Uncertainty in flood frequency analysis of hydrodynamic model simulations

Xudong Zhou[1], Wenchao Ma[1], Wataru Echizenya[2], and Dai Yamazaki[1]

[1]Institute of Industrial Science, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo, 153-8505, Japan
[2]Corporate Planning Department, MS&AD InterRisk Research & Consulting, Inc., 2-105, Kanda Awajicho, Chiyoda-ku, Tokyo 101-0063, Japan

Correspondence: Xudong Zhou: x.zhou@rainbow.iis.u-tokyo.ac.jp

Abstract.

Assessing the risk of a historical-level flood is essential for regional flood protection and resilience establishment. However, due to the limited spatiotemporal coverage of observations, the risk assessment relies on model simulations and is thus subject to uncertainties from cascade physical processes. This study assesses the flood hazard map and also the uncertainties with different combinations of runoff inputs, variables for flood frequency analysis, and fitting distributions based on estimations by the CaMa-Flood global hydrodynamic model. Our results show that deviation in the runoff inputs is the most influential source of uncertainties in the estimated flooded water depth and inundation area, contributing more than 80% of the total uncertainties. Global and regional inundation maps for floods with 1-in-100 year return period show large uncertainty values but small uncertainty ratios for river channels and lakes, while the opposite results are found for dry zones and mountainous regions. This uncertainty is a result of the fitting distributions. In addition, the selected variables are limited but increase from the regular period to the rarer floods, both for the water depth at points and for inundation area over regions. The uncertainties in inundation area also lead to uncertainties in estimating the population and economy exposure to the floods. In total, inundation accounts for 9.1%[8.1–10.3%] of the land area for a 1-in-100 year flood, leading to 13.4%[12.1–15%] of population exposure and 13.1%[11.8-14.7%] of economic exposure for the globe. The flood exposure and uncertainties vary in continents and the results in Africa have the largest uncertainty probably due to the limited observations to constrain runoff simulations, indicating a necessity to improve the performance of different hydrological models especially for data limited regions.

1 Introduction

A flood hazard map (FHM) is a map of flood water depth or inundation area at a specific return period (e.g., 1-in-100 year return period). FHM provides information for flood risk assessment, which is helpful for stakeholders and insurance services (Osti et al., 2008; Luke et al., 2018). However, FHM is a theoretical map of a global-identical reoccurrence (e.g., 1-in-100 year return period), and thus it is not observable. Production of the FHM is based on flood frequency analysis (FFA) with

simulations of flow characteristics (e.g., discharge, water stage, water volume) from flood models (Liscum and Massey, 1980; Wiltshire, 1986; Hamed and Rao, 2019) and a fitting regression to a specific reoccurrence.

Winsemius et al. (2013) established a framework for river flood risk assessment with cascaded global forcing datasets, a global hydrological model, a global flood-routing model, and an inundation downscaling routine. These authors used a single hydrological model (PCR-GLOBWB) to evaluate flood risk in south Asia. However, they recommend that the framework should be extended to a multi-model approach to address any uncertainties. Trigg et al. (2016) analyzed eight global flood hazard models over Africa and China and show that there was only 30–40% agreement in the flood extent and significantly large deviations in flood inundation area, economic loss and exposed population estimates. A similar multi-model approach was applied in Bernhofen et al. (2018) and Aerts et al. (2020). However, because the eight global flood hazard models use different forcing inputs, hydrological models, river routing models and spatial resolutions, it is impossible to attribute how much each process contributes to the uncertainties in the final results or which process is dominant. These authors suggested that component-level comparisons with limited variables could be better able to attribute the uncertainties. Schellekens et al. (2017) therefore controlled the forcing inputs but investigated 10 global hydrological models in terms of evapotranspiration, runoff and soil moisture. However, the flood hazard was not investigated because the river routing model was not applied. Zhao et al. (2017) further evaluated routing models in reproducing the peak river discharge, while uncertainties and results of inundation are not discussed.

Running a flood model requires a large set of model inputs, model parameters, topographic information, and so on. Therefore, implementing flood models at a local or regional scale is much easier than global implementations. The variety of uncertainties has been discussed for specific flood events at a local or regional scale in (Merwade et al., 2008; Bales and Wagner, 2009; Beven et al., 2015). The sensitivity of the inundation to selection of forcing inputs (Ward et al., 2013), Digital Elevation Models (DEM) (Tate et al., 2015), roughness (Pappenberger et al., 2008), spatial resolutions (Merwade et al., 2008), or fitting functions (Kidson and Richards, 2005) has also been analyzed. However, because the regional analysis is highly dependent on the availability of local data, the results and conclusions are not necessarily applicable to other regions or to the global scale. Therefore, we are curious about the magnitude and the spatial patterns of the sensitivity of FHMs to various factors at the global scale.

Zhao et al. (2017) suggested that runoff differences will lead to wide ranges of the uncertainty in peak discharge. Therefore, runoff is selected as an uncertainty source to the FHMs to be investigated in this study. Because length of observations or forcing data is limited, obtaining a FHM with a low-frequency (e.g., 1-in-100 year return period) requires extrapolation based on curve-fitting to the existing data or simulations (Kidson and Richards, 2005). The limitation of FFA is therefore apparent as the fitting based on a priori assumption about the underlying distribution generating flood events. However, because the limited length of the records hardly represents the complete characteristics, a range of more-or-less skewed, relatively complex distributions is always considered to account for the uncertainties. Typical distributions that are used include Pearson type III, Log-Pearson, Gauss, Gumbel and Log-normal distribution (Radevski and Gorin, 2017; Drissia et al., 2019). However, no conclusion is found to support which of the fitting distributions is preferable for most of the regions (Drissia et al., 2019). Therefore, it is recommended that different distributions should be tested with local records.

The FFA can be conducted on any characteristics of river flow, but is mainly used for river discharge and water stage (or water level or water depth) because they can be recorded as gauge observations (Radevski and Gorin, 2017). There is no preference for these two variables and the selection is only determined by data accessibility. The results of the FFA based on the discharge will be slightly different from the results with a water stage because of the loop rating curve relationship between discharge and water stage (Domeneghetti et al., 2012; Alvisi and Franchini, 2013). In addition, Pappenberger et al. (2012) uses river water storage provided from flood models, which is then remapped the fitted water storage to inundation extent. The increase of water storage and water stage is non-linear because of the topographical variety in river channels and floodplains. Therefore, selection of different variables for the fitting is another source of uncertainty for flood estimations.

There are many other uncertainties that can lead to deviations in mapping the floods, including forcing, routing and down-scaling. However, we need to limit the factors to avoid adding too much complexity to the analysis. Therefore, in this study, we will investigate the FHMs along with uncertainties due to selected factors (i.e., runoff generation models, the fitting distributions and the variables to be fitted). Section 2 describes the methods and data that we used. In Section 3, we assess the fitting performance of FFA for all combination of experiments with different flow variables used for FFA, fitting distributions, and the runoff that drives CaMa-Flood. We then present the flood water depth and contributions from different factors over the globe and regional cases for a 1-in-100 year return period. The flood water depth for specific points and the inundation area for specific regions at multiple return periods are discussed, together with their uncertainties. The potential impact (exposure) of the floods on the population and the economy are investigated on different continents. The discussions and conclusions follow in Section 4.

2 Methods and datasets

2.1 Experiment design

The cascade of generating the global flood hazards maps comprises the following steps: 1. global forcing data; 2. global hydrological models; 3. global river routing models; 4. FFA (Winsemius et al., 2013). In this study, we limit the factors to be investigated on the global hydrological models and the FFA. The uncertainties to be investigated in this study are attributed to three major sources, as follows: first, the variables used for the FFA; second, the fitting distributions used for FFA; and third, the runoff inputs to the river routing model–Catchment-based Macro-scale floodplain model (CaMa-Flood). Each experiment is therefore a combination of the three sources (Table 1).

For the variables selection, V1 represents that FFA is based on the numeric results of "river water depth" provided by CaMa-Flood. In V2, the FFA was first conducted on the estimated water storage, which is the only prognostic variable in the CaMa-Flood. Then, at each return period (e.g., 1-in-100 year), the river water depth was estimated based on the storage-water depth relation and the corresponding water storage. Because of the non-linear relation between water level and storage, the fitting will lead to different results. The differences between experiment V1 and V2 denote the uncertainty that results from the selection of the target variables that we used for the FFA. Despite river water depth and water storage, discharge is the variable that is most frequently used in engineering design because discharge is frequently measured. However, with only discharge we

cannot estimate the water level (or water storage) because the relationship between discharge and water level is not one-to-one consistent because of the loop rating curve. While with either river water depth or water storage, we can estimate the flood extent and the floodplain water depth for any target region using CaMa-Flood.

Table 1. Experiments used in this study for uncertainty analysis. There are three groups, as follows: (A) the variables for FFA (B) the fitting distributions and (C) the runoff inputs. Different runoffs are generated by using the same forcing (WFDEI) but with different land surface models or global hydrological models (as specified in the brackets).

