Dear Madam, Sir,

We thank 2 anonymous reviewers for the precise and helpful comments about our manuscript. We realize that a lot of time went into the reviewing process, giving the precise (and justified) remarks. We answered each of the comments, and modified the text and the figures accordingly. The revised version of the manuscript is much clearer, some mistakes was corrected, some new references were added, and some more links to CMEMS products were given, all according to the reviewer's comments.

In the replies to the reviewer #1, we announced to add a new annex with details of the data assimilation method. After completing also the correction of the manuscript for reviewer #2, we think that these details should be provided directly in the concerned section (instead of creating an annex). This is the only change in the reviewed paper, inconsistent with what we replied in the OSD forum. It is written in red in the text below.

We think the revised paper is much better than the original one, and submit it to you for consideration for publication in OS.

The replies to the reviewers were posted in the OSD discussion. They are copied below for your convience.

Reviewer 1
Specific comments:
1) My main concern is that "upscaling" appears to be feasible only if a similar setup is made between the child and the parent models. For instance, the authors use two models based on the same platform, i.e. NEMO, with an exact ration between horizontal grids and I suspect (not written in the text) with an identical vertical grid. All these are OK coinciding with the options to emulate two-way nested simulation. However, within a DA framework one would expect to see more general options, for instance, assimilating pseudo-observations on an entirely different grid (especially vertical for the T, S). The latter would support a more general argument for "upscaling" approaches, using for instance a different setup/grid/platform for the nested model. I leave it up to the authors choice if they wish to perform a DA experiment with a slightly different projection of the pseudo observations. However, I find useful the authors to discuss the limitations of their method.
The reviewer is entirely right that the configuration used to test the upscaling method is based on a nested grid setup using the same model code (Nemo) for both the parent and child grid; and furthermore the vertical grid is also identical (only the horizontal grid is different). This may influence the conclusion compared to a configuration with 2 different model codes. However we think that it is not a fundamental limit of the method a) Concerning the vertical grid, in the "normal" case of assimilating real observations, the latter are on a different grid than the model. Similarly if the child model was on a different vertical grid than the parent model, it would still contain useful information, and be worth to be assimilated in the parent model. What may happen however, is that some observations could be lost (e.g. the lowest model of the child model could be out-of-grid in the parent model)
b) if different model codes are used, the models could represent different processes. Hence, this should be taken into account by modifying the (representativity part of the) observation error covariance matrix. Examples of contributions to the representativity error could be
- different vertical coordinates
- different representations of the surface: rigid lid, free surface (with a linear or non-linear representation e.g. in Nemo)
- hydrostatic model, or not

2) page 2, line 16: "By upscaling the child model into the parent, the latter is brought closer to the former.". The benefits for the child model are obvious, though not so obvious for the parent model. Can the authors provide some guidance for "safe upscaling"? The way this work is constructed, suggests that a forecasting center should only "upscale" in case the child model has a similar modelling setup with the parent, e.g. same platform, vertical discretization, physics, parametrizations etc. The authors should also provide more information in the text about the setup of both models, in order to highlight their differences.

If one considers that the child model is better in its domain than the parent model (e.g. by comparison with real observations), then it would be desirable to upscale it into the parent model. This would be the case is some processes are dependent on resolution, in straits, etc; and is closely linked to the first specific comment in the review. We provide now a table in the annex of the paper giving details about the setup of both models; but upscaling should not be limited to identical parent and child models (see answer to comment 1)

3) page 4, line 23: "these pseudo-observations coming from the nested model are considered independent". This is a very strong assumption, since observations are on C2 OSD Interactive comment Printer-friendly version Discussion paper horizontal resolution 1/80◦ . In DA a common practice to avoid correlated errors is thinning or superobbing. Can the authors justify their option not to apply these techniques?

The reviewer is correct, that the assumption of spatially independant pseudo-observations is very strong. We are actually working on a non-diagonal observation error covariance matrix, but this is a large work that would not fit into the current paper. However, the assumption is partly alleviated by increasing the (diagonal) part of the matrix, in order to compensate for the (missing) non-diagonal elements. Increasing the diagonal elements in the matrix by a factor 3, for example, is similar to thinning observations with a factor 3. This is now stated in the paper.

4) page 4, lines 21-22: "Ensemble Kalman filter" and page 5, line 11: "Ensemble Tranform Kalman Filter variant of the EnKF". Use also in page 4 the word "variant". In addition, the authors should write in this section the DA method in more details. For instance, it should be mentioned that this is a deterministic approach of the EnKF, i.e. pseudo-observations from the child model are not perturbed and the perturbation approach is only applied in the parent to obtain model errors. All these are not apparent to the reader, at least not before start reading the results section.

We added an annex to the paper with the details of the data assimilation filter used in the study.
In the final revised version of the paper, we decided not to enter this information in an annex, but directly in the paper, as suggested by the reviewer.

5) page 6, lines 13-14: "to update directly the tiles from the Mediterranean model restart files, influenced by the nested model, without including the other tiles in the state vector". This is an interesting technical capability of OAK, but if not mistaken that means that there is a crude correction cutoff in the neighboring tiles just outside the nested domain. I would assume that the localization is

enough to constrain the correction in an area slightly broader than the nested domain. Can the authors clarify what is the purpose of this capability?

In the state vector of the parallelized parent model, we include the tiles covered by the nested model, but also the tiles immediately around that area. Therefore, as the reviewer correctly assumes, the correction is not cut off at the margins of the area covered by the nested model, but propagates outside. The extent of the correction outside the area, depends on the radius used in the localization method. This is now better explained in the paper.

6) page 9, lines 7-8: "The ability to ... would be beneficial to constraint the model". This is more a concluding remark, rather than a result of the study. The phrase should be moved in the Conclusions section 5.

We moved the remark to the conclusions

7) Figure 5. The SST is L4 or L3? In section 3.4 it is mentioned as L3.

In the study (section 3.4), L3 satellite images are used. Only in figure 5 is the L4 image used for visual comparison of model and satellite image. This is now clarified in the article.

8) Figure 8. The units are missing from the axes.

added units

9) whole page 16: "Advantages of using upscaling include ...". This is a nice summary of "upscaling" advantages supporting the method. Can the authors provide possible disadvantages (if there are any) and suggest possible remedies?

The reviewer is right that the list of advantages should be accompagnied with a list of possible limitations (or disadvantages). This is now included in the article, and copied here:

a) the child model should be "better" than the parent model

b) exactly as when assimilating real observations, the data assimilation procedure itself uses approximations, and this could degrade the analysis

c) if the parent and child models are very different, the parent model could not manage to ingest the pseudo-observations

d) the coupling with upscaling is not as strong as with real two-way nesting

Potential remedies for limitations b and c would be

b) see all the research about this problem (in the context of assimilation of real observations), such as anamorphosis techniques (when a non-linear relation exists between model variables and observations), particle filters (when the error distribution cannot be considered Gaussian), etc

c) the observation error covariance matrix should be specified carefully to filter out the processes of the child model, that cannot be represented in the parent model

Reviewer 2

Abstract

Line 5: *Therfore* instead of therefor and I would take out "*in practice*"
Changed to Therefore
Line 6: …" to replace the missing model feedback…" I would insert "*child model or high resolution model*". Done
Line 10: I suggest to rephrase something like:
"A basin scale model simulation is compared to one simulation…, and another model analysis which applies the upscaling technique…"
Changed

Introduction

Line 15: "reanalyses, analyses and forecasts" Changed
Line 16: " …by different institutes within the regional monitoring and forecasting centers.." Changed
Line 24: could you insert a reference for this? We rephrased the sentence instead
Line 6 page 2: I would take out "in this artice" Removed

Line 12 page 2: I would substitute *basin-scale* with *regional* Changed

Line 13 page 2: "…in the basin-scale model, … is to obtain" Changed
Line 14 page 2: "(along with…)" Changed
Please rephrase the entire sentence, it seems too informal to me Rephrased
Line 16 page 2: I suggest "…*to the child model will progressively gain consistency with the child model solution within its domain, being beneficial for the child model over time*." Changed
Line 20 page 2: do you have references for this? Added Mason et al 2010, Debreu et al 2012
Line 21 page 2: Is this pertinent? I do not see the connection, please explain. If we assimilate pseudo-observations coming from a nested model, and this improves our (parent) model, then the child model is a pseudo-measurement device which can be seen as a replacement for (costly) real measurement devices.
Line 26 page 2: Other re-initialization techniques have been used blending, through optimal interpolation, coarse resolution operational analyses and coastal observations in so called Rapid Environmental Assessment experiments. Please give a look at *Simoncelli et al.(2011),* they show improvements in the nested coastal model performance using observations. Added
Line 35 page 2: I am not sure that the syntax is correct please check the English. Changed
Line 3 page 3: "Therefore acknowledging that operational…, Schulz-Stellenfleth and Stanev (2016) strongly…" Changed

2.1 Hydrodynamic Model

Line 13: I would substitute *tried out* with *has been implemented* Changed
Line 15: *"created by the junction of the Eastern and Western Corsican Currents"* I would cite *Pinardi et al (2015)* Added
Line 17: I would add some recent references *Pinardi et al (2015), Somot et al. (2016), Simoncelli and Pinardi in von Schuckman et al. (2018).* Added
Line 21: The resolution of MFS is $1/24^{th}$ (Clementi et al. 2017, https://doi.org/10.25423/cmcc/medsea_analysis_forecast_phy_006_013) of a degree since October 2017. The reanalysis (https://doi.org/10.25423/medsea_reanalysis_phys_006_004 Simoncelli et al., 2014, 2016) is still at $1/16^{th}$. Changed

Line 30: please specify for reproducibility issues which analyses has been used, I guess the 1/16? Or the reanalyses? Please clarify and insert the reference. Added

Line 32: ERA Interim is not at a resolution of $1/8^{th}$ of degree, it is 0.75 degrees!!! You might have re interpolated it from $1/8^{th}$ to $1/16^{th}$. Please explain it and add the Dee et al () reference of ERA Interim. Modified and added reference

Line 1 page 4: Which literature? Removed « from the litterature »
Line 2: 5 rivers' data comes from? Various regional websites, not added in the article
Line 4: Please consider the CMEMS has 39 rivers, thus it is not very much coherent. Please look at the CMEMS products descriptions and may be cite the reanalysis instead. Rivers are described in the Simoncelli et al 2016. Added in the text that CMEMS has many more rivers

Figure 1: Please increase the axis font, not readable now. I would show the two models' salinity fields to emphasize the differences due to the daily river outflow, instead of the difference. In fact, you describe the different plumes in the manuscript.
We tried this out, but due tu large salinity variability, a side-by-side plot of salinity does not show clearly the salinity difference. Therefore we chose to show directly the difference

Line 10: please describe more in detail Fig1b eventually, isn't it the difference among the two models' salinity after 1 month of simulation? Please improve also the caption.
Yes Fig 1b is the difference of (nested) model salinity after 1 month, when using climatological Rhone discharge or real, daily discharge. This is now written more clearly in the caption

2.2
Line 16: I would re-phrase something like "*In order to assimilate …,* different set ups could be implemented (adopted, applied) depending on …" Changed

Line 21: please improve the description of the settings, it contains repetitions.
Line 23-25: Start a new phrase please and please say something more about the statement that the T and S pseudo-obs are considered independent. You mean that you assume that even if it is not the case. You assimilate the full resolution 3D T and S fields? None thinning? Please motivate a bit this part. The word *also* could be neglected and substitute ";". It looks like you wrote this in a rush without much care. Improved lines 21-25 as suggested
As the reviewer correctly supposed, we do perform thinning (in the horizontal), and this is now also written in the text. We apologize for forgetting to write it in the article first submitted

Line 31: random not randon Changed

Line 7 Page 5: Why did you select 1 month of spin up time? Please start a new phrase and integrate a bit on that. What you use in the evaluation is the ensemble mean of the 100 members?
In a previous (single) model spin-up, we noticed that the kinetic energy reaches a more-or-less stable value in just a few days, therefor a spin-up of 1 month was considered sufficient. The same is supposed for the ensemble members. As we are now running an ensemble of 100 members, longer spin-up would translate into large computational cost.

