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 ~~at a large scale~~ is essential for regional flood protection and resilience establishment. ~~Due to limitations on the~~ However, due to the limited spatiotemporal coverage of observations, the risk assessment relies on model simulations ~~thus is~~ and is thus subject to uncertainties from ~~various physical processesin the chain of the flood frequency analysis (FFA)~~cascade physical processes. This study ~~assessed the FFA performance as well as~~ assesses the flood hazard map and also the uncertainties with different combinations of ~~FFA variables (river water depth and water storage), fitting distributions and runoff inputs based on the flood characteristics estimated by a~~ runoff inputs, variables for flood frequency analysis, and fitting distributions based on estimations by the CaMa-Flood global hydrodynamic model~~CaMa-Flood. Results show that fitting performance is better if FFA is conducted on river water depth and if Wakeby function is selected as the fitting distribution. Deviations~~ . Our results show that deviation in the runoff inputs ~~are the main source of the~~ is the most influential source of uncertainties in the estimated flooded water depth ~~based on point analysis. This deviation is relevant to the model ability to reproduce the mean state of annual maximum flood extent and it is almost homogeneous for different flood return period . The uncertainty resulted from fitting distributionsincreases~~ 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. ~~The regional investigation of high-resolution inundation area over the lower Mekong River basin shows similar statistics as the point analysis, implying a large uncertainty with 20% deviation in the total inundation area between different runoff inputs. Regional validation of the CaMa-Flood with two other floodhazard maps proves the reliability of the inundation~~

in space and values. Global analysis on the floodplain water depth implies an increasing contribution of uncertainties in fitting distribution to the total uncertainties for rarer floods in almost all land grids. While the changes in contribution of uncertainties in runoff inputs differentiates in regions. The much higher contribution of runoff uncertainty for rarer floods in wet/flat regions necessitates special attention on rainfall-runoff model calibration (or runoff bias correction) if gauge discharge observations are available. Different adaptions to the large floods are needed for regions with different flood water depth and with different inundation agreements among simulations.

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 (**??**). 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.

Flood frequency analysis (FFA) is one way of finding the occurrence and identification of large floods based on limited length of dataset (Hamed and Rao, 2019).It is of vital importance for the analysis of Winsemius et al. (2013) established a framework for river flood control and design of many mitigation projects. The result of FFA can be also used in flood risk assessment which is helpful for stakeholders and insurance services. FFA was introduced more than 30 years ago (Liscum and Massey, 1980; Wiltshire, 1986) , and has been applied to multiple regions in different continents and climates where gauge observations of river status (e. g. , river discharge, river water depth) are availablewith 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.

~~Because most of the gauge observations have been collected for periods of time significantly less than 100 years in GRDC (Global Runoff Data Centre) , the estimation of the "design discharge" (design stage, or water level at a high flood-frequency) necessitates a degree of~~ 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, 2
5 . 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.

10 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 ~~requires~~ based on a priori assumption about the underlying distribution generating flood events. ~~Though, because~~ However, because the limited length of ~~observational records cannot represent~~ the records hardly represents the complete characteristics~~of floods~~, a range of more-or-less skewed, relatively complex distributions is always ~~together~~ considered to account for the uncertainties. Typical distributions that are used include Pearson ~~III type~~ type III, Log-Pearson, Gauss, Gumbel and Log-normal distribution (Radevski and Gorin, 2017; Drissia et al., 2019). However, no conclusion is found ~~whether any~~ to support which of the fitting distributions is preferable for most of the regions (Drissia et al., 2019). ~~Different distributions are recommended to test~~ Therefore, it is recommended that different distributions should be tested with local records~~before selecting the one with the best performance~~.