A	Variables	B	fitting distribution	C	Runoff
V1	$rivdph$	F1	GEV (Generalized Extreme Value)	R1	e2o_anu (W3)
V2	$storage$	F2	GAM (Gamma)	R2	e2o_cnrs (ORCHIDEE)
		F3	PE3 (Pearson type III)	R3	e2o_jrc (Lisflood)
		F4	GUM (Gumbel)	R4	e2o_ecmwf (HTESSEL)
		F5	WEI (Weibull)	R5	e2o_nerc (JULES)
		F6	WAK (Wakeby)	R6	e2o_univk (WaterGAP3)
				R7	e2o_univu (PCR-GLOWB)

Uncertainty due to the fitting distributions used for FFA was evaluated as the resulting differences by applying various fitting functions (i.e., F1 – F6). These distributions are generally used in FFA but for different variables in different fields, and they were treated without priorities in this study. The samples were automatically fitted without any manual modifications in their parameters with L-moments optimization.

The results of the FFA were based on the output of CaMa-Flood associated with different runoff inputs. In our case, the CaMa-Flood was driven by seven different kinds of runoff forcing (i.e., R1–R7) from eartH2Observe (e2o) category (Schellekens et al., 2017). The runoffs were driven by the same WATCH Forcing Data methodology applied to ERA-Interim data (WFDEI, Weedon et al., 2014) but with different land surface or hydrological models. Therefore, the runoff inputs contain the uncertainties in the rainfall-runoff model processes (model structures and model parameters). Therefore, the deviation of the results in the FFA among the seven inputs was the uncertainty caused by the runoff inputs.

2.2 Global river routing model (CaMa-Flood)

The CaMa-Flood is designed to simulate the hydrodynamics in continental-scale rivers. Entire river networks are discretized to irregular unit-catchments with the sub-grid topographic parameters of the river channel and floodplains. The river discharge and other flow characteristics can be calculated with the local inertial equations along the river network map. The water storage of each unit-catchment is the only prognostic variable that is to be solved with the water balance equation. The water level and flooded area are diagnosed from the water storage at each unit-catchment using the sub-grid topographic information. Detailed descriptions of the CaMa-Flood can be found in the original papers by Yamazaki et al. (2011, 2012, 2014).

The major advantage of the CaMa-Flood simulations is their explicit representation of flood stage (water level and flooded area) in addition to river discharge. This facilitates the comparison of model results with satellite observations, either the

altimeters by Synthetic-aperture radar (SAR) or inundation images by optical or microwave imagers. The estimation of the flooded area is helpful in the assessment of flood risk and flood damages by overlaying it with socio-economic datasets.

Another apparent advantage of the CaMa-Flood is its high computational efficiency of the global river simulations. The CaMa-Flood utilizes a diagnostic scheme at the scale of unit-catchment to approximate the complex floodplain inundation processes. The prognostic computation for water storage is optimized by implementing the local inertial equation and the adaptive time step scheme. The high computational efficiency is beneficial for implementations at a global scale. This is critically important because ensemble simulations are frequently applied to account for uncertainties but computation time will be multiplied manyfold. In this study, CaMa-Flood was driven by different runoff inputs (see section 2.1) to achieve the flow characteristics at each unit-catchment at the global scale. The FFA is conducted based on the flow characteristics using CaMa-Flood.

2.3 Flood frequency analysis (FFA)

The runoff inputs are available from 1980 to 2014 (35 years in total). For a specific unit-catchment defined in the CaMa-Flood, the maximum value of the daily river water depth or catchment water storage was obtained for each year and was then sorted. The frequency as the return period (P_m) was calculated with the following equation:

$$P_m = \frac{m}{N+1}, \tag{1}$$

where m is the sorted ranking, N denotes the number of total years (herein 35).

The parameters of the fitting distributions were then calculated based on these sorted annual values with the L-moment method (Hosking, 2015; Drissia et al., 2019). This is defined as a linear combination of probability-weighted moments of the time series. The parameter estimations using L-moment and quantile functions used for different distributions have been described in detail in Hosking (1990). The computation of the parameters was done in the Python *lmoments3* Library. Note that the Wakeby (WAK) is a five-parameter function; the GEV, PE3 and WEI are three-parameter functions; while GAM and GUM are two-parameter functions.

Akaike Information Criterion (AIC, Akaike, 1974; Sakamoto et al., 1986; Mutua, 1994) was used to evaluate the performance of the FFA against the annual values. aic is calculated as

$$aic = 2k + n \cdot log(\frac{\sum (S-O)^2}{n}), \tag{2}$$

where k is the number of parameters needed for the fitting distribution, S represents the simulated values, O represents the observed values, n denotes the number of samples. Smaller aic denotes higher fitting performance because of smaller deviations between simulations and observations. Although there are various performance metrics to measure the goodness-of-fit, the aic is used in our study because it will enlarge the small difference between samples and estimations. We only have 35 samples and these are sorted, therefore the fitting performance should be very high and the fitting results should not have large differences.

2.4 Downscaling to high-resolution inundation map

To reduce the computation cost due to high-resolution simulations, the CaMa-Flood was run globally at a 0.25° spatial resolution, which means that only one unit-catchment was assigned for each 25 km by 25 km grid. The performance of the FFA was evaluated with aic at the global scale to capture the overall features (see Section 3.1).

However, it is difficult to characterise the river water depth or inundation area in detail with local topography at a low resolution (0.25°), and it is difficult to visualize the inundation map at a high resolution (<100m) for the globe. Therefore, high-resolution (3 arcsec, ∼90 m at the equator) regional analysis related to the flood water depth and inundation area with their uncertainties was conducted regionally over the lower Mekong River Basin, where the delta is vulnerable to floods (Shin et al., 2020). Corresponding results on the uncertainties in water depth and inundation area, as well as the flood water depth at specific points will be presented.

The estimated low-resolution storage was downscaled to the high-resolution inundation map with the topography map – Multi-Error-Removed Improved-Terrain DEM (MERIT, Yamazaki et al., 2017) at 90 m. The fundamental assumption is that the movement of water within a unit catchment is instantaneous and that the water surface is flat within the unit catchment at each time step (Zhou et al., 2020). The total water storage under the identical water level should be equal to the water storage estimated in this unit-catchment (see Figure 1-a). The area of lowest elevation is inundated first, until the total water volume approximates the estimated water storage of the unit catchment. The relationship between the water level and water storage or the flooded area is illustrated in Figure 1-b. When the floodplain has been inundated, the small increases of water level corresponds to large changes in the water storage and the flooded area. River water depth can be saturated after inundation (it does not react significantly to the increase of storage after flooding), and this might cause an error in function fitting. The assumption of a flat water surface is not valid for large water bodies (e.g., large lakes or reservoirs with water surface gradient) and steep river segments (e.g., mountainous area). However, the impact of violation is limited at the catchment scale with a grid size of 25 km (which is consistent with the global scale). The inundation area over the mountainous area is also limited compared to that in the floodplains.

The flood exposure of the population and economy is estimated based on the inundation map and population density map (Gridded Population of the World – GPW) in year 2015 (Center for International Earth Science Information Network (CIESIN), 2018), as well as the Gross Domestic Production map in year 2015 (Kummu et al., 2018). The two maps are in 30 arcsec, therefore the 3 arcsec inundation map was aggregated to the 30 arcsec.

3 Results

3.1 Fitting performance

In this section, we will first analyze the fitting performance using aic for all different experiments listed in Table 1. Note that the river water depth and the river water storage are in different units or magnitudes. aic is therefore only applied to the normalized values of water depth or water storage ([0,1]) for each grid (divided by the maximum value of each grid). The

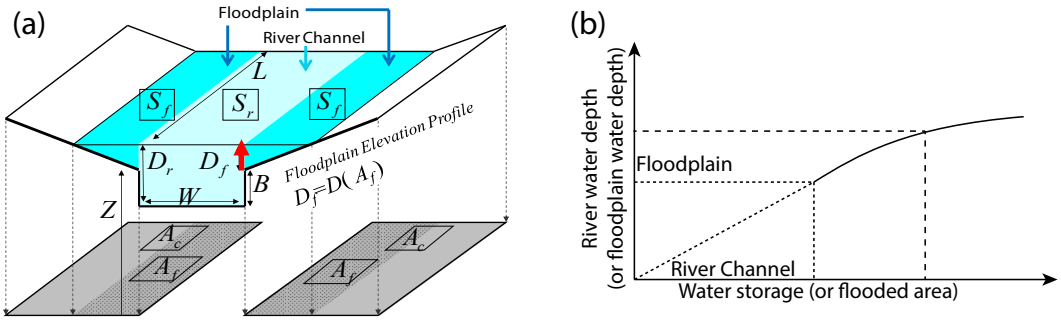

Figure 1. (a) Illustration of a river channel reservoir and a floodplain reservoir defined in each unit-catchment. The water level for the river channel and floodplain is assumed to be the same in each unit-catchment. The denotation of each parameters and its calculation can refer to (Yamazaki et al., 2011). (b) The relationship between the water level and water storage, as well as the flooded area for a specific unit-catchment. The shape of the curve within the river channel is determined by the profile of the river channel and the curve above the river channel is mainly affected by floodplain topography.

fitting performance was evaluated by the aic value (Eq. 2). A lower aic indicates a better fitting performance. Figures 2-a and b display a sample result for e2o_ecmwf (R4) and GEV fitting distribution (F1) for water level and storage, respectively. The difference between the two maps is shown as Figure 2-c.

