

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

Bold: referee’s comment

Not bold: author’s answer

The referee’s comments are copied in this document for ease of reading.

General comment:

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)**

The major aim of this model configuration is to resolve the Rossby Radius on the shelf, therefore the focus was on increasing the resolution from 7 to 1.5 km. All the technical details of the implementation and the validation of the model, without data assimilation, are presented in Graham et al., 2018. That is the precursor work to this paper. The tidal forcing is going to be improved, in the next release of the AMM15 model that will include also the wetting and drying. Experiments are ongoing, using FES2014 and more tidal constituents. The number of vertical levels is the same because the focus of this model is on the shelf (depth < 200m), where 51 z-sigma levels are enough for proving a very high vertical resolution. The resolution is of the order of 20cm the shallower part of the model domain, where the minimum depth is 10. More levels will increase the model vertical resolution in the deepest part of the domain, not on shelf (Siddorn et al., 20016). Another possible approach is using

vertically adaptive vertical coordinates so that you focus resolution on the thermocline. This is done in other models but not here and will be considered in the future configurations.

It's important to have a step by step incremental improvements protocol for updating an operational system, rather than changing many things all at once. This is the very first implementation of AMM15 in operations. This system or part its components will be improved on a yearly base in the future releases.

We have provided more technical information on this topic answering the specific questions here below.

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.

This first implementation of assimilation in the high resolution AMM15 followed the same scheme as used in the AMM7. Although the deep water areas of AMM15 are more limited than AMM7, and so the benefit of assimilating SLA may be small, this is a milestone on the way to assimilating SLA throughout the domain.

However, the assimilation of SLA observations in the deeper water still allows us to constrain the temperature and salinity in those regions, which then provides a better boundary condition for the shallow regions. As discussed in King et al. 2018, the assimilation of altimeter observations and T/S profiles is complementary, and in regions such as the NWS where profile observations are relatively limited, altimeter observations provide a valuable additional constrain on the density structure of the deep water regions.

Suggestions of text changes to improve the paper readability

I would recommend to 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)

We thank the reviewer for this suggestion. We have modified the manuscript moving the “pure” model-model comparison in section 3: “3.1 System performance: AMM5 vs AMM7”.

The sub-section on the “Currents” has been rename as “Currents in the German Bight” and moved up after the “Tidal flow” sub-section, because both based on the same HF radar observations.

The manuscript is now organized as follow:

Abstract

1. *Introduction*
2. *System description*
 - 2.1 *Core model description*

- 2.1.1 *Boundary and surface forcing*
- 2.2 *Assimilation method*
- 2.3 *Operational production*
- 3. *Trial experiments*
 - 3.1 *System performance: AMM15 vs. AMM7*
- 4. *Validation of the experiments*
 - 4.1 *Tides*
 - 4.1.1 *Tidal harmonics*
 - 4.1.2 *Tidal flow*
 - 4.2 *Surface currents in the German Bight*
 - 4.3 *Sea Surface Height*
 - 4.4 *Sea Surface Temperature*
 - 4.4.1 *Comparison with in situ and satellite*
 - 4.4.2 *Variability in SST*
 - 4.5 *Water column*
 - 4.5.1 *Temperature and salinity profiles*
 - 4.5.2 *Moorings in the German Bight*
 - 4.5.3 *Glider transects*
 - 4.5.4 *Mixed layer depth*
- 5. *Conclusion and future developments*

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?)

We added this information in the text.

OLD:

“The latest configuration to be put in operations, an eddy resolving model at 1.5 km (AMM15), replaces the 7km model (AMM7) that has been for a number of years.”

NEW:

“The latest configuration to be put in operations, an eddy resolving model at 1.5 km (AMM15), replaces the 7km model (AMM7) that has been used for eight years to deliver forecast products to the Copernicus Marine Service and its precursor projects.”

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.

The sentence has been reworded to specify that the trial experiments are done with data assimilation:

“Validation of the model with data assimilation is based on the results of two years (2016-2017) trial experiments run with the low”

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.

We have re-worded that sentence as follow:

“The increasing focus on understanding the marine environment in support of sustaining healthy and biologically diverse seas is also a considerable driver in these waters, where human activities like heavy industrial and farming activity, as well as fishing together with climate change effects, may have significant impacts on the quality of the marine environment.”