~~FFA is generally performed for gauge records, while there are many data-poor regions or ungauged regions that also suffering floods disasters, such as the Indus floods in Pakistan (2010) and recent floods in Sri Lanka (2020 May). The characterization of flood behaviour in data-poor regions has received considerable attention (Blazkova and Beven, 2002; Bernhofen et al., 2018)~~
25 ~~while the flood estimations in these ungauged catchments are challenging without enough records (Salinas et al., 2013; Trigg et al., 2016) . There are two families of solutions emerging to facilitate discharge estimation in data-poor regions (Smith et al., 2015) . The first is to relate flood-frequency behaviour in similar catchments with observational records. The second comprises rainfall-driven model cascades that attempt to estimate the river flow through hydrological processes, which is regarded as "continuous simulation" (Bras et al., 1985; Beven and Hall, 2014).~~

30 ~~Continuous simulation is defined as a methodology being developed for estimating flood frequencies where no flood records exist (Kjeldsen et al., 2014). The results of FFA based on continuous simulations are therefore subject to uncertainties propagated from the rainfall, rainfall-runoff models and the routing process which routes the generated runoff to river flow at river profiles of interest (Trigg et al., 2016; Bernhofen et al., 2018). Together with the uncertainties resulting from the fitting distribution, the flow magnitude at the specific "design discharge" will be uncertain. The uncertainties in the FFA are calculated~~

to be the most important source of uncertainty in flood risk assessments which relate the inundation estimations (Merz and Thieken, 2009) .

Associating the FFA analysis and inundation area provides the way to evaluate the potential flood damages for a given magnitude of flood. This first requires a method to estimate the inundation area and links it to the water level. Qi et al. (2009) connected the Poyang Lake area extracted from Landsat images and *in-situ* water level measurements. The relation was then extrapolated to obtain the inundation area according to the frequency of water level. Though, this is only valid for large open water bodies rather for floodplains where inundation is not frequent. Alternatively, the inundation area can be estimated by statistical models (Sarhadi et al., 2012; Odunuga and Raji, 2014) or physical-based models Merwade et al. (2008) which relate the inundation area with calculated floods. The global hydrodynamic river model CaMa-Flood (Catchment-based Macro-scale Floodplain, Y is able to route estimated runoff from various rainfall-runoff models to provide the estimates of flow characteristics (e.g., discharge, water level, water storage in a river channel or floodplain) at all the model points. The inundation area corresponding to a given level of flood (e.g., 100 years return period) can be achieved by downscaling the FFA results to high-resolution maps with bias-corrected topography data.

FFA can be The FFA can be conducted on any characteristics of river flow, but mainly with is mainly used for river discharge and water stage (or named water level or similarly water depth) because they are normally can be recorded as gauge observations (Radevski and Gorin, 2017). There is no preference of for these two variables and the selection is determined only by the only determined by data accessibility. The results of the FFA based on the discharge will be slightly different from the results with the a water stage because of the loop rating curve relationship between discharge and water stage (Domeneghetti et al., 2012; Alvisi and Franchini, 2013). In additionto the aforementioned uncertainty sources from rainfall, rainfall-runoff model, routing processes, fitting and downscaling, the uncertainties of the inundation area corresponding to a certain level of floods will be complex and remain un-investigated. , 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.

In There are many other uncertainties that can lead to deviations in mapping the floods, including forcing, routing and downscaling. However, we need to limit the factors to avoid adding too much complexity to the analysis. Therefore, in this study, Flood Frequency Analysis is applied to the flow estimation by CaMa-Flood and the resulting uncertainties are assessed at various spatial scales on different flood characteristics. Methodologies are introduced in section 2. In section 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, the we assess the fitting performance of FFA is compared among the for all combination of experiments with different flow variables used for FFA, the fitting distributionsas well as fitting distributions, and the runoff that drives CaMa-Flood. Uncertainties resulted from different sources are quantified in the Mekong deltas, with the spatial characteristics shown and agreement of different settings over the inundation estimation evaluated in section 4. A global overview of the uncertainties in floodplain water depth and 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 contribution from sources is provided in section 5. The discussions and conclusions are followed in the end.

2 Methodologies

1.1 CaMa-Flood

The CaMa-Flood (Catchment-based Macro-scale Floodplain) model is designed to simulate the hydrodynamics in continental-scale rivers. The 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. Water storage of each unit-catchment is the only prognostic variable that 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 referred to the original papers by Yamazaki et al. (2011, 2012, 2014).

The major advantage of the CaMa-Flood simulations is the 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 SAR or inundation images by optical or microwave imagers. The estimation of floods on the population and the flooded area is helpful for 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. economy are investigated on different continents. 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 as ensemble simulations are frequently applied to account for uncertainties but computation time will be multiplied manyfold. 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, the CaMa-Flood was driven by the various runoff inputs to achieve the flow characteristics at each unit-catchment at the global scale.

