

The reviewed manuscript has been clearly improved from the previous version, and authors have addressed all the comments with a very thorough revision. Both the introduction and methods sections have been completed with further details on the previous literature which is the basis of the work. The fact of using X-band VOD instead of L-band VOD has been explained and discussed consciously. The manuscript will be ready for publication after addressing some minor comments that may provide a more complete understanding and interpretation of results, as well as after checking other small (text and spelling) issues:

- Discussion on X-band and L-band is appropriate and very thorough. Figure A1 supports it very well. A last point for improving Figure A1 would be to provide the same analysis only for forests (i.e., a third panel in Fig. A1, similar to panel *a*, only containing forest stations).
- In L. 262 you comment that *“For a region in Europe (5 to 15°E and 46 to 51°N), where we generally did observe an increase in all three performance metrics, we find that for GPPvod mainly winter time estimates of GPP are too high compared to GPPfluxcom and GPPmodis (Figure 5). By adding temperature as input to the model, winter observations are markedly dampened and summer observations are only slightly increased.”* This is a very nice result to show, nevertheless it refers only to one region. It will provide more consistency to your results if you include two or three similar regions and see if the winter observations also dampen when adding temperature to the model. This can be included in the supplement.
- Figure A5: this figure provides an interesting conclusion, as you mention in L. 260. In my opinion, this is relevant enough to be moved to the main part of the paper. I suggest to show it as a second panel in Figure 4. Also, in the discussion, you can emphasize the fact that adding temperature improves the relative distribution of GPP.
- Section 4.4: please, add a sentence to highlight that the leave-site-out cross validation shown in Table A2 confirms that, although the benchmark GPP datasets are not fully independent, this does not impact the results.

Other minor comments are:

- L. 75: “have been demonstrated” → “has been demonstrated”.
- L. 79-81: in these lines you describe Figure A1, but data needed for this figure have not been described before. At least you should include a sentence explaining in which sections the reader can find a description of the data.
- L. 93-98: these sentences are redundant. Can you join frequencies (l. 93-94) with time overpasses (l. 95-97) to avoid writing the satellite names twice?
- L. 177: “pyGAM” → “The pyGAM”.
- L. 272: “GPP data set” → “GPP datasets”
- L. 380: “sensitive” → “sensitive to”