The calculations are performed at a basin scale. The phytoplankton rms of the free-control (previously free-reference) starts from an initial value, due to the decrease of the initial conditions of the phytoplankton, low summer values for the free run at the place of high winter values. The phytoplankton evolution gives some vari-

OSD

3, S415–S423, 2006

ability of this distance between the two runs, but very small values are never obtained. The relaxation to the truth is not attained. N instead starts from null values because initial conditions are not changed. These values become higher during the simulation, reaching at the end a nearly flat value about double than that of P. It means that the errors of inorganic nitrogen is after two weeks higher than the P one. This implies that the phytoplankton is an important but not the dominant source of error in the system for this twin experiment. There is also an important part originated from N. Both other variables reach rms lower than P and N variables. The relative errors plotted in Fig. 9 and the following ones are for all the variables the ratio of root mean squared differences. The Fig. 8 has in the new manuscript overall content companion figures of the Figs. 6 a,b (P and Z in the previous text Fig. 7) and 7 a,b (N and D previously Fig.8). Previous Fig 6 is cut.

Technical corrections

1) The title is not appropriate, because it leads the reader to think that there are new developments in the assimilation scheme.

-C- The title now reflects more clearly the work done in MFSTEP: *Simulating biomass assimilation in a Mediterranean ecosystem model using SOFA: setup and identical twin experiments*

2) The abstract should be clarified and made more informative of the content of the paper.

-C- The abstract is modified to introduce general aim and methodologies in the following way: *Assessing the potential improvement of basin scale ecosystem forecasting for the Mediterranean Sea requires biochemical data assimilation techniques. To this aim, a feasibility study of surface biomass assimilation is performed following an identical twin experiment approach. NPZD ecosystem data generator, embedded in one eighth degree general circulation model, is integrated with the reduced-order optimal interpolation System for Ocean Forecasting and Analysis.*

OSD

3, S415–S423, 2006

3) The introduction discusses of problems that are not useful to understand the paper, but fails to introduce the true subject of the paper, which is the setup and the interpretation of twin assimilation experiments.

-C- The introduction now considers some general points about OSSE and ITE, also point 4). New references about the methodology are introduced in the revised version.

4) I think that speaking of OSSE (Observing System Simulation Experiment) is not really appropriate here because one single observation system is being tested. Saying that “the OSSE proposed here provides the quantitative basis for a rational design of subsurface observing systems” is incorrect since only surface data (or near surface data) are assimilated.

-C- Now we introduce OSSEs in terms of the possible improvements of forecasting and/or impact of some fixed network. In some cases when more data are present, the best strategy can be selected; otherwise when data are more scarce the impact can be studied in a controlled way, optimising the parameters of the simulated system or the statistical analysis. Here we consider the second case when ecosystems are controlled, because satellite images can be processed obtaining biomass informations, but the other ecosystem variables (nutrients, zooplankton, etc.) are costly to be measured. The introduction is revised in the following way: *The framework relies on Observing System Simulation Experiment (OSSE). OSSEs are used for two distinct objectives: to define the best observational network for improving forecasting; to determine the impact of an hypothetical dataset upon a simulated system. The OSSE here proposed provides, on the second route, a preliminary quantitative basis for assessing the impact of a synthetical data network of surface biomass on a Mediterranean simulated ecosystem. The methodology of identical twin experiments (ITE) is chosen for understanding the surface biomass data impact on the ecosystem, taking into account the knowledge of the biomass coverage all over the basin.* The statement “the OSSE proposed here provides the quantitative basis for a rational design of subsurface observing systems” is cut because misleading. We meant that, from the setup point of view, the methodol-

OSD

3, S415–S423, 2006

ogy of SOFA is able to manage profiles of biomass or other variables with their vertical statistical informations.

5) The name of the ocean GCM is not given.

-C- At the beginning of the second paragraph the definition of the 3D hydrodynamical model is completed. *The dynamic of the Mediterranean oligotrophic ecosystem is studied through coupling with the Mediterranean basin circulation as simulated by a General Circulation Model (GCM) driven by high frequency forcing. Consistently, an ecosystem description considering the general 3D circulation has been set up and embedded in the implementation of the GFDL-MOM implemented in the Mediterranean Forecasting System Project (Demirov and Pinardi, 2002).*

6) Use one letter, not two, to describe a mathematical quantity (BT for biological tracer is not appropriate), because it can be confused with multiplication.

