RC3: 'Comment on essd-2021-462', Anonymous Referee #3, 16 Mar 2022 reply

The authors present a new Mediterranean Sea regional SST product that reproduces thediurnal cycle. For this, the authors merge the SST from the CMEMS Mediterranean SeaPhysical Analysis and Forecasting product with the SST measurements from SEVERIremote sensor, and they apply a methodology that is presented in Marullo et al. (2014). For assessing the actual capability of the resulting SST product to properly capture the skin SST variations, the authors use a set of drifting buoy SST measurements that are typically acquired at 20 cm depth. This is a clear limitation of the assessment of, not only this product, but all the satellite products that aim at reproducing the skin dynamics, there is not in situ data to compare with. In the absence of in situ skin SST measurements, the quality assessment that the authors present here is clear, and they provide evidence that the product is properly capturing the diurnal cycle, or at least that it is capturing it better than the model. So, I think the manuscript deserves its publication in the Earth Science System Data journal.

I have some minor comments /questions to the authors.

Line 15-16: "The differences between satellite and model SST are free,or nearly free, of any diurnal cycle"->I don't understand this I though model does not reproduce the diurnal cycle while the satellite does

Both the model and the satellite reproduce a diurnal cycle, but the two cycles are not identical. If they were, their difference (our SST anomaly) would not contain any diurnal signal. In the real case a small difference (mainly in terms of amplitude) still exists. This implies that the SST anomaly contains a small diurnal component. In this sense we can say "Free or nearly free", consistently with fig. 2 of Marullo et al. (2014).

However, we do agree that that sentence is rather unclear. It has been removed and re-written as follows: "The choice of using a model output as first-guess represents an innovative alternative to the commonly adopted climatologies or previous analyses, providing physically consistent estimates of hourly SSTs in the absence of any observation or in situ measurement". We also added Section 3.2 in the revised manuscript, which better introduces the optimal interpolation method and the choice of a model output as first-guess.

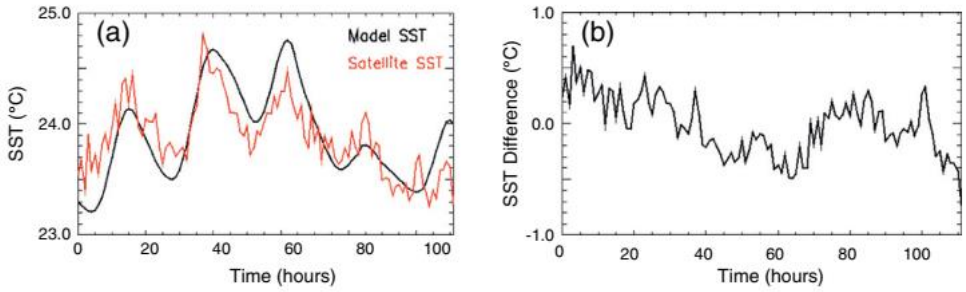

Fig. 2. Example of model first-guess and original SST SEVIRI data over a point in the Mediterranean Sea. (a) Model (black curve) and SEVIRI (red curve) SST time series. (b) Difference between the SEVIRI and model SST.

Line 17: I'm wondering whether these drifting buoys are assimilated in the model or not.

From the model documentation, available via the Copernicus Marine Service website (https://resources.marine.copernicus.eu/product-detail/MEDSEA_ANALYSISFORECAST_PHY_006_013/DOCUMENTATION), one can see that the assimilated variables are:

- In-situ vertical profiles of Temperature and Salinity from ARGO and XBT;
- Sea Level Anomalies (SLA) from available satellites Jason 2 & 3, Saral-Altika, Cryosat; Sentinel-3A/3B.
- Objective Analyses-Sea Surface Temperature (SST) fields, used to correct surface heat fluxes.

Therefore, drifting buoys can be considered as a fully independent validation benchmark.

Line 93: It would be interesting for the reader a comparison between the performance of this skin SST OSTIA and MED DOISST.

We do agree with the reviewer and thus added the global operational diurnal L4 SST OSTIA product in our intercomparison exercise. In particular, as OSTIA ingests the in situ data we used as reference for the validation (which would not be independent for OSTIA), we included OSTIA diurnal in section 4.2.2, which is dedicated to the reconstruction of diurnal warming amplitudes (DWAs) from different sources (DOISST, Model, SEVIRI and In situ data). The first part of the validation (section 4.2.1) is indeed mainly thought to assess the accuracy of the DOISST product against an independent in situ data source, and the inclusion of modelled SST data is thought to evaluate the DOISST performance with respect to the model, which is used as first-guess.

Line 106: Do the authors plan to extend the temporal series backwards?

No such a plan at the moment. Future evolutions will mostly depend on users' feedback.