Line 1 Page 6: it is not clear to me "*...and its addition observation localization*", could you please explain?

It is explained just afterwards (lines 2-3. In means that the model domain is cut into subdomains (water columns in our case) where the analysis is performed separately. Moreover, in each subdomain, only relevant observations are considered (i.e. observations that are far away and will have no impact, are not considered during the computation).

Line 7: many are the sensitivity experiments with different observation errors, maybe you could insert a table.
We agree that there are 4 experiments (each with a value for temperature and salinity observation errors), but these all fit into 1 line in the text. Maybe it is easier to keep them in the text ?

3. Metrics

Line 17: Please add something about: .. these metrics have been computed to compare model solutions. What do you want to show? Upscaling model solution (is it an ensemble mean of the 100 members?) with the basin scale simulation?
Added some explation to Line 17. About the « upscaled » model, we keep one unperturbed ensemble member. The other 99 members of the ensemble are perturbed and are used only to create the model error space. All 100 members are updated daily by the data assimilation procedure.

3.1
Line 23: see previous comment. What do you want to show? Does this metric tell you if the upscaling procedure is driving the upscaled model solution towards the NW-Med? Please integrate a bit.
Yes exactly, we want to see if the upscaling procedure pushes the parent model towards the NW-Med model. This is now explained better at line 23.

3.2
Line 28: again, see previous comments. I would integrate at the beginning of the section, and in each sub-section I would state what exactly is expected from each metrics.

3.3
I do not really understand what you are going to present in the results. Plume length/direction from the free and upscaled models? The comparison model-obs using sat chlla images? "Furthermore,…" What does it mean you do it qualitatively, quantitatively, both?
Yes we compare plume length and directions in between models, and also between model and real observations. This is different compared to the previous metrics where we did not have real measurements. This is now better explained in the text.
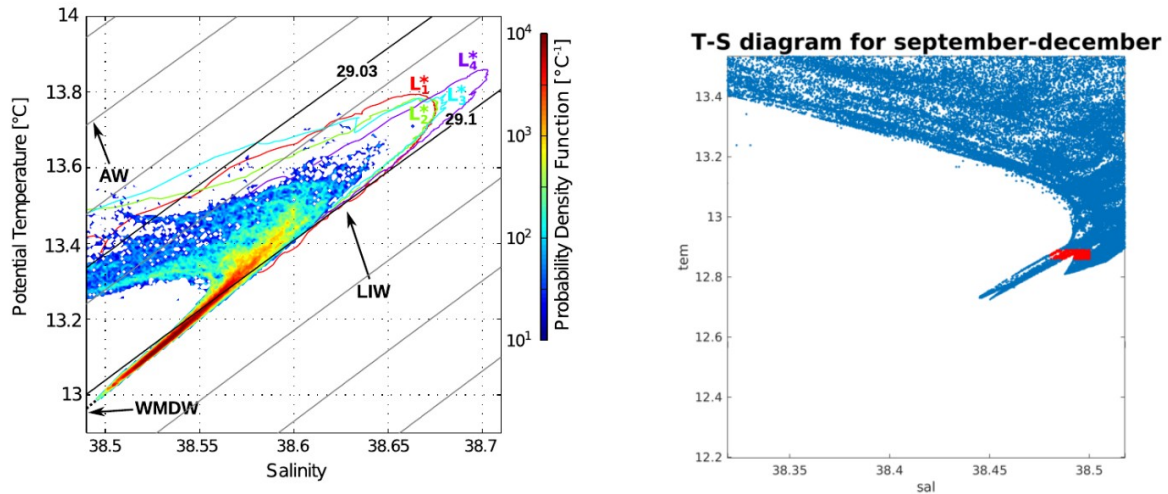
3.4
ok

3.5
Line 15: What do you mean bay "depth is larger than 1000m"? Is it the depth reached by deep convection? Maybe you want to put it at the end of the first phrase at line 13. Here you are describing the WMDW characteristics.
Following Bosse at al (2015), we consider that water is WMDW if it is within a certain T and S range, AND if it is situated below 1000m depth

Line 17: Which tail? How could the river see that? Please explain or describe with some detail.\
Bosse et al (2015) show the T-S diagram and indicated the water masses (left plot below). We reproduce this plot from model results (right plot below) and obtain a longer tail (it goes to ~12.7°C and slightly lower than 38.45 psu).The definition from Bosse at al (2015) would result only in the red rectangle in the right plot below. Therefore we slightly adapt the WMDW definition in order to capture the whole tail. We decided not to include these 2 figures in order to keep the paper more consise.

4. Results

Figure 3: please enlarge the font. Changed
Line 5: The difference of what? Temperature, added in the text and in figure caption
Line 9: Isn't it Fig.1b the difference between unperturbed parent and child model in salinity after 1 month of spin up? Fig 1b is the salinity difference of the nested model, when forced with climatological or with real Rhone river discharge data. Figure 3 is the difference between parent and nested model (for temperature), and the equivalent for salinity is not shown in the article
Line 10: Why should I trust the child model more than the parent? You did not provide any model performance. It is not easy to improve significantly the smooth solution of the coarse parent model (phase errors are common).
Also following recommendation from reviewer 1, we now stated clearly that we *suppose* from the start that the nested model is better (in some sense) than the parent. We do not try to prove the hypothesis is actually valid, and sometimes, it could be wrong. Exactly as the reviewer points out, when the nested model represents small scales that are actually out of phase with reality, it could have higher RMS errors than the parent model that does not represent small scales at all.

But supposing that the nested model is « better », our objective with upscaling is to bring the parent model closer to the nested model (i.e. emulate nesting feedback). This is now written in the introduction.

Figure 4: please include a line to indicate the section location on the map. Done

Line 16: The scope of Figure 5 is to show that the upscaling technique is bringing the upscaled model close to the NW-Med one. RMS difference could be provided towards the satellite SST to show this. However, this rise the question: " why don't you consider to weight the assimilation of pseudo observations according to the misfit with the observed SST, giving more weight where the pseudo-obs are loser to sat obs? Please keep the same notation to call the models (MED, NW-Med…).
This is an interesting suggestion but has also the following limitation. If we have many observations to validate the nested model, then we could as well assimilate these (real) observations into the parent model. Upscaling is interesting mostly when there are few or no real observations.
In the case of SST however, it is a very interesting suggestion. When the nested model is well-validated by SST, maybe we can trust it more also for other variables and assimilate those in the parent model to complement the (real) SST observations. This could be the topic of a follow-up study.

Please switch 4.1 and 4.2 (thus fig. 6 and 7) to be consistent with 3.1 and 3.2. Done, we rather switched 3.1 and 3.2

4.1

Line 21: For which models? Parent model, free and upscaled. Added in the text

Line 2 Page 9: Why the NW-Med model is not shown in Figure 6? Please include it for consistency in presenting the results, since you want to prove that the upscaled model is driven towards the NW-Med solution. Again please keep the same nomenclature to facilitate the reader. Moreover RMSD could be computed as in 4.2 case.
Please see the answer to comment 4.2 (Fig. 7)

Lines 9-12: What is this meaning? Is this correct? Is this coherent with NW-Med?
Yes it can be seen from the figure that both lines are closer one-to-the-other after the first month. There is also a temporal delay in between the 2 curves in August-September. The text merely describes this.

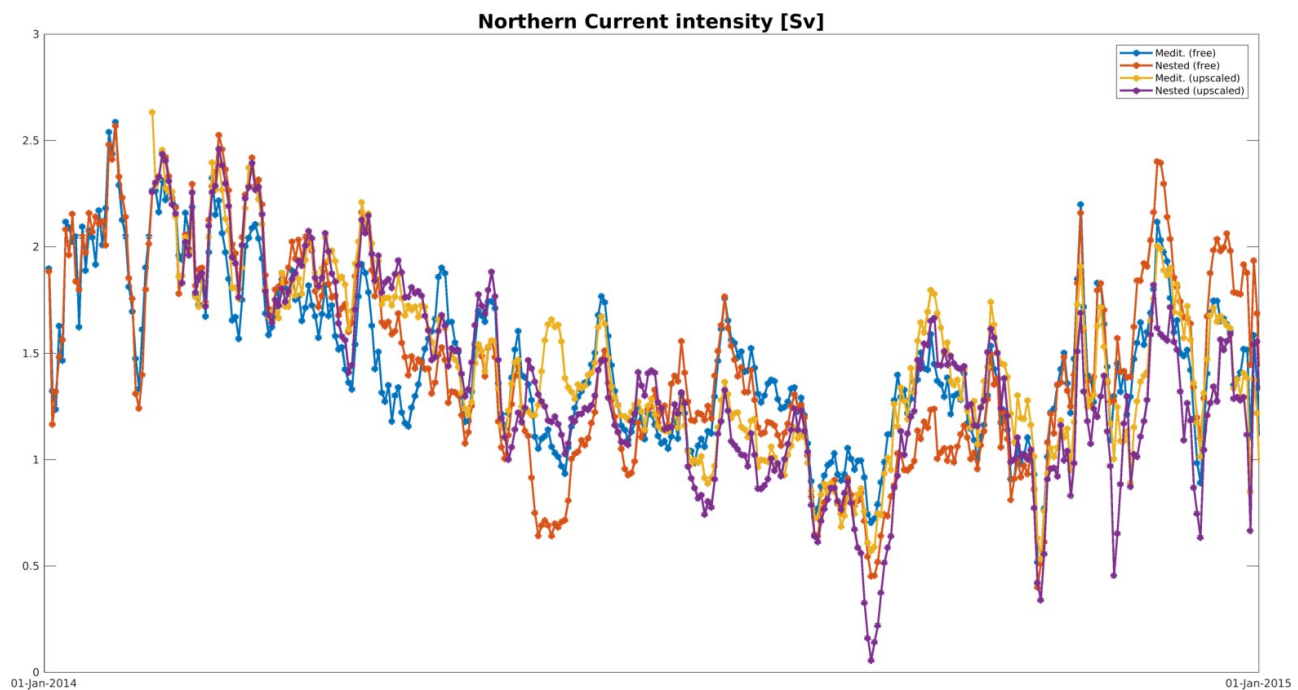Figure 6: please increase the font in the legend and include NW-Med.
We will modify the figure fonts

4.2

Figure 7: Please insert NW-Med and the legend.
Once upscaling modifies the parent model, the child model gets different open-sea boundary conditions from the parent, and starts to be different as well. Thus, for Figures 6 and 7, we could add 2 NW-Med curves to the figures (one corresponding to the free run and one corresponding to the upscaling run). For the Northern Current intensity, it gives this :

Northern Current intensity [Sv]

This figure becomes difficult to interpret. There is no general trend, like for example a systematically more intense Northern Current in nested models. However, what's interesting is that generally speaking, compared to the free Med model (blue curve), the yellow curve seems to get closer to the purple (child model of the upscaled Med model). The discrepancy between yellow and purple is smaller than the discrepancy between blue and red. This can be verified by looking at RMSD errors between parent and child model in the two cases (free model or upscaled model), respectively 0.22 and 0.19 Sv. The RMSD for all metrics have now been included in a table as suggested by the reviewer.

