Review of "*A new global oceanic multi-model net primary productivity data product*", by Ryan-Keogh et al.

This manuscript introduces a new data product that consists of an ensemble mean and associated variances derived from a suite of diverse global satellite net primary production (NPP) models. The data product is generated by application of the various NPP models to an established, merged multi-mission ocean color record that spans the full modern satellite record (1998-2022). The manuscript describes the approach and some of the basic spatio-temporal patterns observed in the product. It is well-written and the graphics are good quality. I recommend this manuscript for publication without major changes. It does not introduce any new science per se, but provides a product that will hopefully be of use to the broader science community. I offer the following points out of general interest and perhaps to better-clarify certain points.

First, the authors acknowledge the need for this product to provide an alternative to other data products already in existence (e.g., those hosted by Copernicus Marine Services or the Oregon State University Ocean Productivity website). However, in doing so, I feel that a massive disclaimer is needed stating that advantages of this 'ensemble approach' may be fully mitigated by combination of estimates of varying quality (as assessed by exercises such as the Primary Productivity Algorithm Round Robin series, Campbell et al. (2002); Carr et al., (2006); Friedrichs et al. (2009); Saba et al. (2011); Lee et al. (2015)). While exercises such as these are not definitive, the community has dedicated tremendous effort to trying to establish the fidelity of these models. This has been challenging and is limited partly by satellite – in situ matchups, but I feel that we were at least converging on a narrative that CAFE>CbPM>VGPM. Personally, I would gladly use the CAFE model applied the OC-CCI dataset as the preferred data product (but I realize not everyone may feel the same :). Perhaps, a bit of a philosophical point, but I think it deserves some discussion. It could easily be placed in the paragraph starting on Line 96 to balance the justification for the current product.

I'm concerned about the global annual integrated NPP values. They are really high, much moreso than we have reported in the original publications and related work. I wonder if the integrals are unduly influenced by a low number of 'spurious' values? They might be easily traceable to something in the input fields (e.g., coastal Chl-a retrievals > 50 mg/m3, or spurious bbp retreivals in the case of the CbPM). It would take more than a pixel or two, but implementing these 'traps' on the input, as well as the resultant NPP, can be important. Also, in the case of the Behrenfeld-CbPM, the formulation inappropriately uses Kd490 to estimate the euphotic depths and mixed layer growth irradiances, both of which can be significantly overestimated in this way. In Westberry et al. (2008), we point out that the global annual NPP is reduced by nearly 2x (from 67 to 35 Pg C) by simply replacing the Kd490 terms with simple Chl-dependent parameterizations (e.g., Morel-type relationships).

Figure 2a, I think you should truncate the CV map at 0.6 or so I order to see more spatial structure? Just a thought ...

Lines 262-265, the convergence or divergence of model NPP over time is interesting. related to the inter-sensor merging of the OC-CCI record. This could be investigated by looking at single missions, with MODIS-Aqua being the obvious candidate because of the record length.

Line 238, typo should read ‘Westberry-CBPM’