The authors present a new Mediterranean Sea regional SST product that reproduces the diurnal cycle. For this, the authors merge the SST from the CMEMS Mediterranean Sea Physical Analysis and Forecasting product with the SST measurements from SEVERI remote 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.

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

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

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

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 20 cms depth and SEVERI measurements are provided in the first mm. Are these differences reflecting in first order the
radiometric errors of SEVERI?

Line 166: Delete “ “ before “.”

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?

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

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

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

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).

Line 262: I would say pointwise difference.

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

Line 265: “tendency”->“predominance”

Fig 6. Is interesting that although the dispersion of DOISST DWA around Drifter DWA has been 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).