# Interactive comment on "A CMEMS forecasting system for the marine ecosystem of IBI European waters" by Elodie Gutknecht et al.

**Anonymous Referee #1**
Received and published: 19 February 2019

Dear Referee,

Please find our comments/responses in blue throughout your text.
We have also attached a new release of the article.
This is a major revision of the initial article: the introduction has been completely revised, the sections are modified, a discussion section has been added. A synthesis of the results has been added by using a table and a Taylor diagram.
All your comments have been taken into account in the new release of the article.
If required we can also provide a version with the word track change, but I don't think that's helpful to you.
Thank you for your useful comments.

Elodie Gutknecht on behalf of co-authors.

## General comments:

The manuscript "A CMEMS forecasting system for the marine ecosystem of IBI" by Elodie Gutknecht et al. aims to describe the skill performances of the CMEMS operational model system "IBI36" for the Iberia-Biscay-Ireland (IBI) area, with specific emphasis on the biogeochemical component. IBI36 is built on a physical-biogeochemical online coupling between NEMOv3.6 physical model and PISCESv2 biogeochemical model, at 1/36° horizontal resolution and with 50 vertical levels. The system is operational since April 2018, and consistently provides 7-days forecasts for ocean physics and biogeochemistry. The validation is applied over the "IBI Extended Domain", and is performed with a 2010-2016 simulation, using a suite of different reference data streams (from satellite and in situ, and 2 BGC-Argo floats), and then on a regional basis with comparison with in situ historical data (Northern Seas, North-East Atlantic waters, Bay of Biscay and Mediterranean Sea). IBI36 results generally consistent with the main properties of biogeochemical variables (chlorophyll, nutrients, oxygen and net primary production), in terms of spatial distribution and seasonal cycles. An indicator concerning oxygen deficiency is also proposed and assessed, which may be of great importance for the broad environmental communities, extending to the general public as well.

The manuscript is clear and well written, with a precise subdivision of the different sections (model configuration, reference data and validation, discussion). However, some major weaknesses can be highlighted:

1. The title specifically refers to the “CMEMS forecasting system”: even if it may be acknowledged the goal of the work, the validation is here performed at seasonal or annual time scale on the period 2010-2016, computing monthly averages (as defined in P8, L29), and there is no reference to the forecasting skill of the operational system, that should be done on daily (or at least, weekly) averages. So, how the described validation may be suitable for the short-term forecasting system assessment? The reader would be interested in see results on how the IBI36 short-term forecast products compare with reference data (or with the background simulation). Which metrics, of the ones proposed, can then be used operationally to quantitatively evaluate the forecasting skill?

**Author answer:**
You are right; the title explicitly refers to the "CMEMS forecasting system". I realize that the title was not judiciously chosen and does not reflect the purpose of this paper. The purpose of this paper is to evaluate the biogeochemical component of the IBI coupled model system. This system has been developed to monitor and forecast the ocean dynamics and marine ecosystems of the European waters. But prior to its operational launch, a pre-operational qualification simulation is necessary to evaluate the performance of the coupled system. Here we evaluate this qualification simulation. It represents a substantial paper, with many figures. But it's necessary and voluntary. This paper represents the first validation of the biogeochemical component of the IBI36 system; it is intended to be complete and detailed, as it will serve as a basis for further studies. So the forecasting capability of the operational system will be studied and described in another paper. To avoid any misunderstanding, we propose to modify the title of the paper by: "Modelling the marine ecosystem of IBI European waters for CMEMS operational applications". We also modified the abstract and the Introduction to make clearer the objective of the paper.

2. The model system products are validated basically following consistency (Figs. 2 to 6, 8, 9, 10ab, 11, 12, 13, 15), providing BIAS and correlation. However, the uncertainty of products (i.e., forecast, see the previous point) seems not addressed. It may be interesting to quantitatively assess the capability of the model to represent the observed variability: did the author consider to evaluate the model uncertainty, e.g. by means of RMSE? The use of quantitative comparison may be even more useful in the NPP analysis, to estimate the spatial gradient. Further, is there any reason why Fig. 7 does not report the linear regression (as done in Figs. 11 and 14)? Moreover, on a higher level, given that the validation assessment is discussed following a regional subdivision (which is basically driven by the availability of data), I think it could be useful to provide the readers with a synthesis table with the mean values of the validation metrics (i.e. BIAS, correlation, uncertainty – e.g. as RMSE) per each region.

**Author answer:**
You are right, in order to complete the skill assessment, the RMSE was added in the Chl-a comparison to satellite estimate. The NPP analysis was also revised. It now presents the mean of

the three NPP products (VGPM, Eppley-VGPM and CbPM), the standard deviation of these three products and the bias between the IBI36 system and the averaged NPP products.

