Geosci. Model Dev. Discuss., 5, C703–C708, 2012
www.geosci-model-dev-discuss.net/5/C703/2012/

Geoscientific
Model Development
Discussions

Interactive comment on "How should sparse in situ measurements be compared to continuous model data?" *by* L. de Mora et al.

Anonymous Referee #1

Received and published: 24 September 2012

General comments

The authors introduce a methodology for the comparison of model results with sparse in situ measurements. The suggested point-to-point method seems to be a canonical method to do this but is firstly not often used and secondly as far as I can see not been systematically investigated. The authors use this method for the comparison of ICES data to their North Sea Model ERSEM/POLCOMS. The region of the data set and the model area do not coincide. Thus it is strongly recommended to do some kind of snapping of data and model results. This is normally done by restricting the model results to be compared to the region where data are available. The authors extended this methodology by a point-to-point snapping in space and time. This leads in most cases of their example to better accordance than the comparison within the

whole region. But this is not a proof that this methodology has a higher evidence. It is true for the here presented model. It is obivious that a mathematical proof of the methodology in a widely generalized way can not easiliy been done. But it could be verified for a combination of model and data where the quality of the model is well known. This can f.e. been done by taking model results and sparse data sets derived from these model results modified by a Gaussian blur and compare these. This could be done for several randomly chosen data sets and a test can be established if the suggested method leads statistically to a better coincidence. An example of such an investigation can be found at the end of Jolliff et al. 2009 where the used data set is a modified model result with controlled statistical parameters.

This might not be the ambition of the authors who might possibly show the results only for the presented model. In this case the title of the paper is misleading. It should be at least extended by the term "an example". If the paper shall focus on this special example, the text should be restructered. The subsections of section 5 should be combined so that one has a clear view on the different representations of the results (tables and figures for the considered variables).

In summary: the paper shall either focus on the example -in this case the title is not appropriate- or the paper should focus on the methodology, then a general -possibly empirical- investigation is needed.

It should be mentioned that similar methods are used even if they are not published, fe: 'Point-to-point matching' is discussed in Genkova et al.(2004) Point-to-point comparison of Satellite and Ground-based cloud properties at the ARM Souther Great Plains Central Facility and Comparison of ASPEN Modeling System Results to Monitored Concentrations at http://www.epa.gov/ttn/atw/nata/mtom_pre.html 'Snap-to-Grid' is a standard method in the ArcGIS product family.

Specific comments

Sec 2 (2314): It is described that the POLCOMS area cover a far larger area than the ICES region from which the in situ data were taken. It is not described whether in the later section of the article the POLCOMS output already has been reduced to the ICES area when talking about the entire model. If this reduction was not done before investigating the 'snap-to-grid' matching, the conclusion of this article is already obvious: 'Model results and in situ measurements always match better if the domains of both are matching better'.

Sec 4.1 (2316) line 21: Building the mean of multiple in situ measurements falling into the same grid cell leads also to a bias. This can easily be understood: Assuming the measurements for the entire region are given by the numbers from 1 to 10. The entire region may be subdivided into two grid cells. The first 3 measurements may snap into the first cell, the rest into the second. The mean of the total region becomes biased: mean(1,2,..,10)=5.5 but mean(mean(1,2,3),mean(4,5,...,10))=mean(2,7)=4.5.

Sec 4.1 (2317) line 5: It is not clear what the reason for the avaraging of the data set is? Averaging over the year or a season does the same with a data set as averaging over a geographical region. Thus, the problems the here presented method shall solve in space will now occur in the time domain.

Sec 4.2 (2317) line 19 and Sec 5: The standard error and the two tailed propability are not further considered in the text. Which information does one get from the change in these numbers?

Sec 5 (2317) line 24: It would be nice to get an overview of the tables as graphs.The different regression lines could be plotted in one graph with different colors, for the seasonal consideration four short lines could be presented (see attached sketch). Generally, the content of the tables should be discussed in more detail.

Sec 5.1 (2318) line 6: The representation of figures 1-3 includes a redundancy. It is obvious that the regression lines fit best in the regions of high data density. This is an intrinsic feature of the linear regression.

Sec 5.1 (2318) line 13: The reproduction seems to be extremly good because of the high number of data near the mean temperature. This only shows that the model does not produce a bias on temperature.

Sec 5.1 (2318) line 24: Here the same problem occurs on the temporal scale the paper focus on on geographical scale. The time average equates here to the average over a large region. Thus, in addition to the here presented point-to-point method a time-to-time method would improve the significance of the comparision.

