https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

Flood Detection Using GRACE Terrestrial Water Storage and Extreme Precipitation

Jianxin Zhang [1,2], Kai Liu [1], Ming Wang [1]

[1] School of National Safety and Emergency Management, Beijing Normal University, 100875 Beijing, China
[2] School of Systems Science, Beijing Normal University, 100875 Beijing, China

Correspondence to: Kai Liu (liukai@bnu.edu.cn)

Abstract. A complete global flood event record helps researchers analyse the distribution law of global floods and better formulate and manage disaster prevention and reduction policies. This study used GRACE terrestrial water storage and precipitation data combined with high-frequency filtering, anomaly detection and flood potential index methods to successfully extract historical flood days quasi-globally between Apr. 1st, 2002, and Aug. 31st, 2016, and further compared and validated the results with Dartmouth Flood Observatory (DFO) data, Global Runoff Data Centre (GRDC) discharge data, news reports and social media data. The results showed that the floods extracted in this study can cover 81% of the flood events in the DFO database and supplement many additional flood events not recorded by the DFO. Moreover, the probability of detection exceeding 0.5 reached 62% in level-4 river basins compared to flood events derived from the GRDC discharge data. These detection capabilities and detection results are both good. We finally provided flood day products with 1° spatial resolution covering the range of 60°S—60°N from Apr. 1st, 2002, to Aug. 31st, 2016. This research provides a data foundation for the mechanistic analysis and attribution of global flood events.

1 Introduction

Flood disasters threaten the lives of millions of people around the world every year, causing more economic losses than any other natural disaster. An increasing number of extreme weather events are occurring frequently under global climate change (Schinko et al., 2017). The latest research on global flood disasters shows that the proportion of the population at risk from floods is increasing each year (Tellman et al., 2021).

Existing global flood data mainly include historical data records, hydrological model simulations and remote sensing observations. Historical data records include those of the Dartmouth Flood Observatory (DFO) (Brakenridge, 2022), international disasters database (EM-Dat) (Em-Dat, 2014), Munich Re's NatCatSERVICE (natcatservice.munichre.com), and Sigma (Re, 2022). The DFO database mainly records large-scale flood events from news and government announcements. This database records not only the country, latitude, approximate scope and start and end time of each event but also the cause and severity level of the event. There have been approximately 4,700 major flood events since 1985. EM-Dat contains basic core data on the occurrence and impact of more than 22,000 large-scale disasters in the world from 1900 to the present. This database is compiled from a variety of sources, including UN agencies, nongovernmental organizations, insurance companies, research institutes and news organizations. NatCatSERVICE is a natural hazard-based disaster loss database with up to 28,000 entries owned by Munich Re's company. Sigma is also a global disaster database comprising anthropogenic and natural catastrophe losses since 1970. In addition to recording basic disaster information, this database also includes the total and insured losses. Flood data are also derived from global hydrological models. For example, the University of Maryland's Global Flood Monitoring System (GFMS) takes real-time integrated precipitation information (Tropical Rainfall Measuring Mission (TRMM) and Global Precipitation Measurement (GPM)) as inputs in a quasi-global (50°N - 50°S) hydrological runoff and routing model to run 1/8° gridded data. Surface water storage statistics are used to derive flood thresholds for each grid location, and the depth above the corresponding threshold is calculated as the flood intensity (Wu et al., 2014; Wu et al., 2012a; Wu et al., 2012b; Wu et al., 2011). Another example is the Floods.Global

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

system (http://floods.global) database, in which GPM Integrated Multi-satellite Retrievals for GPM (IMERG) precipitation data are used to estimate future 72-hour flows, with coverage from 60°N to 60°S and a resolution of 0.1 degrees. With the development of remote sensing satellite products in the 1980s, a cost-effective flood-monitoring method was developed. As long as there are historical areas over which satellites have passed and imaged, there are opportunities to observe flood events; these methods are more realistic and effective than flood models in characterizing actual observed flood areas. Commonly used remote sensing data include optical remote sensing images and microwave remote sensing images, among which microwave remote sensing technologies, especially the commonly applied synthetic aperture radar (SAR), are used. Michaela Rättich et al. (Rättich et al., 2020) developed an automatic procedure to evaluate flood durations and uncertainties using multiple satellites, including Sentinel-1, Sentinel-2, Landsat-8 and TerraSAR-X. The method was successfully demonstrated on the 2019 flood in Sofala, Mozambique and on the 2017 flood in Bihar, India. B. Tellman et al. (Tellman et al., 2021) used 250-m-resolution Moderate-resolution Imaging Spectroradiometer (MODIS) data to extract the inundation extents of a total of 913 global flood events from 2000 to 2018, thus providing data support for vulnerability assessments and flood model improvements. Tong et al. used both Landsat 8 optical imagery and COSMO-SkyMed radar imagery combined with a support vector machine and the active contour without edges model to perform flood monitoring. The results showed high accuracies of 97.46% for optical imagery and 93.70% for radar imagery.

The current databases with the most detailed records come from Re's NatCatSERVICE, EM-Dat, Sigma and DFO. NatCatSERVICE, EM-Dat and Sigma provide data only at the country level. The NatCatSERVICE database covers most large flood events around the world but only a few small flood events in developing countries due to restricted connectivity (De Bruijn et al., 2019). Although there are approximate map locations available in the Sigma and NatCatSERVICE databases, specific location names are not publicly available (Moriyama et al., 2018). Moreover, both the NatCatSERVICE and Sigma databases were developed by reinsurance companies, and the accessibility of the information in these databases is limited (Moriyama et al., 2018; Kron et al., 2012; Huggel et al., 2015). The EM-Dat database records only the numbers of flood events in different countries without corresponding spatial location information. 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. B. Tellman et al. (Tellman et al., 2021) extracted flood extents and analysed the population exposure of 913 large-scale flood events from 2000 to 2018 based on MODIS daily data with a resolution of 250 metres, thus finely delineating the spatial inundation ranges. During this period, there were more than 3,000 flood events recorded in the DFO database, while the number of MODIS-derived floods was less than 30% of that recorded by DFO. Moreover, only some large-scale flood events were recorded by the DFO. The numbers of flood events recorded in China, Russia and other places are obviously lacking. Moreover, flood detection methods based on remote sensing data are mainly aimed at specific flood events in small areas and are influenced by the number of revisit cycles at the same location (especially for SAR images) and bad weather (especially for optical images) (Kussul et al., 2011; Hostache et al., 2018; Manavalan, 2017). Much useful image information is missing, and this affects the flood extraction accuracy. There is a need to fill in the missing flood events with a new observational dataset.

Another remote sensing technique based on gravity satellites, the Gravity Recovery and Climate Experiment Satellite (GRACE), has also been successfully used to detect flood events. J. T. Reager (Reager and Famiglietti, 2009) first proposed the terrestrial water storage capacity and flood potential, creating a precedent for GRACE to assess large-scale flood events. This method was subsequently improved and applied to different river basins. Tatiana Molodtsova et al. (Molodtsova et al., 2016) found an agreement between the flood potential index derived with GRACE and observed floods by using multiyear flood observation data. Diksha Gupta et al (Gupta and Dhanya, 2020) proved that GRACE terrestrial water storage (TWS) and the flood potential index had the capability to assess hydrological extreme events over heterogeneous regions with the occurrences of high- and long-duration floods. They suggested that this method can be useful for flood monitoring when discharge data are rarely available. With the continuous progress of the global GRACE-only gravitational field solution,

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

GRACE daily data products have also been effectively developed and applied (Kvas et al., 2019; Mayer-Gürr et al., 2018). Ben T. Gouweleeuw et al (Gouweleeuw et al., 2018) used a daily solution based on GRACE TWS and daily river runoff data to assess major flood events in the Ganges-Brahmaputra Delta and confirmed the method's potential for gravity-based large-scale flood monitoring. Jinghua Xiong et al. (Xiong et al., 2022) used daily downscaled GRACE data to detect short-duration and high-intensity floods. They found that there is a strong correlation between the high-frequency components of GRACE TWS and runoff.