Lines 4 -6Page 11: I suggest to rephrase *"…this metrics cannot **be** used to compare and validate the models since observations are not available to compute the real NC transport. However, it shows that upscaling of.."*
But the metric *can* be used to inter-compare models, so it might be confusing to say « this metric cannot be used to compare … ». Therefore we rephrased as :
« For the purpose of our study, this metric cannot be used to validate the model since real measurements of the Northern Current transport are not available; but (as for the previous metric), it can be used to compare models, … »

4.3

As for the other metrics I would include the NW-Med to show its consistency with the upscaled model. At Line 15 you mention the nested (NW-Med) model but it is not shown. RMSD could be computed again as more robust argument of your results.
On Fig. 8, the arrows for the upscaled and (not-shown) NW-MED models superpose and are indistinguishable
Line 17: a quantitative estimate of model performance can be provided again by RMSD of MED, and upscaled model towards NW-Med.
We have computed the RMSD and included it in a table

Figure 9: This figure presents MED and NW-Med, why not the upscaled model? I suggest to shaw the three salinity fields.
It is indistinguishable from the nested model

Line 3 Page 12: Is the increase in salinity observed, is it consistent with observations or only with NW-Med? I would include validation with observations, since a lot of them should be available by REP in situ observations fromCMEMS.

We agree that it could be checked whether it is realistic or not. However even if it is *not* realistic, if the nested model predicts it, and upscaling can make the parend model predict it as well, then upscaling is doing what we hoped. Our hypothesis is always that the nested model is « better ».
In other words, if nesting predicts a saltier Corsican Current core, but it is actually not true, what do we hope from upscaling ? That the parent model also increases salinity in the core, or not ?
For this reason, we actually did not check if the nested model is more realistic or not. There are lots of other papers showing the impact of nesting.

4.4

Figure 10 should be mentioned since the beginning of the paragraph to help the reader.
Done
Please harmonize the models' nomenclature in the text but also in the legend and caption of figure 10. The validation (actually the only one that is provided with observations) should cover the entire 2014 year and not only two months.

In the text is written : « A similar plot for the whole of 2014 shows that the situation worsens during summer (errors of $3 \circ$ C) both for parent and child model; the difference in between models is hidden by the temporal variability of the error (not shown). »
If we plot the entire year, due to the scale going to 3°C instead of 1.4°C, we would almost not distinguish the red, black and blue curves anymore at all.

The mean RMSE of the three models should be provided as well to support your results. Upscaling should consider the performance of the child model and assimilate only, or give more weight, in those parts of the domain where the child model is close to observations. Without considering models' performance upscaling could force the parent model towards a wrong solution.

This comment is related to a suggestion higher in the review. We agree that this could be a way forward for SST, but maybe in a follow-up study. Apart from SST, for the other metrics, there are no (real) data to choose whether or not to trust the nested model. The hypothesis of this study is that the nested model **is** better, and the objective is to bring the parent model closer to the nested model.

Lines 6-8 Page 13: why not? REP in situ temperature and salinity profiles are available from CMEMS. I encourage to look at them and compute some validation to make your paper more robust.

This comment is again related to the previous one, and other ones before. In this study, we **only** aim at bringing the parent model closer to the nested model. Validating the nested model, such as doing a QUID for it, is out of our scope (and a whole lot of work).
On this occasion of in-depth temperature, we remembered the reader of the hypothesis and objective of the paper.

Line 8: "Upscaling is able to bring the differences back to the parent model" or the child model?
We rephrased this so that is it more clear

4.5

Here you mention also the child-model of the upscaled model and you speculate without any proof about the largest consistency in the nested upscaled models. I recommend to include a figure to show it. Or if you want to use Fig. 11 please give the RMSD of the model couples or include a specific comment to the figure (Line 7-8?)
We have added a table with the RMSD of model couples, as suggested by the reviewer

Line 10: Therefore Changed

5. Conclusions

Lines 4-9 Page 15: "The underlining hypothesis…" You only state it in the conclusions while it should be written clearly both in the abstract and the introduction. In this way the reader is aware that you only aim to relax the child model solution to the parent model solution, independently from the model performances.
Yes exactly. Sorry for this. It is now written in the introduction

Line 13: please use child instead of nested. changed

Line 15: Please be more precise and refer to the figures when this is shown. changed

Lines 1-6 Page 16: This general statement might confuse the reader since your child models never assimilate observations. Please rephrase to underline this. Rephrased

Line 6: please include a reference since you do not show this in the paper. Changed the text

Line 7-end: The advantage might be real is the upscaled solution has a better skill towards observations. In the last phrase: If the high resolution model is upscaled into the basin-scale one? Please notice that in CMEMS regional systems are the basin-scale ones.
We added explicitely in the text the potential disadvantages of the method, which include the fact that the nested model may be worse than the parent.

Upscaling of regional models into basin-wide models

Vandenbulcke Luc[1,2] and Barth Alexander[3]

[1]seamod.ro, Jailoo srl, Romania
[2]MAST, Université de Liège, Belgium
[3]GHER, Université de Liège, Belgium

Correspondence: Luc Vandenbulcke (luc@seamod.ro)

Abstract. Traditionnally, in order for lower-resolution, global- or basin-scale models to benefit from some of the improvements available in higher-resolution regional or coastal models, two-way nesting has to be used. This implies that the parent and child models have to be run together and there is an online exchange of information between both models. This approach is often impossible in operational systems, where different model codes are run by different institutions, often in different countries. Therefore, in practice, these systems use one-way nesting with data transfer only from the large-scale model to the regional models. In this article, it is examined whether it is possible to replace the missing feedback (coming from the child model) by data assimilation, avoiding the need to run the models simultaneously. Selected variables from the high-resolution forecasts will be used as pseudo-observations, and assimilated in the lower-resolution models. The method will be called "upscaling". A realistic test-case is set up with a model covering the Mediterranean Sea, and a nested model covering its North-Western basin. Under the hypothesis that the nested model is, in some sense, more realisic than the parent model, the upscaling method is implemented. Two simulations of the parent model are then compared: the case of one-way nesting (or stand-alone model), and a simulation using the upscaling technique on the temperature and salinity variables. It is shown that the representation of some processes, such as the Rhône river plume, are strongly improved in the upscaled model compared to the stand-alone model.

1 Introduction

In the present-day operational oceanography landscape, services are provided at different scales by different expert centers. At the European Union level, the Copernicus Marine Environment Monitoring Service (CMEMS) provides reanalyses, analyses and forecasts at global and basin scales. The models for the different basins are run by different institutes and centers within the regional monitoring and forecasting centers. Various regional and coastal oceanographic centers then use the CMEMS products to provide initial and/or boundary conditions to their respective models. These regional and coastal models benefit from the increased experience and specific knowledge of the local teams in their particular area of interest. Furthermore, local models usually run at higher resolution, and may include more accurate data (bathymetry, river discharge data...) and processes of smaller scales, that cannot be easily included into basin-scale models. High resolution observations such as satellite sea surface temperature (SST), and recent ultra-high resolution products (see e.g. Le Traon et al., 2015) have been shown to be best assimilated into nested models, as chances are higher that the observed processes are well represented (Vandenbulcke et

al, 2006). Similarly, high-resolution current observations by high-frequency radars are expected to benefit most to models with a similar high resolution (i.e. nested models).

When large-scale and regional or coastal models are run together (meaning, concurrently and on the same computing platform), it is possible to use two-way nesting; the benefits mentioned above of using a regional or coastal model are then transferred back to the basin-scale model. This has been shown numerous times in the literature, e.g. Barth et al. (2005); Debreu et al. (2012). The beneficial impact of the feedback from regional-scale to the basin-scale forecasts is visible even outside the domain of the regional model. This constitutes the baseline hypothesis of the present study: it is desirable to "copy" the results of the nested model into the parent model.

To emulate this nesting feedback, missing in the operational context, it is analyzed whether forecasts from the regional model can be used as pseudo-observations and assimilated in the basin-scale model. Indeed, data assimilation is not limited to the use of (real) observations by measurement devices. Onken et al. (2005) used data assimilation as a substitute for one-way nesting in a cascade of nested models. Alvarez et al. (2000) used a statistical model to predict SST, which was then assimilated as pseudo-observations in a hydrodynamic model (Barth et al., 2006). In the proposed "upscaling" method, the pseudo-observations come from the nested model. From the point of view of the forecasting centers, a data assimilation scheme is already implemented in the basin-scale model. Hence, implementing the upscaling method requires only to obtain the high-resolution forecasts and assimilate them (along with the real observations) during the analysis phase of the system.

When using grid nesting, problems at the open boundary of the child model include stratification mismatches, artificial waves, artificial rim currents; and ultimately instabilities and model blow-up (Mason et al., 2010; Debreu et al., 2012). By upscaling the child model into the parent model, the latter will progressively gain consistency with the child model solution within its domain, being beneficial for the child model over time. Upscaling can potentially reduce the risk of discrepancies at the open sea boundary.

Upscaling can also be seen as using a regional or coastal model as a "measurement device" that replaces ever-too-sparse (real) measurements. Guinehut et al. (2002, 2004) showed that a coverage of the North Atlantic with a 3°-resolution grid of Argo floats allows to effectively represent the large scales. Using a 5° array reduces the precision of the estimated fields two times. Currently, some CMEMS areas are largely undersampled.

Upscaling can be understood as a complement to downscaling (initialization) techniques such as presented in Auclair et al. (2000, 2001) (VIFOP) or in Simoncelli et al. (2011). The point of these methods is to combine interpolated fields coming from the large-scale model (the background or first-guess field) and existing high-resolution fields, so that small-scale structures present in coastal models are not lost whenever it is (re)-initialized by fields interpolated from the basin-scale model, and the obtained fields are physically balanced with respect to the coastal physics. If upscaling is used to improve the basin-scale fields

and accord them with the coastal model, the "first guess" will be already much better.

Schulz-Stellenfleth and Stanev (2016) is another recent example showing the benefits of two-way nesting, especially in sophisticated modern-day forecasting systems. The study demonstrates that two-way nesting is critical for correct energy transfers between large and small scales (especially in coupled ocean-wave-atmosphere models) for cross-border advection, for the correct use of high-resolution coastal observations that cannot be fed directly into a large-scale model, etc. Ackowledging that operational systems are using only one-way nesting, Schulz-Stellenfleth and Stanev (2016) therefore strongly advocate the research into "upscaling" techniques. The present article tries to develop precisely such a technique.

In this article, the upscaling procedure is tried out in a realistic, nested model configuration covering the Mediterranean Sea and the North-Western basin and simulating the year 2014. The same model, NEMO 3.6 (Madec, 2008), and the same vertical resolution, are used for both configurations (only the horizontal grid is different). This could influence the conclusion compared to a set-up with 2 different model codes. However, this is not expected to be a fundamental limitation of the method. Concerning the vertical grid, in the usual case of assimilating real observations such as vertical profiles, the observations and model forecasts have different vertical resolutions. Similarly, if the child model were on a different vertical grid than the parent model, it would still contain useful information, worth to be assimilated in the parent model. A limitation could be that some of the observations may be lost, e.g. the lowest child model layer may be out-of-grid in the parent model.