2.2 Experiments design

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 (1)~~ , as follows: first, the variables used for the FFA~~, (2)~~ ; second, the fitting distributions used for FFA~~and (3)~~ ; 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 ~~–(*rivdph*) represents the~~ represents that FFA is based on the numeric results of "river water depth" provided by CaMa-Flood. In V2~~–(*sto2dph*)~~, 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., ~~100 yrs, 50 yrs~~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 ~~resulted~~ that results from the selection of ~~target variables used for FFA~~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.

~~The uncertainty~~

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 ~~fitted automatically~~ 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 ~~which is associated with the~~ associated with different runoff inputs. In ~~this~~ our case, the CaMa-Flood ~~were~~ was driven by seven different kinds of runoff forcing (i.e., ~~R1 – R7~~R1–R7) from eartH2Observe (e2o) category ~~(?). The runoff~~ (Schellekens et al., 2017). The runoffs were driven by the same ~~WFDEI~~

~~(WATCH Forcing Data methodology applied to ERA-Interim data, Weedon et al., 2014)~~ WATCH Forcing Data methodology applied to ERA-Interim data (WFDEI, Weedon et al., 2014) but with different land surface ~~and hydrological models, therefore,~~ or hydrological models. Therefore, the runoff inputs ~~have already contained~~ contain the uncertainties in the ~~forcing and that in the~~ rainfall-runoff model processes (model structures and model parameters). ~~The~~ Therefore, the deviation of the results in the FFA among the seven inputs was ~~, therefore,~~ the uncertainty caused by the runoff inputs.

~~Various experiments used in this study for uncertainty analysis. There are three groups as (A) the variables for FFA (B) the fitting distributions and (C) the runoff inputs. Different runoff are generated by same forcing (WFDEI) but different land surface models or global hydrological models (specified in the bracket). A Variables B fitting distribution C Runoff V1 *rivdph* F1 GEV (Generalized Extreme Value) R1 e2o_anu (W3) V2 *sto2dph* F2 GAM (Gamma) R2 e2o_cnrs (ORCHIDEE) F3 PE3 (Pearson 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)~~

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 next section) 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)~~with a spatial resolution of 0.25° (~25km at the equator)~~. 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).

Then the The parameters of the fitting distributions were calculated with the basis then calculated based on these sorted annual values with the L-moment method (Hosking, 2015; Drissia et al., 2019). It This is defined as a linear combination of probability-weighted moments of the time series. The parameters estimation 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 only the Wakeby (WAK) is a 5-parameters functionwhile the others are all 3-parameters functionsfive-parameter function; the GEV, PE3 and WEI are three-parameter functions; while GAM and GUM are two-parameter functions.

2.4 Criterion

Akaike Information Criterion (AIC, Sakamoto et al., 1986; Mutua, 1994) was used to evaluate the performance of the FFA against the annual values. aic is calculation as 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. The performance of fitting is better when the Smaller aic value is lower 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 Study area and downscaling 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 evaluation performance of the FFA performance was evaluated with aic was conducted at the global scale to capture the overall features , corresponding to the results in section 3. (see Section 3.1).

It 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 floodplain flood water depth and inundation area with their uncertainties was conducted regionally over the lower Mekong River basinBasin, where the delta is vulnerable

to floods (Shin et al., 2020). Corresponding results ~~from point analysis to regional analysis~~ on the uncertainties in water depth and inundation area, as well as the flood water depth at specific points will be presented~~in section 4.~~ .

The estimated low-resolution storage was downscaled to the high-resolution inundation map with the topography map ~~MERIT (Multi-Error-Removed Improved-Terrain DEM, Yamazaki et al., 2017)~~ – Multi-Error-Removed Improved-Terrain DEM (MERIT, Yamazaki et al., 2017) at 90 m. The fundamental assumption is that the ~~water~~ movement of water within a unit catchment is instantaneous and that the water surface is flat within ~~each unit-catchment and the~~ 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 ~~should be similar to~~ is illustrated in Figure 1-b~~, as when~~ . When the floodplain has been inundated, the small increases of water level ~~is corresponding~~ corresponds to large changes in the water storage ~~as well as~~ 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 ~~the~~ an error in function fitting. The assumption of ~~the~~ a flat water surface is not valid for ~~long river sections or~~ large water bodies (e.g., large lakes or reservoirs with water surface gradient) and ~~rivers with large slopes~~ 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 (~~in consistency~~ which is consistent with the global scale). ~~Inundation~~ 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 (**?**), as well as the Gross Domestic Production map in year 2015 (**?**). The two maps are in 30 arcsec, therefore the 3 arcsec inundation map was aggregated to the 30 arcsec.

