Responses to reviewer 1 comments on the manuscript ‘High resolution stochastic downscaling method for ocean forecasting models and its application to the Red Sea dynamics’

The authors are thankful to the reviewer for thorough and helpful comments. Our responses are given below. References to line numbers are for the amended MS.

Comment. Figure 2 uses an example in which the grid barely resolves the fields in the x-direction. It may be that it is only in such cases that OI gives significantly better results. Are the authors able to explore that in a little more detail?

Response. It is correct that the maximum enhancement produced by SDD compared to simple interpolation is expected when the parent model barely resolves the field. If the ocean feature is well resolved by the parent model, there is no need for further refinement. For example, if the zonal size of the eddy is increased to 40 km instead of 13 km, and hence it is reasonably well resolved by the parent model with Δx=10km, then the RMSE produced by SDD is similar to that of bi-cubical interpolation. On the other hand, if the parent model misses the features completely, e.g. no eddy permitting, then the SDD method does not have enough information to re-create the smaller scale features. The clarification is given in lines 215-220.

Comment. Second the model field that is being interpolated should not usually be regarded as being precisely correct. More specifically an ocean (or atmosphere) model is well known to be unreliable near the grid-scale. Fields are typically either noisy at the grid-scale or overly damped. It is also a moot point whether the fields represent point values or grid cells means. To what extent it is possible to extract more information near the grid-scale from model fields by using OI needs more investigation. The fact that OI is well suited to interpolation of noisy (stochastic) data makes such an investigation more attractive than it would otherwise be. The cost of increased resolution is so high that such an investigation is worthwhile, though the value of increased resolution is not necessarily in the increased detail in the simulations.

Response. This is correct, the SDD method works better than simple interpolation in a noisy situation. We have added a small subsection ‘Effect of noise in the input data’ and a table showing the errors produced by different methods when the parent model outputs are noisy. This also shows that the coarse field does not need to be regarded as precisely correct. We are thankful to the reviewer for this comment and suggestion to discuss the treatment of noisy data in the MS.

Comment. Abstract: The abstract describes the sort of problem that is being addressed and does introduce the principal idea in lines 20-22. The important point about the lack of a double penalty is clearly made on lines 14 and 27. So the abstract is a reasonably informative summary of the paper.

Response. Thank you.

Comment. Lines 20-21: I found the sentence starting on line 20 somewhat difficult to follow.

Response. The sentence is now re-worded. (Line 20)

Comment. Lines 54 and 55: it might be worth mentioning that the commonly used, more "modern", variational methods are also closely related to OI (Lorenc 1986, QJRMS).

Response. The reviewer's suggestion is implemented in lines 56-57.

Comment. The literature on methods for post-processing of model outputs using Kalman filters should probably be discussed in the introduction. I think the main idea being pursued in this paper is somewhat different from the main ideas in that literature but the techniques are clearly related**.**

Response. The reference to Kalman filtering and other methods are briefly given in lines 47-48 and 56-57. The reference to work by Lorenc (1986) is given where it is shown that the Kalman filter and more modern variational methods are closely linked to the original OI and they can be described using a common Bayes analysis framework.

Comment. One might ask whether the method proposed is a post-processing of model output or a statistical model in its own right. It is described both as a Statistical Model (in SMORS) and a Stochastic Deterministic Downscaling (SDD) method. Personally I would view it as a post-processing method but do not feel strongly about this semantic issue.

Response. We prefer to term the SDD method as part of the model based on how it is implemented in the code. This is shown in the flowchart in Fig.4.

Comment. The introduction does introduce relevant material but at the end of it one does not have much more insight into how the proposed technique works. The structure of the paper is not described.

Response. In the amended MS the structure of the paper is now explained in lines 82-86. The insight into how the proposed technique (SDD method) works is given in lines 66-73.

Comment. Lines 107-109: The primed quantities (that are interpolated) are deviations from monthly means for the Red Sea. These values would not normally be available in real-time. I imagine that deviations from climatology would be a satisfactory alternative.

Response. This is correct. Clarification is given in lines 115-116.

Comment. Line 133: The use of Gaussian and SOAR functions for the autocorrelation function dates back well before Fu et al (2004). In data assimilation the difficulties / uncertainties in the calculation of this function are usually emphasised quite strongly. The textbook by R. Daley (Atmospheric Data Analysis 1991) is a good source of information on the techniques discussed in this paper. Lorenc 1981 QJRMS describes a fairly sophisticated method for retaining consistency of solutions between points.