If different model codes were used, the modes could represent different processes. Hence, this should be taken into account by modifying the (representativity part of the) observation error covariance matrix when performing the data assimilation of the pseudo-observations. Examples of such contributions to the representativity error could be

- different vertical coordinates (see above)

- different implementations of the ocean surface: rigid lid or free surface; for the latter, linear or non-linear representation

- hydrostatic model or not

- different atmospheric forcing fields

- different turbulent closure schemes

- different numerical schemes for advection, horizontal diffusion etc.

The most striking difference between the parent and child models however, remains the horizontal resolution, and therefore, the general conclusions of the paper are expected to be valid, and upscaling should not be limited to the case of parent and child models being identical.

In this study, it is not the aim to verify that the nested model is indeed more realistic, according to some metrics, than the parent model. Rather, this consistutes the baseline hypothesis, and thus it is always considered beneficial to bring the parent

model forecast closer to the child model forecast. It should be noted that some high-resolution processes, resolved by the nested model but not by the parent model, could have large phase errors in the nested model. In this case, the baseline hypothesis would be violated, and the nested model could actually have higher errors than the former. This aspect is not considered in the paper.

The models and data assimilation scheme are described in section 2. Section 3 proposes some metrics to evaluate the system, related to the Rhône river plume, the cross-shelf exchanges, the large-scale current, SST, and the formation of Western Mediterranean Deep Water. Results are given in section 4 and a conclusion in presented in section 5.

2 Model and data assimilation configuration

2.1 Hydrodynamic model

The upscaling technique has been implemented in the North-Western Mediterranean Sea, including the Gulf of Lions and the Ligurian Sea. The latter is characterized by large-scale currents (the Northern Current also called Liguro-Provencal Current, created by the junction of the Eastern and Western Corsican Currents, see e.g. Pinardi et al. (2015)), by intense meso-scale activity and by inertial oscillations. Furthermore, the NW-Med is the siege of formation of Western Mediterranean Deep Water (WMDW), important to the circulation in the whole Mediterranean Sea (e.g. Millot, 1999; Pinardi et al., 2015; Bosse et al., 2015; Somot et al., 2018; Simoncelli and Pinardi, 2018).

A realistic, one-way nested configuration was implemented using the NEMO 3.6 model and the AGRIF nesting tool (Debreu et al., 2008), covering respectively the Mediterranean Sea (MED) with a 6 km horizontal resolution, and the North-Western Mediterranean basin (NW-Med). The parent model resolution is similar to the previous version of the CMEMS Mediterranean Sea analysis-forecasting system (up to October 2017) and to the present reanalysis (1/16°). The bathymetry of the nested model (1/80°) is more realistic than the parent model, at the coastline and more importantly, at the different canyons at the Gulf of Lions shelf break. Supplementary details of the model configuration are given in the Appendix (A1).

When implementing the upscaling method, it is expected that after some time, the feedback from the NW-Med model will modify the Northern Current position and intensity in the MED model, which will in turn influence the NW-Med model through its open-sea boundary. The boundary condition provided by the MED model also influences the stratification of the water column, which is important for the pre-conditionning of the convection (S. Somot, private communication).

Both model bathymetries are interpolated from the GEBCO bathymetry. The temperature and salinity initial condition is interpolated from the CMEMS Mediterranean reanalysis (1/16°) for 01/01/2014 (see https://doi.org/10.25423/medsea_reanalysis_phys_006_004). The model starts from rest. Atmospheric fluxes are computed using bulk formula; the atmospheric forcing fields are obtained from ECMWF ERA Interim with a temporal resolution of 3 hours and a horizontal resolution of 0.75° reinterpolated by the ECMWF server to 0.125° (Dee et al., 2011). In the MED model, the flow between the Black Sea and

the Mediterranean Sea, through the Marmara Sea and the Dardanelles Strait, is modelized as a river, using climatological flow, temperature and salinity values. The salinity of the incoming water has a minimum and maximum of 22.5 psu and 27.5 psu reached in July and March respectively. Five other rivers (Rhône, Po, Ebro, Nile, Drin) are also represented, and monthly climatologic values for the flow and temperature are used, whereas the salinity is put to 5 psu, except for the Drin river where is it put at 2 psu. Using climatologic monthly values is coherent with the operational set-up in the CMEMS Mediterranean system, although the latter represents many more small rivers.

Daily Rhone river discharge measurements at the Beaucaire station were obtained from the Compagnie du Rhône, in order to be used in the nested NW-Med model. Interestingly, the total annual flow computed from the climatology and from the measured values for 2014-2015 are very similar (1% difference). However seasonal and daily values can be very different (see Fig. 1a). In particular, during the considered period, the climatology underestimates the winter discharge, but overestimates the summer discharge. Hence, depending on the dataset used, it is expected that the modelled river plume will also be significantly different. This is illustrated in Fig 1b, showing the difference in surface salinity obtained with the nested model at the end of January 2014. The plume obtained using real river discharge extends much further offshore, almost completely across the Gulf of Lions, whereas the plume obtained with the climatological river discharge is essentially staying at the coast close to the river mouth. This is consistent with the much larger (almost double) discharge values observed in the real river data during January 2014.

2.2 Upscaling experiment description, Ensemble generation, and Data Assimilation scheme

In order to assimile pseudo-observations into the basin-scale models, different setups could be implemented, regarding the choice of the pseudo-observations, the frequency of assimilation, the data assimilation scheme itself, etc. The choices described below are consistent with current-day practices in the CMEMS operational systems. In particular, none of them currently assimilates velocity fields, and all of them use parameterized model state vector error covariances. Only one system (the Arctic system) currently uses an Ensemble Kalman filter, but the other systems are planning to evolve toward ensemble simulations in the future. The following settings were chosen for the current experiment. The filter is an Ensemble Kalman filter, some details are given in the annex. Assimilation will be performed daily. Only temperature and salinity will be used as pseudo-observations but the whole (thinned) 3D fields will be used. Velocity and surface elevation are not updated by the data assimilation procedure. Thinning is realized by taking the average of 5x5 cells of the nested model. The thinned pseudo-observations coming from the nested model are then considered independent, i.e. their error covariance matrix is diagonal. This is still a strong assumption which should be taken into account when determining the (diagonal) part of the observation error covariance matrix. The members of the ensemble have perturbed initial conditions, atmospheric forcing fields and Rhône river discharge, similar to Auclair et al. (2003).

The initial condition is the randomly weighted sum of the real initial condition (01/01/2014), and 6 other initial conditions (1 year, 20 days and 10 days earlier, and 10 days, 20 days and 1 year later). The weight of the real initial condition is a random-normal number chosen in the Gaussian distribution with mean 0.5 and standard deviation 0.2; if necessary, the random number is then limited back into [0.2 0.8], whereas the 6 remaining weights are random numbers chosen uniformly in [0 1],
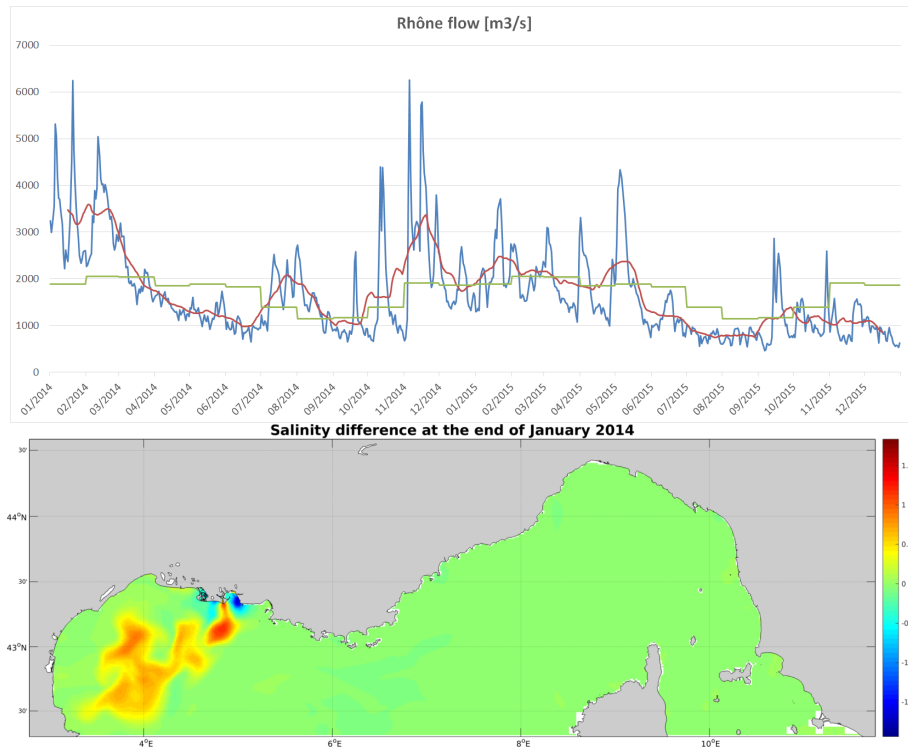
Figure 1. (above) Rhône river discharge from (green) the climatology, (blue) the measurements by the Compagnie du Rhone at the Beaucaire station, (red) the 1-month moving average of the measurements. (below) Difference of surface salinity in 2 different model runs of the nested grid on 28/Jan/2014, when using climatological or measured Rhône discharge values (i.e. using the green or blue curve in the upper panel)

.

and normalized so that the sum of all 7 weights is 1. This procedure ensures that the stability of each member is not modified (for example, the linear combination of 7 stable water columns is still a stable water column).

The atmospheric forcing fields of air temperature at 2m height and wind speed at 10m height are perturbed following the same procedure as in Barth et al. (2011); Vandenbulcke et al. (2017). Point-wise, the forcing fields are decomposed in Fourier series (from 3 hours to 1 year). For each member, a random field is generated, using these Fourier modes and random coefficients which have a temporal correlation length corresponding to the respective mode. This random field is added to the original field. The Rhône river discharge is perturbed using a random walk approach, with the expected perturbation after one year set as 20%. The other rivers are outside the observed part of the domain, and their discharge is not perturbed. With all 3 perturbations, an ensemble of 100 members is then spun up for 1 month.

Data assimilation is performed by the Ocean Assimilation Kit (OAK) package (Barth et al., 2008) implementing different filters such as SEEK and the Ensemble Kalman filter (EnKF). Different variants of the EnKF exist, and are classified in

6

stochastic and deterministic methods. The former require to perturb the observations, adding sampling noise. The latter, also called Ensemble Square Root Filters, do not present this requirement; the perturbation approach is only applied in the model to obtain model errors. Different variants are compared in Tippett et al. (2003). One variant, called the Ensemble Transform Kalman Filter (Bishop et al., 2001; Wang et al., 2004), is used in this study. The filter equations are listed in Barth and Van-

5 denbulcke (2017).

Nerger et al. (2012) summarizes how the spurious long-range correlations can be suppressed using so-called covariance localization, or domain localization and its addition observation localization. OAK uses the latter variant introduced in Hunt et al. (2007). In essence, the state vector is split into subdomains (water columns). In every water column, the analysis is performed independently (domain localization). In addition, for every water column, only nearby observations are used and the inverse

10 of their error variance is multiplied by a localization function (observation localization). In the current setup, the localization function is a radial Gaussian function with an e-folding distance of 30 km.

