Ocean Sci. Discuss.,
https://doi.org/10.5194/os-2020-12-RC2, 2020

Ocean Science
Discussions

Interactive comment on "An approach to the verification of high-resolution ocean models using spatial methods" *by* Ric Crocker et al.

Anonymous Referee #2

Received and published: 1 April 2020

In this contribution, the authors conduct a skill assessment of two operational ocean models running in the North West European Shelf with different configurations and spatial resolutions. Since the increased spatial resolution might require ad hoc metrics to properly reflect the model performance and reduce the impact of so-called double-penalty effects (occurring when using point-to-point comparisons with features present in the model but misplaced with respect to the observations), the present work is welcomed. It addresses this interesting and essential topic by intercomparing models' performances in overlapping regions to infer their respective strengths and weaknesses. Equally, the methodology proposed is consistent and the results obtained are relevant, especially in the framework of the Copernicus Marine Service (albeit not explicitly mentioned in the document).

Printer-friendly version

Discussion paper

Based on my expertise on ocean models validation, I particularly appreciate the proposed approach (named HiRA) since it might be useful in parent-son inter-comparisons in order to quantify the added-value of downstream services such as very high resolution coastal models (embedded into CMEMS regional ocean forecasting systems) that are currently running in port-approach areas.

I am confident this work can attract the interest of the scientific audience, being cited in future works dealing with similar issues. The style was fluent although some parts (mainly the introduction and the references) could be revised and enhanced. Based on my judgment, I deem the manuscript acceptable upon minor revision. In the following lines I provide some comments, which should hopefully strengthen even more the manuscript.

General comments:

1. Since the main purpose of this work is to showcase the potential of the proposed methodology in operational ocean forecasting, I miss a reference to the Copernicus Marine Environment Monitoring Service -CMEMS- (Le Traon et al., 2019)., although the in situ observations used here were downloaded from CMEMS catalogue. Within this context, there are some valuable and concerted initiatives such as the Product Quality Working Group (PQWG) or the North Atlantic Regional VALidation tool -NARVAL- (Lorente et al., 2019) where physical and biogeochemical model intercomparisons are conducted on a regular basis to deliver outcomes to a broad scientific community.

Le-Traon et al., 2019. "From Observation to Information and Users: The Copernicus Marine Service Perspective". Front. Mar. Sci., 22, https://doi.org/10.3389/fmars.2019.00234.

Lorente et al., 2019. "The NARVAL software toolbox in support of ocean models skill assessment at regional and coastal scales". Computational Science, ICCS 2019.

2. Equally, I also miss a reference to GODAE Coastal Ocean and Shelf Seas Task

Team (COSS-TT), where the Met Office is an active member, involved in a wealth of valuable initiatives in terms of ocean model inter-comparisons. In this context, I think that the state-of-art about previous inter-comparison exercises is not thorough and is poorly cited, despite of the abundant literature reported elsewhere. In this work, there are only 28 references (which is insufficient) and nearly the 50% of them were published in 2010 or earlier, so an update is highly recommended. Below I suggest a number of recent works to build upon:

Aznar et al, 2015. “Strengths and weaknesses of the CMEMS forecasted and reanalyzed solutions for the Iberia-Biscay-Ireland (IBI) waters”. Journal of Marine Systems, 159, 1-14.

Mourre et al., 2019. “Assessment of High-Resolution Regional Ocean Prediction Systems Using Multi-Platform Observations: Illustrations in the Western Mediterranean Sea”.

Lorente et al., 2019. “Skill assessment of global, regional, and coastal circulation forecast models: evaluating the benefits of dynamical downscaling in IBI (Iberia–Biscay–Ireland) surface waters”. Ocean Science, 15, 967-996. Doi: /10.5194/os-15-967-2019.

Mason et al., 2019. “New insight into 3-D mesoscale eddy properties from CMEMS operational models in the western Mediterranean”. Ocean Science, 15, 1111–1131.

Hernández et al., 2018. “Measuring performances, skill and accuracy in operational oceanography: New challenges and approaches”. In "New Frontiers in Operational Oceanography", Eds. GODAE OceanView, 759-796, doi:10.17125/gov2018.ch29.

Juza et al, 2015. “From basin to sub-basin scale assessment and intercomparison of numerical simulations in the western Mediterranean Sea”. Journal of Marine System, 149, 36-49, doi;10.1016/j.jmarsys.2015.04.010.

Hernández et al., 2015. “Recent progress in performance evaluations and near real-time assessment of operational ocean products”. Journal of Operational Oceanogra-

phy, 8, Issue sup2: GODAE OceanView Part 2

Rockel et al., 2015. “The regional downscalling approach: a brief history and recent advances”. Curr. Clim. Change Rep., 1, 22–29, https://doi.org/10.1007/s40641-014-0001-3.

Katavouta et al, 2016. “Downscalling ocean conditions with application to the Gulf of Maine, Scotian shelf and adjacent deep ocean”. Ocean Model., 104, 54–72.

And some other older works:

Crosnier, L., and C. Le Provost. 2007. “Inter-comparing five forecast operational systems in the North Atlantic and Mediterranean basins: The MERSEA-strand1 methodology”. Journal of Marine Systems, 65, 354–375.