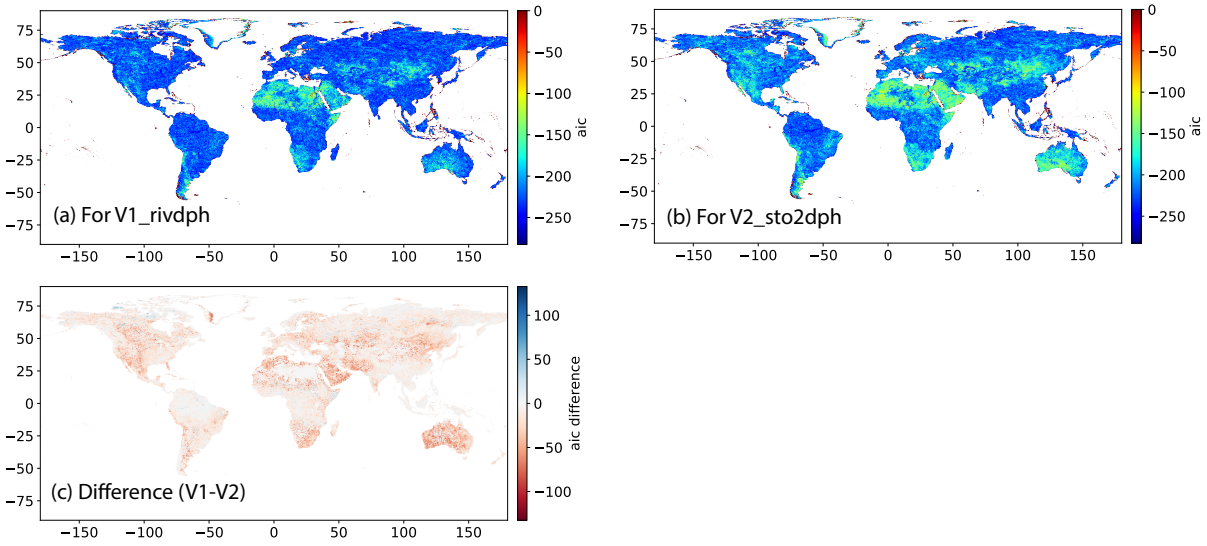

Figure 2. Fitting performance of FFA for (a) V1 (river water depth) and (b) V2 (water storage). The performance was quantified with aic and (c) is the aic difference of (a) and (b). This is only an example for e2o_ecmwf and GEV fitting distribution.

The fitting performance is relatively high with low aic (<-50) in most of the unit-catchments. This is happens because we have only a few samples (35) and the time series was normalized to a range between 0 and 1. The advantage of the aic is that it enlarges the small difference so that we can see a large deviation between different experiments. Relatively low fitting performance is found in the Greenland area and those dry areas in the Sahara, Mongolia and middle Australia (Figure 2-a). The area with low fitting performance (high aic) increases when dealing with the storage, typically in Mongolia, Australia, South Africa, south Latin America and in the western part of North America. These regions are mainly dominated by dry climate or mountainous topography. The accumulative river discharge over those regions is small. The magnitude is thus highly depended on single precipitation events, leading to an unstable relationship between the high floods in different years.

The difference of the aic values for the river water depth and that for the storage is mapped as Figure 2-c. Negative values indicate that the fitting performance is better for water depth than for that for the water storage. Despite the near-zero values, negative values (red scatters) are distributed in the main parts of the world. The places with the largest differences are distributed in the northern and southern Africa, Australia, Northern China, Western America, which is consistent with the high values in Figure 2-b. Although positive values are also found, the values are not large. The results indicate that for most of the lands, the fitting on the data of river water depth is better than the fitting on the water storage. However, this is only the result of a case with e2o_ecmwf runoff input and GEV distribution.

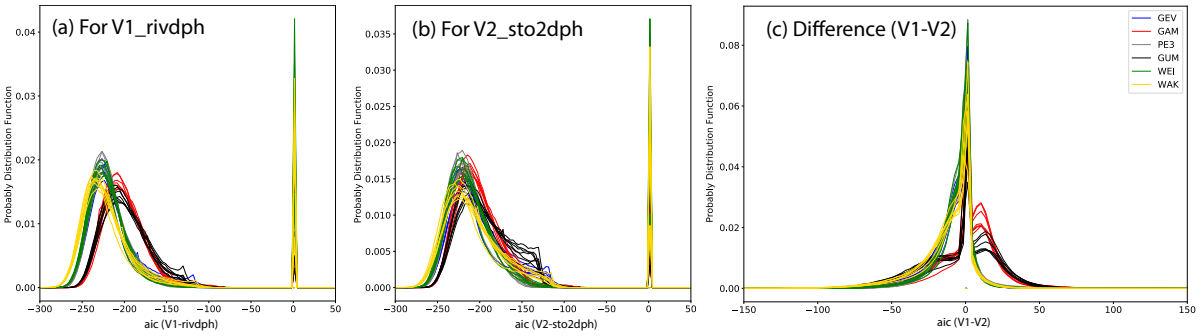

Figure 3. Overall performance of FFA for (a) V1 (river water depth) and (b) V2 (water storage). The performance aic over all the land grids are collected and displayed as the histogram. (c) is the aic difference between (a) and (b). Negative difference indicates better performance of FFA for V1. Different colors represent different fitting distributions, and the multiple lines in a specific color represents results driven by different runoff inputs. The types of the runoff inputs are not specified in these three graphics.

An overall evaluation of all of the distributions and runoff inputs is shown in Figure 3. The probability distribution of the aic values for all the global grids are plotted in Figure 3-a and b using water river depth and storage, respectively. The pdf curves have two peaks, one is normally distributed with mean values around -200 (or -220) and the other is near 0. The latter peak around 0 corresponds to the red scatters in Figures 2-a and b, showing poor fitting performance of the distributions over the coastal regions. The difficulties in representing coastal rivers in CaMa-Flood should be the reason or this. From the variations of curves in different curve in the same color, we find that the performance metric aic is not too sensitive to the

runoff. Regarding the differences among different distributions, WAK (yellow lines) have the lowest aic values with the best performance while GAM (red lines) and GUM (black lines) have the largest values with the poorest performance in Figure 3-a. The other three distributions (GEV, PE3 and WEI) have a similar and moderate performance for water depth. The differences of the fitting performance are mainly due to the degree of freedom of each fitting distribution because the WAK has five parameters, GAM and GUM have two parameters while the others have three. With a higher degree of freedom, the fitting performance will be better. Meanwhile, compared to fitting with the river water depth, the curves were not so distinguishable in Figure 3-b. This indicates that the results are not sensitive to the fitting distribution. As shown in Figure 1-a, the water level is calculated by allocating the water storage to the river channel and floodplain from the bottom to the top. In the channel, the relationship between water level and storage is linear, while it is nonlinear in the floodplains. So, if the maximum water level for the different years locates in both the river channel and the floodplain, then fitting the water level becomes more difficult; especially for GAM and GUM because they have only two parameters. Given that the storage is not affected by the channel shape, fitting the water storage with different fitting functions will not make a large difference.

Figure 3-c shows the difference of fitting performance for water depth and water storage (corresponding to Figure 2-c if e2o_ecmwf and GEV is specified). As in Figure 2-c, negative values indicate that the fitting performance for water depth (V1) is better than that for the water storage (V2). More negative values were found for the distributions of WAK, GEV, PE3 and WEI, especially within the range of [-50, 0]. While for GAM and GUM, more positive values are found within the range of [0, 25], showing better performance for water storage than that for water depth. However, as we see from Figure 3-a and 3-b, the fitting performance of GAM and GUM is not as good as other functions. Because the normalization did not change the relative magnitude of different values, the difference between fitting river water depth and water storage results from their relationship (Figure 1). For the floods (tails of the fitting distribution), the changes in water storage should be larger than that changes in the water level if given a shift of the flood frequency. This leads to the resulting difference in the fitting performance.

3.2 Flood water depth at 1-in-100 year return period

3.2.1 Global flood depth

This section summaries the mean flood water depth and the related uncertainties over the globe at 1-in-100 year return period (Figure 4). The results are based on the original estimations of the FFA, rather than the results after normalization presented in the previous section. For the mean values (Figure 4-a), the floodplain water depth will only exceed 10 m in most of the main channels of large rivers, especially in the Amazon River, large rivers in southern China, southeastern Asia and Siberia. The standard deviation of the flood water depth (Figure 4-b) shares the same spatial patterns with the mean values. The deviation in large rivers can reach 5 m or more, which indicates a high degree of uncertainty in estimating the water depth. However, the spatial patterns of the coefficient of variation (Cv, ratio of the standard deviation to the mean) are opposite because Cv is lower where the mean or deviation is higher, and vice versa. The regions with high Cv are likely to be the dry zones (e.g., Sahara, Central Australia, and Central Asia) and the originating river basins in mountainous regions (e.g., the Rocky Mountains, the Andes, and the Tibetan Plateau).

