This paper documents the main developments that led to the production of the ORAS5/OCEAN5 global ocean reanalysis/analysis system and provides an assessment of this product concerning main key climate parameters. A lot of work has been done to reach this point. The paper is well-written, and being ORAS5 a state-of-the-art ocean reanalysis system that will likely be widely used, I recommend publication of the manuscript after a few, mostly formal, issues are addressed by the authors. The paper will be of great interest for both reanalysis developers and users.

General comments

1) I found the organization of Section4 a bit misleading: i) section 4.1 cannot be really considered part of assessment, it concerns OSEs performed with a low-resolution configuration, without bias correction nor altimetry asssimilation and for a limited period. I don't think it is really relevant for assessing the high-resolution ORAS5 system. I suggest moving it in an Appendix and summarizing the main outcomes in Section 2.3.1 rather than 4. ii) Sensitivity tests (section 4.2 4.3) could be presented in section 4, and start a new Section 5 about the Assessment strictly speaking. iii) Sea/ice section (4.4.3) can benefit of having a comparison symmetrical to 4.4.1 and 4.4.2, namely showing ORAS5 vs ORAS4 and control runs, rather than ORAP5. I think homogenizing the assessment improves its clarity.

2) In many parts of the manuscript, ESA CCI data are considered independent verifying data. I don't really agree with that, since all sensors used by ESA CCI (infrared AVHRR, PMW, altimetry radars) are also at the base of the observational datasets assimilated by ORAS5. This is clearly testified by Figure 4 (for SST) and Figure 22a (for SLA). Suggest dropping the mention to “independent” and consider these datasets as “reference climate data” or similar. P1L12, P29L4, P32L9, P37L32 etc.

3) The developments in Section 2 are often corroborated with tests, each of them performed with different configurations, sometimes even different resolution than the nominal ORAS5. Suggest introducing Section 2 by mentioning that there exists no warranty that the “sum” of improvements leads to the “best configuration”, but obviously this is the standard and only possible procedure (or similar concept).

Specific Comments

Abstract:
L7 (and in P3L7 and Table 1): “1979 onwards, extended to 1958”: for a reader it is not so easy to understand why you don't just say “from 1958 onwards”. If you consider 1958-1978 part of the initialization/spinup strategy, perhaps the backward extension should be drop, or just consider one entire timeseries? Otherwise seems there are two independent streams of production. Better to rephrase and clarify.

L10: “analysis error” never really considered, strictly speaking? Perhaps better to say “reanalysis-observation mismatch” or similar

Intro
P2L23: maybe not important the funding (this should go in the Acknoledg.), but that ORAS5 is part of the C3S service and envelop of products. Is it not also part of CMEMS?

Section 2
P3L26: I guess “observation equivalent background fields” otherwise sounds weird

Table 2: suggest adding the “year of initialization”, explaining in the caption the “capping” and describe qualitatively the meaning of “latitudinal decay”, is it the bias-correction correlation length scales, in units of degrees, latitudinal bands?

Figure 2: Not sure whether the 1975-1988 cooling is realistic or an artifact of the initialization, and likewise the following warming (amplified by the previous cooling?). Perhaps discussing the cooling/warming, also in terms of W/m2, could help the readership to see if this globally integrated signal is trustful, or only upper ocean is trustful?

P7L2 You mean “SST, SSS observations” I guess, if so better to specify.

P8L2 converting in days the SSS restoring term, as for SST, could help

P8L24: Sure ESA SST CCI doesn't use drifters/buoys for calibration?

Section 2.2 and 3: For climate monitoring applications, it will be very beneficial if from time to time (e.g. once per year) the system is rewind and delayed time data are used instead of real-time data as from 2015 on (EN4 and ERA-Interim instead of GTS and NWP). This will produce time-consistent time series not only till 2015. Are the authors considering this? Do they consider the reanalysis strictly speaking ending in 2015? Maybe you could add a sentence about that

Table 3: worth to say if there have (not) been issues with the different sea-ice mask in OSTIA and HadISST2 used for SIC and SST relaxation, respectively

P11L4 worth to add which EN4 data quality flags are used to ingest data, ie all available or only very good quality data?

Figure 8. Different sign of bias (ORA4 vs ORA5) could suggest bias coming mostly from vertical physics rather than forcing, which is the same among the two reanalyses? If the authors have any speculation could be worth adding it.

P17L6 Is a typo and should be 1993-2012, or is there a reason to start from 1996 instead of 1993 (I guess it doesn't really matter though)

P28L18 “This is expected...” This sentence seems to implicitly underline that the majority of RMSE comes from bias^2 and not (standard deviation of innovations)^2, although the authors do not quantify it (just RMSE shown). Since it is probably not the case, it seems to me an over statement. If so, probably better to drop it, unless I miss something.

Section 4.4.1. Comparing Figure 19c with 20a, it seems the main differences of ORAS5 vs ESA CCI comes from the SST dataset ingested for the largest period (HadISST). Perhaps would be worth discussing in more details this aspect, or even showing ORAS5 minus HadISST?

P32L25 this also seems an over statement to me: if the column-integrated density variability is well reproduced in those areas, it doesn't mean they have the smallest errors in general

P34L7 Would you speculate that it s because ¼ degree resolution is still not high enough in the extra-tropics?

Fig 24 maybe the same color palette as Fig 23 helps comparing the two figures.

P37L20: is not “atmospheric analysis” and “NWP forcing” exactly the same?

P38L27: Not sure an observed MDT (SSH and geoid) will be the best for reanalyses, because as the authors showed many times in past works it would lead to unrealistic and abrupt drifts in ~ 1993. Anyway, just a personal comment.

Could be useful in the conclusions to summarize some future directions of the ORAS as already mentioned in the text (ERA5? Stochastic physics? Retuning of BECs?).

Typos

P1L11 system experimentS
P1L12 carried out FOR
P1L16 which ARE possibly
P2L1 improvementS
P3L11 BARNIER (and not Bernard)
P3L15 visco-plastic?
P8L5: remove brackets from “(Titchner...” reference
P8L10 comes FROM HadISST2.1
P11L15 subject
P12L17 due to the assimilation of an evolving....
P21L15 “deliver on...” not sure makes sense, better to rephrase
P37L8 ORAS5 IS a …
P37L25 This result suggestS