P2.1 Include some reference to sustain the paragraph.

We added the following references :

She, J., Allen, I., Buch, E., Crise, A., Johannessen, J. A., Le Traon, P.-Y., Lips, U., Nolan, G., Pinardi, N., Reissman, J.H., Siddorn, J., Stanev, E., Wehde, H.: Developing European operational oceanography for Blue Growth, climate change adaptation and mitigation, and ecosystem-based management, Ocean Science, 12(4) 953-976 <https://doi.org/10.5194/os-12-953-2016>, 2016.

Siddorn, J.R, Good, S. A., Harris, C. M., Lewis, H. W., Maksymczuk, J., Martin, M. J., Saulter, A.: Research priorities in support of ocean monitoring and forecasting at the Met Office, Ocean Science, 12(1), <https://doi.org/10.5194/os-12-217-2016>, 2016.

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

This operational configuration has been implemented for CMEMS. Several CMEMS users are using the product for a wide range of applications or for developing downstream products.

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.

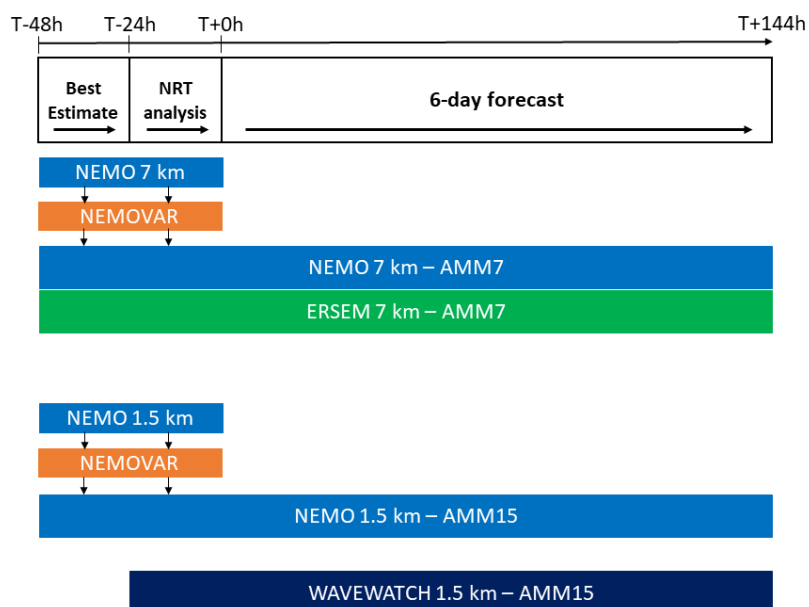
Thanks for the suggestion, we opted for improving the caption of figure1 :

Figure 1: EMODnet bathymetry, in meters (logarithmic scale), showing the NWS high resolution, AMM15, model domain. The red line defines the NWS low resolution, AMM7, model domain. The yellow dotted box is the domain covered by the AMM15 products delivered on a regular grid to the Copernicus users. Figure modified from Graham et al. (2018). The bathymetry colour range has logarithmic scale.

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).

Thanks for this suggestion. We had added this information in figure 2:



We have added the list of the CMEMS products at the end of the introduction:

“The 7km products (AMM7) delivered through Copernicus are:

- NORTHWESTSHELF_ANALYSIS_FORECAST_PHY_004_001_b
(<http://marine.copernicus.eu/documents/PUM/CMEMS-NWS-PUM-004-001.pdf>);
- NORTHWESTSHELF_ANALYSIS_FORECAST_BIO_004_002_b
(<http://marine.copernicus.eu/documents/PUM/CMEMS-NWS-PUM-004-002.pdf>).

The 1.5km products (AMM15) delivered through Copernicus are:

- NORTHWESTSHELF_ANALYSIS_FORECAST_PHY_004_013
(<http://marine.copernicus.eu/documents/PUM/CMEMS-NWS-PUM-004-013.pdf>);
- NORTHWESTSHELF_ANALYSIS_FORECAST_WAV_004_014
(<http://marine.copernicus.eu/documents/PUM/CMEMS-NWS-PUM-004-014.pdf>).

This study is focused on the product NORTHWESTSHELF_ANALYSIS_FORECAST_PHY_004_013 and its inter-comparison with NORTHWESTSHELF_ANALYSIS_FORECAST_PHY_004_001b."

