

Interactive comment on “Design and validation of MEDRYS, a Mediterranean Sea reanalysis over 1992–2013” by M. Hamon et al.

Anonymous Referee #2

Received and published: 19 October 2015

C910

19 October 2015

General comments

The paper presents the description of MEDRYS reanalysis covering the time period October 1992 - June 2013 and its validation results, that considers also the NM-12 hindcast. A long part is dedicated to the description of the atmospheric forcing ARPERA, a dynamical downscaling of ERAInterim reanalysis from ECMWF.

The paper presents old and not precise bibliographic references in the introduction and it does not mention the Mediterranean Sea reanalysis available through the Copernicus Marine Service. There is no the objective of the paper in the introduction and what is its substantial contribution to scientific progress. The scientific approach is superficial and applied methods are not described properly in a dedicated section but inserted in the analysis of the results. For these reasons the paper looks more like a technical report and I suggest to include some scientific analysis, that you already started as mentioned in the conclusion.

The part of the atmospheric forcing could be highly improved with a more precise explanation of the methodology you applied to validate ALDERA and its comparison with other atmospheric model outputs or observational datasets. However you did not prove the added value of introducing the ALDERA forcing instead of ERAInterim in NM12 results, but you only gave it as obvious.

The ocean model description lacks some fundamental detail.

C911

The paper misses a clear description of the validation methodology (one of the topics of the paper) that is hidden between the lines of the results. The data sets considered for the data assimilation and the validation are not the state of the art of the available data sets in the Mediterranean Sea. The author might improve the quality of the paper at least using them for the validation.

The use of a low quality climatology computed for 1998 as a reference for the MEDRYS validation is weak and does not improve the analysis results.

The sections about the high frequency validation should be more robust considering different stations in the basin and covering a longer time period.

It is hard for the reader to follow since there are many missing details given as obvious. In many parts there are repetitions that denote a superficial approach. The repartition in sections and subsections might be simplified. Many figures might be improved in the quality or redone excluding data that are not statistically significant. Captions contain details that are missing in the text. All the figures and subplots should be introduced in the text and motivated before starting the analysis of their content.

Overall the english is good quality even if there are some typing errors that should be corrected.

I suggest a major revision of the paper before publication.

Specific comments

page 1817, lines 7-8: heat loss reference is really old, more recent papers should be referenced. Mariotti 2008 is not the right citation for the the fresh water budget since I could not find it anywhere. Please revise this part of the introduction with updated literature. (Later you use other references)

page 1817, lines 11-25: This part of the introduction is very superficial and not pertinent with the paper. You say : "we must improve our understanding of the water
C912

cycle ..." but the paper does not.

page 1818, lines 7: "Several techniques..." Which ones? Insert some reference here. Please explain better why regional reanalysis are challenging because observational data sets are scarce(r?). You should also specify that the high resolution model needed for the surface fluxes are atmospheric models. The advent of atmospheric reanalyses covering a long time series continuously updated to present time has been a first limiting factor together with others that you should mention. All MyOcean/Copernicus regional reanalyses (exempt the Baltic) did not use high resolution atmospheric forcing and were accepted.

page 1819, lines 8-20: Oddo et al. 2009 does not present any reanalysis. The author should not cite a paper without reading it. What follows should be revised. You did not mention the Mediterranean Sea reanalysis available in Copernicus since April 2014, I suggest to include it here.

page 1819, lines 25-26: ..."giving a close resolution to of NEMOMED16". What is NEMOMED16? Please revise the english.

end of page 1819 The time consistency of the atmospheric forcing is not the only factor that determines the time consistency of an ocean reanalysis, please revise this part.