The regression lines are not discussed, so they are removed in all figures.

We also added a synthesis to summarize the performances for each region. It is presented in a Table for comparison to satellite estimates and in a Taylor diagram for in-situ observations.

3. The synthetic overview of the "IBI waters" (Section 2) in terms of the biogeochemical and ecosystem dynamics is interesting since gathers different regions with different properties in a single framework. However, it is not effectively used to define the areas then adopted in the validation (P9, L19). Authors should refer the definition of the 12 small boxes to this overview and then link the different model performances to the different biogeochemical properties of the specific areas. As an example, the availability of reference data may allow validating the consistency of the IBI36 system (or of a specific product) in a specific area/season. And this may be critical to provide indications on the quality of the model system to simulate eutrophication in the different areas (and again, the use of the synthesis table suggested at point 2 would be beneficial for this purpose).

**Author answer:**
You are right, the "IBI waters" (Section 2) has been revised in order to introduce the areas (the 12 small boxes) adopted in the validation section. Model performances are now better linked to the different biogeochemical properties of the specific areas.

**Specific comments:**

1. Are there other model applications to the IBI waters (also not operational)? Can the authors discuss how does the IBI36 perform in comparison with other operational systems (e.g. CMEMS Global)?

**Author answer:**

Within the framework of CMEMS, three other MFC share a part of their domain with IBI:
- GLO-MFC which covers the world's oceans at 1/4° resolution and is also using the PISCES biogeochemical model,
- MED-MFC which covers the Mediterranean Sea at 1/24° with the Biogeochemical Flux Model (BFM; Vichi et al., 2007a,b),
- NWS-MFC which covers the North-West European Shelf at 1/15° latitudinal resolution and 1/9° longitudinal resolution (~ 7km) with the ERSEM ecosystem model (Baretta et al., 1995).

For the physical component, two intercomparison papers have been submitted in ocean Science - Special issue "The Copernicus Marine Environment Monitoring Service (CMEMS): scientific advances". They are Lorente et al. (2019) and Mason et al. (2019).

For the biogeochemical component, the comparison of the different model applications is a work in progress and will be the subject of a separate paper, including the contribution of the regional in relation to the global, as well as the comparison of three distinct biogeochemical models.

The Introduction has been revised and now includes a state of the art part.

**References:**

Baretta, J. W., W. Ebenhoh, et al. (1995). "The European regional seas ecosystem model, a complex marine ecosystem model." Netherlands Journal of Sea Research 33(3-4): 233-246.

Lorente, P., García-Sotillo, M., Amo-Baladrón, A., Aznar, R., Levier, B., Sánchez-Garrido, J. C., Sammartino, S., De Pascual, Á., Reffray, G., Toledano, C., and Álvarez-Fanjul, E.: Skill assessment of global, regional and coastal circulation forecast models: evaluating the benefits of dynamical downscaling in IBI surface waters, Ocean Sci. Discuss., https://doi.org/10.5194/os-2018-168, in review, 2019.

Mason, E., Ruiz, S., Bourdalle-Badie, R., Reffray, G., Garcia-Sotillo, M., and Pascual, A.: Copernicus (CMEMS) operational model intercomparison in the western Mediterranean Sea: Insights from an eddy tracker, Ocean Sci. Discuss., https://doi.org/10.5194/os-2018-169, in review, 2019.

Vichi, M., Pinardi, N., and Masina, S.: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: theory, J. Marine Syst., 64, 89-109, https://doi.org/10.1016/j.jmarsys.2006.03.006, 2007a.

Vichi, M., Masina, S., and Navarra, A.: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part II: numerical simulations, J. Marine Syst., 64, 110-134, https://doi.org/10.1016/j.jmarsys.2006.03.014, 2007b.

A link with the global model is included only in Section 5 (P15, L15), where some more details (and relevant references) could be added. Also, the comment to the global performance at the eastern boundary (Skagerrak and Kattegat, P12, L16) could be more supported by some reference (again, the CMEMS Global).

**Author answer:**
The link with the global model is made in Section 3.2 as it is used at initial and open boundary conditions of the IBI system. The comment to the global performance at the eastern boundary (Skagerrak and Kattegat) has been modified. There is not particular evaluation of the Global system is this region. The CMEMS Baltic Sea regional configuration instead of the global product should be tested at the eastern open boundary of the IBI36 system.

2. GODAE Class metrics are widely recognized as references for the skill performance assessment of ocean operational systems and adopted also within the Copernicus community. Is there a specific reason why the authors did not include any mention to them?