Greenberg et al, 2007. “Resolution issues in numerical models of oceanic and coastal circulation”. Cont. Shelf Res., 27, 1317–1343.

Hernández, 2011. “Performance of Ocean Forecasting Systems–Intercomparison ProjectsÂů. Book: Operational Oceanography in the 21st Century, Chapter 23.

3. In section 1 (Introduction), a preliminary paragraph about why model inter-comparisons are necessary would be convenient. Equally, a brief description of the types of inter-comparisons exercises would be pertinent: i) between two different forecasting systems in the overlapping region to check the consistency of each model solution; ii) between two versions of the same system, in order to evaluate the added-value of the upgraded one before it is transitioned into fully operational status; iii) a parent-son inter-comparison, to evaluate the quality of the downscaling approach adopted; iv) a comparison between both the forecast and the reanalyzed solutions of the same model in order to infer the primary role of both the grid resolution and the atmospheric forcing, especially in coastal areas (see Aznar et al., 2015, for further details).

4. In section 2.1 (Data and Methods: Forecast), I strongly suggest adding a table to provide a general overview of the two model′s main features in a more synthetized way:

version of model, geographic domain, grid resolution, number of depth levels, number of forecast horizons, open boundary conditions, tidal forcing, atmospheric forcing, river forcing, assimilation scheme, bathymetry, etc. Although most of this information is already provided in the text, I think a table would be rather useful as a summary.

5. In section 2.1 (Data and Methods: Forecast), neither river forcing is mentioned, nor river freshwater discharge is taken into account when describing the general considerations. The study-area comprises several rivers estuaries (Seine, Rhine, even Loire) with significant freshwater runoff that might eventually impact on the SST field in coastal areas. Figure 2 shows that some stations are located quite close to those rivers mouth. Please clarify this point, why the river forcing is out of the discussion. In particular, Graham et al (2018) suggested that AMM7 configuration might be more diffusive than AMM15 within river plumes, allowing freshwater input from the Rhine to be advected offshore.

6. In the same line, an event-oriented inter-comparison (with a focus on river plumes and abrupt SST drops due to impulsive-type riverine discharges) would allow you to better infer the ability of each system to capture small-scale coastal processes (with and without HiRA approach). This process-based validation approach, albeit commonly used in meteorology and weather forecasting, is rather novel in operational oceanography and mostly devoted to extreme sea level and wave height episodes. I am not asking to provide new and complementary analysis but please take it as a kind suggestion for future works.

7. With regards to the double-penalty effect, I was somehow expecting a multi-parameter analysis, with a special focus on altimetry products, sea level anomalies and mesoscale eddies. Did you have the chance to test HiRA approach with other variables? If so, could you add a comment about it, even if you only obtained preliminary results? If not, I think this task should remain as a priority for future works and thus be explicitly mentioned in the text.

8. Likewise, I miss a deeper discussion respect to the previous works by Tonani et al (2019) and especially that one by Graham et al (2018) where a "traditional point-to-point SST validation approach" was performed with the new AMM15 system. I think that the fact of contrasting results from both papers / both methodologies could benefit the discussion section, particularly when dealing with on-shelf and off-shelf differences as far as Graham et al (2018) proved the reduction in seasonal SST bias was greater off-shelf than on-shelf when using AMM15 (which supports the results exposed in Figures 9 and 10 of the present work). Again, on-shelf results were worse and you succinctly listed river mixing as a potential source of uncertainties, but no additional information was provided about the role of river forcing (as I aforementioned in point 5). I guess that the river fluxes could have been altered between the two models (being one configuration fresher and cooler than the other).

Minor comments:

Abstract:

I recommend explaining briefly (in two lines) the double penalty effect as part of the potential audience might not be familiarized with this concept. For instance: "[. . .] referred to as the double-penalty effect, occurring in point-to-point comparisons with features present in the model but misplaced with respect to the observations."

Keywords:

I suggest adding "skill assessment", "validation" and/or "double-penalty".

Figure 1:

As previously indicated by the anonymous reviewer 1, a more contrasted color bar is required to highlight the spatial SST differences. Bathymetric contours would be also welcomed.

Figure 8:

Printer-friendly version

Discussion paper

Albeit rather obvious, please indicate that masked regions are in grey color.

Introduction:

Lines 58-60: That sentence sounds odd. Could you rephase it, please?

Line 61: please replace “suggested” by “suggesting”

Line 65: please replace “more like” by “more similar to”

Section 2.1. Forecast

Lines 106-108: I guess that hourly instantaneous values are provided for the sea surface and daily averages for the rest of the water column. Please, could you clarify it?

Line 117: Why the study period comprises from January to September 2019? Any chance to expand the analysis to cover the entire 2019 year? That would be interesting to infer seasonal differences between both model configurations...

Lines 132-133: please comment that semi-diurnal M2 is one of the predominant tidal constituents in this region (that is the reason to compute means over 25 hours in order to remove the tidal signal).

Section 7: Discussion and conclusions.

Lines 538-539: as previously indicated, provide further insight into on-shelf and off-shelf differences, contrasting the results obtained with those reported in Graham et al (2018).

Lines 540-545: is there any adopted rule or any agreed proposal to wisely select the neighborhood sizes?

Interactive comment on Ocean Sci. Discuss., https://doi.org/10.5194/os-2020-12, 2020.