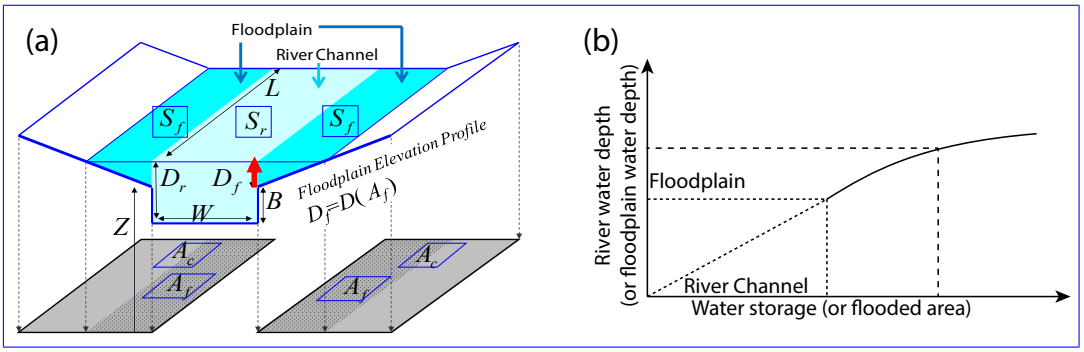
3 ~~Comparisons among different experiments~~

~~In this section, the performance of the FFA will be inter-compared among different experiments given in Table 1. The lower *aic* indicates the better fitting performance. Comparisons are conducted in three groups. In the first *variables* group, the two variables on which the FFA was based are compared. In the second group, the fitting performance is compared among using different fitting distribution. In the last group, the FFA performance determined by different runoff inputs is compared.~~

3 Results

3.1 ~~Comparison between using different *Variables*~~

~~Water storage is the prognostic variable in the CaMa-Flood that transfers water from the upstream to the downstream. It is estimated by the inflow to this catchment, added runoff within the catchment and outflow to the downstream catchment. Whereas, river water depth is co-determined by the river water storage, river channel cross-section and floodplain topography~~



~~(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.~~

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.

~~profile. Estimating the river water depth hence includes extra information as well as the uncertainty from the topography. This section evaluates how the fitting distributions work for the two different variables. Because~~

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 ~~not in the same unit or same magnitude, they were normalized to the range of~~ 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 ~~for each unit-catchment. The fitting distributions (i.e., F1 – F6, Table 1)were applied to fit the modelled time series~~of each grid). The fitting performance was evaluated by the aic value ~~. The estimated~~ (Eq. 2). A lower aic ~~values for the two variables were compared and one sample~~ indicates a better fitting performance. Figures 2-a and b display a sample result for e2o_ecmwf ~~and GEV function is shown in~~ (R4) and GEV fitting distribution (F1) for water level and storage, respectively. The difference between the two maps is shown as Figure 2-c.

~~Because the time series were normalized to ranges of 0 and 1, the~~ The fitting performance is relatively high with low aic (<-50) in most of the unit-catchments. ~~Low~~ 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