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

This system is just for research purposes at present, as written in the text "*coupled ocean-wave-atmosphere research system*".

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

We added this sentence to the manuscript:

" , because the production time exceeds the 24-hr"

2. System Description

P3.10 Use CMEMS instead Copernicus.

Done

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?

The increased horizontal resolution moving from AMM7 to AMM15 allows a step change in the ability to represent small-scale processes, but there remains work to be done to address biases in the vertical representation in the shelf seas (detailed in Graham et al. 2018a). These are influenced by many factors including the vertical mixing scheme, advection, light attenuation scheme, and wave-mixing parameterizations. Simply increasing the vertical resolution was not expected to lead to any improvements without first addressing these physical mechanisms (see also answer to the first general question).

P4 No reference in the text to Table 1?

Thanks, we added the reference to Table 1 in the manuscript.

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 1/12° 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).

We thank the author for this comment. There is a typo in the manuscript, the tidal constituents are 11, not 12 (We have corrected the manuscript).

The first implementation of AMM15 has been set up as much as possible as AMM7 for understanding the impact of the increased resolution. The number of tidal constituents has not been increased.

AMM15 has less tidal constituents than AMM7 due to the different model used (see the updated Table 2 here below).

It is worth to note that the tidal boundaries in AMM are in the deep (off shelf) region for the most part (excepting short stretches where they cross the continental shelf). The higher modes are less important in the deep. They are significant only in the shallows where a large component of them are going to be locally generated by the interaction of the primary constituents with the bed and coastline rather than remotely forced at deep water boundaries.

Yes, the referee is correct, it is important to improve the tidal forcing of AMM15, in terms of tidal constituents and atlases. Research activities are ongoing to validate the impact of using a different model, FES2014, with many more tidal constituents.

The updated version of Table 2 is:

Forcing	AMM7	AMM15
Surface forcing	Met Office Global Unified Model (MetUM) Atmospheric model NWP analysis and forecast fields, calculated in the MetUM using COARE4 bulk formulae (Fairall et al. 2003).	ECMWF Integrated Forecasting System (IFS)-Atmospheric Model High Resolution (HRES) operational NWP forecast fields using CORE bulk formulae (Large and Yeager 2009)
Surface forcing resolution	Horizontal grid: ~10 km (2560 x 1920 grid points) Frequency: 3 hourly mean fluxes of long and short wave radiation, moisture, 3 hourly mean air surface temperature but hourly 10m winds and surface pressure	Horizontal grid: ~14 km (0.125°x0.125°). Frequency: 3 hourly instantaneous 2m dew point temperature, surface pressure, mean sea level pressure, and 2m air temperature. 3 hourly accumulated surface thermal and solar radiation, total precipitation, and total snow fall.
River run-off	Daily climatology of gauge data averaged for 1950–2005. Climatology of daily discharge data for 279 rivers from the Global River Discharge Data Base (Vörösmarty et al., 2000) and from data prepared by the Centre for Ecology and Hydrology as used by (Young and Holt, 2007).	Daily climatology of gauge data averaged for 1980–2014. UK data were processed from raw data provided by the Environment Agency, the Scottish Environment Protection Agency, the Rivers Agency (Northern Ireland), and the National River Flow Archive (personal communication by Sonja M. van Leeuwen, CEFAS, 2016). For major rivers that were missing from this data set (e.g. along the French and Norwegian coast), data have been provided by the same climatology used by AMM7 (Vörösmarty et al., 2000 and Young and Holt, 2007).
Tidal constituents	M2, S2, N2, K2, K1, O1, P1, Q1, M4, MS4, L2, T2, S1, 2N2, MU2, NU2 (15) from a tidal model of the North-East Atlantic (Flather, 1981).	M2, S2, N2, K2, K1, O1, P1, Q1, M4, MS4, MN4 (11) from Topex Poseidon cross-over solution (Egbert and Erofeeva, 2002; TPX07.2, Atlantic Ocean 2011-ATLAS).
Lateral boundaries	Met Office FOAM North Atlantic (1/12°; 6 hourly fields) and CMEMS Baltic Sea (2km, 1 hourly fields). AMM7 and AMM15 have Atlantic and Baltic boundaries in a different geographical location.	