Figures related to Sec 5.1 (2331, 2332, 2333): The information of figures 1-3 is not clear in respect to the here suggested approach for comparing model results against in situ measurements. Understanding them as two dimensional histrograms of already matched model results versus in situ measurements, they could be a base for a discussion on the quality of the model itself but not to justify the here presented method.

Sec 5.2 (2319) line 3: Here two additional averaging processes have been included: annual mean and depth average. Again the question is how these averaging processes produce a bias and lead to misleading results.

Sec 5.2 (2321) line 15: It is argued that the quality of matching may depend on small difference in the timing of the bloom and thus, in the chlorophyll data. This show again the necessity of a time-to-time method. It should mentioned that according to section 2 the seasons are defined as JFM,AMJ,JAS,OND. This results in the fact that in some coastal parts of the ICES area the spring bloom may occur in the winter period.

Figure 10 (2340): This figure gives the impression that the in situ data of chlorophyll have preferably been measured in high salinity regions over the years. The message of this figure is not clear within the context of this paper.

Sec 5.3 (2322): The here presented results of the 'target diagram' method give confirmation of what the authors already have shown in the sections before: Spatial matching of model results to in situ measurements lead to better correlation of both when it comes to comparison and validation. When utilising target diagrams the results should be discussed in more detail, e. g.: why does matching apparently increase the correlation for the physical variables (temperature and salinity) but not for the 'biological' variables (nitrate, phosphate, chlorophyll)? Is it a consequence of the matching model or a more or less singular result of this model. As mentioned this subsection should be combined with the subsections before.

Sec 6 (2323-2324) line 26 ff. - 2324 line 3: The third part of the conclusion somehow contradicts what the authors did state in the first and second part (2323 line 15 ff. and 2323 line 20 ff.). The matching of model results to in situ measurements clearly does not modifiy or improve the model.

Technical comments

Tables (2327-2330) and Figures (2334-2339) related to Sec 5.2: Table 1 and Table 2: Although mentioned earlier in the article, the two tables are missing the column for the autunm period. It is suggested to add the autumn column. Moreover, these two tables could be merged to one table or if this is not appropriate it would be better to have the more or less physical quantieties (T, Sal) in one table and the ecological in the other.

Table 3 and Table 4: Same as for table 1 and 2 (see above). Moreover, it would be helpful if the shown variables would be consistent. Fe, all variables evaluated as 'annual' or all variables evalutated for one season (summer, spring, etc.). It is difficult to estimate if here only the "best" results are shown or if the missing are "boring".

Figures 4 - 9 (2334-2339) are recommended to be adapted accordingly. The figure captions should contain a cross refrence to the columns of the tables.

Sec 5.3 (2322) line 16: It is assumed that winter nitrate instead of winter nitrogen is meant.

Interactive comment on Geosci. Model Dev. Discuss., 5, 2311, 2012.

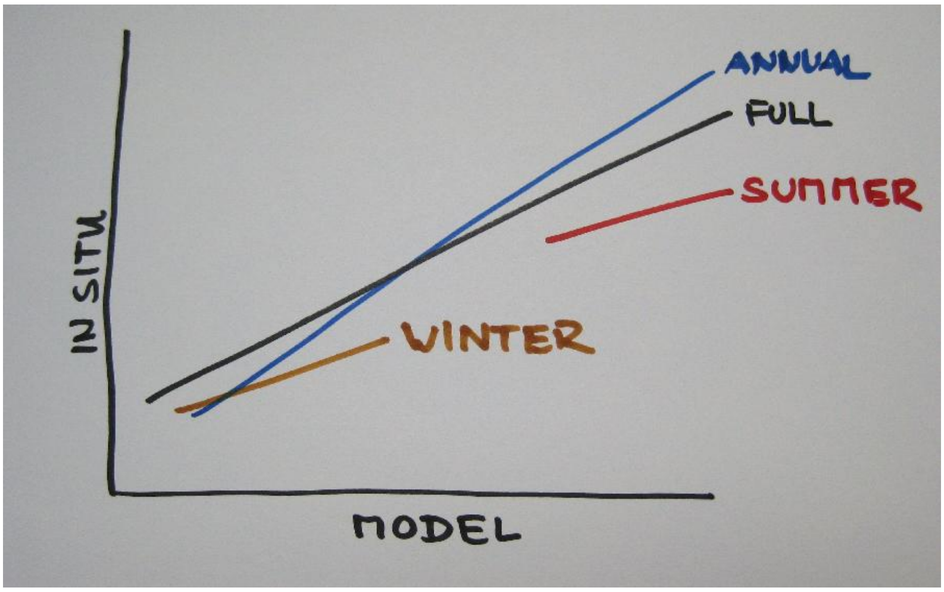

Fig. 1. Sketch Fig. 1 to 3