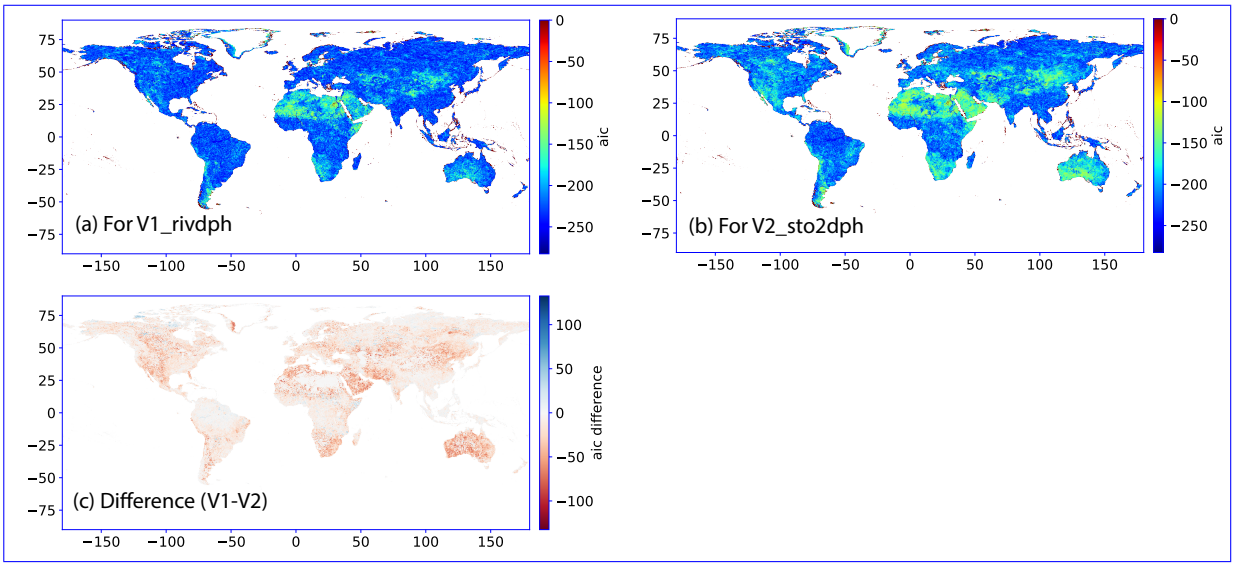


Figure 2. ~~Performance~~ Fitting performance of ~~flood frequency analysis~~ FFA for (a) V1 ~~_*rivdph*~~ (river water depth) and (b) V2 ~~_*sto2dph*~~(water storage). The performance was quantified with aic and (c) is the aic difference of (a) and (b). ~~Negative difference indicates better performance of FFA for V1_*rivdph*.~~ This is only an example for e2o_ecmwf and GEV fitting distribution.

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 ~~west~~ western part of North America. These regions are mainly dominated by dry climate or mountainous topography. The ~~relatively low river discharge could be the reason for low model performance in the fitting~~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. ~~In which, negative~~ 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, ~~in high consistency~~ 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. ~~Though~~ However, this is only the ~~results~~ result of a case with e2o_ecmwf runoff input and GEV distribution.

An overall evaluation ~~on all~~ of all of the distributions and runoff inputs ~~are~~ is shown in Figure 3. The probability distribution of the aic values for all the global grids are plotted in Figure 3-a and 3~~-b for V1_*rivdph* and V2_*sto2dph*~~b using water river depth and storage, respectively. ~~We found that the fitting distribution determines the *aic* values as the pdf curves for the same~~

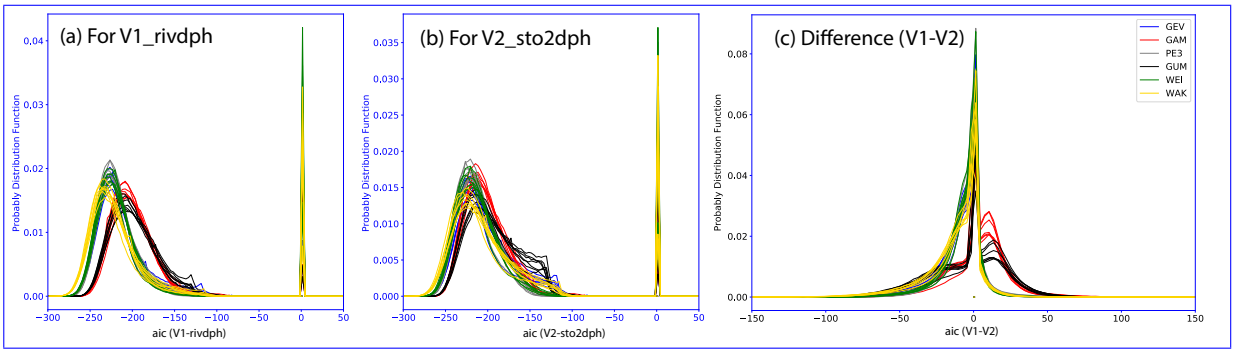


Figure 3. Overall performance of FFA for (a) V1 (river water depth) and (b) V2 (water storage). The performance aic over all the land grids are collected and displayed as the histogram. (c) is the aic difference between (a) and (b). Negative difference indicates better performance of FFA for V1. Different 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.

