Dear editor,

We have difficulty identifying which discussions are missing in our rebuttal about general or serious objections both reviewers made. For that reason, we list below those that we interpret as general or serious and re-address them in an effort to mitigate any omission in our original rebuttal.

We took this opportunity to update a reference in line 95 and replace a duplicate term in line 265.

Best regards.

João Encarnação et al.

Reviewer 1

General Review: [...] One general comment is that the figures are not of high quality – the legends and axes are all very difficult to read.

We replotted all figures and increased the font size.

Further, there are many places where multiple figures could be combined into a single figure with subpanels.

We re-arranged Figures 4-6 into a single figure (now Figure 3).

It is unclear to me why the authors decided to use this approach for C20.

We followed the reviewers' suggestion and resorted to SLR-derived C20. We choose the model kindly provided by Dr. Bryant Loomis from GSFC because that model is on a weekly basis and meets our needs to interpolate over GRACE/GRACE-FO and Swarm epochs.

The authors state it is a mystery as to why the oceans have larger errors than over land. Is this really true?

We have followed the reviewers suggestion to look at the latitudinal dependence of the errors and added figure 15 to demonstrate it does not change significantly.

Reviewer 2

As a methodological paper, the paper somewhat lacks a hypothesis. What is expected from the outset? That could have been described better.

Our hypothesis is that the Swarm models describe mass transport on a monthly basis at the 1500km spatial scale over land with accuracy comparable to GRACE (at the same spatial scales). It is mentioned in the abstract as well as in the conclusions.

What message is conveyed in view of other LEO missions that could be used for gravity retrievals?

The Swarm satellites are not dedicated gravimetric satellites. They are equipped with sensors that have particularities that make that application possible but less than ideal. Even with those characteristics, we are unaware of other non-dedicated LEO satellites that are able to observe Earth’s temporal changes in gravity. Additionally, going deep into this discussion does not address any characteristic of the models we are presenting, which is our main objective.

Should one have different orbits, instruments, what did we learn now for the next LEO mission?

The topics the reviewer suggests is well outside a data description article. They rely exclusively on simulation (i.e. no real data) and requires numerous assumptions (e.g. sensor noise characteristics and background force models errors).

The big issue for this reviewer is whether the authors were well-advised to submit to ESSD. [...] Here, the focus is clearly on the methodology of generating the data and neither its use nor reuse, and I guess other journals are more appropriate.

We disagree this is a methodological paper, quite far from that. We mention methodologies in order to illustrate how the data was produced. The data's use and reuse is implicitly to be determined by the audience of this article, who resort it to learn their particularities and the assumptions leading to their production.

While the authors motivate their study with the need of the community to rely on data sets for studying "glacial cycles and long-term trends", the GRACE-GRACE-FO gap is 10 months and this is the period where these data will be relevant, in addition to few monthly gaps.

We recognised that the motivation for the Swarm models was lacking and added the application to Sea-level studies and the improvement of the low-degree coefficients.

It seems like a huge effort and the groups are to be congratulated, but they don't show what Earth Science applications will be enabled now that were not possible.

Swarm is not a dedicated gravimetric mission. Any model derived from Swarm's data is invariably going to be of less quality than any dedicated mission. For the time being, GRACE-FO is providing that data and Swarm cannot compete with it. We took great care in presenting our results in the context of the GRACE/GRACE-FO data quality exactly to illustrate this fact to the users of the data. Swarm gravity field models may not enable new Earth Science applications (as far as we know), but it is providing independent data that can be used for validating models derived from dedicated missions. We demonstrate this in a few hydrological basins and illustrate a few deviations from the climatological model that are observed by GRACE-FO and Swarm. We essentially show that GRACE, GRACE-FO and Swarm are observing the same Earth.

We don't learn from their results for the understanding of processes. In their words, the "consequences of the 10-month gap" are not outlined and it is not clear what we gain.

These are examples of applications for our models, to be undertaken by the user community, as was done by (e.g.) Meyer et al. (2019b), although different models were used.

The other very major problem for this reviewer is that apparently no independent validation is provided, except comparisons to GRACE. We don't have GRACE them for the gap, but the authors could compare their low-degree results to satellite laser ranging (SLR) solutions, e.g. for single coefficients or for the ocean mass change.

We are unaware of monthly global gravity field models, other than those produced from GRACE/GRACE-FO. We choose to restrict our analyses to those in the article to limit the length of the paper. The ocean mass changes has already been object of study in Luck (2018), as mentioned in Section 3.4.2. The topic of observing C20 with gravimetric satellites, in particular, stands on its own; any discussion could not be done in an acceptable way as a sub-

section on a data description paper. It is a good example of an application for the user community.

The authors mention only one SLR time series for C20 but that's just one data set and more exist. Others provided C02 timeseries and it is not clear to this reviewer why they rely on GRACE extrapolatons for this.

We have replaced the C20 with SLR-derived time series, also in response to R1.

Moreover, the authors, e.g. in Berne or Potsdam, could easily use their Swarm models in SLR range residual analysis as a validation but it is not done.

This was done by the various colleagues (references in Table 1) pertaining to the quality of their respective kinematic orbits. Those articles are much more adequate for the analyses the reviewer suggests because they are done at the level of observations (kinematic orbits), which are most compatible with the SLR observations. We know the information from those orbits has been correctly synthesised in our models because of the agreement with GRACE.

And several other validation techniques and data sets were developed for GRACE but nothing is used here. This is somewhat disappointing.

We could not understand to which validation techniques and data sets the reviewer refers. In any case, as already mentioned, the length of the article discourages additional validations.