What is the improvement of MEDRYS reanalysis with respect of the previous Mediterranean Sea reanalysis? Considering your approach in the introduction you might include some considerations in the conclusions. What is MEDRYS added value?

page 1820 The title "Experimental set up" is not appropriate you never talk about experiments.

page 1820 Why did you title a section "Ocean model configuration: NM12-Free" to describe your reference hindcast? Up to now you mentioned this hindcast only in the abstract. Is the MEDRYS model configuration the same of the free run? What are the lateral boundary conditions? Is it the model nested within a global model, how? Please revise this section and include some more details that are not obvious for the reader.

page 1821 line 9 What is MEDATLAS1979? How did you compute it? I believe that a citation of Rixen et al 2005 here is not enough since you computed a different product. What is ORAS4? How did you blend the two data sets? The following description of the buffer zone is not clear to the reviewer, could you please specify it?

page 1821 line 19 How many rivers did you implement totally? What does it mean that the dataset of Ludwig is split in two parts? ..."the others (how many?) rivers are gathered in each sub-basin (which one?)...The Black Sea is not included in the NM-12-FREE configuration but..." What about the MEDRYS? The Marmara Sea is in your domain? Please explain better.

page 1822 line 20 What are bi-periodization (11) and the relaxation (2x8) zones?

page 1823 lines 5-6 .." the sea ice limit (Black Sea)..." what does it mean??? Why sea ice is needed? The reader should not go on the MEDCORDEX webpage and look for the definition of the cordex domains and then the Mediterranean. A figure of the MEDRYS and ALDERA domains with LION buoy location and others geographical references mentioned in the text would be useful for the reader.

page 1823 lines 10-end What is the bulk formula used in ARPERA2? ..."To our knowledge..." I would avoid this kind of statements. I would use consistent instead of

C914

"homogeneous" which is misleading. Which conservative interpolation scheme? The last two sentences are not pertinent here, moreover you mention 2 tables and two figures without any explanation.

Section 2.2.1 should start with the second paragraph, which is the continuation of the previous section! I suggest to describe at the beginning the data sets you consider and the methodology. I suggest to revise all this section and the table captions, they contain informations that should go in the text.

Section 2.2.2 The author statement at the end of page 1824 and the beginning of page 1825 is approximative maybe some analysis could be added. An introductory phrase to justify the sub-section title could help the reader to follow. I cannot understand how the phrase ending at line 5 of page 1825 is linked to the previous one. At line 6 you talk again about inter annual variability of the heat flux components. This is confuse, there are numbers that the reader cannot follow, are you still referring to table 1 or 2? Please revise it.

page 1825 lines 17-end You introduce the trends in the last sentence of the section, without any explanation of the methodology or without showing any plot. I suggest to insert some plot to compare ARPERA fluxes components to ERAInterim, which is the parent model.

Section 2.2.3 This section is superficial. The last phrase is in contrast with the title and maybe it is the only needed. I'm not sure about the usefulness of both figure 1 and figure 2, they show the same thing.

Section 2.4 Why didn't you use high resolution SST products specific for the Mediterranean Sea free available in Copernicus marine service catalogue? You developed a 12km horizontal resolution atmospheric forcing but you assimilated a 0.25deg resolution SST which is well known in literature it is not the best product for the

C915

Med Sea, please justify it. The Copernicus marine service provides delay mode and reprocessed satellite SST datasets specific for the Mediterranean Sea that have been created for reanalysis and validation purposes, why didn't you consider them neither for assimilation and validation purposes?

page 1828 line 11 Which specific along-track SLA did you use? Reprocessed, delay mode, real time? Please specify better which data set you used, there are many available.

page 1828 lines 18-end Why didn't you use the existing specific and higher resolution MDT for the Med Sea? Ex. Rio et al 2007, 2014? Your product is at 1/4 of a degree, very smooth. Please justify it. I checked the CNES-CLS09 documentation and it looks like it does not contain the Mediterranean Sea, please see AVISO website here and clarify. What is GOCE? Which reanalysis data are used to adjust your mean surface reference?

page 1829 line 5 How many data have been discarded from observation thinning, which is the percentage referred to the overall dataset? Why did you apply a data thinning?

page 1829 line 10 How many observations have been discarded from the further SAM2 check? Which climatology do you use? Seasonal climatology check might discard a lot of useful observations.

Section 3 I believe that the first sentence is not appropriate for a scientific paper. Which is the methodology you applied? You should also start saying that you use the hindcast in your validation. The title of the section is Validation and scientific assessment, but you start immediately talking about the assimilation statistics. Innovation is not the observation minus the forecast, could you please insert a formula?