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 decision?

Yes, we agree with the referee that we need to run impact studies to assess the impact of the different resolution in the atmospheric forcing. These studies were not part of the pre-operational implementation, therefore are not discussed in this paper. These experiments are carried out in the frame of a UK project. As soon as available these results will be shared with CMEMS. If the impact of the coarser spatial resolution is significant we could switch to the Met UM forcing in a future evolution of the operational system.

The ECMWF-IFS analyses are not used, we use only 3-hr forecast fields. We are planning to switch to hourly ECMWF-IFS products by the end of 2020.

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.

This first implementation of assimilation in the high resolution AMM15 followed the same scheme as used in the AMM7. Although the deep water areas of AMM15 are more limited than AMM7, and so the benefit of assimilating SLA may be small, this is a milestone on the way to assimilating SLA throughout the domain.

However, the assimilation of SLA observations in the deeper water still allows us to constrain the temperature and salinity in those regions, which then provides a better boundary condition for the shallow regions. As discussed in King et al. 2018, the assimilation of altimeter observations and T/S profiles is complementary, and in regions such as the NWS where profile observations are relatively limited, altimeter observations provide a valuable additional constrain on the density structure of the deep water regions. Without additional experiments, the contribution of altimeter observations is difficult to quantify.

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

Thanks, we swapped table 3 and table 4 and the corresponding cross-references.

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 (GHRSS):”

Thanks for this comment. We modified changed Table 4 taking into consideration this comment. The column “Data source” is now “Data source/Products”:

Type	Fields	Platforms/Satellite	Data source/Product
IN SITU	SST Temperature and salinity profiles	<ul style="list-style-type: none"> • Ships • Drifters • Fixed moored arrays • Gliders • XBTs • CTDs • ARGO • Ferry boxes • Recopesca buoys* • Thermosalinograph 	GTS; http://www.wmo.int/pages/prog/www/TEM/GTS http://marine.copernicus.eu/ INSITU_GLO_NRT_OBSERVATIONS_013_030
SATELLITE	SLA Along Track* <i>*along with the corrections necessary for the use with a wind and pressure forced, tidal coastal model. SLA are assimilated only in deep regions (> 700m).</i>	<ul style="list-style-type: none"> • Cryosat-2 • AltiKa • Jason 3 • Sentinel 3a 	http://marine.copernicus.eu/ SEALEVEL_EUR_PHY_ASSIM_L3_NRT_OBSERVATIONS_008_043
	SST L2p/L3c	<ul style="list-style-type: none"> • NOAA-AVHRR • MetOp-AVHRR • SEVIRI • VIIRSG • AMSR2 	Group for High-Resolution Sea Surface Temperature (GHRSS) www.ghrsst.org

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?).

We store all the information on the quality check in a set of files called “feedback files” (each for each type of observations: sub-surface profiles of temperature and salinity; SLA; SST). We are working with the CMEMS Product Quality Cross-Cutting working group to identify a CMEMS standard for conveying this information to the data producers.

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?

The two models providing the boundaries are not consistent and this is the reason why we don’t force the Baltic Boundary with SSH (The Baltic model is not constrained by data assimilation while the North Atlantic is). The Atlantic model providing the boundaries and AMM are both constrained by the assimilation of the same SLA data, this should avoid major discrepancies at the Atlantic boundary. The forcing at the boundaries is an active research topic for our model and hopefully in the future we can improve the parametrisation we are using now even if the problem is more

worrying when producing reanalysis or climate simulations spanning over a much longer number of years than a short-term forecast system.

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?

These operational systems are running on the Met Office HPC – Cray XC40 super computer. The following table describes the number of nodes and processors used by each component:

System	Component	# of nodes	# of processors
AMM7	NEMO	8	256
	XIOS	--	--
	NEMOVAR	2	64
AMM15	NEMO	48	1536
	XIOS	8	256
	NEMOVAR	48	1536

XIOS is for the NEMO I/O. The small size of AMM7 model grid doesn't require dedicated nodes for this task.

We added this information to the manuscript.

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.