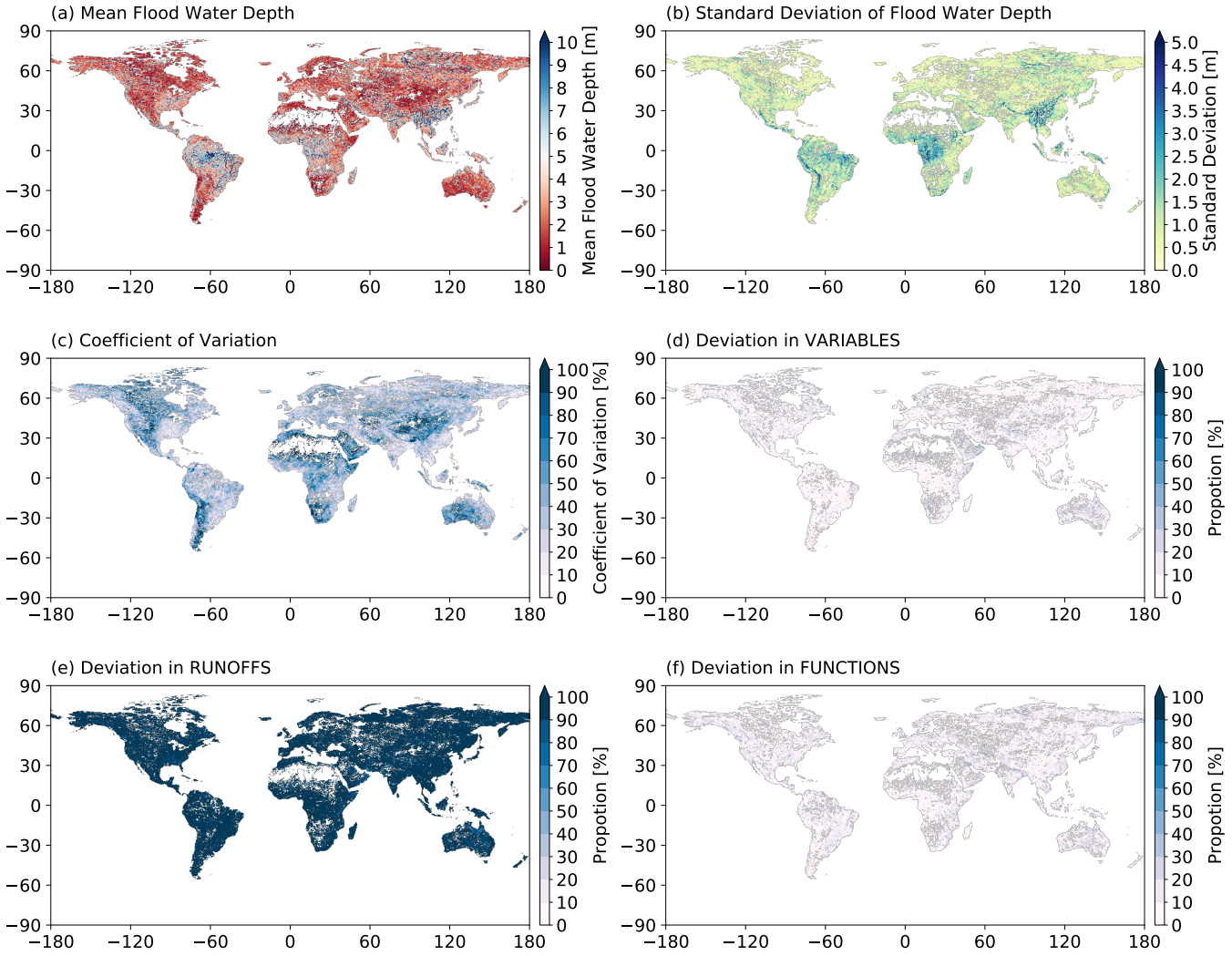

Figure 4. The mean and uncertainties of the flood water depth for the 1-in-100 year return period. The mean floodplain water depth (a), the standard deviation (b) and the coefficient of variation (c) are estimated based on all of the experiments. The deviation proportion to the overall standard deviation (b) is displayed in (d)–(f) for different variables, runoffs and fitting functions, respectively.

This deviation in the flood water depth can be caused by various factors, including the used variables, runoffs and the functions listed in Table 1. Figures 4-d, e and f show the proportion of the standard deviation due to each factor to the total standard deviation in Figure 4-b. A larger proportion indicates the deviation due to a certain factor contributes more to the total standard deviation. Therefore, for most of the global grids, runoff deviation from different land surface models or global hydrological models is the major contributor, taking a proportion larger than 80% (Figure 4-e). Schellekens et al. (2017) evaluated the monthly anomalies with the signal-to-noise ratio (SNR) among all runoff inputs that are used in this study. Their results suggested that the runoff has a larger spread over cold regions (e.g., high latitudes in Asia and North America, and the Tibetan Plateau) and dry zones (e.g., Sahara, and Central Asia). However, the spatial patterns of runoff spread are not seen in the deviation ratio of the flood water depth Figure 4-e). This suggests the spread of flood water depth due to runoff is not sensitive to the climate zones.

The deviation among different variables (Figure 4-d) or functions (Figure 4-f) contributes similarly with a very small proportion to the total deviation. The difference is the deviation due to variables is scattered and likely to have larger values in dry regions or coastal areas. While a larger deviation among different fitting functions is primarily found along the large rivers. A difference indicates that the flood water depth will be more sensitive to the functions while less sensitive to selected variables in large rivers (higher water level or larger water storage). Therefore, more attention is needed to select the fitting function when evaluating the flood risks for large river basins.

3.2.2 Regional flood water depth

The global analysis is at the spatial resolution of 0.25°, which is insufficient to show enough spatial details. In this section, we evaluate the uncertainty range in the water level and inundation at a higher spatial resolution (i.e., ∼90m) after applying the downscaling (see Section 2.4). The analysis presented in the main text is for the lower Mekong region where the delta is vulnerable to floods. We also provide results and analyses for other large river basins (e.g., Amazon, Yangtze, Mississippi, Lena, and Nile) in the supporting material.

Figure 5-a displays the flood water depth for the 1-in-100 year flood at 90 m for the lower Mekong. The largest water depth (>10.0 m) is found in the centre of Tonle Sap Lake and the main channel of the Mekong River. A large extent in the lower Mekong delta suffers relatively low inundation water depth (in dark red). Low water depth also occurs along the boundaries of lakes and main channels. The river tributaries have low average water depth in all of the experiments.

Figure 5-b shows the uncertainties resulted from different experiments listed in Table 1. In general, the uncertainty range is higher where the estimated water depth is deeper (Figure 5-a) because the largest uncertainties are found in the main channel of Mekong with magnitude higher than 2.0 m, while the lowest uncertainties are found in the deltas. The uncertainty in the Tonle Sap Lake is homogeneous with a magnitude around 1.0 m. The coefficient of variation (Figure 5-c) is higher where the mean flood water depth and the deviation is smaller. The overall uncertainties mainly result from the runoff inputs (Figure 5-e) and from the fitting distributions (Figure 5-f) and the variables (Figure 5-d). This is consistent with the conclusions from the global analysis.

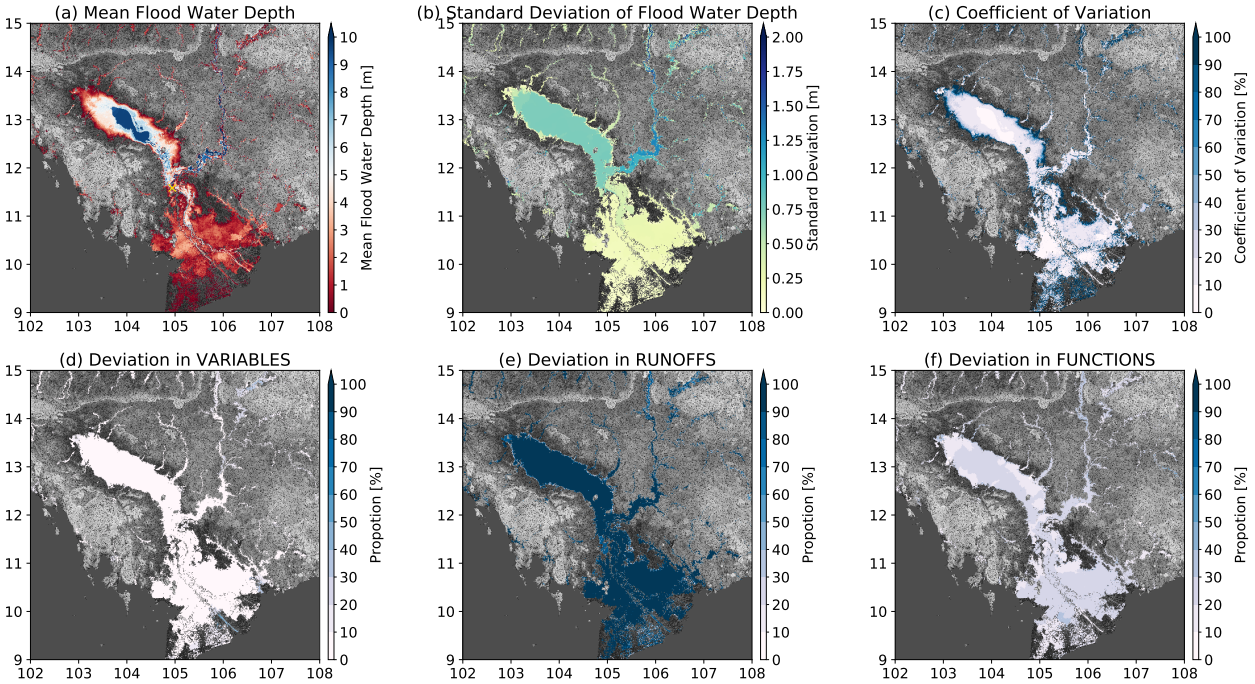

Figure 5. The mean and uncertainties of the flood water depth for the 1-in-100 year return period in the lower Mekong River Basin. The mean floodplain water depth (a), the standard deviation (b) and the coefficient of variation (c) are estimated based on all the experiments. The deviation proportion to the overall standard deviation (b) is displayed in (d)–(f) for different variables, runoffs and fitting functions, respectively. Area with floodplain water depth less than 0.01 m are masked out. We use Multi-Error-Removed Improved-Terrain DEM (MERIT DEM) as the terrain model. The cross in yellow in (a) is the representative GRDC gauge to be analyzed in the next subsection.