C916

Then you mention Crosnier and Le Provost paper, containing Class1-Class4 diagnostics developed in GODAE. You should describe which one you use and you are going to present together with the formula and their meaning. Please insert here a comprehensive description of the validation methodology and the reference data sets you use and why.

page 1829 last phrase I believe that you cannot say that you assess a 21 years of reanalysis at high frequency only showing 2 months of hourly data (since one month of data are not trustable due to sensor failure) at one location in the Mediterranean. This is a weak part of the paper, I suggest to add more stations in different part of the domain and provide some comprehensive statistics.

page 1830 lines 5-16 Could you explain better why the linear decrease of the mean innovation should be related to your volume correction? Could you explain why the seasonal signal should be related to the runoff? What do you mean that the RMS of observations is 8 cm? Which observations, referred to what?

page 1830 line 20 You are showing that almost none in situ SST value has been assimilated, why??? Is this because of your thinning procedure or the climatological check?

page 1830 line 25-27 If the statistics are not significant before 2005 why do you show them? I recommend to split Fig.4 for satellite and in situ validation. The bias between in situ and satellite data could derive from your SST satellite data correction? Before applying the satellite SST correction did you compare it with CORA4 in situ data?

page 1831 lines 3-4 The statement that there are few profiles deployed in the

C917

Mediterranean Sea before 2005 is wrong and misleading, especially if not supported with numbers. You should again explain why your assimilation or preliminary quality check or thinning procedure discard almost all in situ data.

page 1831 I would avoid repeating the word innovation, mean innovation should be the MEDRYS bias. I suggest to revise this part that is confusing and superficial. What are the maximum and minimum temperature RMS values? FIG 5. is misleading since the color palette is saturated and hides the maximum values. RMS should be reasonable with the bias by default, what about the relationship with the specified observation error? Please explain it better. Did you consider the significance of the statistics?

Section 3.2 Why did you select a global product at 1/4 of a degree to validate the model MSSH? Moreover you use CNES-CLS09 in your assimilation of SLA, thus this is not a validation but rather a verification of our assimilation procedure. I believe you should repeat this part with a Med Sea specific product. Moreover you spend 14 lines talking about the free run and just 3 for the MEDRYS. I suggest to focus on the MEDRYS which is the topic of the paper.

page 1833 line 17 The author should introduce Fig.8 and the analysis of the mean kinetic energy. The outcome is in lines 25-26, the assimilation introduces smaller scale features increasing the energy, as expected.

sub-section 3.2.2 The author should introduce EN3 and IMEDEA reference data sets in section 3. EN3 gridded products is not climatological, please clarify it. Why didn't you use EN4? What is IMEDEA reconstruction? The reference is not in the bibliography and the reviewer did not find it on the web! Lines 8-9-10 are a repetition and demonstrate a superficial approach. Fig. 9-10 starts from 1980, why? I suggest to focus on the MEDRYS time period which is the topic of the paper. Did you consider

C918

the mapping error?

The author should investigate more deeply the salinization in early 2000s, is it due to stronger evaporation, as you state at line 1 of page 1834 or is it a spurious positive anomaly (line 8) due to the fact that in 2000 you start assimilate a lot of temperature profiles? What about the temporal consistency of MEDRYS since the number of data assimilated varies enormously during the considered time period?

Considering what you say at line 15 I believe that your validation on the deepest layer is not solid.

sub-section 3.2.3 Here you introduce a MEDATLAS1998 climatology computed over three years of time period as a reference product for a MEDRYS validation. I believe that this is not a good reference field, since it has to be very smooth due to the lack of data. What's the meaning of this choice? I would avoid this and concentrate on the CLASS4 diagnostic analysis. Could you briefly specify in the validation methodology what is a CLASS4 diagnostic and the different meaning with the innovation diagnostic previously presented? The formulas are welcome.

page 1835 line 22 You say that MEDRYS is very close to observations, what is the mean MEDRYS RMSD? CLASS4 is meant to measure the performance of the MEDRYS and its capability to reproduce the ocean state consistently with observations. Does it? Which observations did you use for this validation? This should go in section 3 or at least at the beginning of this subsection.