distribution always gather together. This is more distinguishable for the water depth than that for the water storage. The pdf curves have two peaks, one is normally distributed with mean values around -200 (or -220) and the other one is near zero. The later peak is near 0. The latter peak around 0 corresponds to the red scatters in FigureFigures 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 smallest 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 the water depth. Although the lines 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, the sequence of the fitting performance for different distributions is the same as for the water depth. 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.

Overall performance of flood frequency analysis for (a) V1_*rivdph* and (b) V2_*sto2dph*. The performance aic over all the land grids are collected and displayed as the histogram. (c) is the aic difference of (a) and (b). Negative difference indicates

~~better performance of FFA for V1_*rivdph*. Different colors represent different fitting distributions, and the multiple lines in a specific color represents results driven by different runoff inputs. The type of the runoff inputs are not specified in these three graphies.~~

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). ~~Same as~~ As in Figure 2-c, negative values indicate that the fitting performance for water depth (V1) is better than that for the water storage ~~. For~~ (V2). More negative values were found for the distributions of WAK, GEV, PE3 and WEI, ~~more negative values were found~~ 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. ~~But~~ 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. ~~We, therefore, can conclude that the fitting is better applied to the water depth (V1_*rivdph*) rather than the water storage (V2_*sto2dph*). Since the~~ 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 ~~causes~~ leads to the resulting difference in the fitting performance.

3.2 ~~Comparison between different *Fitting distributions*~~Flood water depth at 1-in-100 year return period

~~In order to find the better fitting distribution for FFA, we ranked the fitting performance by in distributions according to the *aic* values at each unit-catchment. The distribution with the best performance (the smallest *aic* value) was scored 6 and the function with the worst performance (the largest *aic* value)was scored 1. The other distributions were scored from 2 to 5 according to the sequences of *aic* values. The scoring results for runoff e2o_ecmwf with V1_*rivdph* are discussed in this subsection as an example~~

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 **??**) .

~~The type of fitting distribution corresponding to (a) the top ranking and (b) the lowest ranking of the FFA performance according to the *aic* criterion.~~

~~The WAK is scored as the best function in most of the global grids (Figure **??**-a) except the dry areas in~~ 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 ~~Asia~~ Australia, and Central Asia) and the originating river basins in