In this study, we focus on extracting global historical flood events based on daily GRACE and precipitation data. Apart from being affected by battery management in some months, GRACE's gravity measurements cover most months and are not affected by varying kinds of weather conditions. This study mainly extracted all flood days in the historical time series caused by extreme precipitation, regardless of whether the flood event caused severe damage. Finally, we provide global flood days with a resolution of 1° during the period from Apr. 1st, 2002 to Aug. 31st, 2016. These data replenish the missing flood events in the historical record and provide a new and complete flood dataset, thus providing a sufficient data foundation for research on the inducement of global floods.

2 Data

2.1 GRACE

The GRACE constellation is a pair of twin satellites that can measure changes in Earth's gravitational field. There is a precise radar rangefinder between the two satellites. When Earth's gravitational field changes slightly, it can be detected by either of the two satellites. The distance signal between the two satellites is amplified to measure the state change at the current moment relative to the previous moment (Cazenave and Chen, 2010). The short-term gravitational field changes of Earth are mainly caused by changes in terrestrial water storage, atmospheric water vapour, ocean tides, etc. When these signals are deducted, the change in the entire terrestrial water storage can be inverted (Wahr et al., 1998). The daily GRACE data selected in this study come from daily solutions obtained using Kalman smoothing by Torsten Mayer-Gürr et al of Graz University of Technology (Mayer-Gürr et al., 2018). The time period spans from Apr. 1st, 2002, to Aug. 31st, 2016, and the resolution is 1°. A third-order autoregressive (AR) model was used to stabilize the daily solution. A set of spherical harmonic coefficients for the various degrees (n=2…40) was estimated. When GRACE data were not available for a specific day, daily solutions were delivered through an adjustment process (Bergmann-Wolf et al., 2015; Dill, 2008; Dill et al., 2008).

2.2 Precipitation

This study used Global Precipitation Measurement (GPM) data to calculate extreme precipitation. GPM is an international satellite mission launched by the National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA). It is the next-generation global rain and snow satellite observation network after the TRMM and is a set of high-quality global precipitation and snow observation datasets that can provide important data bases for scientific researchers to understand the Earth's water resources and energy cycles and improve their ability to predict extreme events (Huffman et al., 2015). The resolution of these data is 0.1°, mainly covering the range of 60°S—60°N, and both north-south latitudes of 60—90° have partial coverage. This study selects the IMERG Final Run product, which uses global microwave precipitation data, infrared data, precipitation station data and other potential precipitation indicators to cross-calibrate, fuse and interpolate TRMM and GPM data at refined temporal and spatial scales. It is an officially recommended product and can be obtained from the following website: https://gpm.nasa.gov/data/directory. To remain consistent with the GRACE resolution and maintain extreme precipitation signals, we take the maximum values of the precipitation data under the GRACE grid coverage to further calculate the flood potential index and the number of extreme precipitation days.

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

2.3 Dartmouth Flood Observatory

The DFO dataset records large flood events from various news reports and government websites. It contains the start and end times of each flood, the country where it occurred, the approximate scope, the cause of the flood and the degree of damage. It is a rare and useful product for studying global historical floods. This data product has been widely used in flood hazard science research (Tellman et al., 2021; Hagen et al., 2010; Winsemius et al., 2013; Idowu and Zhou, 2019). This study focuses on precipitation-induced floods. A total of 2380 flood events in the 60°S—60°N range were caused by heavy precipitation between Apr. 1st, 2002 and Aug. 31st, 2016. This product was primarily used to validate the flood data extracted in this study.

2.4 MODIS-derived Flood Inundation Data

The flood inundation data used in this study came from a total of 807 flood inundation datapoints recorded in the 60°S—60°N region between Apr. 1st, 2002 and Aug. 31st, 2016, extracted based on MODIS data by Tellman et al. (Tellman et al., 2021). This product was produced using atmospherically corrected Terra (MOD09GA/GQ) and Aqua (MYD09GA/GQ) MODIS images and threshold analysis methods (including 3-day standard, 2-day standard, 3-day Otsu and 2-day Otsu methods) as well as slope constraints (slopes greater than 5° were masked out) to extract inundations at a 250-m spatial resolution according to the flood events recorded by the DFO. The extraction results were compared and verified for coincidence with the 30-m-resolution inundation data derived from Landsat 5, 7 and 8 imageries, and flood map quality control analysis was also performed. This product relies on the Google Earth Engine platform (Gorelick et al., 2017), which can be obtained from the following site: https://developers.google.com/earth-engine/datasets/catalog/GLOBAL_FLOOD_DB_MODIS_EVENTS_V1. These data were further refined on the basis of the approximate range of the DFO and provided reliable spatial inundation dataset for verification in our study.

2.5 GRDC Discharge Data

The Global Runoff Data Centre is an international data centre operating under the auspices of the World Meteorological Organization. It was established in 1988 to support research on global climate change and integrated water resource management. The corresponding dataset is a discharge product that records the mean discharge, country, longitude, latitude and river name associated with each flood event. This study selects the records from Apr. 1st, 2002—Aug. 31st, 2016, as the validation dataset to verify the extracted floods. To facilitate the comparison with the GRACE-derived flood days and eliminate the effects of random errors, we considered the spatial average values of the HydroSHEDS Basins Level 4 data (Lehner and Grill, 2013; Lehner et al., 2008) covering discharge measurements and converted them to flood events using the statistical method described below. These data can be obtained from https://www.bafg.de/GRDC/EN/Home/homepage_node.html.

3 Methods

Figure 1 shows the technical route of this study, which mainly includes three parts: data preparation, flood day extraction, and result verification. Daily precipitation and daily GRACE TWS data were used for the flood data extraction step, and daily discharge, DFO, MODIS-derived flood inundation and social media data were used for the flood validation step. The second part of the flood extraction step was mainly based on the use of high-frequency TWS signals and the flood potential index to obtain the preliminary possible flood days; then, extreme precipitation constraints were used to obtain the

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

final flood days. The third part of the flood validation step included comparisons with the flood extents extracted from DFO records and MODIS images, comparisons with floods derived from GRDC discharge data, and, finally, comparisons with significant flood events recorded on social media.

