We would like to thank the reviewers for the thorough reading of the manuscript and the valuable remarks that helped us to improve the manuscript. We have revised the manuscript carefully according to the reviewer's comments, and have incorporated the suggestions into the revised manuscript.

The notes below provide a point-by-point response to each comment from the referees. The texts with blue font are the reviewer's original comments, the texts with black font are authors' responses.

Major Commentsï¼

1. How do you consider the detected flood (extreme values in TWSA) over the glacier regions? I do not see the demonstration for these regions.

Response: In this study, we detect flood events by identifying change in high frequency signals of GRACE. The high-frequency signals capture the impact of extreme precipitation which has high magnitude and relatively short duration. While the melting of glaciers is usually a slower process and cannot be identified by high frequency signals. Therefore, in this paper, the floods we extracted are mainly precipitation-induced floods. Glacier-melting-induced floods are not considered in this study. To avoid misunderstanding, we have modified the title of this paper as follows:

Precipitation-induced Flood Detection Using GRACE Terrestrial Water Storage

2. Which term does the “high frequency signal” with the STL decomposition for the GRACE TWS? Does the final signal include a seasonal signal? It is not clear in Section 3.1, Please provide detailed explanations. I think sometimes the seasonal signal can also relate to floods.

Response: We thank reviewer's suggestions. The "high frequency signal" means the "remainder" in the STL decomposition. The "remainder" is the remaining part of the original data after the STL decomposition method which eliminates the seasonal and trend components. The TWS seasonal signal is also associated with flood mainly when the peak of the seasonal time series encounters heavy precipitation, which may trigger flood. Since the impact signal exerted by heavy precipitation is already included in the high-frequency signals of TWS, the above situation can also be detected.

3. How do you extract the flood day for different cases, I think the authors should show some figures (maybe time series) for a detailed demonstration. I do not think Section 3.4 is clear enough.

Response: We thank reviewer's suggestions. Figure 1 shows the intermediate process of flood detection for randomly selected grid. The steps are as follows. Firstly, we extracted the high-frequency signal of TWS using the STL method; Secondly, we calculated the possible flood days using the GESD method; Thirdly, we use flood potential index to supplement possible flood days in case the daily GRACE TWS data have lost useful high-frequency signals due to the interpolation process; Finally, we constrained the pre-selected floods using the extreme precipitation days derived from daily and cumulative precipitation. To explain the method more clearly, we will add a more detailed additional description of the flood days extraction process in the revised manuscript.