-C- The biological tracer is now expressed in the formula as B.

7) The paper provides technical details that are not useful here (computer trademark, number of nodes, memory size, CPU power, names of computation libraries, introduction of O2 optimization for the compilation).

-C- The technical details are cut in the revised text. *The version 3.0 of System for Ocean Forecast and Analysis (SOFA) by De Mey and Benkiran (2002) is used for assimilating surface biomass data. SOFA, using temperature and salinity as tracers (Raicich and Rampazzo, 2003), has been integrated with the NPZD-based ecosystem model. After optimization on SP4, the integrated system execution time is reduced approximately by three from about 600 s to 227 s for one day simulation with time-step of 900 s.*

8) Very often, new concepts or acronyms are mentionned before being defined or introduced. This gives the feeling that the paper has been built from pieces of text coming from elsewhere, that were put together without sufficient checking. For instance, what

OSD

3, S415–S423, 2006

Interactive Comment

are areas DS4, DJ7, DH3 that are mentionned in the second paragraph of section 3. They are not defined, and not used anywhere else. In the same section, the concept of “preconditionning period” is used before defined. Since the description of this “preconditionning period” is the purpose of section 3, it is only at the end of the reading of this section that the reader knows what it is about. I insist that these are only a few examples. The text should be carefully checked and better organized.

-C- The definitions are now as follows. *Mean nitrate summer conditions are extracted from the MEDAR climatology (Manca et al., 2004). Southern Balearic (DS4), northern Ionian (DJ7) and Cretan Passage (DH3) areas are selected and averaged, in correspondence to the stations in which phytoplankton data were acquired. The interpolation at the levels of the model is shown in Figure 2, large panel (diamonds). This profile initializes nitrate variable for all the Mediterranean basin.* For what regards “preconditioning period” it is substituted by the clearer “dynamical adjustment stage”.

9) In the result section there is a discussion on technical aspects of data preprocessing, which are not useful. It is even misleading because it mentions “profile data” several times, giving the impression that you are assimilating profile data (even if it is stated a few lines above that that biomass data characterizing the 2 upper levels are assimilated). I understand that it is only the technical way by which the data are processed but this makes the text very unclear. That discussion must be dropped.

-C- This discussion is cut because main information are given at the end of Par. 3.

10) I do not understand the sentence: “AR zooplankton behaviour is very close to the free one, even if only phytoplankton biomass is assimilated in our experiment.”

-C- The new text is:*AR zooplankton behaviour is very close to the control one and it begins to grow after the fourth assimilation cycle. This is due to higher grazing, determined by the increased phytoplankton present after assimilation. At the end of the simulation both experiments remain well beneath CR.*

11) There is a discussion beginning at the end of page 515, continuing on page 516, which is difficult to relate to the results that are presented. Please explain the relation or drop the discussion.

-C- The results about the fluxes are cut, because it is difficult to relate them to the identical twin experiment interpretation. The estimates are introduced in the MFSTEP Final Report.

12) The conclusions should certainly be modified according to the first of the specific comments above.

-C- The potential recovery of the true state of the system, using only "satellite" information under complete average information at the end of each week, is rewritten in this way: *The results of this preliminary feasibility study show advantages and limits of the chosen univariate methodology. The advantages are: efficiency of the reduced-order optimal interpolation in reducing the relative error of the phytoplankton, which is assimilated once at the end of each week; ability of the ecosystem model in spreading the assimilated information from phytoplankton variable to the other biochemical ones.* Instead the role of inorganic nitrogen and total N, system biochemical energy, are critical. The loss of more than 0.75 teragrams N which is half of the maximum energy trapped in the phytoplankton compartment during winter bloom in our simulation is hardly recovered. Another important point is the role of inorganic nitrogen errors versus phytoplankton one, assimilated variable. *The limits of the methodology consist in: inorganic nitrogen errors reduced mainly at surface by the surface biomass assimilation and poorly at an overall scale; slight corrections of the basin total nitrogen; anomalies, i. e. nonperformances, in the surface total nitrogen.*

Interactive comment on Ocean Sci. Discuss., 3, 503, 2006.