We also investigated the flood water depth for other rivers, including Amazon, Yangtze, Mississippi, Lena, and Nile (see Figure S1-S5). Floods will cause a large inundation area in the deltas although the flood water depth is small. Higher uncertainty in water depth with lower coefficient of variation is found in the river channels. While lower uncertainty of water depth with higher coefficient of variation is found for the delta plains. The uncertainties are still mainly caused by the runoff inputs. The selected variables and fitting functions will not lead to large deviations compared to the runoff inputs.

3.3 Flood water depth for multiple return periods

3.3.1 Point analysis

In addition to the global and regional pattern at a single return period (1-in-100 year as shown in previous section), we are also curious to understand how the uncertainty varies at different return periods. We selected the Phnom Penh (Latitude: 11.5617,

Longitude: 104.9317, yellow cross in Figure 5-a), which is a representative GRDC gauge at the confluence point of the outlet of the Tonle Saple Lake and the main Mekong River. The estimated mean water depth and the uncertainty range (standard deviation) for different conditions are plotted as the solid line and shaded area, respectively, in Figure 6. The overall mean value of the estimated water depth is shown in Figure 6-a. The water depth at 1-in-2 year return period is 8.14 m and it is 9.58 m for the 1-in-100 year flood. The overall standard deviation is large up to 0.69 m and it is generally the same for different return periods.

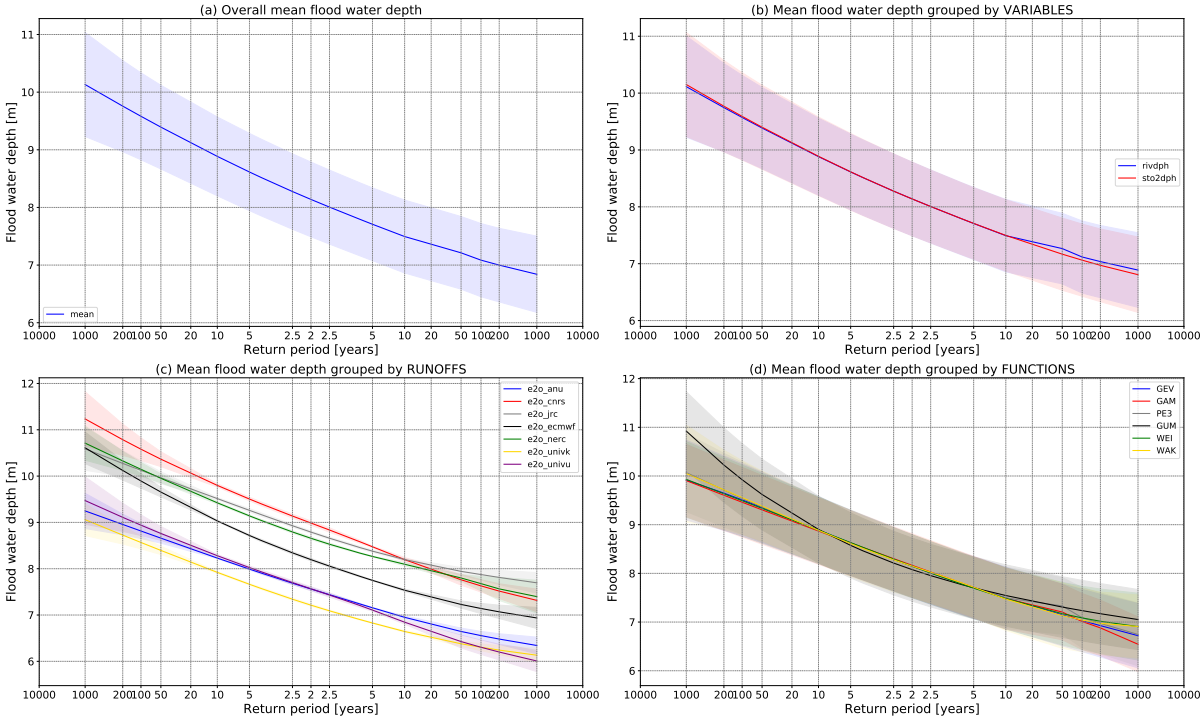

Figure 6. Uncertainties in the estimated floodplain water depth at Phnom Penh (104.9317°E, 11.5617°N) in the Mekong River Basin in different groups. a) the mean floodplain water depth and overall uncertainty; b) the mean and uncertainty in groups of different variables for FFA, the uncertainty is then not related to the selected variable; c) the mean and uncertainty in groups of different runoff inputs; d) the mean and uncertainty in groups of different fitting distributions.

In Figure 6-b, the differences between mean floodplain water depth using river depth (V1) and storage (V2) is very small. The uncertainty range is still as large as that in Figure 6-a. This indicates that the uncertainty receives little contribution from the variables for FFA but is large for other sources. Similarly, subtracting the uncertainty from fitting distributions does not apparently decrease the uncertainty range (Figure 6-d). This indicates that the uncertainty that resulted from the selection of fitting distribution is still small. In particular, the mean value for GUM function in the tails of the floods (more than 1-in-20 year flood) is higher than results of other functions, indicating that GUM may provide a relatively deviated estimate of mean

floodplain water depth for the extreme flood events. This difference of GUM mainly happens because GUM only has two degrees of freedom. The uncertainty ranges of other uncertainties except GUM are similar, which indicates that the uncertainty from experiments excluding the fitting distribution is still large.

Figure 6-c separates the uncertainties of the runoff inputs from the overall uncertainties. It is notable that the mean values significantly vary from different runoff inputs (solid lines). For the 1-in-100 year flood, the mean water depth ranges from 8.57 m in e2o_univk to 10.58 m in e2o_cnrs (2.01 m in difference). As for each of the runoffs, the uncertainty caused by other sources (variables and fitting distributions; the shaded area in Figure 6-c) is now very small, especially within the normal period covered by the modelled simulations (35 years in this study). Meanwhile, the uncertainty range starts to increase for the extreme floods. The uncertainty range increases to 0.3-0.5 m for 1-in-100 year flood (on average 5% of the total uncertainty) and 0.4-0.6 m for 1-in-200 year flood, although the uncertainty range is still much smaller than the deviations of the mean values. The increasing uncertainty is similar at the other end of the tails. Similar results are found for other specific points in other river basins and further details can be found in the supporting information.

3.3.2 Inundation area

The uncertainties are also reflected in the inundation area which can be used for assessing the flood exposure of population or economic losses. Figure 7 displays the results for the lower Mekong River Basin at all return periods. The mean values (solid line) and also the uncertainty (standard deviation, colored shades) are displayed in different groups. The mean inundation area increases from 52098 km^2 for a normal flood (1-in-2 year return period) to 59330 km^2 corresponding to a 1-in-100 year flood (Figure 7-a).

Similar to the features of floodplain water depth at Phnom Penh (Figure 6), the magnitude of uncertainty range in the inundation area is similar for all the return periods (Figure 7-a). The uncertainty range for the two ends of tails is a little bit larger. The uncertainties also mainly resulted from the deviation of means values in different runoff inputs (Figure 7-c), rather than the variables (Figure 7-b) or the fitting functions (Figure 7-d). The predicted inundation area for a 1-in-100 year flood ranges from 54000 km^2 to 64000 km^2 in different experiments, indicating a 20% difference to the largest extent. The standard deviation of the inundation area for a 1-in-100 year flood is around 2000 km^2, which increases to 3000 km^2 for a 1-in-200 year flood.

In general, the uncertainties for the inundation area share similar patterns to the results for specific points. These results demonstrate that runoff input is the primary source of uncertainty in river water depth simulation. This uncertainty is mainly due to the systemic bias in the runoff inputs. While for a specific runoff input, the uncertainty is small, especially during the normal period when the estimated values are available (35 years simulation in our case). That extrapolation is applied to FFA in the tails, where the uncertainty range is increased, mainly due to the different tail shape of various fitting distributions. However. the uncertainty range is still smaller than the deviation between results driven by different runoff inputs. Therefore, for impact assessment over the extreme events, the runoff inputs or the average state of the extremes should be evaluated first with observed information, if allowed. Attention can then be given to the selection of different fitting distributions if observations of large floods can be used to optimize the fitting performance, especially in the tails.