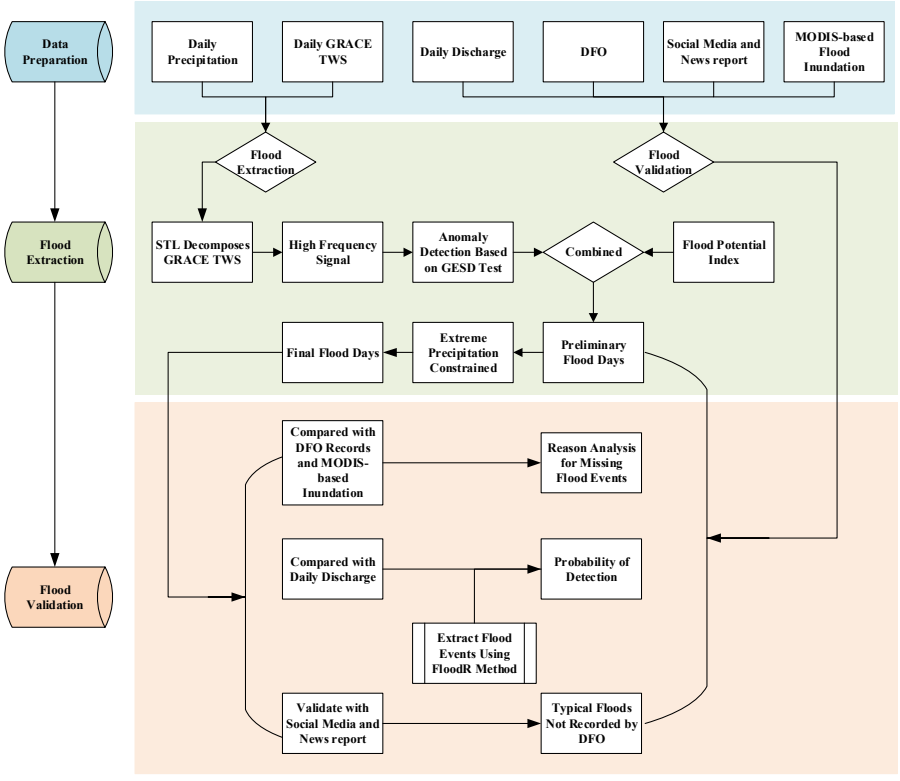


Figure 1 Workflow of the global flood extraction method based on GRACE and precipitation data.

3.1 Seasonal and Trend decomposition using Loess (STL)

Extreme precipitation has sudden characteristics, and its signals are reflected in the high-frequency signals of GRACE (Xiong et al., 2022; Gouweleeuw et al., 2018). Seasonal and trend decomposition using loess (STL) (Robert et al., 1990) is a filtering process as well as a general and robust time series decomposition and forecasting method used to decompose time series variables to obtain season, trend and remainder information for further forecasting. This process can handle data with any type of seasonality as well as high-frequency seasonal data. It also allows seasonal components to vary over time and is robust to outliers. In this study, we selected this method as a high-frequency filtering tool to process GRACE TWS and obtain high-frequency signals for subsequent analyses. In this study, the STL function in the R language "stats" package was used to process all grid point time series corresponding to the GRACE TWS period (Apr. 1st, 2002—Aug. 31st, 2016), where the parameter "t.window" was set to 31 days based on the methods of Gouweleeuw and Xiong et al. (Xiong et al., 2022; Gouweleeuw et al., 2018) and "s.window" was set according to the period of the time series dataset itself. The period was determined using a Fourier transform to convert to the frequency domain to obtain the frequency (i.e., period) corresponding to the maximum amplitude.

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

3.2 Anomaly Detection Based On Generalized Extreme Studentized Deviate Test

The generalized extreme Studentized deviate (GESD) test (Rosner, 1983) is a simple and effective statistical method for detecting one or more outliers in univariate data that follow an approximately normal distribution. The method requires only that an upper bound for suspected outliers be specified and determines the number of possible outliers based on hypothesis testing. The basic assumptions of GESD are as follows:

H_0: There are no outliers in the dataset.

H_a: There are at most r outliers in the dataset.

The corresponding statistic is calculated as follows:

$$R_i = \frac{max_i |x_i - \bar{x}|}{\sigma} \quad (1)$$

where $\bar{x}$ and σ are the sample mean and standard deviation, respectively. After each iteration in which the largest $|x_i - \bar{x}|$ value is removed, the remaining statistics are calculated, and the above process is repeated until at most r outliers are removed.

Consistent with the r test statistics, the r critical values are computed as follows:

$$\lambda_i = \frac{(n-i)t_{p,n-i-1}}{\sqrt{(n-i-1+t^2_{p,n-i-1})(n-i+1)}} \quad i = 1,2,\dots,r \quad (2)$$

$$p = 1 - \frac{\alpha}{2(n-i+1)} \quad (3)$$

where $t_{p,v}$ is the 100p percentage point in a t-distribution with v degrees of freedom and α is the significance level. The number of final outliers is then determined by the corresponding maximum i value in $R_i > \lambda_i$ (Rosner, 1983).

Considering that the GRACE high-frequency signal contains both random errors and useful signals, we used the GESD test to control the number of outliers so that they were not affected by subjective thresholds. In this study, the "AnomalyDetection" package (https://github.com/twitter/AnomalyDetection) (Aggarwal, 2013; Chandola et al., 2009; Rosner, 1983; Vallis et al., 2014) was used to extract GRACE HPF data. This package not only encapsulates the GESD algorithm but can also specify the direction of detected outliers. As we considered extreme weather events caused by heavy precipitation in this study, important information was contained in the peak. The main parameter "direction" was set to the position, and the maximum possible number of abnormal days "max_anoms" was set to 0.1 to cover the maximum number of abnormal days among the global time series comprising every grid cell.

3.3 Flood Potential Index

The derivation of the flood potential index with GRACE data was first proposed by J. T. Reager et al. (Reager and Famiglietti, 2009) in 2009 and has since been widely used to evaluate flood events (Gupta and Dhanya, 2020; Molodtsova et al., 2016; Reager et al., 2014). Its basic assumption is that the regional water storage capacity can be approximated by the maximum time series data value in the historical record; the water storage capacity at the current time can be calculated by subtracting the TWS at the previous time from the time-series maximum value. The proposal of this method was based on monthly data, but this does not affect the application of GRACE daily data. The specific definition is as follows.

The water storage capacity of the current day can be expressed as the temporal difference between the maximum time-series value and the previous-day value, and the formula is expressed as follows:

$$TWS_{DEF}(t) = TWS_{MAX} - TWS(t-1) \quad (4)$$

where $TWS_{DEF}(t)$ represents the maximum allowable relative water storage change on the current day, TWS_{MAX} represents the maximum value over the entire time series, and $TWS(t-1)$ represents the TWS value of the current day relative to the previous day. A low storage deficit and high precipitation mean a high probability of flooding, i.e., the occurrence of floods should be based on the mismatch between the extreme precipitation level and the increase in water storage:

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.


$$F(t) = P_{day}(t) - TWS_{DEF}(t) \quad (5)$$

where $P_{day}(t)$ represents the daily precipitation and $F(t)$ represents whether the current precipitation matches the water storage capacity. When $F(t) > 0$, flooding may occur. This study uses the 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.

3.4 Flood Detection Based on GRACE TWS and Precipitation Data

This study used the GRACE HPF data and the flood potential index combined with the GESD algorithm to preselect the possible flood days pixel by pixel. Then, we further used the number of extreme precipitation days to constrain and obtain the final results. This study focuses on flood events caused by heavy precipitation. Considering that floods are affected not only by single-day precipitation but also by cumulative precipitation, we calculated the extreme precipitation days based on the one-day precipitation, 3-day cumulative precipitation and 5-day cumulative precipitation. Regarding extreme precipitation, the most commonly used metric is the percentiles of the precipitation time-series data, including the n-th quantiles of the entire time series data; alternatively, the n-th quantile of wet days (daily precipitation > 1 mm) can be considered (Myhre et al., 2019; Pendergrass, 2018; Shi et al., 2021). Based on the principle of extracting as many flood events as possible with as few errors as possible, we choose the 95th quantile of the entire time series as the condition to constrain the results extracted from GRACE data. We also provide products obtained with the 90th and 99th quantiles of the entire time series data and wet days in supplementary data files.

3.5 Flood Event Separation Based on Daily Discharge Data

To verify the reliability of the extracted results, this paper used the global discharge data products released by GRDC and the statistics-based automated flood event separation (FloodR) method to extract flood events. FloodR is a statistical-based flood event separation method proposed by Svenja Fischer et al. (Fischer et al., 2021). It can automatically separate flood events using a univariate daily discharge time series, and it integrates expert knowledge tools to manually and quickly validate and correct the separation results. The results extracted by this method include information such as the start and end times of each flood, the flood peak date and the flood baseline, thus providing an important data foundation for verifying the time-series comparison ability of this study.