The observation errors for temperature and salinity are set respectively at 0.3°C and 0.09 psu. These values were determined after a sensitivity experiment with observation errors of (0.5°C, 0.15 psu), (0.3°, 0.09 psu), (0.2°C, 0.05 psu) or (0.1°C, 0.03 psu); as a trade-off between generating a close emulation of two-way nesting (hence very small observation errors), and generating

15 fields as balanced as possible, that will not cause adjustment shocks into the model (hence larger observation errors). With the latter 2 choices for the observation error, the obtained assimilation increment was not much larger than with the final choice of (0.3°, 0.09 psu), but qualitatively, unrealistic small scale variations started to appear.

From a technical point of view, OAK allows to use a multi-variate multi-grid state vector. As the Mediterranean model is paralized in 64 tiles, the multi-grid feature allows to update directly the tiles from the Mediterranean model restart files, influenced

20 by the nested model, without including the other tiles in the state vector. The procedure thus allows to skip the reconstruction of the complete Mediterranean restart files. It should be noted that the tiles of the parent model, considered in the data assimilation procedure, are the ones covering the nested-model area, but also the neighbouring ones which are influenced by the data assimilation.

3 Metrics

25 To assess the upscaling method, five metrics were defined, that allow to compare the nested model to both the basin-scaled model without upscaling (free model) and with upscaling. If upscaling is succesfull, the basin-scale model with upscaling will be closer to the nested model, than its counterpart without upscaling.

3.1 Cross-shelf transport

The penetration of off-shore water on the GoL (or inversely when negative transport values are obtained), is critical for the

30 circulation on the shelf, for the shelf-open sea exchanges, etc. It is obtained by integrating the current over the boundary shown in Fig. 2. This metric is useful to compare basin-scale models (free and upscaled) and check whether the upscaling procedure
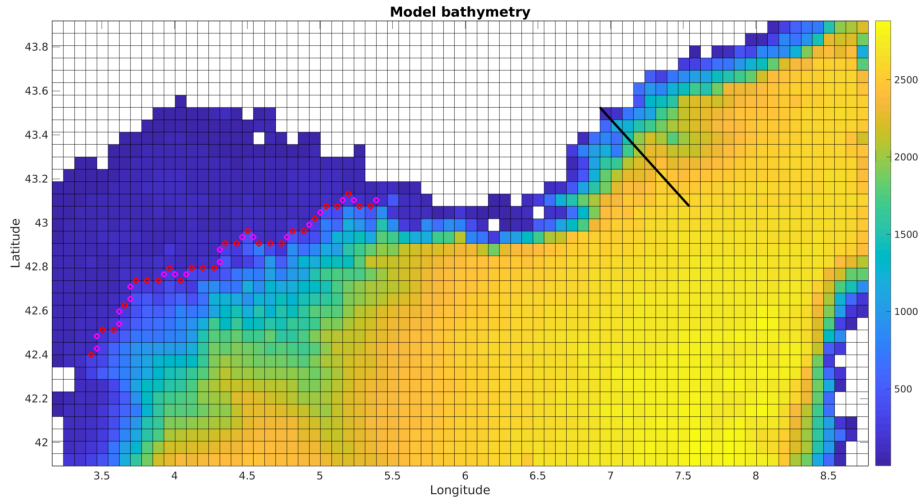
Figure 2. Zoom of the MED model grid, with the positions for computing the Northern Current metric (black line) and the cross-shelf transport metric (magenta and red dots)

is able to drive the solution toward the nested model solution. The intensity of the cross-shelf transport cannot, however, be compared to real measurements (by lack of them).

THE ORDER OF THE FOLLOWING TWO SUBSECTIONS WAS INVERTED

3.2 Northern Current intensity

5   The Northern Current (NC) is the most important large-scale feature of the region of interest. It is considered to have a width of 40-50km during summer and 20-30km during winter; but the most offshore currents do not modify the transport much. Similarly, the NC is considered to be 100-200m deep in summer and 250-400m in winter. Following Alberola et al. (1995), its intensity is obtained by integrating the currents normal to a line from Nice to the location (43.0756°N, 7.5415°E), 214 km to the South-East, indicated in Fig. 2. As for the previous one, this metric only allows to inter-compare different models.

10 ### 3.3 Rhône river plume

The plume of the Rhône River is measured by selecting all points around the river mouth with a salinity smaller than 37psu, and then choosing the most distant one from the river mouth. This provides the plume length and direction, although it may be an approximation: the plume can be curved, in which case its real length is larger than the estimation, or it can cover a large area, in which case the algorithm still obtains an azimuth although in reality it is not well defined.

15 This metric can be used quantitatively to compare models. Furthermore, it can be used to compare model results to real measurements. Indeed, although the real Rhône river plume length and direction are not measured directly, they can be estimated from satellite chlorophyll-a images. The model-observations comparison is then qualitative.

3.4 Sea surface temperature

This metric is the root mean square (rms) difference between the model and observed SST. For the latter, the L3 images are used. Furthermore, in order to examine the position of features such as fronts and eddies, the rms difference of the norm of the spatial gradient of the SST is also computed.

3.5 Western Mediterranean Deep Water formation

Following Bosse et al. (2015) and references herein, the formation zone of Western Mediterranean Deep Water (WMDW) is comprised in 41-43°N, 4-6°E. WMDW forms an easily identifiable area in the temperature-salinity diagram, and the present-day characteristics are: temperature of 12.86~12.89°C, salinity of 38.48~38.50 psu, depth is larger than 1000 m. The nested model (NW-Med) southern boundary is at 42.3°N, and hence only a part of the formation zone is included in the area of MED covered by pseudo-observations. Furthermore, the WMDW tail in the modeled temperature-salinity diagram is longer, with the following characteristics: temperature of 12.7~12.89°C and salinity of 38.4~38.5 psu. The metric measures the total volume [m^3] of WMDW in the domain covered by the NW-Med model, and is used to compare the different models.

4 Results

The temperature difference between the (unperturbed) parent and child models at the end of the spinup (31 January 2014) is represented in Fig. 3 on the child model grid. There are large temperature differences at the shelf break of the Gulf of Lions (the canyons are much better represented in the regional model); which extend all the way from the surface to the bottom of the Gulf of Lions. Other large differences appear in the Eastern and Western Corsican Currents, and their junction resulting in the Northern Current, as well as at the southern open boundary. The difference in salinity (not shown) has large values around the Rhône river plume (over 1 psu), and in a lesser extent in the Eastern Corsican Current. It appears that after a month, the differences are already significant, and if one trusts the regional model more, then it would be beneficial to bring these differences back to the basin-scale model.

At the end of the spin-up, the spread of the ensemble of models (Fig. 4) is very visible over the basin, at all river mouths, but also in other regions (Alboran Sea, Tunesian coastal zone...) as all 3 perturbations are applied at once. The ensemble spread is also visible in depth (i.e. deeper than when only the river discharge is modified).

As an example, the first data assimilation cycle is shown in Fig. 5 depicting SST. Qualitatively, it appears that upscaling changes important features: the Rhône river plume is oriented offshore instead of being mostly along-shore; fronts seem to be more well-defined; and the Northern Current flows along the shelf break instead of covering a large part of the shelf. The nested model, and the "upscaled" model, seem to be in closer agreement with the satellite image, than the free model.
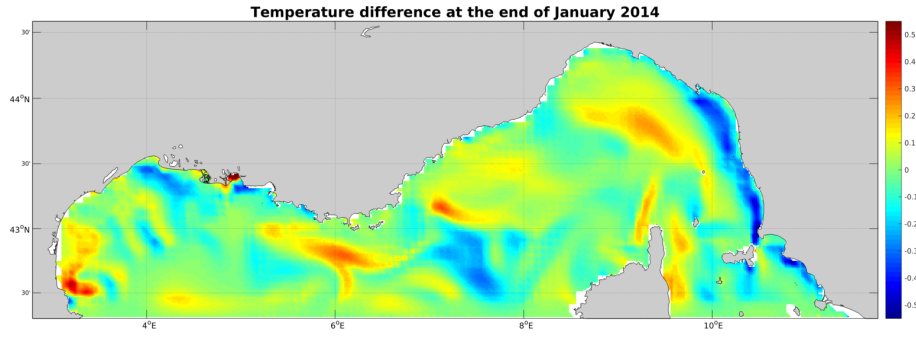
Figure 3. Temperature difference between the parent and nested models at 31/01/2014, projected on the nested model grid.
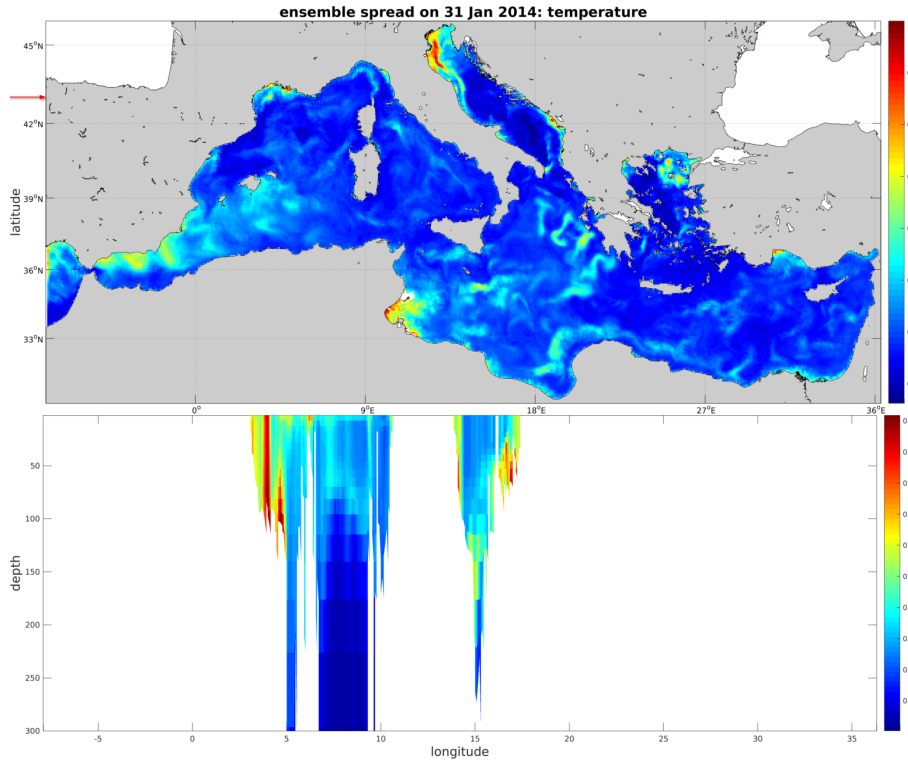

Figure 4. Spread of the ensemble of MED models at 31/01/2014: (upper panel) surface temperature, (lower panel) section at 43°N, indicated by a red arrow on the upper panel