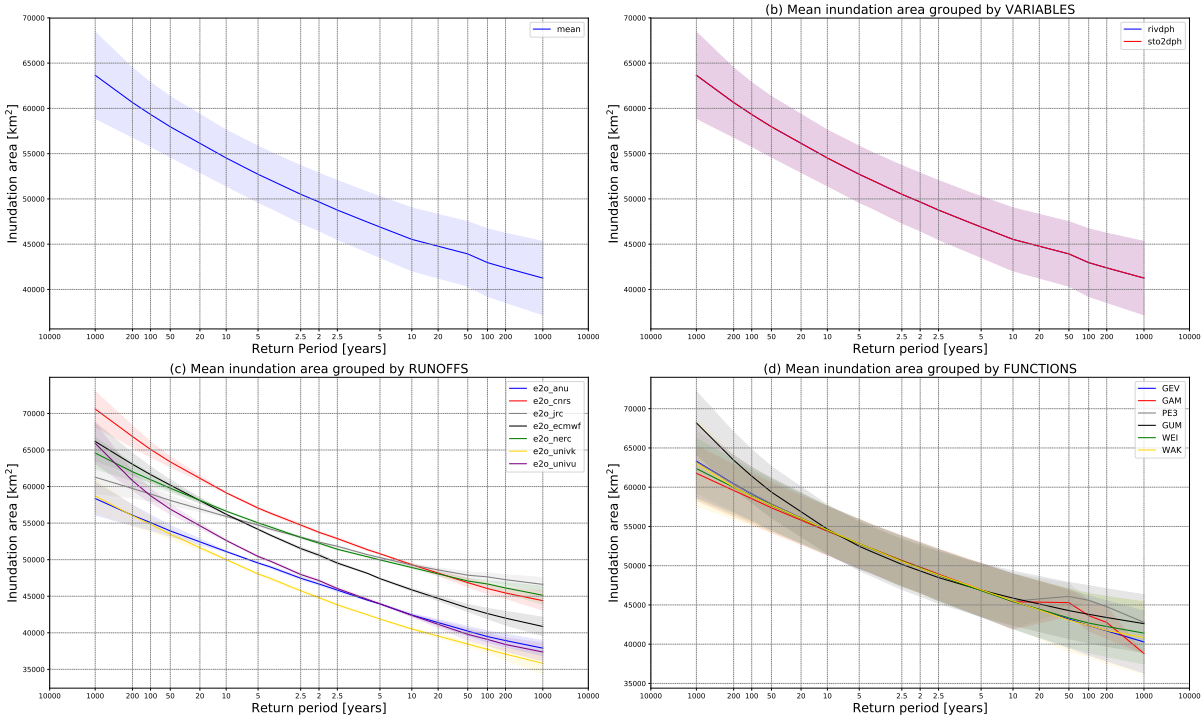

Figure 7. The uncertainties in the estimated inundation area for the study area: a) shows the mean inundation area and the overall uncertainty, (b)-(d) show the mean and uncertainty in different groups by variables, runoff inputs and the fitting functions, respectively.

3.3.3 Population and economic exposure to floods

Previous results show that the inundation area varies in floods with different return periods in the Lower Mekong River Basin. This inundation will lead to migration and economic losses, although the impact should be with uncertainty because of the uncertainties in inundation estimations. In this section, we evaluate the exposure of population and economy to the floods at a global scale. The results are summarized for each continent (see Figure S11 for location map). The global population density (period per km^2 and the economic development (GDP, USD per km^2) can be found in Figure S12 and S13. Given that runoff is the major source of these uncertainties, we did not show uncertainty ranges due to sources other than the runoff in Figure 8, although the uncertainty ranges can be found in Figure S14.

In total, the inundation area for floods at 1-in-100 year return period reaches $13\times10^6\mathrm{km}^2$, accounting for 9.1% of the global area (excluding Antarctica). The ratio for different runoff inputs ranges in 8.1–10.3%. Regarding the population exposure, the gross number is 1.17 billion, accounting for 13.4% [12.1%–15%] of the total population. The potential impact on the GDP will reach up to 14.9 trillion US dollars (USD) in average with the proportion of 13.1%[11.8%–14.7%] of the total values.

Among all the continents, Asia will suffer the largest flood extent and also the largest population exposure (above 0.6 million) and economic exposure (above 6 trillion USD) to the floods. Although these values are not the final flood damages, the potential

impact of the flood in Asia (AS) is the highest. The area with high population density and economic development is highly consistent with the flood-prone area (e.g., the Yangtze, Mekong, Ganges and Indus). Compared to AS, North America (NA) will also suffer large flood extent, while the population and economic exposure is relatively small because the area with high population density or economic development (i.e., the eastern coastal area of the US) is not consistent with the flood-prone area (central plain or Mississippi area). The other continents will suffer smaller inundation area and lower total exposure of the population and economy to the floods. However, it is better to compare the relative values (compared to the specific continent) rather than absolute values because of the area difference.

Regarding the relative inundation and flood exposure (the second row in Figure 8), the inundation area accounts for 12.5–15% of the continental area in NA while the flood exposure of population and economy is around 7–9%. In comparison, the inundation area accounts for 3–8% of the continental area in Africa (AF), the population exposure ratio is 7–13% and economic exposure is 10–18%, indicating a high vulnerability in AF to the floods. The ratio of population exposure in AS (12–19%) is higher than that in AF due to the high consistency of population distribution in Southern Asia and the flood-prone areas. The economy in AS (12–20%) is less fragile than that in AF given a relatively larger flood inundation (7.5–10%). The inundation ratios and flood exposure ratios in other continents are similar, which suggests an even distribution of population and economy in the flood-prone and other regions.

The deviations of curves in the same color reflect the uncertainties. It is notable that the uncertainties in AF for the economy is the largest. For instance, the highest economic exposure to 1-in-100 year floods approaches 19% for a certain runoff, while it is 13% for the lowest with a up to 6% difference. The economic exposure for a 1-in-2 year flood for the former runoff (>15%) is already higher than that for the latter 1-in-100 year flood. This deviation is primarily caused by the various processes in the land surface models or hydrological models. However, the parameterization in AF is not well solved among different models compared to other continents, which is probably due to the complexity of the topography and climate zones in AF. This high degree of uncertainty makes it difficult to accurately assess the economic impact of the floods in the current situation and also for the future projections.

4 Discussions and conclusion

4.1 Discussions

This study assessed the FHM based on simulations with a global hydrodynamic model (CaMa-Flood). The analysis of flood hazards can be uncertain because of the multiple choices of runoff inputs, fitting distributions for FFA and the variables for FFA. Our results show that variation in runoff derived from different land surface models and hydrological model is the primary factor behind the uncertainties in flood water depth and the inundation area, as well as the flood impact on population and economy.

The FHM and FFA only uses the annual maximum water level (or water storage); therefore, the variety only demonstrates the performance of rainfall-runoff models in reproducing the peak discharge. Separation of surface runoff and subsurface runoff, and the evaporation rate during the extreme raining events can lead to the deviations in total runoff and the hydrography after

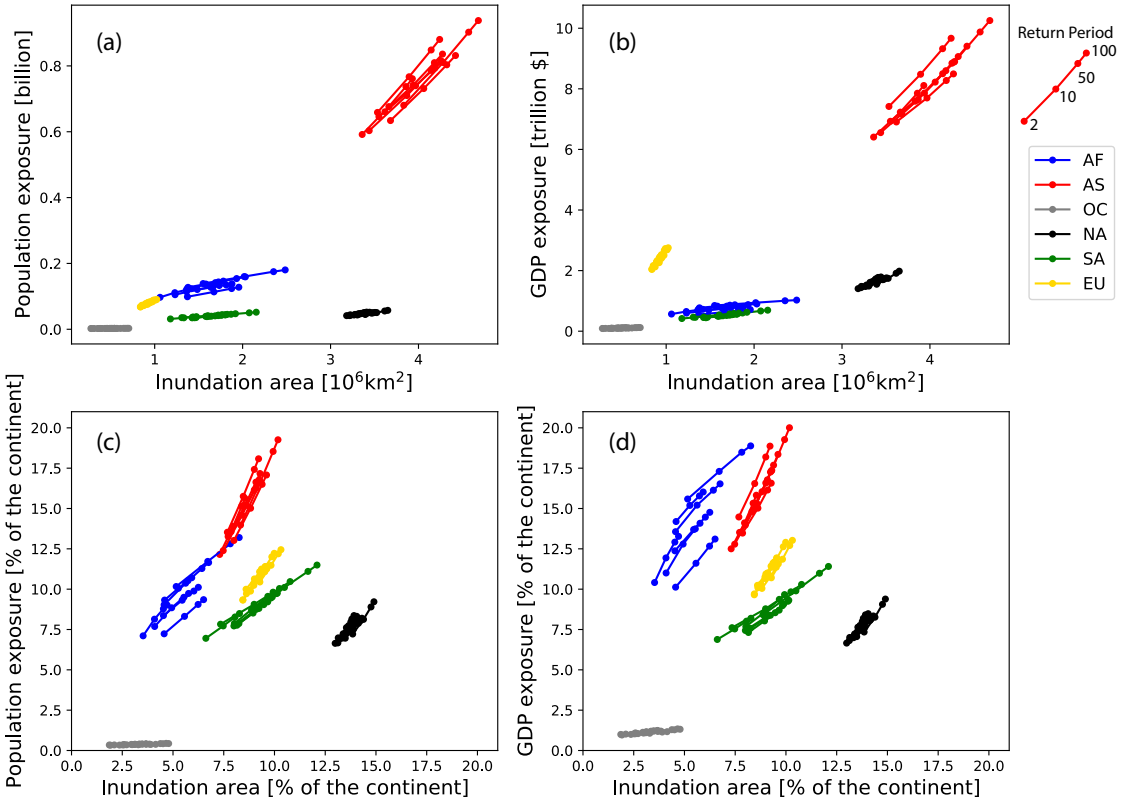

Figure 8. Population and economic exposure to floods in different return periods and the uncertainties due to runoff. The first row shows the absolute values, while the second shows the proportion to total values in the specific continent. The left-hand column shows the population, while the right-hand column shows the Gross Domestic Product. Different colors represent different continents and the values for four different return periods (i.e., 1-in-2, 1-in-10, 1-in-50 and 1-in-100 year) are marked. Different curves in the same color represent the results for various runoff inputs. The names of different runoff inputs are not specified.