4 Results

4.1 Flood days and Events Based on GRACE TWS and Precipitation Data

This study considers the results obtained under the constraint of the 95th percentile of the entire time-series dataset. Figure 2 shows the global cumulative flood days and flood events from Apr. 1st, 2002—Aug. 31st, 2016, and Figure 3 shows the histograms of flood days and flood events corresponding to Figure 2. Although the number of flood days extracted above cannot accurately reflect how many flood events occurred, we can simplify the results such that a number of consecutive detected flood days, or the interval between two consecutive flood days, no more than 3 days or 5 days can be considered a flood event. The principle involves roughly calculating the spatial distribution of global flood event occurrences. Consistency was found between the global spatial distributions of flood events and flood days. We found that 99.8% of the grid cells around the world experienced fewer than 400 flood days except in Southeast Asian countries and countries at the junction of North and South America, which experienced more than 400 flood days from Apr. 1st, 2002—Aug. 31st, 2016.

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

In addition, the areas with the most flood days and events were mainly located in the tropics. Island countries, western Africa, India, the Himalayas, southern China, etc., were also prone to floods. From the perspective of the divided flood events, the numbers of grid cells with fewer than 100 events accounted for 96.15% (Figure 2 b), 97.11% (Figure 2 c) and 97.52% (Figure 2 d) of all grid cells.

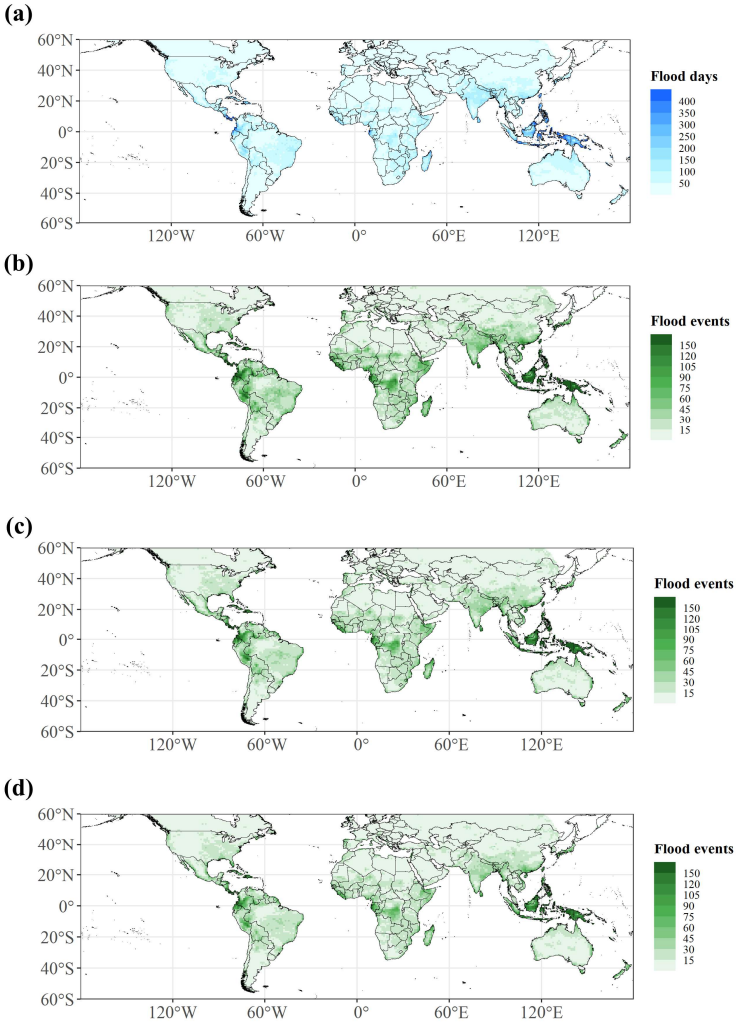


Figure 2 Spatial distribution of global floods, including (a) global flood days from Apr. 1st, 2002—Aug. 31st, 2016; (b) global flood events based on consecutive flood days; (c) global flood events based on the interval between two consecutive flood days not exceeding 3 days; and (d) global flood events based on the interval between two consecutive flood days not exceeding 5 days.

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

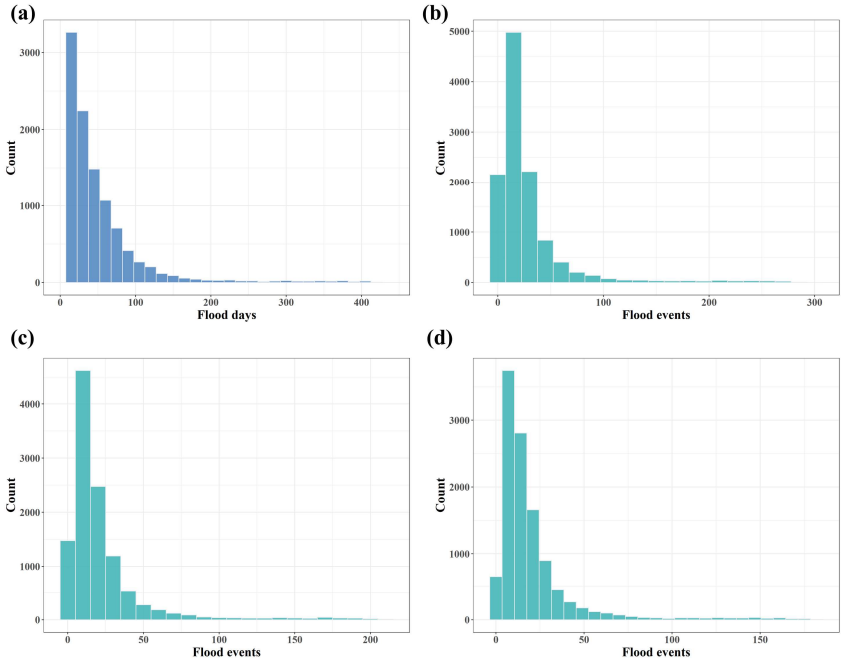


Figure 3 Histograms of cumulative flood days and flood events from Apr. 1st, 2002—Aug. 31st, 2016. (a) Histogram of global flood days from Apr. 1st, 2002—Aug. 31st, 2016; (b) histogram of global flood events based on consecutive detected flood days; (c) histogram of global flood events based on the interval between two consecutive flood days not exceeding 3 days; and (d) histogram of global flood events based on the interval between two consecutive flood days not exceeding 5 days.

We also calculated the average flood days in the same month in each year from 2002—2016 to identify seasonal characteristics. As shown in Figure 4, the global flood distribution reflected obvious seasonal characteristics, and differences between the Northern and Southern Hemispheres are clear. More flood days were identified in the Northern Hemisphere in summer (approximately June-September), while flood days in the Southern Hemisphere are concentrated from December-March.

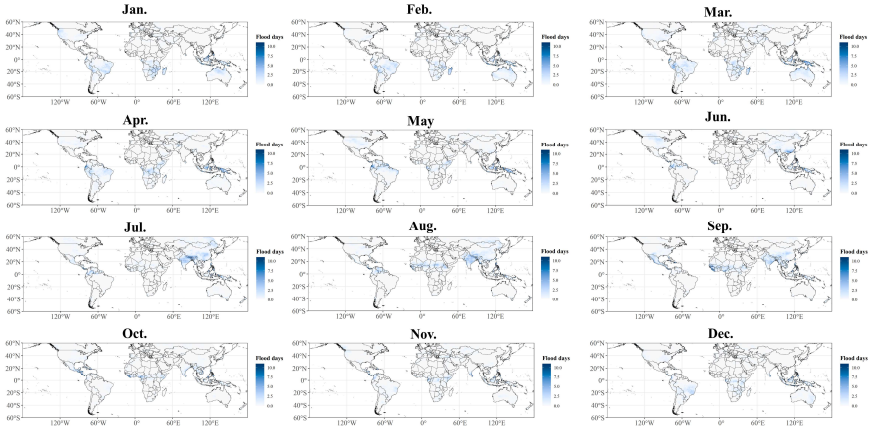


Figure 4 Average flood days in the same month each year from 2002 to 2016.