Response. We agree. The text is amended as requested and an additional reference to Daley (1991) is added in lines 145-146.

Comment. The idealised case has a=4.1km and the parent grid has Δx = 10 km. So the sinusoidally varying field in one direction is really close to the 2-grid point wavelength. It is very impressive how well the OI solution handles this problem (Figure 3 of the paper and line 201 -204). A brief summary of results for some less extreme interpolation cases would probably be informative.

Response. The efficiency of SDD in comparison to different interpolation method is added as a new sub-section in lines 225-252.

Comment. Figure 6: The lack of a double penalty is certainly an interesting result and the comparison with OSTIA data seems sound to me (though I'm not an expert in this issue). It's not clear to me what explains the lack of a double penalty. Is the model SST a relatively smooth field in which case OI and linear interpolation might give relatively small differences? I think there needs to be some further quantification of the double penalty. For example, one could calculate the rms of f' at all points on the high resolution grid that do not coincide with low resolution grid points for the OI, bi-linear and bi-cubic fields. How do these rms values compare with those in figure 6? One might also ask how these rms values compare with the rms of the values at the original (lower resolution) gridpoints. This calculations would help to shed some light on the lack of a double penalty.

Response. Double penalty phenomenon is more evident in the high resolution models (Gilleland, E., Ahijevych, D., Brown, B. G., Casati, B. and Ebert, E. E.: Intercomparison of spatial forecast verification methods. Weather Forecast, 24(5), 1416-1430, 2009). The SDD method honours the data on the parent coarse grid and hence the spatial structure is anchored onto the coarse grid, therefore there is no additional spatial shift and no additional double penalty effect compared to the parent model. Clarification is given in lines 337-341 of the revised MS.

Comment. Lines 296-298: It seems strange to use nearest neighbour values in the ARGO inter-comparison. With the OI method one can do much better interpolations! Some readers may be concerned that the nearest neighbour method could somehow account for the lack of a double penalty (see previous paragraph).

Response. We use the nearest neighbour method for compatibility reasons, as it is used for validation of MyOcean / Copernicus Marine Environment Monitoring Service products, see e.g. Delrosso, D., Clementi,E., Grandi,A., Tonani,M., Oddo, P., Feruzza, G., Pinardi,N. 2016. Towards the Mediterranean forecasting system MyOcean v5: numerical experiments results and validation, 2016. INGV technical report, No 345, ISSN 2039-7941. Clarification and additional reference are given in line 358.

Comment. Lines 311-318: This point that the interpolation will reproduce the field exactly at the parent grid points (to within truncation errors) is an important one. Many readers would find it helpful to mention this earlier in section 2.1.

Response. We agree. Clarification is added in lines 160-163.

Comment. Some results have already been presented in section 2. So the section title seems strange. Results for vorticity or Vorticity diagnostics would be a better title.

Response. The results presented in section 2 have now been moved to section 3, and new sub-sections are created: 'Eddy and mean kinetic energy' and 'Analysis of vorticity and enstrophy'.

Comment. Line 353: describing the two models as eddy-permitting and eddy-resolving seems contentious at this point.

Response. Analysis of the efficiency of the SDD method for eddy-permitting and eddy-resolving models is now added throughout the text.

Comment. Line 360: Could you confirm that Figure 9 shows area mean values of vorticity? Lines 363-364 suggest it does. The phrase "Absolute values of vorticity" in line 360 gives the reader some cause for uncertainty on this point. I suggest you remove "Absolute" from that phrase.

Response. The text amended as advised (lines 419-420)

Comment. Figure 14 is a nice illustration of the potential of this method. I wonder whether there are any plotting packages which use this type of approach.

Response. We are not aware of any plotting packages which use this type of approach.

We thank Reviewer 1 for helpful comments

Responses to reviewer 2 on the manuscript 'High resolution stochastic downscaling method for ocean forecasting models and its application to the Red Sea dynamics'

General.

Comment. A clear limitation is an assumption (lines 95-96) that the coarser-resolution wider-area model is accurate at all its grid points…. However, the assumption leaves no scope for adjusting values at the coarser-model grid points. Thus the limited accuracy of the coarser model is "built in" and the method is strictly interpolation, albeit allowing for statistical properties of finer-scale fluctuations (anomalies). It seems to me that this is reflected in the validation (section 2.4) that the comparisons with OSTIA and ARGO data show very similar bias and RMS error for the coarser and finer models.