4.1 Cross-shelf transport

The flow accross the shelf break is represented in Fig. 6, for the basin-scale model in the free and upscaled cases. Although alternating periods of inflow and outflow appear, the transport seems to show a chaotic behavior. Yet it can be seen that while both models are generally similar, some periods exist where the simulated transport is very different. During the first month
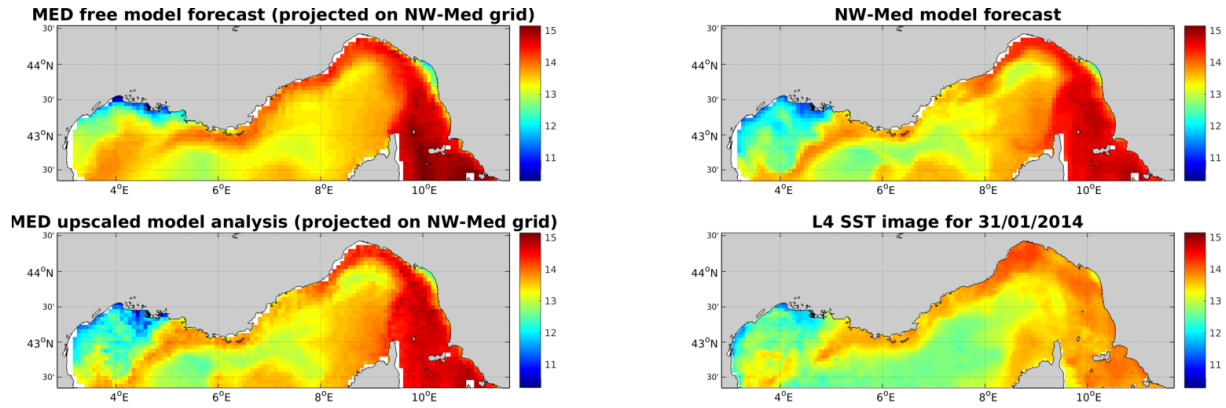
Figure 5. Sea surface temperature after the first upscaling step (31/01/2014), in the free model (upper left), nested model (upper right), upscaled model (lower left) and satellite observation (lower right)

(February 2014), the free model predicts a net outflow during the first 2 weeks, followed by a net inflow during the last 2 weeks. The nested model (not shown) and hence also the upscaled model actually predicts the exact opposite. The reasons for the nested model to behave differently than the parent model may be an effect of wind interaction with the (different) bathymetries, or related to the different resolution. The actual transport is not measured or available; but the result of interest

5   here is that the upscaling method is able to align the (parent model) currents with the ones from the nested model, and hence emulate two-way nesting, although only temperature and salinity pseudo-observations are used.

During the remainder of the year, the upscaled model predicts somewhat larger transports (both inward and outward). Generally speaking however, the two transport curves are closer than in February (or at least they are not of opposite signs anymore). Noticeably, in August-September, the upscaled model predicts a period of large inflow on the Gulf of Lions. The free model

10 also predicts this inflow, but delayed by about 2 weeks.

4.2 Northern Current

The transport by the Northern Current off Nice is represented in Fig. 7. Over the whole period, the root mean square difference between parent and nested models is 0.22 Sv for the free model, and 0.19 Sv for the upscaled model. The same qualitative observations can be made as for the cross-shelf transport. Both models generally agree, but periods exist with relatively impor-

15 tant differences. Interestingly, a large difference appears in August-September, when the free model predicts a larger transport than the upscaled model. This is also the period when the transport accross the shelf break presents a temporal shift in between models.

For the purpose of our study, this metric cannot be used to validate the model since real measurements of the Northern Current transport are not available; but (as for the previous metric), it can be used to compare models, and to show that our goal is

20 reached and upscaling of scalar fields is able to modify the velocity field of the parent model although only temperature and salinity are observed.
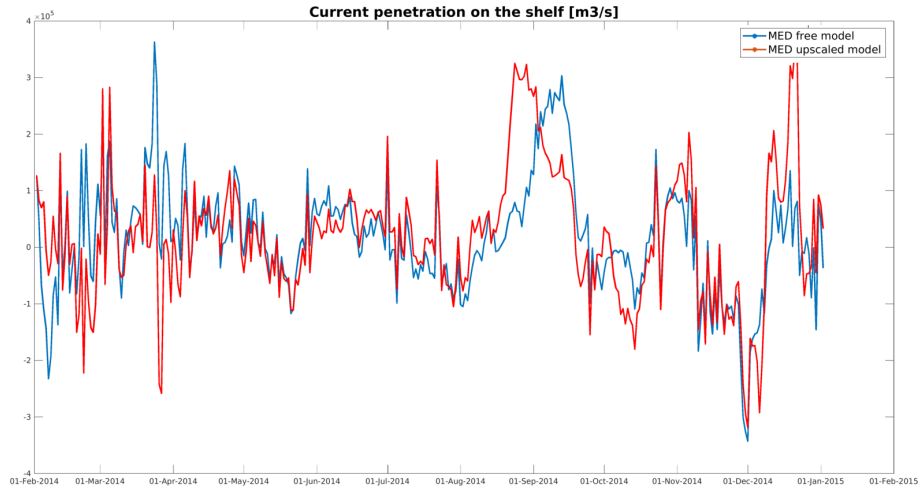
11

Figure 6. Water transport accross the shelf break during 2014, as obtained by the free (blue curve) and upscaled (red curve) parent models. Positive values indicate a net on-shelf transport.
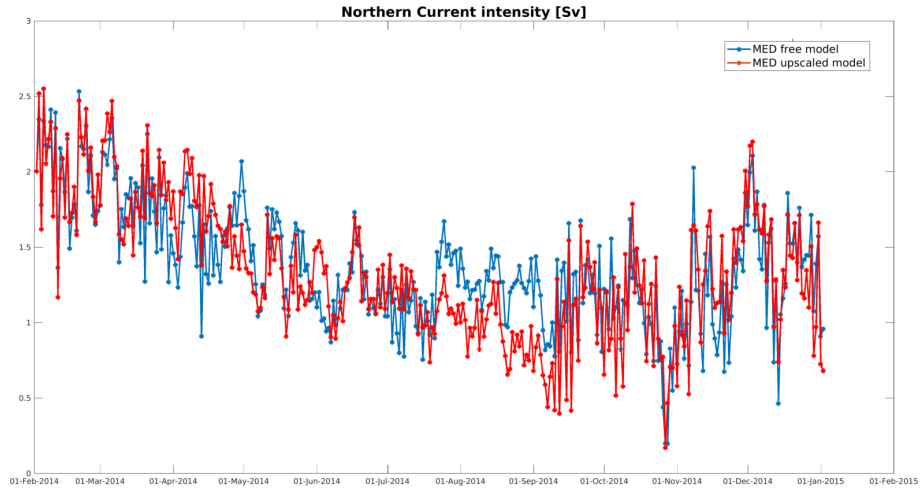

Figure 7. Water transport by the Northern current off Nice (France) during 2014, as obtained by the free (blue curve) and upscaled (red curve) parent models.

4.3 River plume

The Rhône plume metric is perhaps the metric most significantly altered by upscaling. During the first month of the upscaled simulation, the free model usually places the plume along-shore, to the North-East, whereas the nested model usually orients the plume off-shore to the South-East (see Fig. 8). On top of the resolution-related differences between parent and nested models (in particular the bathymetry and the interaction of the water masses with the wind), both models have different freshwater
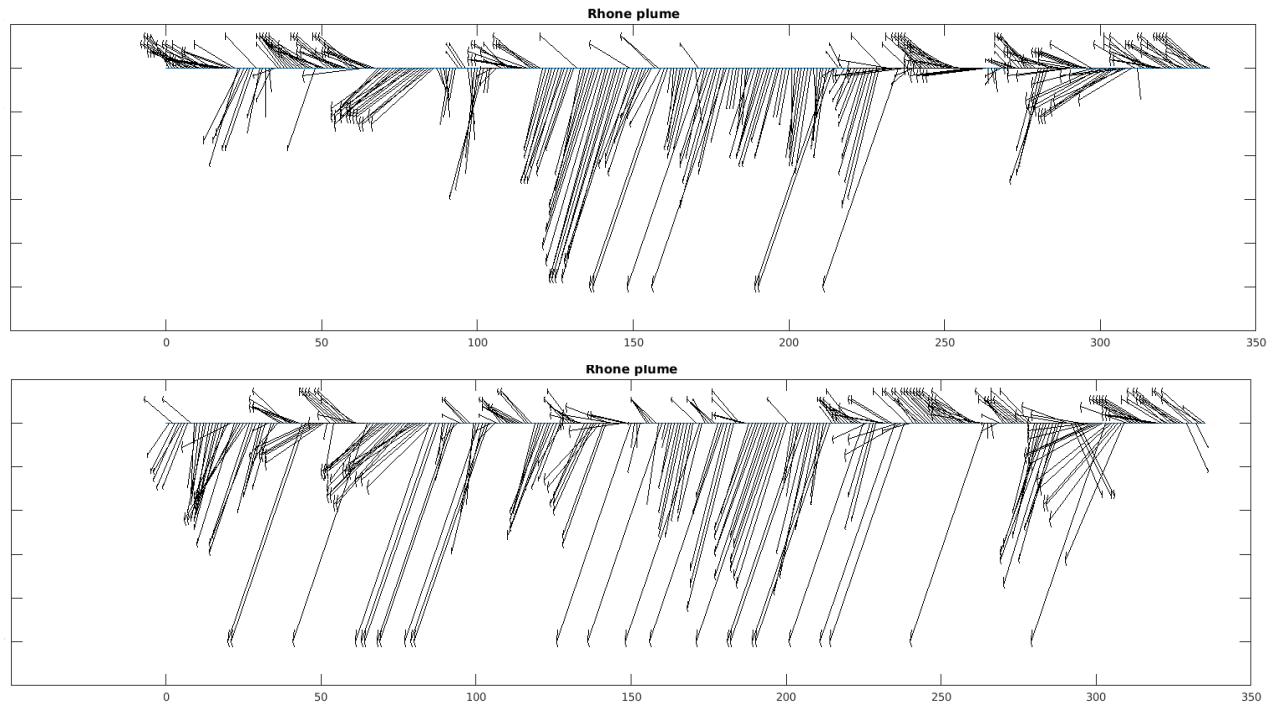
5

Figure 8. Rhône river plume direction and length (upper panel) for the free and (lower panel) upscaled MED models. The horizontal scale represents the days after the start of the upscaling experiment (Feb/2014).

discharge values, which is usually much higher and has also a much larger variability in the nested model during February 2014. The upscaling method is clearly able to make the parent model ingest the different plume dynamic coming from the nested model. During the remainder of 2014, differences between both models are smaller, with another period (late August - early September) where the opposite case occurs: the free model plume is oriented off-shore, but the nested (and upscaled) model predicts an along-shore plume.