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

4.2 Validation with DFO and MODIS-derived Flood Data

Figure 5 shows the spatial distribution of 2,380 precipitation-type floods recorded by the global DFO from 60°S~ 60°N. Judging from the floods recorded by the DFO, floods have occurred in most parts of the world except in the Sahara Desert, the Great Victoria Desert and the northern part of North America.

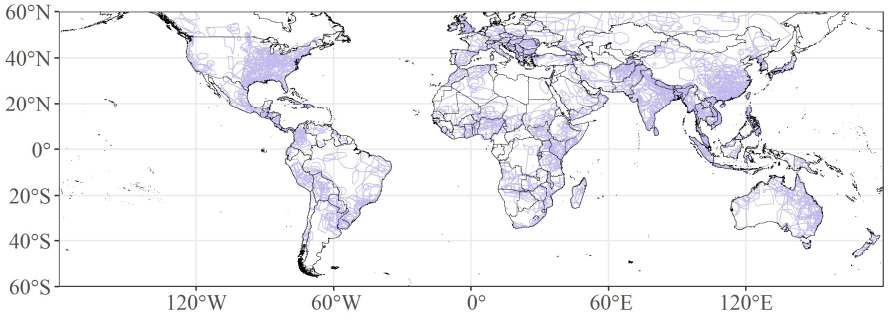


Figure 5 Spatial distribution of DFO polygon features

In this study, the temporal length of the DFO database was compared throughout the Apr. 1st, 2002—Aug. 31st, 2016 period. According to the database attributes, flood events caused by heavy precipitation were extracted as the validation dataset for this study. Given that the temporal and spatial DFO recording characteristics are approximate and considering the effect of advanced or delayed times on short-duration records, the start and end times of the DFO records were extended forward and backwards by three or five days, respectively, when being compared with the flood day results. Similarly, when the extent of the DFO polygon was less than 3 degrees, we appropriately built a buffer (3 degrees) to compensate for the positioning errors. Then, we detected every event in the DFO record to determine whether flood days could be identified from its temporal and spatial coverages.

Figure 6 shows that the polygon features marked by the DFO are allocated to grids with the same resolution as that of GRACE data according to their spatial positions; each grid exhibits the number of times it was identified in the DFO polygon features. The figure shows that the eastern part of North America, the northern part of South America, the central and southern parts of Africa, western Europe, northern India and southern China are all areas with high-frequency flood events. Except for the archipelagic countries in Southeast Asia, the entire spatial distribution is consistent with our results.

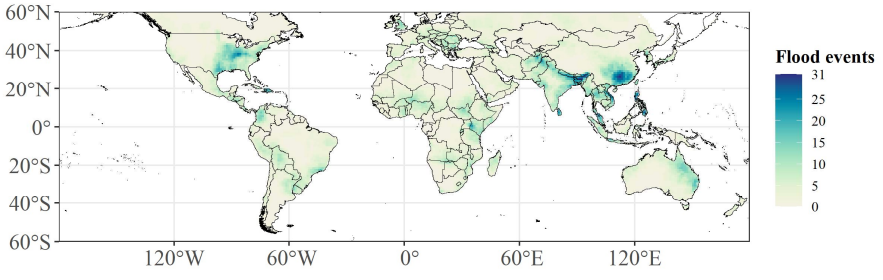


Figure 6 Polygon features marked by the DFO are allocated to grids with the same resolution as GRACE data according to their spatial positions

We compare the 2380 precipitation-type flood events recorded by the DFO one by one with the real inundation range extracted by B. Tellman1 et al. based on remote sensing images (Tellman et al., 2021). Figure 7 shows part of the flood event comparison results derived based on GRACE, MODIS and DFO data. The dark blue polygons show the approximate

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

flood ranges delineated by the DFO; red pixels are the flood inundation areas extracted based on MODIS data; and light blue regions show the flood days (≥1 day) extracted using GRACE TWS and extreme precipitation data during the period recorded by the DFO. MODIS-based inundations were calculated according to the DFO time period and the union of DFO polygons and HydroSHEDS Basins Level 4 data (Tellman et al., 2021; Lehner et al., 2008; Lehner and Grill, 2013). We also used the DFO-recorded time series as a reference to filter flood days in each grid cell and obtained the spatial flood distributions in specific areas. Sometimes the flood duration in the DFO record was longer (by much more than 1 month), so the spatial distribution area of flood days detected by GRACE in a relatively long time period will be larger than that recorded by DFO. This study focuses on whether the flood events recorded by the DFO can be detected by GRACE, so as long as the number of flood days (≥1 day) extracted by GRACE could be found at the time and space specified by DFO, the effectiveness of the method could be demonstrated. Considering the results of 2380 events, 463 flood events were not detected, resulting in a detection rate of 81%. Among the undetected events, 85 events went undetected due to low precipitation (and were thus not cases of extreme precipitation), while 69 went undetected due to the lack of GRACE data in certain months, resulting in the inability to obtain effective high-frequency signals. Among the remaining 309 undetected floods, the omission of 184 floods may have been caused by the fact that the maximum daily precipitation was less than 50 mm, causing GRACE to fail to identify a flood signal. The other 125 undetected flood events may have been caused by GRACE itself failing to identify flood conditions. To view the spatial distributions of precipitation-type floods and the corresponding situation obtained from flood inundation data extracted from GRACE and MODIS, please refer to the corresponding supplementary materials (https://doi.org/10.5281/zenodo.6831105 (Zhang et al., 2022a)).

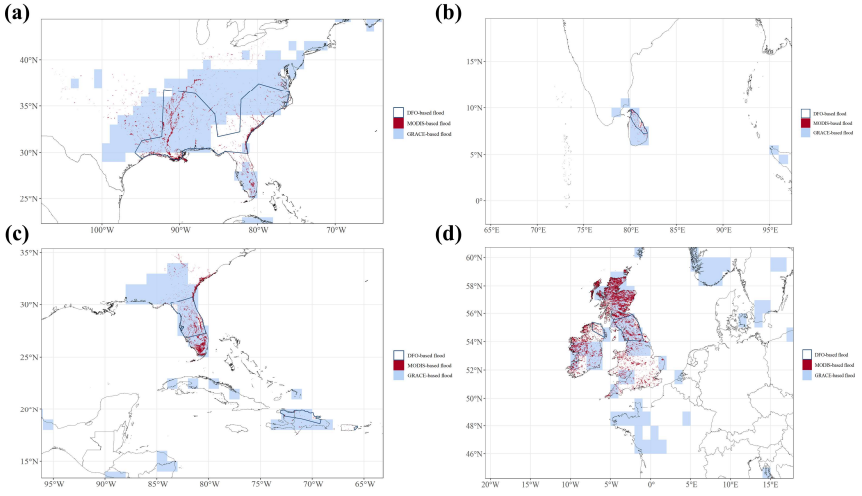


Figure 7 Flood inundation information recorded by the DFO (dark blue polygon), MODIS (red pixels) and GRACE (light blue). (a) The ID-2167 flood event that occurred from Feb. 22nd—Mar. 17th, 2003; (b) the ID-2601 flood event that occurred from Dec. 11th—Dec. 23rd, 2004; (c) the ID-2566 flood event that occurred from Sep. 15th—Oct. 1st, 2004; and (d) the ID-4319 flood event that occurred from Dec. 5th, 2015—Jan. 26th, 2016.