**Author answer:**
You are right; it's an oversight on our part. GODAE metrics are introduced in Section 4.

3. Do you have any explanation of why the model anticipates the spring-summer bloom development and decrease in the north Atlantic (P9, L14-18)? Is it linked with physical forcing or with the biogeochemical parameterization?

**Author answer:**

The model anticipates the spring-summer bloom development and decrease in the north Atlantic. The bloom onset is in phase in the south part of the domain, but spreads more rapidly to the north as compared to satellite data. The summer decrease after the bloom is then earlier and sometimes more pronounced in the model. In fact, modelled chlorophyll-a is found deeper during summer with the formation of a Deep Chlorophyll Maximum while maximum Chl-a remains at the surface in BGC-Argo estimates.

I think that both phenomena (timing of the bloom and vertical migration of the maximum chlorophyll) are linked. As phytoplankton starts growing earlier in the model, the mixed layer becomes nutrient-depleted at the end of spring (oligotrophic conditions) in the model and phytoplankton migrates just below. This behavior is also present in the global model.

Several experiments on the physical and biogeochemical components are being carried out. The analysis of the BGC-Argo opens new doors for us to understand the vertical dynamics of phytoplankton. Another area for improvement is opened up with the assimilation of water color data.

This problem is now discussed in the Discussion section.

4. Fig. 4 and comment at (P9, L19-24): referring to "high correlation coefficient between the model and the data" I suppose the authors consider correlation values (r) larger than 0.7. However, box 12 (Gulf of Cadiz) has a correlation r=0.55, which is actually lower than the one for the English Channel (box 6, r=0.59). Please comment.

**Author answer:**
High correlation coefficient refers to values (r) larger than 0.7.
Boxes were slightly revised. As this box was not discussed, it has been removed in the revised version.

Further in Fig. 4: how the associated error to ESA OC-CCI product is estimated?

**Author answer:**
ESA OC-CCI distributes an estimation of the error associated to the chlorophyll-a concentration in sea water. The name of this variable is "CHL_error". But the term "error" is misleading. This is actually a standard deviation of the daily data when calculating the monthly average. It has to be analyzed jointly with the number of daily files that contained data incorporated into the monthly average. So to avoid any misunderstanding, it has been removed in the revised version.

5. Which specific products have been used from EMODnet? Please clarify the kind of “regional aggregated products” used (P7, L31-33). Further, you refer to EMODnet database as “in situ” (P10, L23; P13, L10, and P14, L25), but this looks in contradiction with the “aggregation” previously referred. Please clarify and reformulate.

**Author answer:**
In the two projects Seadatanet and EMODnet chimie, regional datasets are developed. They are "aggregated datasets" or "observation collections" by region, including one product in the North Atlantic and one in the Mediterranean basin. These datasets receive additional quality control of metadata and data.
They can be found here:
http://www.emodnet-chemistry.eu/products/catalogue#/search?fast=index&_content_type=json&from=1&to=20&sortBy=popularity&any=aggregated%20datasets

EMODnet website explicitly uses the term "aggregated dataset", and not “aggregated product” as referred in the paper. So to avoid any misunderstanding, the term “aggregated product” has been changed to “aggregated dataset” in the revised paper.

6. Fig. 10 is potentially very interesting, especially for communities more linked to environmental management. I suggest to enrich the discussion and provide some more information on this indicator: how is computed the minimum in Fig.10a and b (is this defined considering at least 1 daily value lower than the threshold as in P11, L17? is the minimum found along the vertical profile, or at the sea bottom? That should be defined at the beginning of the discussion at P11)? Do the two maps show the absolute minimum over the whole investigated period for each cell shown? What is shown in the time series of Fig.10c? Please clarify. Further, the deficiency threshold (at 187 umol/l) could be added to Fig.10c (e.g. with a dashed line): this may help to appreciate the model skill in representing the oxygen deficiency, and the fact that in 2011 and 2015 the model predicts higher levels than observed. To measure the skill of the model to overcome a specific threshold in comparison with observed data, authors may also consider using the statistics based on the relative operating characteristic (ROC), as for example shown in Sheng and Kim (J. Mar. Sys. 76, 212-243, 2009).

**Author answer:**
Fig. 10a present the absolute minimum oxygen concentrations for each 1°x1° grid point from ICES dataset (Fig. 10a left) and from IBI36 daily outputs co-located to ICES dataset (Fig. 10a right).
Fig. 10b presents the time series of the absolute minimum oxygen concentrations for each date from ICES dataset (red line) and from IBI36 daily outputs co-located to ICES dataset (black line).
So, in Fig. 10a, and b, we only catch low oxygen concentrations reported to the ICES database. We miss all other low oxygen events. But the objective is here to evaluate/validate the model estimates. The model performs in reproducing the spatial distribution and seasonal evolution of low oxygen concentrations.