Response. This limitation has been removed in the revised manuscript by adding a new sub-section 2.3 Effect of noise in the input data. The calculations when the parent model is noisy (i.e. not accurate at all of its grid points) show that SDD method gives much better approximation to the true field that 'strictly interpolating' methods such as bi-linear or bi-cubic interpolation of the coarse mesh data. For example in case of 10% noise in coarse grid, the RMSE error between the fine grid data generated by the SDD method and the true field is the same 10%, by bi-linear it is 25%, and by bi-cubic it is 19%. Moreover, we present an example where the coarse mesh is eddy permitting and the fine mesh is eddy resolving, with a resolution doubled in each spatial dimension. In this situation, mesoscale eddies are embryonic in the coarse mesh and can be restored into the fine grid.

Comment. Probably a finer-resolution model (impractical – the point of the manuscript) would be more accurate and give different results at the coarser-model grid points.

Response. The assumption that a finer model would be more accurate is not always the case, at least when standard point-wise metrics are used like RMSE and bias. The following quote from (Crocker et al. 2020) explains the situation (emphasis added):

> One of the issues faced when assessing high-resolution models against lower-resolution models over the same domain is that often the coarser model appears to perform at least equivalently or better when using typical verification metrics such as root mean squared error (RMSE) or mean error, which is a measure of the bias. **Whereas a higher-resolution model has the ability and requirement to forecast greater variation, detail and extremes, a coarser model cannot resolve the detail and will, by its nature, produce smoother features with less variation resulting in smaller errors. This can lead to the situation that despite the higher-resolution model looking more realistic it may verify worse (e.g. Mass et al., 2002; Tonani et al., 2019).**
>
> This is particularly the case when assessing forecast models categorically. If the location of a feature in the model is incorrect, then two penalties will be accrued: one for not forecasting the feature where it should have been and one for forecasting the same feature where it did not occur (the double-penalty effect, e.g. Rossa et al., 2008). **This effect is more prevalent in higher-resolution models due to their ability to, at least, partially resolve smaller-scale features of interest.** If the lower-resolution model could not resolve the feature and therefore did not forecast it, that model would only be penalised once. Therefore, despite giving potentially better guidance, the higher-resolution model will verify worse.
>
> The manuscript was amended to include the clarifying text and references (Lines 322-342)

Comment. Another assumption is that the distribution of fluctuations (anomalies, at any one depth) is statistically uniform and isotropic horizontally (line 129). This is inherently a limitation on the area of the (sub-)region where interpolation for finer resolution is desired. It may imply avoidance of nearby coasts, other distinct topography or water-mass boundaries (for example), despite the optimisation of weighting coefficients allowing for coasts.

Response. The assumption is actually that the distribution of fluctuations is statistically uniform and isotropic horizontally only locally, within the radius of computations given by Eq (8), not over the whole area. Clarification is added in Line 137. Such assumption is not unusual. Modern data assimilation schemes assume statistical uniformity/isotropicity in the horizontal. For example, “The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4” section 4.6.2 “Length scales” reads: “The horizontal background-error correlations for X = T , S U and η U are assumed to be isotropic poleward of a given latitude φ_L , with an identical length-scale $L_\lambda = L_\varphi = L$ used for all variables and at all depths”. (https://www.ecmwf.int/en/elibrary/11174-nemovar-ocean-data-assimilation-system-implemented-ecmwf-ocean-analysis-system-4) Some data assimilation schemes allow for non-homogeneity in the length scale, but *local* homogeneity is still required. This means that when computing the correlation matrix, homogeneity is assumed within the computation radius of every node.

Comment. Abstract. It is important that the abstract is clear and easily understood. Please clarify:

Line 14. What is the “double penalty” effect?

Response. The double penalty effect is described in the literature as follows. If the location of a feature in the model is incorrect, then two penalties will be accrued: one for not forecasting the feature where it should have been and one for forecasting the same feature where it did not occur (the double-penalty effect, e.g. Rossa et al., 2008). Double penalty phenomenon is more evident in the high resolution models (Crocker, R., Maksymczuk, J., Mittermaier, M., Tonani, M., and Pequignet, C.: An approach to the verification of high-resolution ocean models using spatial methods, Ocean Sci., 16, 831–845, https://doi.org/10.5194/os-16-831-2020, 2020). The SDD method honours the data on the parent coarse grid and hence the spatial structure is anchored onto the coarse grid, therefore there is no additional spatial shift and no additional double penalty effect compared to the parent model. Clarification is given in lines 322-342 of the revised MS.