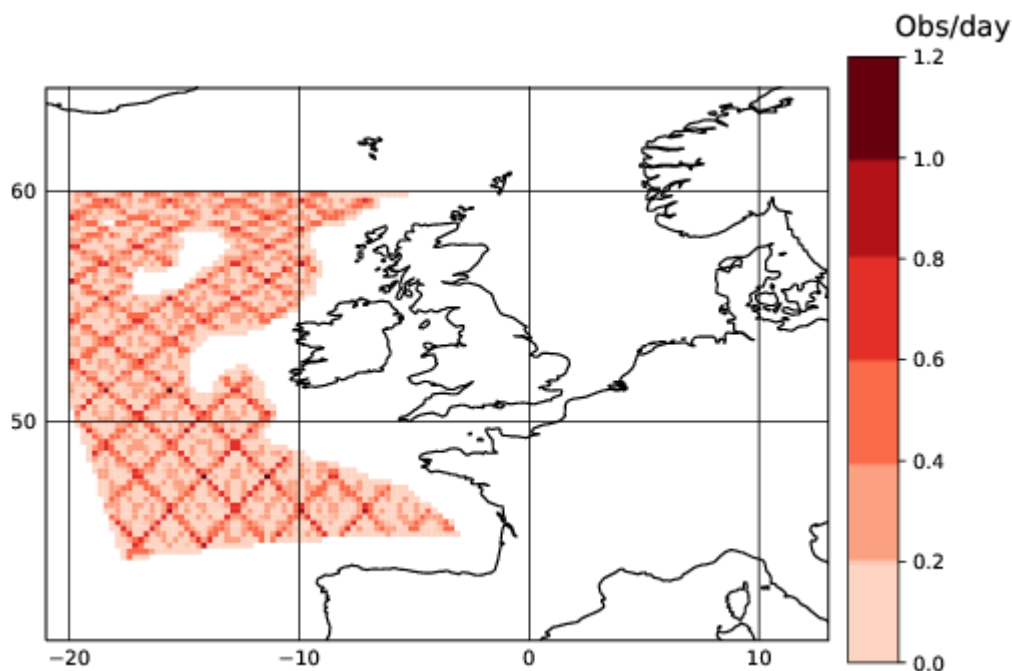
Yes, the figure has been updated, please see comment P2.24

We use only the forecast fields from ECMWF-IFS, due to the coarse time resolution of the analysis (6-hr), as specified in table 2 and at P6.2 of the manuscript: *"The IFS analysis is available only at a low temporal resolution (6 hours) therefore the decision was made to force the system using forecast fields only (3 hourly), from the 00:00 UTC forecast base time"*.

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.

Thanks for the comment, the new version of the Figure 3 shows now only the assimilated SLA obs:



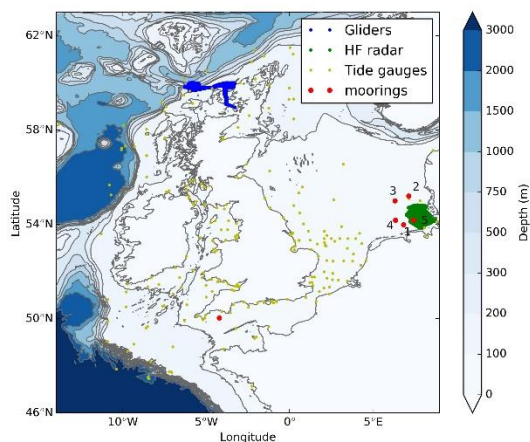
The authors should consider the possibility to include in this section the analysis of the differences between the dynamical patterns modeled by the 2 different model systems, currently included in the Validation Result section. This point is suggested below.

Yes, thanks for the comment. We have taken this suggestion into consideration as described in detail at the beginning in the general comments.

4.Validation of the experiments

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

Thanks, Figure 4 has been updated with the location of all the tide gauges used for the tidal validation (yellow dots).



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 finds 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.

Thanks for pointing out this inconsistency. Yes, it's correct, the tide gauges data are not assimilated. We have updated Figure 4 adding the tide gauge location (see the figure in the comment above).

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.