After this positive evaluation, we are confident in the model results, so we can estimate the total surface area vulnerable to oxygen deficiency over the continental shelf (bathymetry <= 200m). In Fig. 10c, we do not restrict to the IBI36 daily outputs co-located to ICES dataset. We do not use ICES dataset anymore. But instead we use the full outputs of the simulation to give estimation over the whole IBI domain. We consider that a model grid point is an area vulnerable to oxygen deficiency if at least one daily value decrease below the threshold of 6 mg $l^{-1}$ (or 187.5 µmol $l^{-1}$) during the simulation length, as in Ciavatta et al. (2016). The minimum is found along the vertical profile over the continental shelf (bathymetry <= 200m).

More information or clearer information is now given to describe Fig. 10 in the revised paper. The deficiency threshold is now added to Fig.10b as a dashed line.

The model predicts higher levels than observed in 2011 and 2015 (Fig. 10b), but they come from a few measurement points very close to the coast in the vicinity of river mouth, not captured by the IBI36 system.

The ROC (relative operating characteristic) score is a very interesting skill assessment method to measure the skill of the model to overcome a specific threshold in comparison with observed data. I didn't know this metrics, and I think the biogeochemistry modelling community is not used to seeing this metrics. It therefore needs to be presented and explained in detail, as in Sheng and Kim (J. Mar. Sys. 76, 212-243, 2009). This would add complexity to the paper and this section. But I keep this metrics in mind carefully for a study dedicated to oxygen vulnerable areas.

7. Figs. 14 and 15 show the comparison with BGC-Argo floats data in Atlantic and West Mediterranean areas covered by the IBI system. May the authors add any comment on the systematic behavior of the model in comparison with the float also for nitrate and oxygen (as done for Chl-a at depth, see P14, L2, and then in the discussion, P15, L13)? Do the authors suspect that this behavior is related to possible errors of the model (due to representativeness or biogeochemical parameterization), or may be also related to uncertainty associated to float sensors? Further, does the white line in Fig. 15 represent the MLD? If so, it should be added to the caption, and briefly explained how it was computed.

**Author answer:**

Low oxygen concentrations are systematically overestimated and high nitrate are systematically underestimated. This systematic behavior of the model in comparison with the float is also observed when compared to other observational datasets. It can be related to possible errors of the model (due to representativeness, physical or biogeochemical parameterization).

But uncertainty associated to float sensors can not be excluded. An important work on the quality control is currently being done by the BGC-Argo teams, and the quality check process is evolving very rapidly. Temporal drifts, constant or even non-constant vertical bias, and negative concentrations (see the nitrate in Fig. 15d) are still observed in the BGC-Argo data.

Nitrate and oxygen are now more discussed in the revised paper.

Yes, the white line in Fig. 15 represents the MLD. It is now added to the caption, and its computation is now explained.

8. In the discussion (P15, L1-6) why do the authors write that the higher biases on the continental shelf are "no surprising"? Is this related to larger uncertainty given by the river input? Or is there any limits due to the physical drivers? If it seems obvious, please clarify. Secondly, is there any literature basis for the NPP (e.g. with in situ measurements) that can be used to establish which of the 3 referred products we should trust more?

**Author answer:**

Continental margins are very productive regions and play an important role in the biogeochemical cycle of nutrients and carbon. They are the site of complex interactions between physical, chemical and biological processes that include exchanges between shelf and the open ocean, sediment-water interactions, air-sea fluxes, and land-ocean freshwater inputs. In addition, coastal systems are locally strongly affected by human activities. All these interactions make the continental shelf a challenge to obtain realistic models.
It is now clarified in the revised paper.

Regarding the NPP, we are currently collecting the estimates reported in the literature in order to build an in-situ database. This work is very time-consuming, and unfortunately cannot be included in this paper.

9. At the end of the discussion, there is a reference to OMI (P16, L8-11): this might not be interesting for the general reader. However, since there is a specific indication of the "IBI-MFC biogeochemical forecast service", this should be properly validated before providing OMI (see major point n.1).

**Author answer:**
The text has been modified to avoid confusing the reader.

**Technical/other corrections:**

In the webpage of OSD (https://www.ocean-sci-discuss.net/os-2018-161/#discussion), the correct author's surname should be "Sotillo" and not "Sottilo".

The author's surname has been corrected.

P2, L8 and P4, L31. The "IBI-MFC Team" has not been defined.