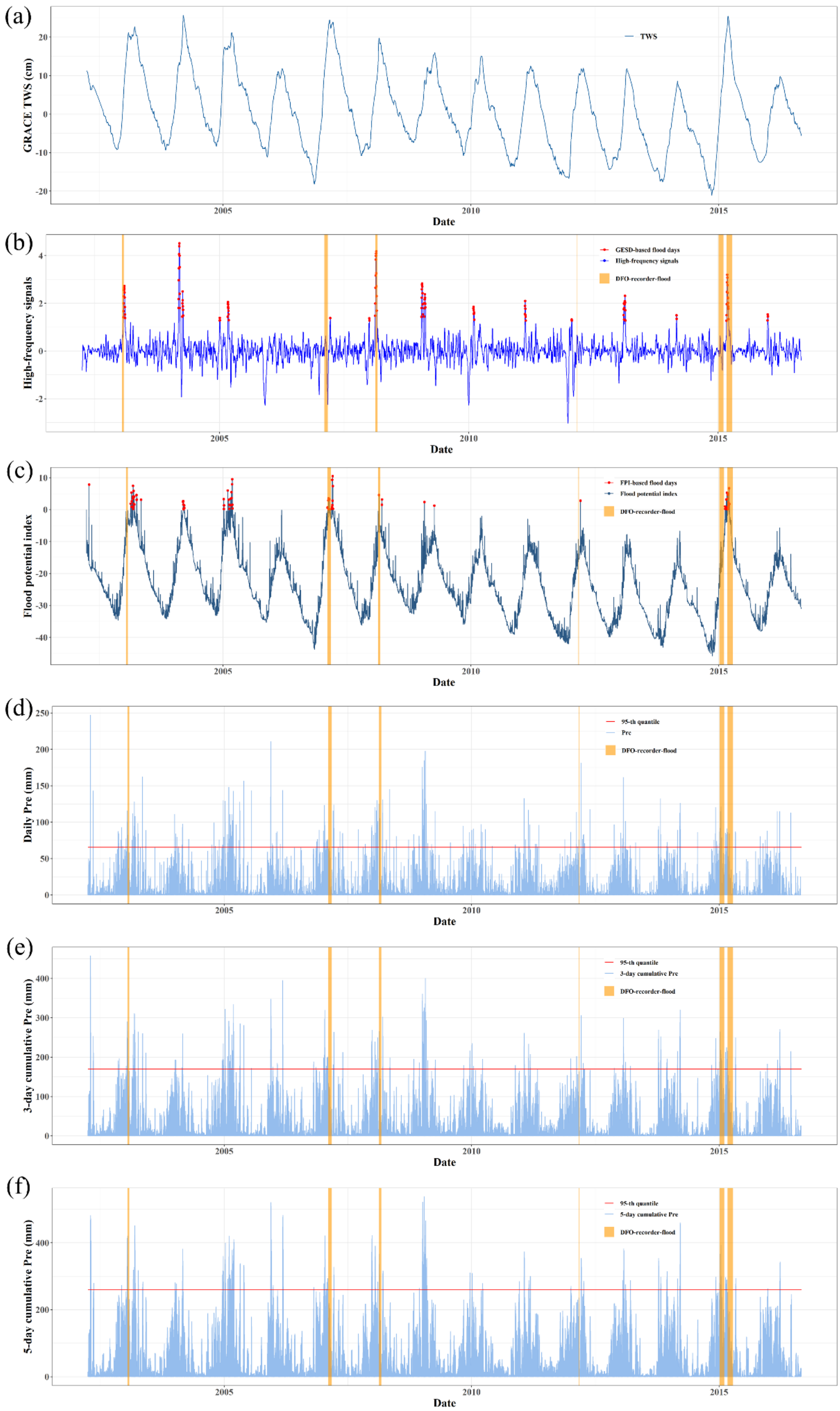

Figure 1 Schematic diagram of the intermediate variable time series processing for flood extraction based on GRACE and extreme precipitation for a randomly selected grid from Apr. 1st, 2002 to Aug. 31st,

2016. (a) GRACE TWS times series; (b) high-frequency signals time series derived from GRACE and flood days extraction based on GESD; (c) flood days extraction based on flood potential index time series; (d) daily precipitation time series and extreme precipitation days based on 95-th quantile; (e) 3-day cumulative precipitation time series and extreme precipitation days based on 95-th quantile; (f) 5-day cumulative precipitation time series and extreme precipitation days based on 95-th quantile.

4. Which version of ITSG was used in this study? Please list it in Section 2.1.

Response: We thank reviewer's suggestions. We chose the GRACE daily solution based on ITSG-Grace2018 gravity field model. The ITSG-Grace2018 gravity field model, which offers unconstrained monthly and Kalman-smoothed daily solutions, is the most recent GRACE-only gravity field model computed in Graz. It is a reprocessing of the whole GRACE time series starting from 2002-04. We will add this in the revised manuscript.

5. Section 3.2, how do you evaluate/consider the reliability of the GESD test, please give a more detailed demonstration.

Response: We thank reviewer's suggestions. GESD test is a commonly used univariate anomaly detection algorithm. It has been widely used in the field of hydrological anomaly detection (Saghafian et al., 2014; Clark and Zipper, 2016).GESD test is mainly used in this study to extract possible flood days corresponding to the high-frequency signals. In this study, the method to be selected for extracting flood information from the high-frequency signals should ensure that it was affected by the random error in the high-frequency signal as little as possible and that the flood signals were extracted as much as possible. The GESD test method overcomes the primary limitation of the Grubbs test and the Tietjen-Moore test that the suspected number of outliers, k, must be specified exactly. The GESD test only requires that an upper bound for the suspected number of outliers be specified (Rosner, 1983). These advantages of the GESD test method were highly compatible with our needs.

Moreover, in the practical operation for our dataset, the method performed well. On the one hand, the method effectively extracted the important anomalous peaks (i.e., the pre-extracted flood days), and the extracted flood days for each spatial grid were stable and did not increase with the increase of the preset up bound. As shown in Fig. 2, the histogram distribution of the pre-extracted global flood days for each spatial grids was concentrated around 200 days, not more than 500 days. It was not more than 200 flood events after converted from flood days to flood events. The order of magnitude was consistent with the number of historical flood events. Finally, the pre-extracted flood based on GESD test was an important part of the final extracted flood days, and the good validation of the final product also confirmed that GESD test method played a reliable and important role in the intermediate processing. We will add a more detailed description of the GESD test method in the revised manuscript.

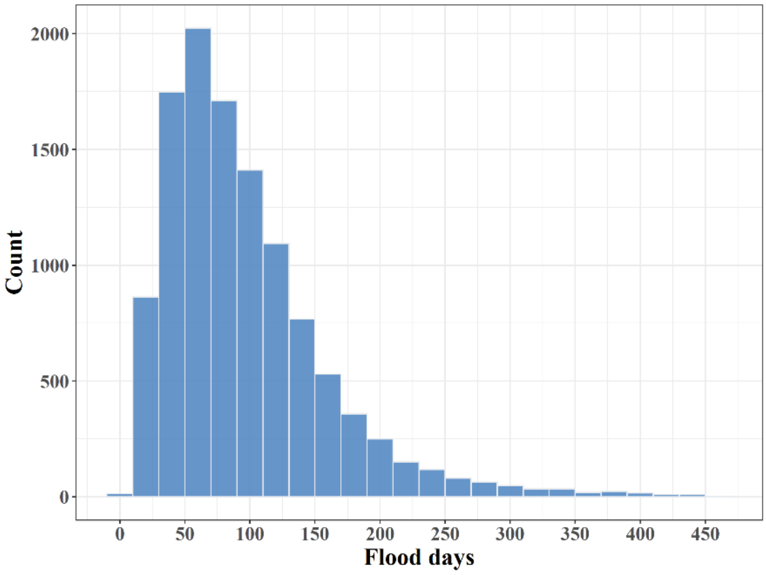

Figure 2 The histogram of pre-selected flood days based on high-frequency signals of TWS using GESD test method.