Comment Lines 20-21. “areas smaller than the Rossby radius, where distributions of ocean variables are more coherent”. If the point about “more coherent” is necessary then what is more coherent with what? Maybe small structures have internal coherence but their occurrence and scales are more likely to be stochastic, not coherent.

Response. Ocean fields are more coherent within 1-2 Rossby radii than between more distant points, so that mesoscale eddies of that size are sometimes called oceanic coherent structures, see e.g.(G.I. Barenblatt et al (eds), 1992. Coherent structures and self-organisation of ocean currents. M.Nauka, 198pp. In Russian: Г.И.Баренблатт и др. (ред). 1992. Когерентные структуры и самоорганизация океанических движений : М. : Наука, 198 с. ISBN 5020008079; F. J. Beron-Vera, M. J. Olascoaga, and G. J. Goni,2008. Oceanic mesoscale eddies as revealed by Lagrangian coherent structures. Geophysical Research Letters, Vol. 35, L12603, doi:10.1029/2008GL033957;
P.F.J. Lermusiaux and F. Lekien, 2020. Dynamics and Lagrangian Coherent Structures in the Ocean and their Uncertainties,
http://web.mit.edu/pierrel/www/talk/pfjl_lekien_final_oberwolfach05.pdf) and the references in the MS . Clarification is added to the text (LINE 21, 490).

Comment. Line 23. 1/24th degree from 1/12th degree is only a factor of 2 and begs the question of how much refinement the method works for.

Response. An increase of resolution by a factor of 2 (in each horizontal direction) increases the computational cost be a factor of 10 or more. The number of nodes is quadrupled (2x2) and the time step should be made 2 times smaller to comply with the Courant–Friedrichs–Lewy stability condition, which give the increase of number of computations by 2x2x2=8. The computation would require a larger number of computing cores and the overhead adds 20%-30% or more due to non-linearity of scaling. The cost of a relatively small HPC cluster is about £100K, so the purchase of 10 times larger computer can be a game-stopper. The SDD method adds a small number of calculations which can be performed even on a laptop computer. The efficiency of the SDD method is discussed in Section 2.2 of the revised MS.

Comment. Line 25. “. . cost function which represents the error between the model and true solution.” In practical use the true solution is not available.

Response. "True solution" or "true state" is standard parlance in data assimilation when calculating a cost function. In many formulations, variables for the true solution are included, even if that true solution is never known (see for example R. N. Bannister, A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, 2008. Quarterly Journal of the Royal Meteorological Society, https://doi.org/10.1002/qj.339) . Clarification is given in Line 180.

Comment. Lines 167-168. “The correlation matrix is calculated . . . for each grid node on the fine mesh.” This is possible where the true field is known (as here) but not in practical application unless there are data with resolution as good as on the fine mesh. Such data cannot come from the coarser model.

Response. The correlation matrix can be computed in practice using one of several methods (e.g. Hollingsworth and Lonnberg, 1986) which do not require knowing the true field. For instance, in H-L method, the true state is removed and only the errors remain by subtracting the background values from the observations. In this case, the only requirement is that the data is unbiased, the true state is not needed. In our paper eq. (7) is presented as a parametrised approximation by Fu et al. The text is additionally clarified below Figure 1.

Comment. Line 170. Surely the “final stochastic downscaling is carried out using” Equation (1) with the now-known pi. Eq. (7) was used earlier to calculate the correlation matrix.

Response. The text is corrected as advised (reference to eq(7) is replaced with Eq(1)).

Comment. Lines 245-247. Regarding the comment on lines 167-168, actual data for Eq (5) only exists at nodes of the coarser grid. Do the other 75% of points on the finer grid invoke the assumption that deviations f' are statistically uniform and isotropic in the horizontal plane? Please clarify.

Response. This is correct. In common with the theory of 2D turbulence (eg Rhines, P.B. 1975. Waves and turbulence on a beta-plane. J. Fluid Mech. 69, 417–443), the SDD model requires that all deviations are locally (within the search radius) isotropic and homogeneous. Additional clarification is added in Lines 136-137

Comment. Line 256. “previously considered” meaning nearest adjacent (node) already solved for?

Response. This is correct. We have amended the Ms to incorporate this clarification (line 305).

Comment. Lines 324-325. I think that one cannot argue from the accuracy of the idealised experiment in view of the question about data at nodes on the fine mesh (lines 167-168 comment). In the Red Sea example the finer-resolution model has accuracy very close to the coarser-resolution model and may well have more small-scale features (as figure 9 – yet to come – suggests) but it is not yet clear that “it also improves the accuracy of simulation.”