4.3 Further Validation with News Reports

Considering that the flood event records in the DFO database are not complete, we further validate floods that were not recorded by the DFO but were detected in this study. We selected the commonly used international social media platform Twitter, international news reports, Weibo in China and other channels to verify the validity of the extracted results. Jens A. de Bruijn et al (De Bruijn et al., 2019) developed an online database of historical and real-time flood events based on the global mainstream social media platform Twitter from July 2014 to the present. We used Jens A. de Bruijn et al.'s research

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

results because their results have been strictly verified, are updated in real time, and are published in the form of an online database (www.globalfloodmonitor.org) for use by scholars (De Bruijn et al., 2019). According to the spatiotemporal information released by Twitter, potential flood areas can be reflected in real time, thus providing part of the data basis required for verification. In addition to the Twitter social media platform, China's Weibo is also a very active social media platform. Considering that people in China cannot use Twitter, we select Weibo data and corresponding news reports to perform a mutual verification to determine the accuracy of the spatial location and time information characterizing flood events to the greatest possible extent. The floods recorded by Twitter are mainly located in first-order administrative divisions. Therefore, we compared the correspondence between flood events recorded by social media and the news and our results at the global first-order administrative division scale as well as at the national scale. Figure 8 shows some of the global flood events not recorded by the DFO, including floods that occurred in different time periods and areas such as the eastern United States, northern and southern South America, Mozambique in Africa, France, India, China, Malaysia, Indonesia and Australia. These important flood events can be identified in our results.

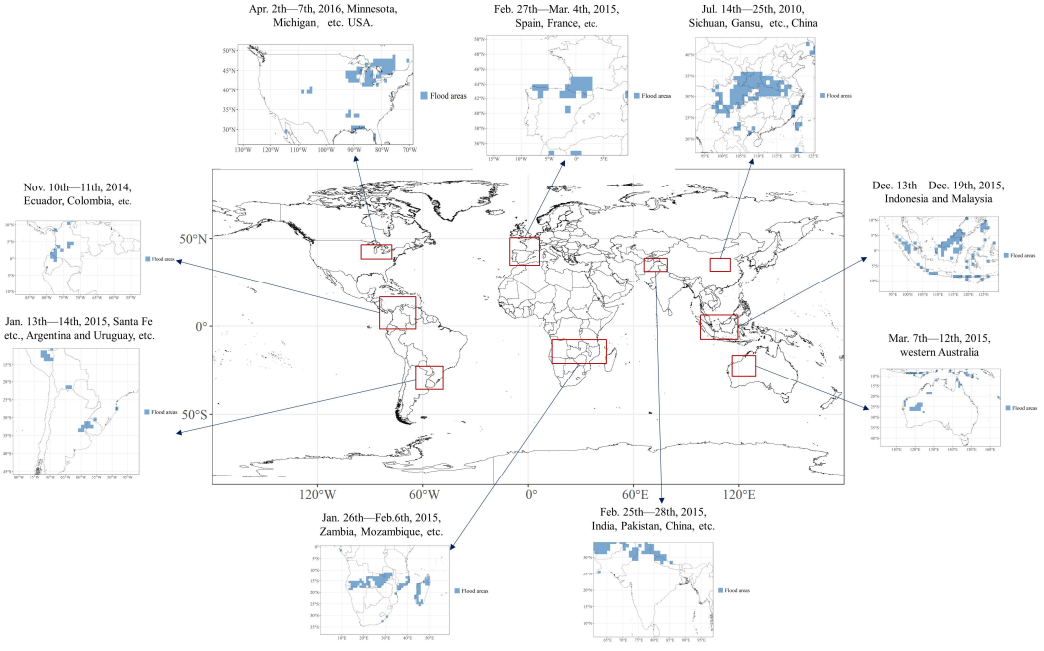


Figure 8 Validation of some flood events recorded by social media and news but not by the DFO.

4.4 Validation with Discharge Data

We also compared our results with discharge data to assess our detection ability. We used the FloodR method of Svenja Fischer et al. (Fischer et al., 2021) to extract flood events from the GRDC discharge data to serve as a basic reference standard when verifying the accuracy of the results extracted in this study. We focused on precipitation-type flood events and similarly constrained the results derived from the discharge data with extreme precipitation data. We considered the spatial average of the discharge data under the coverage of the HydroSHEDS Basins Level 4 data (Lehner and Grill, 2013; Lehner et al., 2008) to eliminate the random errors in individual measurements. The accuracy index used for comparison in this study is the probability of detection (POD) (Yang et al., 2021), i.e., whether the floods extracted from the discharge data correspond to the flood days extracted in this study. Although FloodR cannot guarantee that all flood events will be extracted,

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

and although some of the discharge data themselves were missing or unevenly distributed, the discharge data time series provide us with necessary data to support the reliability of the time series verification process. The spatial distribution of the discharge data is shown in Figure 9. The data distributions in North America, South America, Europe and south-eastern Australia were relatively dense, while data were seriously missing in central, northern and eastern Asia. The flood events in level-4 river basins derived based on discharge data using the FloodR method are shown in Figure 10. We find that most flood events reflected by the discharge data are mainly located in the eastern United States, central South America, western Europe and New Zealand.

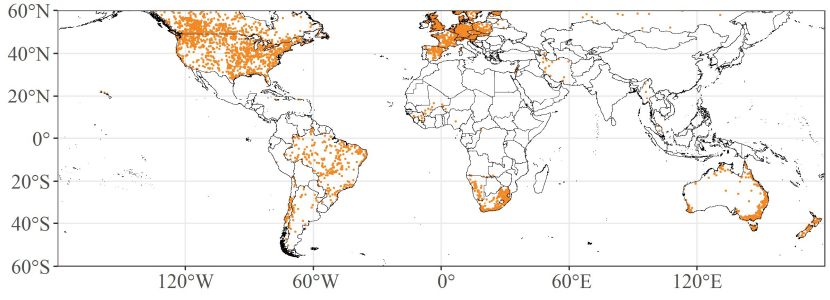


Figure 9 Spatial distribution of GRDC discharge data

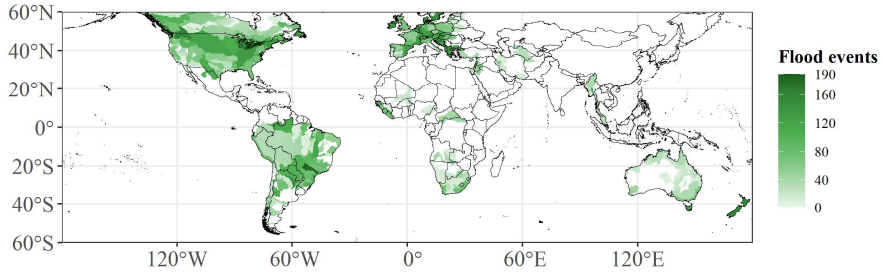
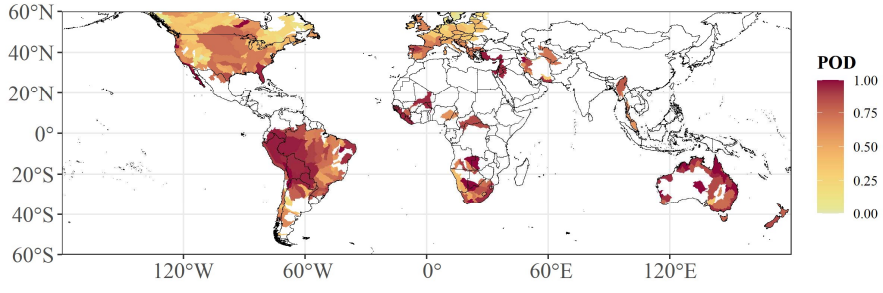