Lines 128-130: I don't understand this. Why are the differences between SEVERI SST and drifters larger during nighttime than in daytime? I would expect larger differences during daytime because drifter measurements are acquired at 20cms depth and SEVERI measurements are provided in the first mm. Are these differences reflecting in first order the radiometric errors of SEVERI?

First, these validation results were produced by OSI-SAF, the data provider. From our side, we do not actually see large differences between nighttime and daytime statistics. Indeed, the bias is practically identical during nighttime and daytime, resulting in -0.1K and -0.09K respectively. This means that the bias remains unchanged during the 24 hours, thus revealing a good stability of the SST retrieval. Similar considerations are for the RMSD with the only difference being that the error during daytime (0.56K) is slightly higher than that at nighttime (0.53K).

Line 166: Delete " " before "."

Deleted.

Lines 188-191: I don't understand this paragraph: 1) Why are you using differences between satellite and model instead of satellite measurements directly? I don't understand the point of the reduction of one order of magnitude of the difference. 2) Do you mean that for generating hourly products you are considering all observations around the model in +/- 24 hours? Have you assessed the impact on the final product of considering different (reduced) temporal windows?

1) The optimal interpolation (or statistical interpolation) method determines the optimal solution to the interpolation of a spatially and temporally variable field with data voids, where "optimal" is intended in a least square sense (see e.g. Bretherton et al., 1976). Optimal interpolation requires two datasets, observations (as satellite-based SST measurements) and first-guess estimates (as model output SSTs). The optimally interpolated variable, or analysis (Fa), is obtained as:

$$F_a(x,t) = F_b(x,t) + \sum_{i,j=1}^{n} W_{i,j}(F_{obs,i}(x,t) - F_b(x,t))$$

In practice, the analysis F_a(x,t) at a particular location (x,t) (in space and time) is obtained as a correction to a background field $(F_b(x,t))$, estimated as a linear combination of the observation anomalies ($F_{obs}-F_b$), where the coefficients $W_{i,j}$ are obtained minimizing the analysis error variance.

The reduction of one order of magnitude means that the difference between satellite and model SST (our SST anomaly) contains only a small amount of diurnal signal.
These concepts are now included in the Section 3.2 of the revised manuscript.

2) Yes, each hourly product is generated by using 24 SEVIRI L3C SST maps, following the approach proposed by Marullo et al. (2014). This choice was actually the result of several trials, and this has been clarified in the revised text (line 307).

Line 203: I would specify here also the model spatial grid.

This has been corrected

Line 204: I would specify at which grid the regridded is performed.

The remap is performed over a 0.0625° regular grid. We added in the revised text the reference to Table 2.

Line 256: Estimates of the correlation with in situ may also provide useful information.

We added estimates for correlation coefficient (see Table 3 in the revised manuscript).

Line 258: Have you assessed SEVERI SST? It would be interesting for the reader the comparison between SEVERI and MED DOISST performances (not only in the DWA).

This is a good comment. We decided not to include, at least in this subsection, a comparison with SEVIRI for different reasons. First, the inclusion of modeled SST data in the first part of our validation (section 4.2.1) is mainly thought to evaluate the DOISST performance with respect to the model, which is used as first-guess. This is the same reason for which we included OSTIA diurnal in the second part of the validation (Section 4.2.2). Then, the inclusion of SEVIRI data would have reduced the number of matchups since these data present gaps (data voids). Additionally, the validation for SEVIRI data has already been performed in two previous papers, namely in Marullo et al. 2014, 2016, and we often refer to these papers. However, we added in section 2.1 (lines 142-144) the main results obtained from Marullo et al. (2016) that quantify the bias and RMSD for SEVIRI SSTs over the Mediterranean Sea, and compared them to those obtained for DOISST in the Summary and Conclusions (lines 554-556).

Line 262: I would say pointwise difference.

Corrected.

Fig. 2: Perhaps it would be interesting to separate the map into daytime and night time.

Interesting suggestion. We separated the computation between day and night but the two corresponding maps are practically identical, indistinguishable. This might not be so surprising since the bias between DOISST and in situ temperatures is always very low, both during nighttime and daytime (see Table 4 of the revised manuscript). For this reason, we think it is preferable to keep just a single map.

Line 265: "tendency"->"predominance"

Corrected.

Fig 6. Is interesting that although the dispersion of DOISST DWA around Drifter DWA hasbeen significantly reduced with respect to the one of SEVIRI DWA, the maxima DWA events seem to be better captured with SEVERI than with DOISST (that they seem to be a bit underestimated).

Yes, this is quite reasonable since SEVIRI provides a direct measurement of sub-skin SST while DOISST performs a blending of these temperatures with modelled data.