The river plume can qualitatively be compared to real observations by using satellite observations of chlorophyll. During the first month of simulation, where the most significant differences appear, only a few level-3 satellite images are not almost entirely obscured by clouds. An example is given in Fig. 9 for 12 February 2014. One can clearly see the off-shore plume from the chlorophyll observations, whereas the free model plume is mostly along-shore. The nested and upscaled models correctly place the plume off-shore. Apart from the surface salinity used to characterize the Rhône plume, upscaling also modifies in-depth salinity. In particular, the cores of the Eastern and Western Corsican Currents are saltier in the upscaled model, with differences of about 0.15 psu during the first assimilation cycle.
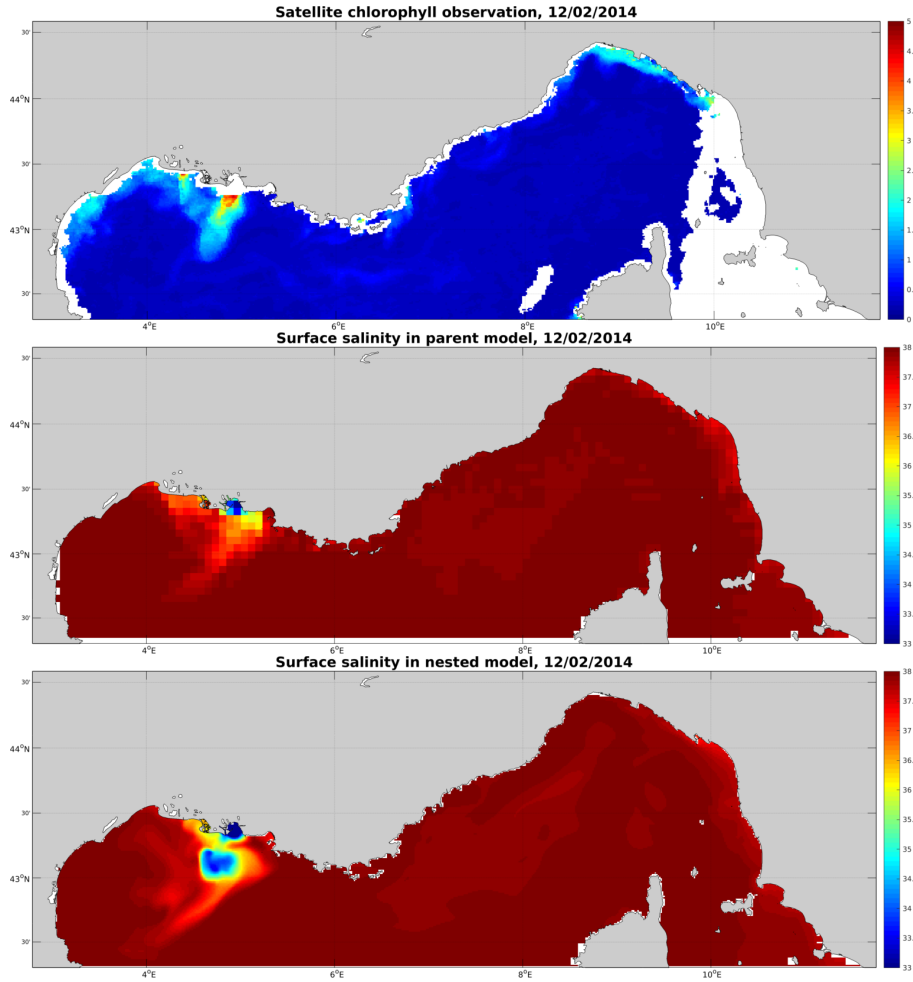
Figure 9. Comparison of the Rhône river plume on 12/Feb/2014 (upper panel) satellite chlorophyll image (middle panel) free parent model salinity (lower panel) and nested model salinity

4.4 SST

The sea surface temperature metric allows to quantify the model error by comparison with satellite images. Level 3 images are used for computing the metric (a level 4 image shown in Fig. 5 is used only for visual comparison). Results are shown in Fig. 10. It appears that generally, the (parent) model is already in very good agreement with observations, even though no

5   data assimilation is performed. Usually, the nested model is better still in some areas (e.g. coastal waters), and the upscaling procedure brings back these local improvements to the parent model. However, the area-wide RMS error is not influenced very much by upscaling, as large areas are essentially unmodified (parent and child models use the same atmospheric forcing fields). Some days, some processes appear to be missed by the models (both parent and nested), so that the RMS error is relatively large. In this case again, upscaling does not influence the RMS error of the parent model very much, as the nested model is not
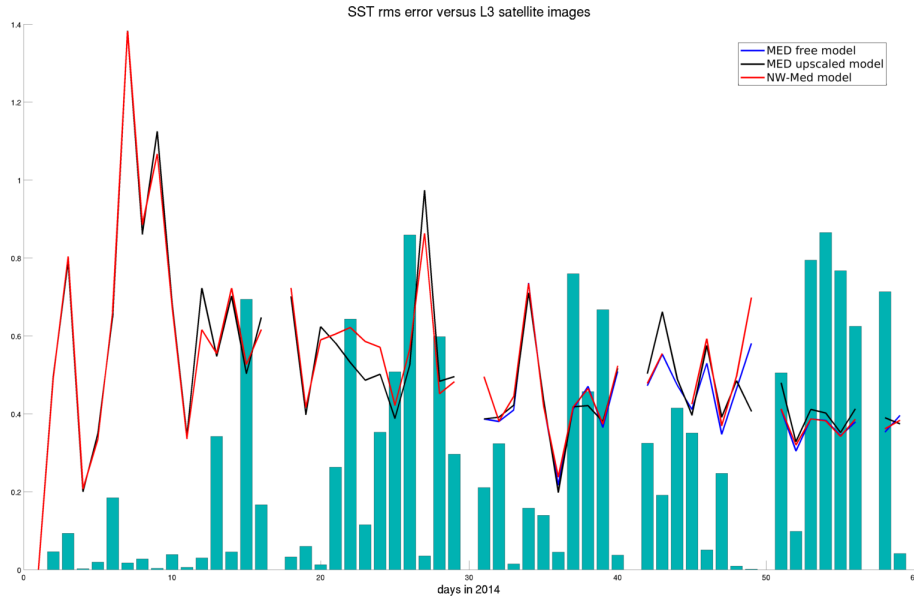
Figure 10. SST RMS error in the free model (black curve), nested model (red curve), upscaled model (blue curve) during the first 2 months of simulation. The bars represent the proportion of unclouded points in the L3 satellite image.

representing these processes any better than the parent model.

In both cases, this does not imply that the upscaling method is flawed, but rather that, in the current setup, the nested model is not able to generate an RMS error significantly lower than the parent model; hence upscaling does not have much to feed on. Fig. 10 shows the RMS error during the first 2 months of simulation. A similar plot for the whole of 2014 shows that the

5   situation worsens during summer (errors of $3°C$) both for parent and child model; the difference in between models is hidden by the temporal variability of the error (not shown). In any case, the upscaled model is still very close, and slightly better, than both the (free) parent and the nested models.

The model temperature in depth can be only punctually evaluated against observations (when e.g. drifter observations are available). In any case, the goal of the current study is to check whether upscaling is able to bring the parent model closer to

10   the nested model, under the hypothesis that the latter is "better" in some sense. Differences between the parent and the nested model are locally important, e.g. on the bottom of the Gulf of Lions, or in the Eastern Corsican and Northern Current cores (with differences of up to $0.3°C$), and upscaling is able to push the temperature field in the parent model toward the nested model solution.

4.5   WMDW

15   The total amount of Western Mediterranean Deep Water in the free model (blue curve in Fig. 11) and the nested model (green curve) is periodically important (10^3km^3), and both models do not appear to converge during the simulation. On the contrary, a period of large discrepency appears during most of the second half of the year. Upscaling largely modifies the parent model,
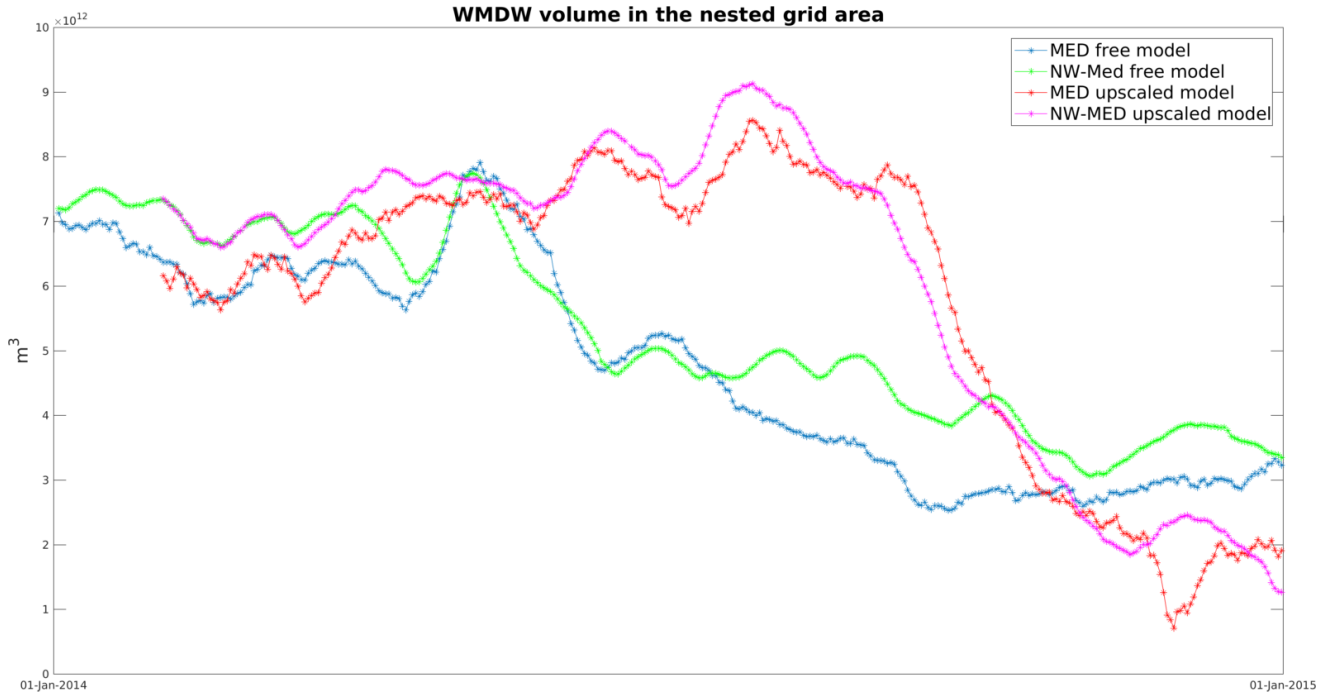
Figure 11. Time serie of total amount of WMDW in the area covered by the nested model: (blue) free parent model (green) nested model in the free model (red) parent model with upscaling (magenta) nested model in the upscaled model

which in turns provides modified boundary conditions to the nested model, so that after a while, the upscaled model and its child model significantly diverge from the free models. Without measurements and due to the choice of the model domain, it is not possible to assert which pair of models is more realistic. However, as for other metrics, the discrepency between parent and child model is reduced in the upscaled pair of models, which is certainly a desirable characteristic. This can be explained by the fact that the data assimilation also modifies the parent model solution outside the nested area (in the limit of the localization radius used in the data assimilation procedure). Therefore, the water immediately outside the nested domain is modified and made more coherent with the nested solution. East and West of Corsica, the Corsican currents will reintroduce this water into the domain, and one can see how this repeated procedure will ultimately reduce discrepencies between parent and nested models. The root mean square differences between parent and child models is summarized in Tab. 1 for all metrics.

5   Conclusions

When a nested model is available, it usually benefits from higher resolution, and improved representation of some relevant processes. However often, and particularly so in the operational oceanography context, there is no feedback from the nested model to the parent model. Data exchanges are limited to the parent model providing initial and/or boundary conditions to the

16

Metric	Free model	Upscaled model
Cross-shelf transport [m^3/s]	99.6 10^3	85.2 10^3
N.C. intensity [Sv]	0.22	0.19
Rhône Plume [km]	95.1	88.0
SST [°C]	1.3	1.3
WMDW [km^3]	1563	1422

Table 1. Root Mean Square difference between parent and child model for the case of the free parent model and the upscaled parent model, for the defined metrics

nested model. Thus, the benefit of having a nested model is lost to the parent model.

The upscaling method consists in assimilating results from a regional model into a basin-wide model, in order to emulate the feedback of two-way nesting. The underlying hypothesis is that the nested model is more realistic than the parent model.

The method was tried out using a nested model configuration of the Mediterranean Sea and the North-Western basin, with a resolution ratio of 5. Data assimilation was performed using a localized ensemble Kalman Transform filter; as pseudo-observations, thinned 3D fields of temperature and salinity were used. The aim of this study is limited to verifying whether nesting feedback could be emulated by data assimilation; without trying to verify whether the nested model is indeed more realistic than the parent one.