6. When I check the supplementary materials, I find that there are many differences among the spatial patterns of DFO-based, MODIS-based, and GRACE-based floods. More discussion should be drawn to improve this MS, the work in Gouweleeuw (2018) also showed that ITSG might not catch the flooded area well. Besides, how do you consider the different representation between terrestrial water storage (TWS) and runoff for one flood event?

Response: We thank reviewer's suggestions. The DFO database mainly records large-scale flood events from news and government announcements and misses some flood events. Although DFO records the start and end times and the approximate spatial locations of flood events, the temporal range is sometimes long (more than one or two months), and the spatial locations are only roughly delineated according to news reports, which contained large uncertainties in spatial location and duration. Moreover, although flood detection methods based on MODIS data could finely delineated the spatial inundation ranges, it was mainly aimed at specific flood events in small areas. Only a limited amount of flooding can be identified by remote sensing images due to the influence of bad weather. MODIS resolution is 1 km, which is a higher resolution product compared to GRACE TWS (1°, ~100km). The flood products we provided can only indicate the presence of flood events under the ~100km grid coverage, and specific detailed locations require further identification by high-resolution satellite remote sensing image. Therefore, we can see differences among the spatial patterns of the three datasets. However, as we show in Figure 8-9 and supplementary figures, there is still an overwhelming majority of results indicating a consistent spatial pattern of flood inundation. Our products can provide a sufficient data foundation for large-scale research on the spatiotemporal distributions and attributions of global floods. We will further deepen the discussion and analysis of this result in the subsequent revision.

As for the different representation between terrestrial water storage (TWS) and runoff for one flood event, we extracted the flood days based on GRACE. To facilitate the comparison between GRACE-based floods and discharge-based floods, we first extracted the corresponding flood events with the discharge data using the FloodR method, and then compared whether each flood event period contained the flood days we extracted. This was used to calculate the effectiveness of our product in comparing to floods derived from discharge.

Minor Comments:

1. Figure 3, the tick values of the x-axis should be denser.

Response: We thank reviewer's suggestions. We have made the x-axis denser.

2. P2L69ï¼□ what is the "other places"?

Response: We thank reviewer's suggestions. We have changed the "other" to a more critical description, which reads as follows.

The numbers of flood events recorded in China, Russia and Canada are obviously lacking.

3. "Diksha Gupta et al (Gupta and Dhanya, 2020)"… in all the MS, these citation formats are strange, maybe wrong.

Response: We thank reviewer's suggestions. We have modified it to the correct format

4. P3L109, what is the difference between the two citations of "Dill, 2008; Dill et al., 2008"?

Response: We thank reviewer's suggestions. One was misquoted here and has been removed.

5. P10L296, “Tellman1” to “Tellman”

Response: We thank reviewer's suggestions. We have modified it.

6. The two references share the same data name, I think there is no need to put them on two websites. In addition, there is a mistake in the name of the authors.

“Zhang, J., Liu, K., and Wang, M.: Flood detection using GRACE Terrestrial Water Storage and Extreme Precipitation (1.0.0). Zenodo. https://doi.org/10.5281/zenodo.6831105, 2022a.

Zhang, j., Liu, K., and Wang, M.: Flood Detection Using GRACE Terrestrial Water Storage and Extreme Precipitation (1.0.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.6831384, 2022b.”

Response: We thank reviewer's suggestions. Since each uploaded file generates a unique DOI, we couldn't put them on one website. We have modified the mistake in the name of the authors.

7. I do not think the reference format is right. “Mayer-Gürr, T., Behzadpour, S., Kvas, A., Ellmer, M., Klinger, B., Strasser, S., and Zehentner, N.: ITSG-Grace2018: monthly, daily and static gravity field solutions from GRACE, 2018.”

Response: We thank reviewer's suggestions. We have modified this reference format in subsequent revisions.

8. The references should be checked carefully.

Response: We thank reviewer's suggestions. We have checked all citation formats of this paper.

References:

Clark, Elyse V and Zipper, Carl E: Vegetation influences near-surface hydrological characteristics on a surface coal mine in eastern USA, Catena, 139, 241-249, 2016.

Rosner, Bernard: Percentage points for a generalized ESD many-outlier procedure, Technometrics, 25, 165-172, 1983.

Saghafian, Bahram, Golian, Saeed, and Ghasemi, Alireza: Flood frequency analysis based on simulated peak discharges, Natural Hazards, 71, 403-417, 2014.