routing. In this study, the runoff and the river discharge estimated by CaMa-Flood are not yet calibrated against observations, although calibration will ruin the designed sensitivity test with different runoffs.

There is a lack of studies that have assessed the sensitivities of runoff selection to the flood inundation at a global scale. Although Hirabayashi et al. (2013) assessed the global inundation and population exposure with multiple runoff inputs, their results are simulations for each year (not for a low-frequency flood) with much wider uncertainty range. For regions, the estimated inundation area ranges from 3.5%–9% for the 1-in-100 year return period in Africa (Trigg et al., 2016). While it is 4.5% for experiment "GloFRIS", which is the same as our experiment R7 but with different routing DynRout, approximating our results in Africa 4.4%[3.5%–5.2%]. This suggests that the deviation due to routing models (i.e., DynRout and CaMa-Flood) is limited, while deviation due to forcing can lead to comparable deviations from the hydrological models. However, in Zhao et al. (2017), different routing model leads to 34%–85% in bias of annual peak discharge for global GRDC gauges. Meanwhile, when using a single routing model (i.e., CaMa-Flood), the bias decreases to 39%–50%. This shows that the selection of routing model may also lead to deviation of the inundation area, although the magnitude deserves further studies.

Although the uncertainty sources have been discussed in this study, there are some other factors that are associated with inundation estimation. (Tate et al., 2015) revealed the choice of digital elevation data to be an important factor as for the whole Amazon River basin, vegetation removal from DEM results in an increase in flooded area of 27.5%, and 9.3% for the Congo basin, while the impact is not effective in other basins. There are also regional sensitivity tests of the inundation to roughness (Pappenberger et al., 2008; Jung and Merwade, 2012) and spatial resolution (Yu and Lane, 2006; Horritt and Bates, 2001), although their results needs validation over a global scale. Thomas Steven Savage et al. (2016) systematically investigated the sensitivities of various factors (e.g., inflow hydrograph, channel friction parameter, floodplain friction parameter, spatial resolution) in different phases of flood events and concluded that the channel friction parameter is the most influential factor during peak flood inundation. However, because the ranges of uncertainty are determined within a standard deviation of the mean values, they cannot reflect real sensitivity if using entirely different forcing. With improvement of the computation capability, the sensitivity analysis over the global scale becomes much easier.

One limitation of our study is that we lack validation because the FHMs are not measurable. However, from comparison of long-term water frequency with Landsat and GIEMS data, we noticed that there are some limitations in the current CaMa-Flood that will lead to different results in the uncertainty evaluation. CaMa-Flood does not include flood defense projects (e.g., levees, dams), which will lead to overestimation of the flood inundation in the floodplains and the uncertainty, but lead to underestimation of flood water depth and uncertainties in the river channel. Meanwhile, representing the flood defenses remains a big challenge because the global data for flood defenses are strongly limited (Sampson et al., 2015). Attempts to improve CaMa-Flood by integrating the dam regulation (Shin et al., 2020) and levees (Tanaka and Yamazaki, 2019) have been tested at a regional scale.

4.2 Conclusions

This study assessed the uncertainties in the FHMs from uncertainty sources, including the variables for FFA, fitting distributions and the runoff inputs which drive the routing model for estimating the water depth. Among all o the uncertainty sources,

deviations in runoff inputs contribute the most to the total uncertainty; mainly due to the deviated mean values of extreme water depth. This suggests the importance of rainfall-runoff model calibration (or runoff bias correction) if gauge discharge observation is available. The FHM for the global and specific river basins show the distribution of the mean flood water depth and the uncertainties. Larger deviation values are found in wet regions and along the river channels, while a larger deviation ratio (uncertainty in percentage) is found in dry zones and mountainous regions. Analysis of the flood water depth at specific points and inundation areas for regions displays the uncertainty changes in different return periods. Higher uncertainty is found for a rarer flood compared to normal floods, which is mainly caused by the deviation in the tail shapes of various fitting distributions. Uncertainties in inundation area leads to uncertainties in population and economic exposure to the floods. Globally, 9.1% of the inundation area for 1-in-100 year floods with 2.2% uncertainty leads to 13.4% population exposure (2.9% in uncertainty) and 13.1% economic exposure (2.9% in uncertainty). The uncertainty is the largest in Africa, among all continents, which suggests a large deviation in the structures or parameters of hydrological models that are applied in Africa. Overall, model calibration/validation with advanced tools (assimilation of remote sensing products) and also model improvement by taking into account human interventions are needed to reduce the various uncertainties.

Data availability. The latest global hydrodynamic model CaMa-Flood (v4) is available from https://github.com/global-hydrodynamics/CaMa-Flood_v4. The topography data MERIT is available from http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/index.html. The estimated floodplain water depth and related source codes are available from the authors upon request. The library *lmoments3* for L-moments parameters estimation is available from https://github.com/OpenHydrology/lmoments3.

Author contributions. XZ and DY conceived the study. XZ, WM, WE and DY contributed to the development and design of the methodology. XZ analysed and prepared the paper with review and analysis contributions from WM, WE and DY.

Competing interests. The authors declare no competing interests.

Acknowledgements. This study was supported by "KAKENHI" 20H02251 & 20K22428 by JSPS and also by "LaRC-Flood project" by MS&AD Holdings. The computation is run at the server in Yamazaki Lab in IIS, The University of Tokyo.

References

Aerts, J. P. M., Uhlemann-Elmer, S., Eilander, D., and Ward, P. J.: Comparison of estimates of global flood models for flood hazard and exposed gross domestic product: a China case study, Natural Hazards and Earth System Sciences, 20, 3245–3260, https://doi.org/10.5194/nhess-20-3245-2020, 2020.

Akaike, H.: A new look at the statistical model identification, IEEE transactions on automatic control, 19, 716–723, 1974.

Alvisi, S. and Franchini, M.: A grey-based method for evaluating the effects of rating curve uncertainty on frequency analysis of annual maxima, Journal of Hydroinformatics, 15, 194–210, https://doi.org/10.2166/hydro.2012.127, 2013.

Bales, J. D. and Wagner, C. R.: Sources of uncertainty in flood inundation maps, Journal of Flood Risk Management, 2, 139–147, https://doi.org/10.1111/j.1753-318X.2009.01029.x, 2009.

Bernhofen, M. V., Whyman, C., Trigg, M. A., Sleigh, P. A., Smith, A. M., Sampson, C. C., Yamazaki, D., Ward, P. J., Rudari, R., Pappenberger, F., Dottori, F., Salamon, P., and Winsemius, H. C.: A first collective validation of global fluvial flood models for major floods in Nigeria and Mozambique, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aae014, 2018.

Beven, K., Lamb, R., Leedal, D., and Hunter, N.: Communicating uncertainty in flood inundation mapping: A case study, International Journal of River Basin Management, 13, 285–295, https://doi.org/10.1080/15715124.2014.917318, 2015.

Center for International Earth Science Information Network (CIESIN): Documentation for the gridded population of the world, Version 4 (GPWv4), Revision 11 Data Sets, 2018.

Domeneghetti, A., Castellarin, A., and Brath, A.: Assessing rating-curve uncertainty and its effects on hydraulic model calibration, Hydrology and Earth System Sciences, 16, 1191–1202, https://doi.org/10.5194/hess-16-1191-2012, 2012.

Drissia, T. K., Jothiprakash, V., and Anitha, A. B.: Flood Frequency Analysis Using L Moments: a Comparison between At-Site and Regional Approach, Water Resources Management, 33, 1013–1037, https://doi.org/10.1007/s11269-018-2162-7, 2019.

Hamed, K. and Rao, A. R.: Flood frequency analysis, CRC press, 2019.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., and Kanae, S.: Global flood risk under climate change, Nature Climate Change, 3, 816–821, https://doi.org/10.1038/nclimate1911, http://dx.doi.org/10.1038/nclimate1911, 2013.

Horritt, M. S. and Bates, P. D.: Effects of spatial resolution on a raster based model of flood flow, Journal of Hydrology, 253, 239–249, https://doi.org/10.1016/S0022-1694(01)00490-5, 2001.

