Comments on the manuscript: *"The impact of a new high-resolution ocean model on the Met Office North-West European Shelf forecasting system". by M. Tonani et al.*

The manuscript presents a detailed description of a new MetOffice ocean forecasting system at 1.5Km. This new high-resolution AMM15 system, together with the previous existing one AMM7 (at 7Km resolution), complete the physical ocean model system used to produce the CMEMS North-West-Shelf ocean forecast and analysis product. A comprehensive validation of both systems is provided. To achieve this validation, a comparative assessment of both model systems has been performed, using trial runs over 2 years period.

As the authors mention in the manuscript, in some occasions it is not an easy task to demonstrate the significant improvement of higher resolution model performances. The difficulties to assess the differences between both model systems, the higher and lower resolution ones, are mainly related to the scarcity of adequate observational data sources. Nevertheless, the present paper aims to do it, and it presents a complete general validation work. Besides, it is shown some additionally examples of model validation with very specific (but geographically limited) observational data sources, such as gliders and HF radar sites.

Despite the general scientific interest of the manuscript may be enhanced, the proposed paper is of interest in the context of the present CMEMS OS special issue. The complete description and exhaustive assessment of the new High-Resolution model system with respect to the previous existing one, is of interest for future, scientific and non-scientific, CMEMS NWS end-users, using products derived from the model systems here presented. Therefore, I do recommend publication of the manuscript after revision of some points.

For instance, I would ask the authors to justify in the revised manuscript some of the decisions taken to build the new 1.5Km model system set-up. The authors should address in more detail some of the choice made related to:

1) the model configuration (i.e. why the authors keep in the high-resolution system the same tidal forcing (using the same 12 harmonics) then in the lower resolution system, why they use the same vertical grid distribution)
2) the data assimilation scheme used in the AMM15 system (i.e. why in a shelf model system, as the AMM15 is, it is assimilated SLA only outside the shelf; and how do the authors face the challenge of assimilating altimetric observations in high tidal environments).

Providing more info on these points, of interest for ocean shelf modelers, the authors certainly will enhance the scientific interest of the paper for the ocean data assimilation and modelling community.

Apart of the points to be address in more detail by the authors, there are also some suggestions of text changes to improve the paper readability. For instance, I would recommend moving any "pure" model-model comparisons, that is with no observational data source used as reference, from the Section 4, dedicated to validation results. Thus, the final comparisons shown on the surface currents patterns may be moved to another earlier section. I find these results very illustrative and give a good measure of the differences that we can expect from the new increased resolution model system (so, they

should be included in the manuscript to show the different model performance achieved), but they do not provide any model validation (so, this text should be out of Section 4).

As it is said before, I recommend publication of the manuscript after revision of the following points listed below.

Abstract:

P1.15 "… (AMM7) that has been used for many years". Please, specify the context (CMEMS?, before Copernicus?)

P1.18 "Trial experiments run with the low and high resolution systems in their operational configuration". Please, specify if this operational configuration includes Data Assimilation, or means just forecast runs.

Section 1. Introduction:

P1.35 In this paragraph, the authors mention human activities (industrial, farming, fishing) with climate change as source of impacts in the quality of water environments. All of them have certainly an impact, but I would suggest re-drafting the sentence, separating the impacts from climate change and the human-related activities, since they are at different levels.

P2.1 Include some reference to sustain the paragraph.

P2.13 The (CMEMS?) operational forecasting for the North-West European Shelf (NWS). Other applications?

P2.15 To describe the geographical domain, the Figure 1 is referred. However, when a reader goes to this Figure, sees 2 different model domains: the AMM15 & AMM7, not mentioned yet and with not defined acronyms. It is a bit confusing for the reader at a first reading. The authors should improve this point: 1) moving after in the text the citation of this Figure 1, or 2) improving the figure caption to give more information on the features shown.

P2.24 From this line up to the end of this Introduction Section, the authors mention different components of the CMEMS NWS system (i.e: ocean physical model, data assimilation system, together with the biogeochemical model coupled into it). Also, there is a mention to a wave model system, and to an ocean-wave-atmospheric coupled system. In order to enhance the understanding of the systems and its multiple connections with other applications, here outlined, the authors should include a figure showing a schematic view of the CMEMS NWS operational forecast system, here described. This extra figure suggested may be included as part of the present Figure 2. This way, the number of figures is not increased and the present Figure 2, what currently provides certainly very few information, is enhanced.

I would also miss in this part of the manuscript some reference to the CMEMS operational products generated through the model systems here described (with citation to their documentation).

P2.30 With respect to the latest system mentioned in this paragraph: the ocean-wave-atmosphere coupled system, is this currently an operational one? or is it in a pre-operational phase? Or only for research purposes?

P2.37 Please, include some quantitative numbers or estimations to support the adjective "prohibitive".

Section 2. System Description

P3.10 Use CMEMS instead Copernicus.