The “IBI-MFC Team” is now defined in the Introduction.

P3, L14-15. “The annual primary production is then limited its seasonal variations are limited”: not clear, please correct or reformulate.

The sentence has been reformulated to : “The annual primary production is then limited and so is its seasonal variations”.

P3, L29. References about N:P ratio around 20 in the Western Mediterranean are also found in the recent work of Lazzari et al. (Deep-Sea Res. I 108, 39-52, 2016; see their Fig. 7), that may be added.

The reference Lazzari et al. (2016) is now added.

P5, L25-27. “Although PISCES was originally designed for global ocean applications, the distinction of two phytoplankton size classes and the description of multiple nutrient co-limitations allow the model to represent ocean productivity and biogeochemical cycles in the major ocean biogeographic provinces (Longhurst, 1998).” The reference seems not appropriate: PISCES is not used in Longhurst (1998), which only refers to the bio-provinces. Please reformulate.

The reference Longhurst (1998) is now removed to avoid any misunderstanding.

P6, L1-2. “To respect the conservation of the tracers, the coupling between biogeochemical and physical components is done every other time.” Not clear, please clarify: what do the authors consider as “conservation of the tracers”? Is it mass conservation?

Yes, we refer to mass conservation. The sentence has been modified.

This question also relates to the reference to Aumont et al. (2015), at L19-L34: it would be clearer for the reader to add more details about the river inputs (which are very important in coastal regions). Please describe the specific inputs (also reporting from Aumont, Section 4.9.2, and the one from EEA), list the names of the rivers (or their total amount), specify what you consider as “natural” and what as “anthropic”, and describe their impact in the investigated region. Finally, please clarify the use of “reminder” the last sentence “For the other variables, a reminder of the initial conditions is given.”: does this mean the initial condition values are maintained constant during the simulation?

More details and clearer explanation is now given about the external inputs of nutrients linked to the river runoffs. We briefly describe the impact of additional nutrient inputs in the investigated region. The terms “natural” and what as “anthropic” have been removed to make the text clearer

and avoid any misunderstanding. The location of the 33 rivers considered in the IBI36 system is available in Maraldi et al. (2013).
Yes, the other variables are maintained constant during the simulation at the open boundary conditions that are the river points. This last sentence “For the other variables, a reminder of the initial conditions is given” is now removed. It is not useful because we are only talking about nutrient inputs here.
We are aware that river discharges could be greatly improved. But it is hampered by the lack of in situ observations.

P6, L8. “is” should be “are”.

“is” has been changed to “are”.

P7, L20. As done with VGPM and CbPM, please add a reference also to the “Eppley version”.

A reference is now added in the revised paper.

P8, L6-18. Could you provide which fields from the BGC-Argo repository have been used and whether additional calibrations were performed after?

The APEX float observations in the Atlantic are adjusted following Johnson et al. (2017). The three variables are "delayed mode" data. For the PROVOR float observations in the Mediterranean, oxygen and nitrate are "Real time" data and Chl-a is "adjusted" data. They are adjusted following Mignot et al. (2018).

P8, L18. “Johnson et al.” (dot missing).

The dot was added.

P11, L13. “Breitburg et al.” (dot missing).

The dot was added.

P11, L19. It should be “surface area in winter”.

It is now corrected.

P11, L20. The authors refer to “the west coast of France” but in Fig.10 a and b it looks like only the north-western French coast is covered by data, and with value larger than 200 umol/l. Please check.

As detailed above, Fig. 10a only catches low oxygen concentrations reported to the ICES database. All other low oxygen events are missed. In Fig. 10c, IBI36 outputs over the whole IBI domain are used to estimate the total surface area vulnerable to oxygen deficiency. So areas not detected by the ICES database may appear vulnerable due to simulation. It is now better explained in the text.

P12, L4. "The seasonal cycle is in phase for Chl-a but out of phase for ammonium" it is also possible to refer to Fig.7e, where correlation for NH4 is r=0.2.

The reference to Fig. 7e is added.

P14, L25. Instead of "is described here", it should be clearer to indicate the corresponding Section or Figure (Figs. 2, 4 for Chl-a and 5, 6 for NPP). Or are the authors here referring to the general approach of the manuscript? Please clarify. Further, it should be "are described".

"is described here" has been changed to "are described in this paper".

P17, L3. The fact that a co-author is also acknowledged sounds somehow strange. Please check.

The acknowledgement is removed.

P29, L3. Correct "Eppley".

Eppley is corrected.

P33, Fig. 10. Add the units for O2 minimum (umol/l).

The units was added to the legend of the figure.