Hosking, J. R. M.: of Distributions using Linear Analysis and Estimation of Order Statistics Combinations by equating the Yet moment-based The alternative approach described here is based on quantities, Journal of the Royal Statistical Society, 52, 105–124, https://doi.org/10.2307/2345653, 1990.

Hosking, J. R. M.: L -Moments , Wiley StatsRef: Statistics Reference Online, pp. 1–8, https://doi.org/10.1002/9781118445112.stat00570.pub2, 2015.

Jung, Y. and Merwade, V.: Uncertainty Quantification in Flood Inundation Mapping Using Generalized Likelihood Uncertainty Estimate and Sensitivity Analysis, Journal of Hydrologic Engineering, 17, 507–520, https://doi.org/10.1061/(asce)he.1943-5584.0000476, 2012.

Kidson, R. and Richards, K. S.: Flood frequency analysis: Assumptions and alternatives, Progress in Physical Geography, 29, 392–410, https://doi.org/10.1191/0309133305pp454ra, 2005.

Kummu, M., Taka, M., and Guillaume, J. H.: Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015, Scientific Data, 5, 1–15, https://doi.org/10.1038/sdata.2018.4, 2018.

Liscum, F. and Massey, B. C.: Technique for estimating the magnitude and frequency of floods in the Houston, Texas, metropolitan area. Final report., Tech. rep., U.S. Geological Survey, Water Resources Division, 1980.

Luke, A., Sanders, B. F., Goodrich, K. A., Feldman, D. L., Boudreau, D., Eguiarte, A., Serrano, K., Reyes, A., Schubert, J. E., Aghakouchak, A., Basolo, V., and Matthew, R. A.: Going beyond the flood insurance rate map: Insights from flood hazard map co-production, Natural Hazards and Earth System Sciences, 18, 1097–1120, https://doi.org/10.5194/nhess-18-1097-2018, 2018.

Merwade, V., Olivera, F., Arabi, M., and Edleman, S.: Uncertainty in flood inundation mapping: Current issues and future directions, Journal of Hydrologic Engineering, 13, 608–620, https://doi.org/10.1061/(ASCE)1084-0699(2008)13:7(608), 2008.

Mutua, F. M.: The use of the Akaike Information Criterion in the identification of an optimum flood frequency model, Hydrological Sciences Journal, 39, 235–244, https://doi.org/10.1080/02626669409492740, 1994.

Osti, R., Tanaka, S., and Tokioka, T.: Flood hazard mapping in developing countries: Problems and prospects, Disaster Prevention and Management: An International Journal, 17, 104–113, https://doi.org/10.1108/09653560810855919, 2008.

Pappenberger, F., Beven, K. J., Ratto, M., and Matgen, P.: Multi-method global sensitivity analysis of flood inundation models, Advances in Water Resources, 31, 1–14, https://doi.org/10.1016/j.advwatres.2007.04.009, 2008.

Pappenberger, F., Dutra, E., Wetterhall, F., and Cloke, H. L.: Deriving global flood hazard maps of fluvial floods through a physical model cascade, Hydrology and Earth System Sciences, 16, 4143–4156, https://doi.org/10.5194/hess-16-4143-2012, 2012.

Radevski, I. and Gorin, S.: Floodplain analysis for different return periods of river Vardar in Tikvesh Valley (republic of Macedonia), Carpathian Journal of Earth and Environmental Sciences, 12, 179–187, 2017.

Sakamoto, Y., Ishiguro, M., and Kitagawa, G.: Akaike information criterion statistics, Dordrecht, The Netherlands: D. Reidel, 81, 1986.

Sampson, C. C., Smith, A. M., Bates, P. B., Neal, J. C., Alfieri, L., and Freer, J. E.: A high-resolution global flood hazard model, Water Resources Research, 51, 7358–7381, https://doi.org/10.1002/2015WR016954, 2015.

Schellekens, J., Dutra, E., Martínez-De La Torre, A., Balsamo, G., Van Dijk, A., Sperna Weiland, F., Minvielle, M., Calvet, J. C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., Van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W., and Weedon, G. P.: A global water resources ensemble of hydrological models: The eartH2Observe Tier-1 dataset, Earth System Science Data, 9, 389–413, https://doi.org/10.5194/essd-9-389-2017, 2017.

Shin, S., Pokhrel, Y., Yamazaki, D., Huang, X., Torbick, N., Qi, J., Pattanakiat, S., Ngo-Duc, T., and Duc Tuan, N.: High Resolution Modeling of River-floodplain-reservoir Inundation Dynamics in the Mekong River Basin, Water Resources Research, p. e2019WR026449, https://doi.org/10.1029/2019wr026449, 2020.

Tanaka, Y. and Yamazaki, D.: The automatic extraction of physical flood protection parameters for global river models (in Japanese), Journal of Japan Society of Civil Engineers, 75, 1099–1104, 2019.

Tate, E., Muñoz, C., and Suchan, J.: Uncertainty and Sensitivity Analysis of the HAZUS-MH Flood Model, Natural Hazards Review, 16, 04014 030, https://doi.org/10.1061/(asce)nh.1527-6996.0000167, 2015.

Thomas Steven Savage, J., Pianosi, F., Bates, P., Freer, J., and Wagener, T.: Quantifying the importance of spatial resolution and other factors through global sensitivity analysis of a flood inundation model, Water Resources Research, 52, 9146–9163, https://doi.org/10.1002/2015WR018198, 2016.

Trigg, M. A., Birch, C. E., Neal, J. C., Bates, P. D., Smith, A., Sampson, C. C., Yamazaki, D., Hirabayashi, Y., Pappenberger, F., Dutra, E., Ward, P. J., Winsemius, H. C., Salamon, P., Dottori, F., Rudari, R., Kappes, M. S., Simpson, A. L., Hadzilacos, G., and Fewtrell, T. J.: The credibility challenge for global fluvial flood risk analysis, Environmental Research Letters, 11, https://doi.org/10.1088/1748-9326/11/9/094014, 2016.

Ward, P. J., Jongman, B., Weiland, F. S., Bouwman, A., Van Beek, R., Bierkens, M. F., Ligtvoet, W., and Winsemius, H. C.: Assessing flood risk at the global scale: Model setup, results, and sensitivity, Environmental Research Letters, 8, https://doi.org/10.1088/1748-9326/8/4/044019, 2013.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing data methodology applied to ERA-Interim reanalysis data, Water Resources Research, 50, 7505–7514, https://doi.org/10.1002/2014WR015638, 2014.

Wiltshire, S. E.: Identification of homogeneous regions for flood frequency analysis, Journal of Hydrology, 84, 287–302, https://doi.org/10.1016/0022-1694(86)90128-9, 1986.

Winsemius, H. C., Van Beek, L. P., Jongman, B., Ward, P. J., and Bouwman, A.: A framework for global river flood risk assessments, Hydrology and Earth System Sciences, 17, 1871–1892, https://doi.org/10.5194/hess-17-1871-2013, 2013.

Yamazaki, D., Kanae, S., Kim, H., and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, Water Resources Research, 47, 1–21, https://doi.org/10.1029/2010WR009726, 2011.

Yamazaki, D., Lee, H., Alsdorf, D. E., Dutra, E., Kim, H., Kanae, S., and Oki, T.: Analysis of the water level dynamics simulated by a global river model: A case study in the Amazon River, Water Resources Research, 48, 1–15, https://doi.org/10.1029/2012WR011869, 2012.

Yamazaki, D., Sato, T., Kanae, S., Hirabayashi, Y., and Bates, P. D.: Regional flood dynamics in a bifurcating mega delta simulated in a global river model, Geophysical Research Letters, 41, 3127–3135, https://doi.org/10.1002/2014GL059744, 2014.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., and Bates, P. D.: A high-accuracy map of global terrain elevations, Geophysical Research Letters, 44, 5844–5853, https://doi.org/10.1002/2017GL072874, 2017.

Yu, D. and Lane, S. N.: Urban fluvial flood modelling using a two-dimensional diffusion-wave treatment, part 1: Mesh resolution effects, Hydrological Processes, 20, 1541–1565, https://doi.org/10.1002/hyp.5935, 2006.

Zhao, F., Veldkamp, T. I., Frieler, K., Schewe, J., Ostberg, S., Willner, S., Schauberger, B., Gosling, S. N., Schmied, H. M., Portmann, F. T., Leng, G., Huang, M., Liu, X., Tang, Q., Hanasaki, N., Biemans, H., Gerten, D., Satoh, Y., Pokhrel, Y., Stacke, T., Ciais, P., Chang, J., Ducharne, A., Guimberteau, M., Wada, Y., Kim, H., and Yamazaki, D.: The critical role of the routing scheme in simulating peak river discharge in global hydrological models, Environmental Research Letters, 12, https://doi.org/10.1088/1748-9326/aa7250, 2017.

Zhou, X., Prigent, C., and Yamazaki, D.: Reasonable agreements and mismatches between land-surface-water-area estimates based on a global river model and Landsat data, Earth and Space Science Open Archive, p. 31, https://doi.org/10.1002/essoar.10504917.1, https://doi.org/10.1002/essoar.10504917.1, 2020.