

Earth Syst. Sci. Data Discuss., referee comment RC5
<https://doi.org/10.5194/essd-2022-242-RC5>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on **essd-2022-242**

Anonymous Referee #4

Referee comment on "Flood detection using Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage and extreme precipitation data" by Jianxin Zhang et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-242-RC5>, 2022

Flood Detection Using GRACE Terrestrial Water Storage and Extreme Precipitation

by Jianxin Zhang, Kai Liu and Ming Wang

The paper titled "Flood Detection Using GRACE Terrestrial Water Storage and Extreme Precipitation" by lead author Jianxin Zhang and co-authors explored capacity of the combination of satellite products, the GRACE total water storage and GPM precipitation, for detection of flood events between 60°S and 60°N. The authors have done an considerable efforts for validation of obtained results (flood location and duration) against existing flood records collected by Dartmouth Flood Observatory, MODIS-derived flood product, social media records as well as flood events derived by the authors from the GRDC river discharge records. The idea is interesting and the theoretical workflow seems to be consistent with the objectives.

My major concern with the paper is the lack of clarity and consistency in the description of the datasets used and in validation of new flood product. The results of validation of the GRACE derived product are mostly qualitative and does not allow for evaluation of the overall quality. Moreover, the structure of sentences, unnecessary repetitions and the vocabulary used make the reading and understanding difficult. I would recommend to re-write the manuscript with the help of a native English speaker who is also an expert in the Remote Sensing. I am also not convinced of the value of the flood event product of such a low 1°x1° spatial resolution.

Overall, I think the paper needs a major revision.

General comments.

Please, re-write the description of the datasets and products used. Some sentences give an impression that the authors did an additional transformation of GRACE and MODIS products, but I could not understand what was done by teams elaborated/provided the products and what was added by the authors of current manuscript. What was the source of the GRACE daily product? Please, provide the link. Does the link for the DFO flood dataset exist? The GRDC discharge data description is inadequate. The social media flood events database description is given in the Result section and its description lacks the details (number of recorded floods, spatial and temporal coverage etc). I also advice to pay an attention on the titles of sections (for example, sec. 2.1, 2.3) . The section 3.3 needs to be carefully re-written. In the section 3.5 an important information about FloodR is missing (approach used in the package, realisation, validation accuracy). Moreover, the authors often confuse the concepts of method, approach and product. For the citations, the use of surname, name and affiliation followed by (Name et.al., 20XX) is not a common practice. Please, consult the ESSD journal citation model.

Several software, tools or codes were used for processing of data. The description of these tools is inappropriate. A short physical/ mathematical description of mentioned parameters (such as "t.window" and "s.window", "direction") is required.

Regarding the validation of the obtained results, the authors compare the number of detected GRACE-precipitation derived flood events with the DFO database and provide the figures and some light statistics. They also refer to the comparison of their flood event retrievals with the MODIS-derived dataset. A large archive of figures is downloadable as supplementary materials, however the statistical evaluation of comparison of GRACE and MODIS flood events is missing in the text. I would appreciate a summary table with statistics for comparison with all 4 reference products (DFO dataset, MODIS flood events, discharge-derived flood events and social media flood records). Important validation information is scattered throughout the text, but the quality of the language does not allow to understand well the authors' logic in many subsections.

Specific comments.

Line 55. Repeated information.

Lines 63 and 66. Replace the word "range".

Line 68. Repeated information

Line 72. " Much useful image information" - be more specific. What kind of information is missing?

Line 78-79. What the " multiyear flood observation data" is ?

Line 81 and 85. Specify the method in the "this method"

Lines 112-115. The sentence is too long.

Lines 120. What does it mean " under the GRACE grid coverage" ?

Lines 134-137. The sentence is too long. What is the "Otsu" ?

Line 137 and in all other places. Please, replace the phrase "extraction results" with more specific terms.

Line 154. The sentence is too banal.

Line 158. Find another term for "preliminary possible flood dates"

Line 159. Not clear the comparison of what with what was done during the validation described here. If I understood well the section 2.3, the flood extent is not provided in DFO dataset.

Line 168. The phrase "to obtain the season" is the scientific slang.

Line 169 . Please, explain what the "high-frequency seasonal data" is?

Lines 198-201. Please, provide the physical meaning for term "direction". The phrase ..."direction" was set to the position..." is not clear.

Lines 203-205. Repeated information.

Line 206. Rephrase the first part of the sentence.

Line 199. The issue with the missing of the high-frequency signal in the daily GRACE product needs more details in the section 2.1. and some discussion in the Discussion section.

Lines 222-223. Please, do not repeat the banal phrases in the beginning of each section, provide more specific for each section information instead.

Line 224. The floods are not "affected" by precipitation, but can be "caused".

Line 234. Please, replace the word "separation"

Line 240. What is the flood "baseline"?

Figure 5 and in the text. Replace the "polygon feature" with the name of parameter/phenomena represented by these polygons.

Line 279. Some part of the sentence is missing.

Line 287-288 and the Figure 6. are not clear.

Line 347. "...the spatial average of the discharge data under the coverage of the HydroSHEDS Basins Level 4 data". The spatially averaged discharge is something new in hydrology. I would ask the authors to provide a reason to invent this parameter, explain its physical meaning and give a solid base for its application in the context of the study. For me the "spatially averaged discharge" is something meaningless.

Line 349. The estimation of POD is not clear. This is an important part of the study related to the evaluation of the accuracy of obtained results. And this part is given in vary loose and general manner. Line 366. The time series of what?

Line 173. What the "accuracy rate" is?

Lines 376-378. I would be surprised if the "precipitation-type" floods are not related to the "extreme precipitation".

Line 401. derived products of what and what?

Line 404. Better than what?

Line 405. The use "for the first time" is not correct as both DFO dataset and MODIS-derived flood events product have global coverage.

Line 411. Please, explain what is the value of the flood event product of such low ($1^{\circ} \times 1^{\circ}$) spatial resolution? Discuss also the cases of false detection of flood events, i.e. the events not supported by datasets/products used for verification of the obtained results. Provide their statistics. How do you separate false flood detection from non recorded flood cases?