Thanks for the comment, it's not appropriate calling "buoys" the tide gauges data. We have corrected the manuscript as follow:

"The number of tide gauges taken into consideration for AMM15 and AMM7 is the same, therefore the coastal data,"

The tide gauges data are from BODC

(https://www.bodc.ac.uk/data/hosted_data_systems/sea_level/uk_tide_gauge_network/) and from

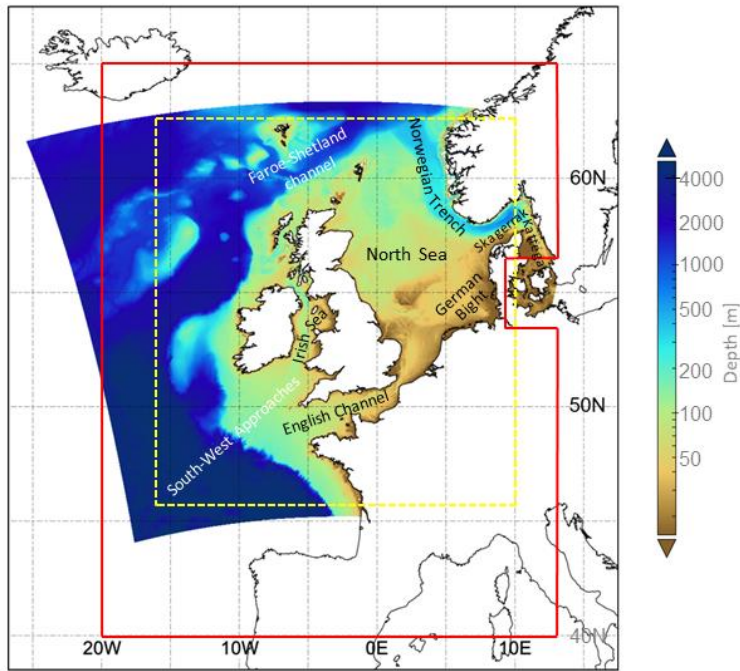
the North West Shelf Operational Oceanographic Service, NOOS, data portal

(<http://noos.eurogoos.eu/>). All the data are from tide gauges.

We have included these two web sites into the manuscript.

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.

Thank for the comment, done



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.

We decided to focus our high resolution model validation on small areas and short time period, for both the glider and the HR radar observations. Both observations could be available for longer period and in different areas, but we have decided to stay focus on a short period to understand the impact of the high resolution and what is an adequate protocol to assess the quality of a high resolution model. As written in the manuscript at P1211 “...have the benefit of providing an understanding of the impact of the high resolution locally on small area and short time scales”. This validation will be extended to assess the future evolution of the AMM15 system, since we proved it’s useful and complementary to the standard validation protocol.

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.

Thanks for this comment, we added the requested information in the manuscript,

“One month, March 2017, of HF radar surface current velocity data were used to compare AMM7 and AMM15 in the German Bight where the bathymetry is shallow (Figure 4) and AMM15 is expected to performed better. The total surface velocity data from the COSYNA (Coastal Observing System for Northern and Arctic Seas) observing network (Gurgel et al., 2011), available through the EMODnet Physics data portal, are computed from radials of three HF radars installed on the islands of Sylt and

Wangerooge, and in Büsum (as shown on Figure 6). Data are averaged every 20 minutes on a grid of resolution of ~3 km. At the operating frequencies used, the total surface velocities represent an integrated velocity over a depth between 1 and 2 m. Relative error provided with the dataset was used to keep only data with error smaller than 15%. Model output were interpolated at the time and locations when and where observations were available to avoid applying gap-filling technics. Temporal coverage over the domain is larger than 75% everywhere except along the base line between Büsum and Wangerooge where the temporal coverage is ~29%."

Yes, the three names correspond to the HF radar sites, thanks for this comment. The caption of the picture includes now this information.

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.

Thanks for this comment, we have corrected the manuscript as follow:

*"The in-situ measurements are from different instruments, as detailed in **Error! Reference source not found.**"*

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

Thanks for the comment. We have modified the manuscript as follow, including a reference and enhancing the explanation:

"A Butterworth filter (Butterworth, 1930) has been applied to the hourly model and observed SST data, using a cut-off for the filter at 5 days which removes the large scale synoptic and seasonal signals, leaving the internal dynamics and the wind driven signals, as well as the tidal frequencies".

S. Butterworth, "On the Theory of Filter Amplifiers," Experimental Wireless and the Wireless Engineer, Vol. 7, 1930, pp. 536-541.

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.