Figure 10 Flood events based on the discharge data in the HydroSHEDS Basins Level 4 dataset

The POD calculation results are shown in Figure 11. The redder the colour is, the higher the corresponding flood detection accuracy is. We found that the overall accuracy performed well; the detection accuracies obtained for the central parts of the United States, western South America, southern Africa and Australia are relatively high. The percentage of sites with POD values greater than 0.5 reached 62%. This finding showed that our extraction results also reflected relatively high accuracies in the time-series comparison.



https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

Figure 11 POD values of the discharge data in the HydroSHEDS Basins Level 4 dataset

4.5 Uncertainty Analysis

The uncertainty analysis performed in this study mainly focused on the selection of the extreme precipitation threshold. The most common method for determining the extreme precipitation threshold is to use the quantile of the analysed time series, considering either the quantile of the entire time series (QETS) data or the quantile of wet days with daily precipitation greater than 1 mm (QWDTS). This study compared the accuracy rates recorded in the DFO database under different quantile threshold selections. We selected the 90th, 95th and 99th quantiles for the two methods described above. Figure 12 shows that the selection of different thresholds in the two extreme precipitation-detection methods influenced the flood extraction accuracy with contributions ranging from 72.4% to 81.4%. This shows that precipitation-type floods are closely related to extreme precipitation and that the selected thresholds affect the detection of approximately 9% (approximately 214) of flood events. We also provide the 6 products derived based on these two constraints for further analysis and use by researchers.

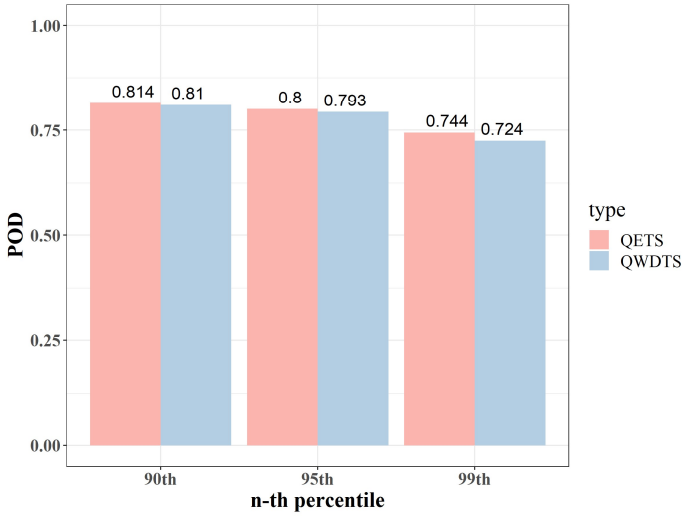


Figure 12 Influence of the selected threshold and extreme precipitation standard on the results

5 Data Usage Instructions

The data obtained herein are provided using the polygon shapefile (SHP file) format. A separate file is provided for each day, and each file represents the global flood day distribution. The spatial resolution is 1°, covering the range between the 60° south and north latitude lines. The SHP files have two fields, namely, "ID" and "Value"; "ID" represents the index number of the 1° grid, and "Value" is a binary variable (with a value of 0 or 1) indicating whether a flood occurred on the specific day. This flood day product can be used to analyse the spatial distributions of historical flood days within 60° north-south latitude and extract specific flood events in combination with historical data from observation sites. At the same time, the products are obtained based on observed data and can be used as to verify flood model results. Considering that the El Niño-Southern Oscillation affects both drought and flood events in different parts of the world, these data can be used to further analyse the impacts of the El Niño-Southern Oscillation on flood days around the world.

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

6 Data availability

The flood day product produced in this study can be obtained from https://doi.org/10.5281/zenodo.6831384 (Zhang et al., 2022b).

7 Conclusion and Discussion

This study successfully extracted global flood days using GRACE TWS and extreme precipitation data recorded between 60°S and 60°N from Apr. 1st, 2002 to Aug. 31st, 2016. The results were compared in time and space with the flood events recorded by the DFO, the most commonly used global flood database, and the results show that our detection results not only identify 81% of the flood events recorded by the DFO but also complement a large number of flood events not recorded by the DFO. To further verify the reliability of the derived products, we compared them with the flood events extracted from global GRDC discharge data, and the probability of detection exceeding 0.5 reached 62%. Moreover, we selected representative flood events not recorded by the DFO but recorded in social media and news in different regions of the world as verification examples. These results also show that our products can better identify flood events.

This study provides, for the first time, quasi-global flood day data obtained based on GRACE TWS and precipitation data, compensating for the lack of a global flood database based on remote sensing observation data. We also acknowledge that there are some limitations to these data. On the one hand, we use extreme precipitation to constrain the data, and the detection ability of some small floods is thus insufficient. Moreover, the results are affected by ocean signals around island countries due to the coarse data resolution, and researchers should be careful when using these data in such areas. However, the value of these results as an important flood observation dataset is still undeniable. Although the dataset has a relatively low resolution of 1°, it represents substantial temporal continuity and spatial accuracy improvements compared to the DFO databases that record only discrete flood events. Moreover, our results provide not only a reference for large-scale quasi-real-time flood event monitoring with the development of GRACE-FO and the improvement of GRACE daily data but also important data support for follow-up studies of the spatiotemporal distributions and attributions of global flood events.

Acknowledgements

The research in this article was supported by the National Natural Science Foundation of China [grant number 41771538]. The financial support is highly appreciated. We also thank the data support of the GRACE daily solution from the Institute of Geodesy, Graz University of Technology.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

Author Contributions

Conceptualization, K. L., J. Z., and M. W.; Methodology, J. Z., K. L., and M. W.; Validation, J. Z. and K. L.; Formal Analysis, J. Z. and K. L.; Investigation, J. Z., K. L., and M. W.; Resources, J. Z.; Data Curation, J. Z.; Writing-Original Draft Preparation, J. Z.; Writing-Review & Editing, K. L.; Visualization, J. Z; Supervision, K. L.; Project Administration, K. L.; Funding Acquisition, K. L. All authors have read and agreed to the published version of the manuscript.

References

Aggarwal, C. C.: Outlier ensembles: position paper, ACM SIGKDD Explorations Newsletter, 14, 49-58, 2013.
Bergmann-Wolf, I., Forootan, E., Klemann, V., Kusche, J., and Dobslaw, H.: Updating ESA's Earth System Model for Gravity Mission Simulation Studies: 3. A Realistically Perturbed Non-Tidal Atmosphere and Ocean De-Aliasing Model, 2015.
Brakenridge, G. R.: Global Active Archive of Large Flood Events. Dartmouth Flood Observatory, University of Colorado,