Response. We meant that the SDD model has the ability to forecast greater granularity, variation, and extremes with respect to simpler interpolation schemes shown (bilinear, bicubic and spline). We have amended the Ms to clarify this point. (Line 250-253).

Comment. Line 373. What is the basis for “underestimates”?

Response. We meant that the coarse model shows lower values of gradients. The text is amended as requested.(LINES 431-432)

Comment. Line 381. “vorticity” should be “enstrophy”?

Response. Yes. The text has been amended (LINE 440).

Comment. Lines 416-417. Same comment as on lines 324-325.

Response. Please see our response to comment for lines 324-325.

Comment. Line 429. Repetition: “optimal . . optimised”

Response. Thanks. The text has been amended to avoid repetition (Line 489).

Comment. Line 430. “short range, comparable with the resolution of the parent model”; is this a limitation on the refinement from coarser to finer?

Response. SDD has been designed to improve on the results from Eddy permitting models. From this point of view, it is not a limitation of the model but rather its desired area of applicability.

Comment. Lines 475-480. I think the origin of “greater granularity” in the finer model should be further discussed. Conceivably the statistics are related to those determining the correlation matrix etc. but is there any deterministic element in the small-scale (c.f. lines 521-522), or (more likely) in the seasonal variation of their statistics (line 479)?

Response. Yes, there is seasonally (monthly) variation of the statistics in the norm that is used to compute the innovations. The text is amended to emphasize this (Lines 541-543)

Comment. Lines 489-495. The sign of vorticity can be biased (e.g. in coastal eddies) but does not show in enstrophy. Has SMORS a basis for showing such bias? How does any such bias in its output compare with the best available evidence? More enstrophy is likely in the finer-resolution model but does its increase take it significantly closer to the "truth" – is there evidence to test that? Certainly the finer resolution in figure 14 presents a more convincing picture but it appears to add little except interpolation; all the features are embryonic in the coarse-resolution figure.

Response. We did not notice any bias in the sign of enstrophy in our calculations. It is correct that more enstrophy is likely in the finer-resolution model. More enstrophy is closer to the truth where the true filed is known as it was shown in the idealised experiment in Sections 2.2 and 2.3. In section 2.2 and 2.3 it is shown that the SDD method is significantly more efficient in recreating smaller scale features that common interpolation methods such as bi-linear or bi-cubic. It is correct that the SDD method is designed mainly to improve the eddy-permitting models where the smaller scale structures such as mesoscale eddies are only embryonic.

We thank Reviewer 2 for helpful comments.

Responses to community comments on the manuscript 'High resolution stochastic downscaling method for ocean forecasting models and its application to the Red Sea dynamics'

Comment. The mentioned "double penalty" effect should be expected in downscaling due to the fact that the coarse model has first of all a coarse bathymetry compared to the higher bathymetry of the downscaled model. Therefore, the spatial displace of hydrodynamical features should be appeared especially in areas with not smooth bathymetry.

Response. It is an interesting thought. We have found that more detailed bathymetry and the coastline result mainly in the differences in the area integrated vorticity. Clarification is given in lines 420-423 of the revised MS.

In our calculations the double penalty was more of traditional nature as explained in Crocker et al. 2020: 'If the location of a feature in the model is incorrect, then two penalties will be accrued: one for not forecasting the feature where it should have been and one for forecasting the same feature where it did not occur (the double-penalty effect, e.g. Rossa et al., 2008). This effect is more prevalent in higher-resolution models due to their ability to, at least, partially resolve smaller-scale features of interest.' Clarification is added in lines 322-342 of the revised MS.

Comment. I propose to the authors to look the recent Red Sea paper here below and add the relevant citation: Hoteit, I., Abualnaja, Y., Afzal, S., Ait-El-Fquih, B., Akylas, T., Antony, C., Dawson, C., et al. (2020). Towards an End-to-End Analysis and Prediction System for

Weather, Climate, and Marine Applications in the Red Sea. Bulletin of the American Meteorological Society, 1-61. https://doi.org/10.1175/bams-d-19-0005.

Response. We are thankful for bringing our attention to the paper by Hoteit et al (2020). It is an interesting and comprehensive paper. It covers a wide range of topics many of which are outside the scope of our manuscript. We plan to work on some of the issues covered in this paper and add relevant discussions and citations in our future publications.

We thank Dr Zodiatis 2 for helpful comments