Thanks for the comment. Since SST has a very good satellite data coverage, it's not easy to exclude one single mooring from the set of observation assimilated and consider that observation completely independent. This is a typical dilemma while running an operational system. We are trying to assimilate all the available data to improve the quality of our products, but this implies that we reduce significantly the number of independent observations available for the validation. Due to our choice, we don't have a time series for a non -assimilated mooring.

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).

The power spectra shown in Figure 9 is for the FINO 3 buoy (number 2 in Figure 4). The buoy is in the German bight, where the bathymetry is shallow (~20m). The 12h energy peak overestimation is remarkable in SON (wrongly marked as DJF in the manuscript, now corrected), at the end of the summer when probably the two models have different stratifications. The water column is moved by the tides (M2 in the predominant tide) and this could bring to differences in the SST variability. The stratification in this area could also be enhanced by the fresh water contribution of two major rivers, Elbe and Weser. This hypothesis is supported also from the analysis of the map of SST gradients (not shown in the paper) where AMM15 shows stronger gradients than AMM7. Further studies are needed to understand better the SST variability in AMM5.

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.

Thanks for the comment. E-Hype is the hydrological model for the European areas developed by the Swedish Meteorological and Hydrological Institute (SMHI), <http://hypeweb.smhi.se/explore-water/geographical-domains/#europehype> We have included this information in the manuscript at P23.1 but not in the section where forcing are described because neither AMM15 nor AMM7 (in the trial run described in this study) are forced by E-Hype in the experiments described in this study. AMM7 was forced by E-Hype in a version in operation before April 2017.

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).

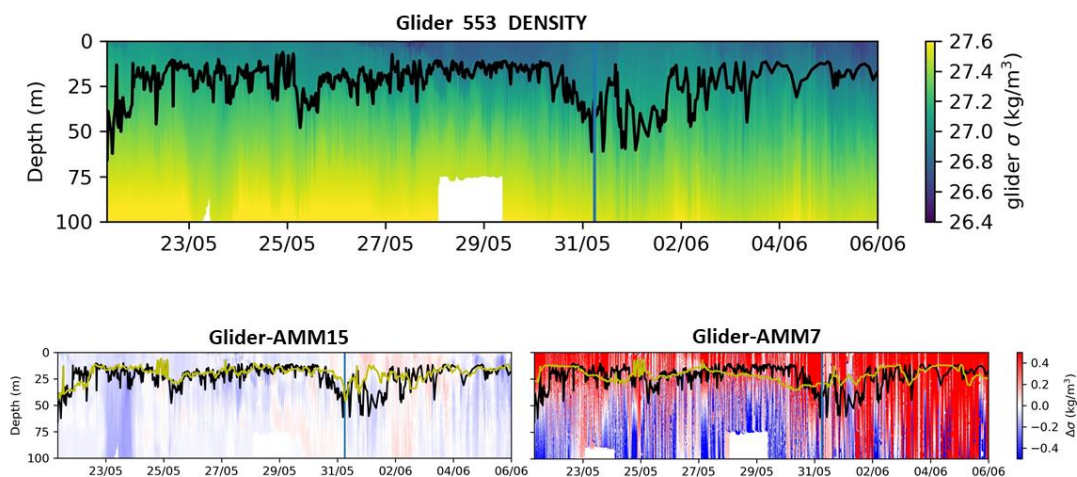
Thanks, done.

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.

Thanks for the comment, we have added the model MLD to these figures (Yellow line for AMM15 and AMM7 respectively. The black line represents the MLD from the observations).

We added this information in the manuscript, P27.17:

"...with AMM15 and AMM7 in the corresponding locations (yellow line in Figure 15)" and in the caption of the figures.