P3.19 The new AMM15 system uses the same vertical grid resolution than the AMM7 one. Why it was not considered to increase the vertical resolution consistently with the horizontal one? Is the present vertical resolution with 51 levels enough? Have the authors performed any sensitivity test to evaluate the impact of an enhancement of vertical resolution? Or the decision to keep the vertical resolution unchanged is more a matter of computational resource availability? Any comment on this point?

P4 No reference in the text to Table 1?

P5.9 Are 12 tidal harmonic constituents enough to rightly reproduce the tides in a region such as the one covered by the high resolution AMM15 model, that is marked by shelf shallow waters with very high tidal environments? Can the authors justify why the same 12 harmonics are used in both systems? Since the objective is to model the region at a very high resolution, it would not be worthy to count with an improved higher resolution tidal forcing (the original TPX harmonic are at a 1/12º resolution). Furthermore, please, include in the manuscript the list of the 12 harmonics used (this list of harmonics can be provided directly in the text, or in Table 2).

P5.18 The authors mentioned that ECMWF IFS data is used as forcing in the AMM15 system, whereas the AMM7 uses the MetUM forcing. The move to the ECMWF forcing is justified as a requirement of the CMEMS service. However, the IFS data have lower resolution than the MetUM (around 14 Km in the former, instead of the 10 Km of the later). Apart of this "service" reason, can the authors comment on the impacts that move from a higher resolution forcing to a lower one has in the ocean model solution?
Furthermore, later the authors mention that using IFS there is a lost in terms of analysis frequency availability (from 3h to 6h). Can the authors provide some quantification of the impact related to the change in the forcing? It is certainly not very intuitive for a reader to understand how when a new higher resolution model system is being set up, it is decided to use a lower resolution atmospheric forcing. Can the authors explain any positive impact of the change in the atmospheric forcing to support the decission?

P7.11 It is stated that there is SLA assimilation both in MM7 and MM15 systems, and in both cases, for regions with bathymetric depths > 700m. In the case of the MM7 configuration this option can make sense, since extended deep water areas are covered. However, on the contrary in the case of the AMM15 shelf model system, this set-up option seems to result in a SLA data assimilation limited to a very narrow area (and very close to the open boundaries!). Can the authors explain in more detail the impact of the SLA data assimilation approach performed on the AMM15 shelf system? Can the authors provide a measure of the benefit of assimilating SLA data assimilation on such a limited (and so close to the boundaries) area? The authors should explain better the potential gain of using such limited SLA data assimilation with respect to a free non-assimilative approach.

P7.13 Table 4 cited before Table 3. Please, try to respect the order in Figure and Table citation.

Table 4. In the column of Data source: 1st arrow: "CMEMS –INS-TAC" may be substituted by "CMEMS-INS-TAC Product:" The same for "GTS" ("GTS Product:"?); 2nd arrow: "CMEMS-SL-TAC Product:"
3rd arrow: "Product from the Group for High Resolution Sea Surface Temperature (GHRSST):"

P9.15 What is it done with the info on the profile quality check performed? Any communication established with the observational data producers? (a kind of blacklisting?).

P9.19 Do the authors foresee any problem in using OBCs from different model data sources? Are they consistent? Can be a source of problems due to volume conservations issues?

P9.37. The production process takes approximately 4 hours. How many CPUs are used during the process? Can the authors include here a computational cost estimation?

Figure 2: Include here info on the ECMWF IFS forcing (analysis/forecast) used. Complete this Figure, as suggested in previous comment, showing a schematic view of the CMEMS NWS operational forecast system described.

Section 3. Trial experiments

Figure 3. Number of observations used for assimilation. The panel on the SLA show effectively the satellite SLA observations available. However, this panel can mislead the reader, since the data assimilation is applied only on areas with depths > 700m. I suggest the authors will identify in the plot the area where SLA is effectively assimilated in AMM15 system.

The authors should consider the possibility to include in this section the analysis of the differences between the dynamical patterns modelized by the 2 different model

systems, currently included in the Validation Result section. This point is suggested below.

Section 4. Validation of the experiments

Figure 4. specify also here the locations where observations from coastal tide gauges are available for the tidal validation.

2 figures are dedicated to display location of observational data sources used in the paper. The Figure 3 shows those observations used in the data assimilation. On the other hand, the Figure 4 displays other observational data sources used in the validation process. Where are the coastal tide gauges? I guess they are not assimilated, however they are not depicted neither in Figure 4.

Furthermore, the reader founds later in Table 6 (where results from the validation of different variables are shown) results for the M2 tidal harmonic and there it is said that validation is done for the full domain. However, no info on the location of the tide gauges used is provided up to that moment. Later, already in Section 4.1, in the Figure 5 there is a map of model-obs differences in M2 amplitude and phase. Please, clarify a bit the geographical information on the tide gauge locations.