Whether upscaling was able to emulate two-way nesting, was measured using 5 metrics related to processes relevant in the study domain: the intensity of the Northern Current, the cross-shelf transport, the position of the Rhône river plume, sea surface temperature, and the quantity of Western Mediterranean Deep Water. These metrics show that the upscaling method is indeed able to emulate two-way nesting and bring the parent model closer to the child model. By assimilating only temperature and salinity, velocity and transport metrics were also improved in the parent model. The ability to constrain the cross-shelf transport by T/S assimilation is also an indication that the data from a high-resolution glider fleet would be beneficial to constrain the model. Finally, concerning the Rhône river plume, upscaling was able to strongly modify the plume direction and length in the parent model. Qualitatively, when real chlorophyll observations were available, the nested and upscaled parent model seemed to be more realistic than the free parent model.

Advantages of using upscaling include the following. Most importantly of course, the parent model takes advantage of improvements in the nested model. In the current study, these improvements may be due to higher resolution, better representation of local processes, and the use of more realistic river discharges. In general, they may also have other causes, such as assimilation of local and/or very high resolution measurements (e.g. HF radar observations), atmospheric fields from a regional weather forecasting model, or other more realistic boundary conditions. Another advantage is that over time, discrepencies between parent and nested model are attenuated. The parent model then provides more consistent boundary conditions to the nested model, and artefacts such as wave reflexion at the boundary may be avoided.

In the operational context, a supplementary advantage may appear. If a user is interested in a particular region not entirely cov-

parameter	parent model	child model
horizontal resolution	8km	1.6km
vertical resolution	31 levels	31 levels
bathymetry	interpolated from GEBCO	interpolated from GEBCO, smoothed at the commen open-sea boundary
surface module	Nemo MFS bulk formula	Nemo MFS bulk formula
surface forcing data	ECMWF Era Interim	ECMWF Era Interim
nudging (damping)	/	/
advection scheme	TVD	TVD
horizontal diffusion	bilaplacian	bilaplacian
vertical diffusion scheme	TKE	TKE
rivers	6 (climatological discharge)	Rhône (daily measured discharge)

Table A1. Model parameters for the parent and child models

ered by a regional model, it may be difficult for him to merge 2 products (the large-scale model, and the finer model not entirely covering the area of interest). By default, the user may then use only the coarser model. If the regional model is upscaled into the large-scale model, this is the only product the user needs to consider.

5   The most important limitation of the method is that the child model should be more realistic than the parent model. Furthermore, the coupling with upscaling is not as strong as with real two-way nesting. Other limitations are linked to data assimilation methods and are not different from the assimilation of real observations: (i) the data assimilation procedure itself uses approximations, and this could degrade the analysis; (ii) if the parent and child models are very different, the parent model could be unable to ingest the pseudo-observations. These limitations are investigated in the litterature in the context of assimilation

10   of real observations, and potential solutions include (i) anamorphosis techniques (when a non-linear relation exists between model variables and observations), particle filters (when the error distribution cannot be considered Gaussian), etc; and (ii) carefull specification of the observation error covariance matrix (and more specifically the contribution of the representativity error) to filter out processes of the nested model that cannot be represented in the parent model.

Appendix A: Model configuration

15   NEMO model parameters for the parent and nested models are given in Table A1.

Competing interests.   No competing interests are present.

References

Alberola, C., Millot, C., and Font, J.: On the seasonal and mesoscale variabilites of the Northern Current during the PRIMO-0 experiment in the western Mediterranean Sea, Oceanologica Acta, 18, 1995.

Alvarez, A., Lopez, C., Riera, M., Hernandez-Garcia, E., and Tintore, J.: Forecasting the SST space-time variability of the Alboran sea with genetic algorithms, Geophysical Research Letters, 27, 2000.

Auclair, F., Casitas, S., and Marsaleix, P.: Application of an inverse method to coastal modeling, J. Atmos Oceanic Technol, 17, 1368–1391, 2000.

Auclair, F., Marsaleix, P., and Estournel, C.: The penetration of the Northern Current over the Gulf of Lions (Mediterranean) as a downscaling problem, Oceanologica Acta, 24, 529–544, 2001.

Auclair, F., Marsaleix, P., and De Mey, P.: Space-time structure and dynamics of the forecast error in a coastal circulation model of the Gulf of Lions., Dyn. Atmos. Oceans, 36, 309–346, 2003.

Barth, A. and Vandenbulcke, L.: Ocean Assimilation Kit documentation, Universite de Liege, GHER, http://modb.oce.ulg.ac.be/mediawiki/index.php/Ocean Assimilation Kit, 2017.

Barth, A., Alvera-Azcarate, A., Rixen, M., and Beckers, J.: Two-way nested model of mesoscale circulation features in the Ligurian Sea, Progress In Oceanography, 66, 171–189, 2005.

Barth, A., Alvera-Azcarate, A., Beckers, J.-M., Rixen, M., and Vandenbulcke, L.: Multigrid state vector for data assimilation in a two-way nested model of the Ligurian Sea, Journal of Marine Systems, 2006.

Barth, A., Alvera-Azcarate, A., and Weisberg, R. H.: Assimilation of high-frequency radar currents in a nested model of the West Florida shelf, Journal of Geophysical Research, 113, 2008.

Barth, A., Alvera-Azcarate, A., Beckers, J.-M., Staneva, J., Stanev, E., and Schulz-Stellenfleth, J.: Correcting surface winds by assimilating high-frequency radar surface currents in the German Bight, Ocean Dynamics, 61, 599–610, 2011.

Bishop, C., Etherton, B., and Majumdar, S.: Adaptive sampling with the Ensemble Transform Kalman Filter part I: the theoretical aspects., Monthly Weather Review, 129, 420–436, 2001.

Bosse, A., Testor, P., Mortier, L., Prieur, L., Taillandier, V., d'Ortenzio, F., and Coppola, L.: Spreading of Levantine Intermediate Waters by submesoscale coherent vortices in the northwestern Mediterranean Sea as observed with gliders, Journal of Geophysical Research: Oceans, 120, 1599–1622, 2015.

Debreu, L., Vouland, C., and Blayo, E.: AGRIF: Adaptive grid refinement in Fortran, Computers and Geosciences, 8, 8–13, 2008.

Debreu, L., Marchesiello, P., Penven, P., and Cambon, G.: Two-way nesting in split-explicit ocean models: Algorithms, implementation and validation, Ocean Modelling, 49-50, 1 – 21, https://doi.org/https://doi.org/10.1016/j.ocemod.2012.03.003, 2012.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Guinehut, S., Larnicol, G., and Le Traon, P.: Design of an array of profiling floats in the North Atlantic from model simulations, Journal of Marine Systems, 35, 2002.

Guinehut, S., Le Traon, P., Larnicol, G., and Philipps, S.: Combining ARGO and remote-sensing data to estimate the ocean three-dimensional temperature fields - A first approach based on simulated observations, Journal of Marine Systems, 46, 85–98, 2004.

Hunt, B., Kostelich, E., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, Physica D: nonlinear phenomena, 230, 112–126, 2007.

5 Le Traon, P.-Y., Antoine, D., Bentamy, A., Bonekamp, H., Breivik, L., Chapron, B., Corlett, G., Dibarboure, G., DiGiacomo, P., Donlon, C., Faugère, Y., Font, J., Girard-Ardhuin, F., Gohin, F., Johannessen, J., Kamachi, M., Lagerloef, G., Lambin, J., Larnicol, G., Le Borgne, P., Leuliette, E., Lindstrom, E., Martin, M., Maturi, E., Miller, L., Mingsen, L., Morrow, R., Reul, N., Rio, M., Roquet, H., R., S., and Wilkin, J.: Use of satellite observations for operational oceanography: recent achievements and future prospects, Journal of Operational Oceanography, 8, 2015.

10 Madec, G.: NEMO ocean engine, Note du Pôle de modélisation, Institut Pierre-Simon Laplace(IPSL), France, No 27, ISSN No 1288-1619, 2008.

Mason, E., Molemaker, J., Shchepetkin, A. F., Colas, F., McWilliams, J. C., and Sangrà, P.: Procedures for offline grid nesting in regional ocean models, Ocean Modelling, 35, 1–15, https://doi.org/10.1016/j.ocemod.2010.05.007, 2010.

Millot, C.: Circulation in the Western Mediterranean Sea, Journal of Marine Systems, 20, 423–442, 1999.

15 Nerger, L., Janjic, T., Schroter, J., and Hiller, W.: A regulated localization scheme for ensemble-based Kalman filters, Q. J. R. Meteorol. Soc, 138, 802–812, 2012.

Onken, R., Robinson, A. R., Kantha, L., Lozano, C. J., Haley, P. J., and Carniel, S.: A rapid response nowcast/forecast system using multiply nested ocean models and distributed data systems, Journal of Marine Systems, 56, 2005.

Pinardi, N., Zavatarelli, M., Adani, M., Coppini, G., Fratianni, C., Oddo, P., Simoncelli, S., Tonani, M., Lyubartsev, V., Dobricic, S., and

20 Bonaduce, A.: Mediterranean Sea large-scale low-frequency ocean variability and water mass formation rates from 1987 to 2007: A retrospective analysis, Progress in Oceanography, 132, 318 – 332, https://doi.org/https://doi.org/10.1016/j.pocean.2013.11.003, http://www.sciencedirect.com/science/article/pii/S007966111300222X, 2015.

Schulz-Stellenfleth, J. and Stanev, E.: Analysis of the upscaling problem - A case study for the barotropic dynamics in the North Sea and the German Bight, Ocean Modelling, 100, 109–124, 2016.

25 Simoncelli, S. and Pinardi, N.: Water mass formation processes in the Mediterranean sea over the past 30 years, Journal of Operational Oceanography, 11, S1–S142, https://doi.org/10.1080/1755876X.2018.1489208, 2018.

Simoncelli, S., Pinardi, N., Oddo, P., Mariano, A., Montanari, G., Rinaldi, A., and Deserti, M.: Coastal Rapid Environmental Assessment in the Northern Adriatic Sea, Dynamics of Atmospheres and Oceans, 52, 250–283, https://doi.org/10.1016/j.dynatmoce.2011.04.004, 2011.

Somot, S., Houpert, L., Sevault, F., Testor, P., Bosse, A., Taupier-Letage, I., Bouin, M.-N., Waldman, R., Cassou, C., Sanchez-Gomez, E.,

30 Durrieu de Madron, X., Adloff, F., Nabat, P., and Herrmann, M.: Characterizing, modelling and understanding the climate variability of the deep water formation in the North-Western Mediterranean Sea, Climate Dynamics, 51, 1179–1210, https://doi.org/10.1007/s00382-016-3295-0, https://doi.org/10.1007/s00382-016-3295-0, 2018.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S.: Ensemble Square Root Filters, Monthly Weather Review, 131, 1485–1490, 2003.

35 Vandenbulcke, L., Beckers, J., and Barth, A.: Correction of inertial oscillations by assimilation of HF radar data in a model of the Ligurian Sea, Ocean Dynamics, 67, 117–135, 2017.

Wang, X., Bishop, C. H., and Julier, S. J.: Which Is Better, an Ensemble of Positive–Negative Pairs or a Centered Spherical Simplex Ensemble?, Monthly Weather Review, 132, 1590–1605, 2004.