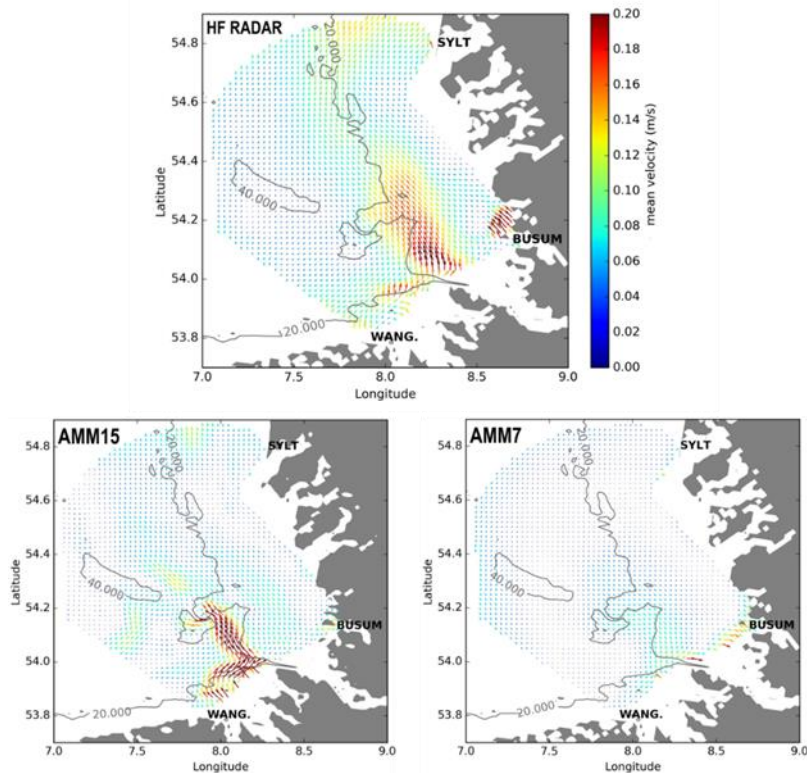
Section 4.5 is devoted to show some results from currents compared with HF-Radar data. As in the previous case for the tides, only a month of data (March 2017) is shown. Please, justify why a so short temporal coverage for the validation.

Please see answer to comment P16.5

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.

Thanks to this comment we realised that we used in the manuscript the map of velocities before the cleaning of data instead of after. The reviewer is right, we can't ensure that the high currents depicted at the border of the HF radar coverage are reliable. We substituted Figure 16 with the corrected Figure.



We have also added the following information in the manuscript:

“The HF radar surface currents were also used to investigate the sub-tidal circulation in the German Bight. The strong tidal signal in the shallow German Bight results in Kelvin waves propagating eastward on the southern boundary along Germany and northward at the eastern boundary along Denmark. However, this cyclonic circulation may not dominate as other processes are also influencing the circulation such as topographic effects from the shallow basin, wind and stratification resulting from freshwater input mostly from the Elbe and Weser river discharge. Wind tends to also produce a residual cyclonic circulation (Schrum 1997, Dick et al 2001, Port et al 2011). During the month of March 2017, a weak cyclonic circulation was observed in the mean HF radar surface currents along the German and Danish coasts (Figure 6). It is also observed in the AMM15 simulations and as a weaker flow in AMM7. The strong flow out of the Elbe estuary is evident in AMM15 currents pattern, even if shifted to the west. AMM7 shows an intensification of its currents in this area, but with a speed much smaller than the observations and AMM15 (Figure 6). Generally....”

Yes, we added the following references:

Dick S, Eckard K, Müller-Navarra S, Klein H, Komo H: The operational circulation model of BSH (BSHcmod)— model description and validation. Berichte des Bundesamtes für Seeschifffahrt und Hydrographie (BSH) 29, BSH, 2001.

Port, A., Gurgel, K. W., Staneva, J., Schulz-Stellenfleth, J., & Stanev, E. V.: Tidal and wind-driven surface currents in the German Bight: HFR observations versus model simulations. *Ocean Dynamics*, 61(10), 1567-1585, 2011.

Schrum C: Thermohaline stratification and instabilities at tidal mixing fronts. Results of an eddy resolving model for the German Bight. Cont Shelf Res 17(6):689–716, 1997.

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.

Done, as described at the very beginning of this document.

5. Conclusions

P30.7 typo: temporal

Corrected

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).

Thanks for this comment. We added in the manuscript the following paragraph:

“The assimilation scheme used in AMM15 is broadly unchanged from that used in AMM7. While the short correlation length-scale is now ~5km (compared to ~20km), the observation and background error covariances, and the observation types assimilated, remain unchanged. In this initial implementation of AMM15 we have not attempted to improve the use of observations in the assimilation scheme. We are currently investigating how to adapt our assimilation scheme to assimilate SLA observations in stratified water and will be re-estimating the observation and background error covariances for this new higher resolution system.”

P31.18 The AMM15 ocean (system?).

Corrected