P14.11 Tide gauges observations from BODC. "The number of tides gauges taken into consideration for AMM15 and AMM7 is the same, therefore the coastal buoys"; are the authors here referring to tide gauges? Or to buoys? Can the authors provide more details about the tidal observations offshore, where do they come from? (From platforms?, pressure sensors?). More explanation about the tidal measurements from the BODC it may help the reader.

P15.13 Suggestion to ease readability: in one of the maps, for instance in Figure 1, the authors should detail all the geographical names mentioned in the text (i.e. German Bight). This reference to geographical features will ease the reading of the paper to those potential readers not familiar with the regional geography.

P16.5 The authors shown in the paper (section 4.1.2) some results of the model validation with a HF Radar system. The results are only for 1 month (March 2017). If the authors have 2 years model runs, why do they perform/show a so short temporal coverage of the model- HF Radar validation? Due to observation availability? Please, explain reasons in the manuscript.

Figure 6 shows the results of the model-HFRadar validation. In this figure, it is shown some statistics fields (RMSD, Bias, Veering) limited to the HF Radar spatial coverage. However, the reader have no information about the number of observational data that support these statistics. Do the HF-Radar system provide exactly the same number of observations everywhere? If yes, please detail what gap filling methodology is being applied. If not, please, show the % of HF-Radar data availability.
I guess the 3 names referred in this figure 6 corresponds to the HF Radar sites. Please, detail in the Figure Caption.

P18.13 The in-situ measurements are from buoys and ships of opportunity. Please, detail if "buoys" means fixed moorings, surface drifters or ARGO profilers.

P19.4 A Butterworth filter. Please, explain in more detail or add a reference.

P19.20 AMM7 and AMM15 models provide very similar values of SST, probably due to the data assimilation of SST that brings models close to the observations. Can the authors include in the paper any SST timeseries analysis as the one here shown for the 3 proposed sites, but in a station, whose SST observational data would not be assimilated? See some independent validation would certainly be of interest for readers and potential users of the model products.

Figure 9. It is quite remarkable the overestimation of the 12-h energy peak in AMM15. Any relation with the harmonic bias in M2? It is also interesting the notorious AAM15 peak around 6-h. Can the authors comment on it? Any explanation? May it be linked to the meteorological forcing? (different in both model systems).

P23.1 E-Hype. What is E-Hype? No mention to this name in the section where forcing are described. Please, introduce complete name of the source or reference.

P24.24 Please, check the date: 23rd March or 23rd May (as referred in the Figure 12 caption; in this Fig 13 caption, correct typo: 23rt).

In Section 4.4.4 it is discussed about model differences in MLD and it is referred to the Figure 15, where only the MLD computed from the glider is depicted. Why the MLD computed from the models are not shown in the panels Glider-MM15 & Glider-MM7 together with the one derived from the glider data? Include the MLDs from both models in the plot can enhance the analysis in this section dedicated to MLD.

Section 4.5 is devoted to show some results from currents compared with HF-Radar data. As in the previous case for the tides, only a month of data (March 2017) is shown. Please, justify why a so short temporal coverage for the validation.
Figure 16 shows monthly values of the HF-Radar and from the 2 models, interpolated to the observational field. However, no information on how many observations support the resulting monthly value is provided. Please, include the % of data availability for the month shown. It will be also useful to have some information on the validation of the HF-Radar measurement, as well as on the gap filling methodology used (if someone is used). The explanation/discussion of the comparative results is quite poor. Please, provide some more description of the features depicted. For instance, it will be interesting that the authors describe the high currents feature existing in front of Wang and Busum stations, reproduced by the AMM15, but not for the AMM7 model. Likewise, any explanation or comment about possible border effects in the HF-Radar field shown would also be pertinent. Can the authors ensure that all the high currents depicted at the border of the HF Radar coverage are reliable? Please, include some info in the text (a reference would also help) on the existing validation of the HF Radar data used and about the possibility of border effects in the observational data used.

The analysis of the AMM7 & AMM15 model currents provided from P 28.14 till the end of the Section 4.5 (including reference to Figure 17) is not referred to any model validation. It is not used any observational data source used as reference. Therefore, I would suggest taking this analysis out from this Validation section.

I found the analysis interesting, and it illustrates quite well the dynamical differences existing between both model solutions.

If the authors want to keep this analysis in the manuscript, I would suggest moving this part of the text and the figure to the end of the Section 3 (where Trial experiments are described). This analysis of the dynamical patterns obtained gives a good idea of how different the 2 model solutions are and it may give a good introduction to the reader to the validation results that come later in Section 4.

Section 5. Conclusions and future developments.

P30.7 typo: temporal

Please, include in the conclusion section some reference to the Data Assimilation performed in the AMM15 system, with mention to potential future plans to enhance the assimilation process (and very specially for SLA on the shelf).

P31.18 The AMM15 ocean (system?).