Each figure and subplot should be introduced, referenced, motivated and analyzed. Please revise the text accordingly. The text is confusing the reader, I suggest to describe fig 11, temperature bias and RMS then fig. 12 salinity bias and RMS. I suggest to introduce a table summarizing mean bias and RMS with the relative standard deviations to demonstrate the MEDRYS performance. A comparison with Copernicus reanalysis results could be useful for the reader and to show the added value of MEDRYS.

C919

Why the number of data is so high (O104) in the deep layer? Are those sample profiles? Numbers are not clear please enlarge the font.

page 1836 lines 10-13 Your results suggest that the salinization is not in the gridded reference products (which are at this point) but in the dataset you use here. Please clarify this aspect.

page 1836 lines 15-end The analysis of the 6 RMS time series should be more detailed. Compare the RMS of MEDRYS to MEDATLAS98 does not make sense to me and I recommend to take it out.

page 1836 lines 26-27 You assess the ability to reproduce the high frequency surface temperature and salinity at one location over a 2 months period. This sub-section should present the same statistics for other fixed stations in other parts of the domain over a more extended time period, results might be very different in other parts of the domain. Where is the lion buoy? I suggest to take out the part of the time series affected by the sensor failure and present only bias and RMS (please add RMS) over the trustable period. Please revise accordingly. The anomaly could show better the difference between the two model solutions in figure 13.

Section 3.2.5 The first phrase up to line 6 of page 1838 should go in the introduction. Formulas could explain better how you compute the transports.

page 1838 line 15 The reader should know approximately the ranges of values present in literature, could you please include it?

page 1838 lines 22-23 Temperature and salinity are consistent since your relaxation to ORAS4 in the buffer zone, I would not use this to corroborate your results. I would also avoid to compare MEDRYS to a simulation without including details of it,

C920

are they relaxing to ORAS4 temperature and salinity as well?

page 1839 lines 19-20 The referee does not agree with the statement that MEDRYS uses all available observations of CORA 4 dataset, you excluded a lot of observations without providing a clear explanation in the text. Please clarify this aspect introducing more details in section 2.4 and here.

page 1839 lines 25-end The time consistency of ALDERA does not assure the time consistency of the MEDRYS since there are other factors affecting it that the authors should take into consideration. Please clarify this aspect.

page 1840 lines 1-7 I do not agree, the verification of the MSSH presented in section 3.2.1 says that your assimilation methodology modify the solution with respect to the reference hindcast. Moreover the reference field you selected is a coarse resolution global product used in the assimilation and not an independent and specific one. The choice of the reference fields for the validation of integrated temperature and salinity is questionable, as you report also at page 136 line 3, and not documented in literature (IMEDEA reference is missing).

page 1841 lines 1-15 You present here preliminary results of additional work, this is not needed.

page 1841 lines 16 In spite what you say I suggest to repeat the analysis in this paper with the best version of MEDRYS and to include some of the scientific results you already identified in the Med-CORDEX/HyMeX framework to improve the quality of the paper that now looks more a technical report.

Technical corrections

C921

Abstract

page 1816, line 13: you talk about "The first version" of MEDRYS but the paper never talk about different versions I suggest to eliminate "first version".

page 1816, line 17: I would eliminate the word "simulation", it is misleading, MEDRYS is a reanalysis and what you call free simulation is an hindcast.

page 1816, line 19: ... "at intermediate layers" instead of "in"

page 1816, line 21: please correct the \tilde{w} with the correct symbol

page 1819, line 20: Please start a new line.

page 1822 line 24 Please take out "conditions".

page 1822 lines 23-end Please revise this sentence. The last line is a repetition you already said that ALDERA covers the period 1979-2013.

Section 3.2 I would eliminate subsections titles 3.1 and 3.2 that repeat the same titles of the parent section 3, they are redundant.

Fig. 9-10 start from 1980, why? I would insert a legend. The central blue line should be solid, I suggest to redo it only for the period of MEDRYS. In some figures MEDRYS becomes MEDRYSV1, please correct. In figure 13 the legend should use the same acronymous used in the paper.

Fig. 14 should be a table.

C922

page 1839 lines 22-25 I suggest to rephrase.

C923