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

USA. http://floodobservatory.colorado.edu/ Archives/ (Accessed Jan.1, 2022). 2022.
Cazenave, A. and Chen, J.: Time-variable gravity from space and present-day mass redistribution in theEarth system, Earth and Planetary Science Letters, 298, 263-274, 2010.
Chandola, V., Banerjee, A., and Kumar, V.: Anomaly detection: A survey, ACM computing surveys (CSUR), 41, 1-58, 2009.
de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J., and Aerts, J. C.: A global database of historic and real-time flood events based on social media, Scientific data, 6, 1-12, 2019.
Dill, R.: Hydrological model LSDM for operational Earth rotation and gravity field variations, 2008.
Dill, R., Thomas, M., and Walter, C.: Hydrological induced Earth rotation variations from standalone and dynamically coupled simulations, Proceedings of the Journées,
EM-DAT, C.: Centre for research on the epidemiology of disasters, EM Dat The international Disaster Database https://public.emdat.be/ 2014.
Fischer, S., Schumann, A., and Bühler, P.: A statistics-based automated flood event separation, Journal of Hydrology X, 10, 100070, 2021.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote sensing of Environment, 202, 18-27, 2017.
Gouweleeuw, B. T., Kvas, A., Gruber, C., Gain, A. K., Mayer-Gürr, T., Flechtner, F., and Güntner, A.: Daily GRACE gravity field solutions track major flood events in the Ganges–Brahmaputra Delta, Hydrology and Earth System Sciences, 22, 2867-2880, 2018.
Gupta, D. and Dhanya, C.: The potential of GRACE in assessing the flood potential of Peninsular Indian River basins, International Journal of Remote Sensing, 41, 9009-9038, 2020.
Hagen, E., Shroder Jr, J., Lu, X., and Teufert, J. F.: Reverse engineered flood hazard mapping in Afghanistan: A parsimonious flood map model for developing countries, Quaternary International, 226, 82-91, 2010.
Hostache, R., Chini, M., Giustarini, L., Neal, J., Kavetski, D., Wood, M., Corato, G., Pelich, R. M., and Matgen, P.: Near ‐ real ‐ time assimilation of SAR ‐ derived flood maps for improving flood forecasts, Water Resources Research, 54, 5516-5535, 2018.
Huffman, G. J., Bolvin, D. T., Nelkin, E. J., and Tan, J.: Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation, Nasa/Gsfc Code, 612, 2019, 2015.
Huggel, C., Raissig, A., Rohrer, M., Romero, G., Diaz, A., and Salzmann, N.: How useful and reliable are disaster databases in the context of climate and global change? A comparative case study analysis in Peru, Natural hazards and earth system sciences, 15, 475-485, 2015.
Idowu, D. and Zhou, W.: Performance evaluation of a potential component of an early flood warning system—A case study of the 2012 flood, Lower Niger River Basin, Nigeria, Remote Sensing, 11, 1970, 2019.
Kron, W., Steuer, M., Löw, P., and Wirtz, A.: How to deal properly with a natural catastrophe database–analysis of flood losses, Natural Hazards and Earth System Sciences, 12, 535-550, 2012.
Kussul, N., Shelestov, A., and Skakun, S.: Flood monitoring from SAR data, in: Use of satellite and in-situ data to improve sustainability, Springer, 19-29, 2011.
Kvas, A., Behzadpour, S., Ellmer, M., Klinger, B., Strasser, S., Zehentner, N., and Mayer ‐ Gürr, T.: ITSG ‐ Grace2018: Overview and evaluation of a new GRACE ‐ only gravity field time series, Journal of Geophysical Research: Solid Earth, 124, 9332-9344, 2019.
Lehner, B. and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, Hydrological Processes, 27, 2171-2186, 2013.
Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos, Transactions American Geophysical Union, 89, 93-94, 2008.
Manavalan, R.: SAR image analysis techniques for flood area mapping-literature survey, Earth Science Informatics, 10, 1-14, 2017.
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.
Molodtsova, T., Molodtsov, S., Kirilenko, A., Zhang, X., and VanLooy, J.: Evaluating flood potential with GRACE in the United States, Natural Hazards and Earth System Sciences, 16, 1011-1018, 2016.
Moriyama, K., Sasaki, D., and Ono, Y.: Comparison of global databases for disaster loss and damage data, Journal of Disaster Research, 13, 1007-1014, 2018.
Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillmann, J., Schaller, N., Fischer, E., and Schulz, M.: Frequency of extreme precipitation increases extensively with event rareness under global warming, Scientific reports, 9, 1-10, 2019.
Pendergrass, A. G.: What precipitation is extreme?, Science, 360, 1072-1073, 2018.
Rättich, M., Martinis, S., and Wieland, M.: Automatic flood duration estimation based on multi-sensor satellite data, Remote Sensing, 12, 643, 2020.
Re, S.: Sigma: Insurance research.available at http://www.swissre.com/sigma/ (last access: 18 Jun. 2022), 2022.
Reager, J. T. and Famiglietti, J. S.: Global terrestrial water storage capacity and flood potential using GRACE, Geophysical research letters, 36, 2009.
Reager, J. T., Thomas, B. F., and Famiglietti, J. S.: River basin flood potential inferred using GRACE gravity observations at several months lead time, Nature Geoscience, 7, 588-592, 2014.
Robert, C., William, C., and Irma, T.: STL: A seasonal-trend decomposition procedure based on loess, Journal of official statistics, 6, 3-73, 1990.

https://doi.org/10.5194/essd-2022-242
Preprint. Discussion started: 1 August 2022
© Author(s) 2022. CC BY 4.0 License.

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

Schinko, T., Mechler, R., and Hochrainer-Stigler, S.: A methodological framework to operationalize climate risk management: managing sovereign climate-related extreme event risk in Austria, Mitigation and Adaptation Strategies for Global Change, 22, 1063-1086, 2017.

Shi, X., Chen, J., Gu, L., Xu, C.-Y., Chen, H., and Zhang, L.: Impacts and socioeconomic exposures of global extreme precipitation events in 1.5 and 2.0° C warmer climates, Science of The Total Environment, 766, 142665, 2021.

Tellman, B., Sullivan, J., Kuhn, C., Kettner, A., Doyle, C., Brakenridge, G., Erickson, T., and Slayback, D.: Satellite imaging reveals increased proportion of population exposed to floods, Nature, 596, 80-86, 2021.

Vallis, O., Hochenbaum, J., and Kejariwal, A.: A Novel Technique for {Long-Term} Anomaly Detection in the Cloud, 6th USENIX workshop on hot topics in cloud computing (HotCloud 14),

Wahr, J., Molenaar, M., and Bryan, F.: Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, Journal of Geophysical Research: Solid Earth, 103, 30205-30229, 1998.

Winsemius, H., Van Beek, L., Jongman, B., Ward, P., and Bouwman, A.: A framework for global river flood risk assessments, Hydrology and Earth System Sciences, 17, 1871-1892, 2013.

Wu, H., Kimball, J. S., Mantua, N., and Stanford, J.: Automated upscaling of river networks for macroscale hydrological modeling, Water Resources Research, 47, 2011.

Wu, H., Adler, R. F., Hong, Y., Tian, Y., and Policelli, F.: Evaluation of global flood detection using satellite-based rainfall and a hydrologic model, Journal of Hydrometeorology, 13, 1268-1284, 2012a.

Wu, H., Adler, R. F., Tian, Y., Huffman, G. J., Li, H., and Wang, J.: Real‐time global flood estimation using satellite‐based precipitation and a coupled land surface and routing model, Water Resources Research, 50, 2693-2717, 2014.

Wu, H., Kimball, J. S., Li, H., Huang, M., Leung, L. R., and Adler, R. F.: A new global river network database for macroscale hydrologic modeling, Water resources research, 48, 2012b.

Xiong, J., Wang, Z., Guo, S., Wu, X., Yin, J., Wang, J., Lai, C., and Gong, Q.: High effectiveness of GRACE data in daily-scale flood modeling: case study in the Xijiang River Basin, China, Natural Hazards, 1-20, 2022.

Yang, Y., Pan, M., Lin, P., Beck, H. E., Zeng, Z., Yamazaki, D., David, C. H., Lu, H., Yang, K., and Hong, Y.: Global Reach-Level 3-Hourly River Flood Reanalysis (1980–2019), Bulletin of the American Meteorological Society, 102, E2086-E2105, 2021.

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.