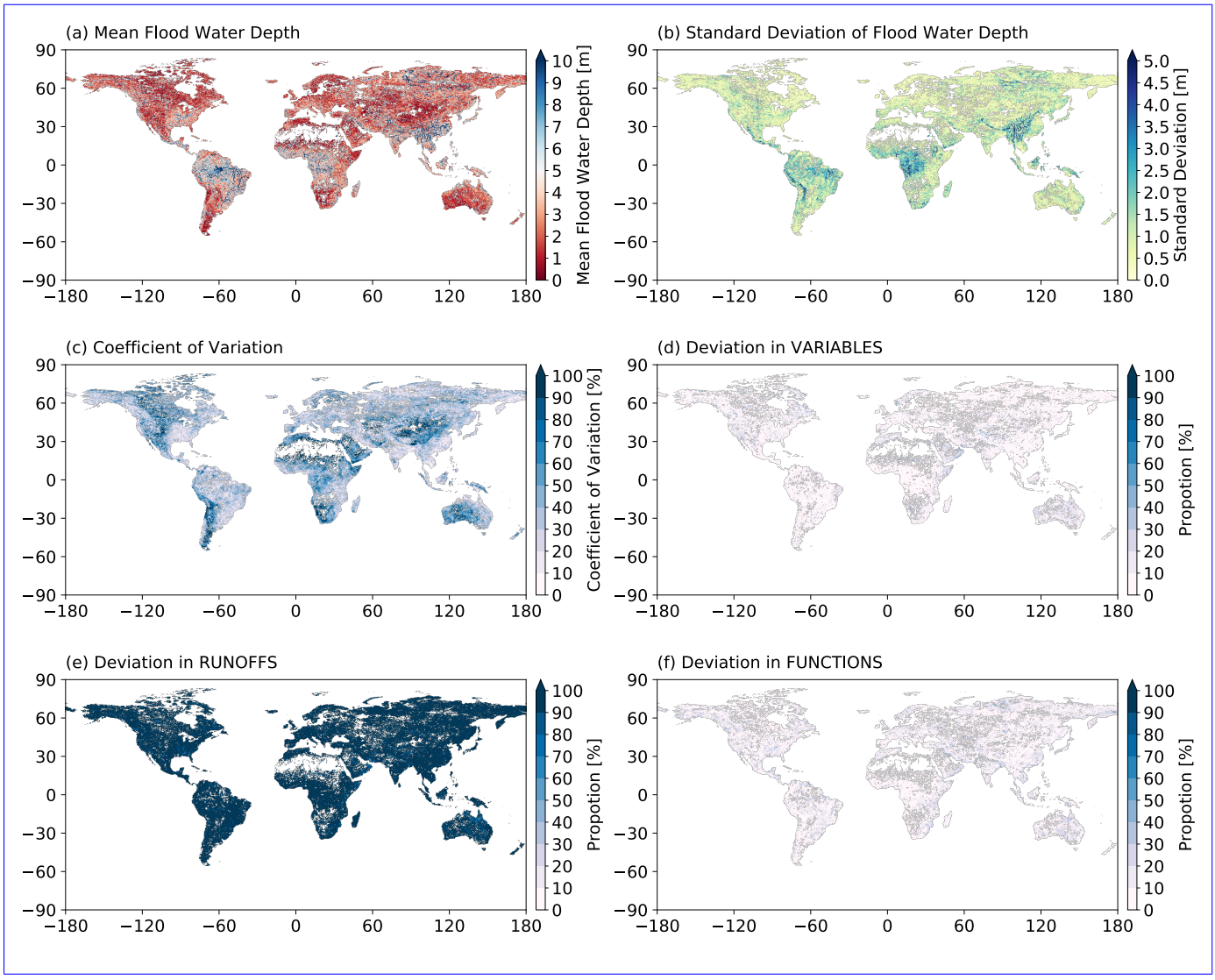
mountainous regions (e.g., the Rocky Mountains, the Andes, and ~~middle southern Australia.This is mainly because WAK is a 5-parameters function and might be overfitted while all other functions only have three parameters.Despite the best distribution, the PE3, GEV and WEI are marked as the second-best functions in different parts of the globe. While the GAM and GUM are generally ranked the last implying that the two functions are not suitable for FFA on the water depth (Figure ??-b). The same results have been shown in Figure 3, as aic for GAM and GUM are in a large probability higher than aic with other fitting distributions.~~ the Tibetan Plateau).

~~The average score for all the global catchments shows the same results~~ 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 **??**). ~~No matter the average water depth ($rivdph$) is, the WAK ranks the first.WAK, PE3 and WEI have similar performance when the river water depth is less than 1.0m, corresponding to the land grids where this is a low probability to suffer heavy floods.For the grids with average river water depth larger than 1.0m, the performance becomes more distinguishable as WAK outperforms other functions but the differences among GEV, PE3 and WEI become small. The GAM and GUM always have the worst performance for all ranges of the water depth.~~ 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 same comparison is not conducted for other runoff inputs or V2_$sto2dph$ because from Figure 3 we can conclude that the performance is mainly determined by the fitting distribution. The differences between different runoff will not change the fitting performance, so the ranking scores.~~

3.3 ~~Comparison between different *Runoff inputs*~~

~~In the aforementioned analysis,~~ 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 ~~distributions mainly associate the shape of the distribution of high values. While the runoff inputs mainly determine the average states and have small impact on the fitting performance.Figure **??** shows the ranking of runoff inputs in terms of the mean values of the annual maximum river water depth in the original CaMa-Flood outputs. The spatial variation of the ranks is more complicated than Figure **??**, thus the coverage of each runoff input is displayed besides the map in the unit of percentage~~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



~~The averaged ranking among the different fitting distributions. The results are grouped by grids with different river water depth.~~

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.

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.

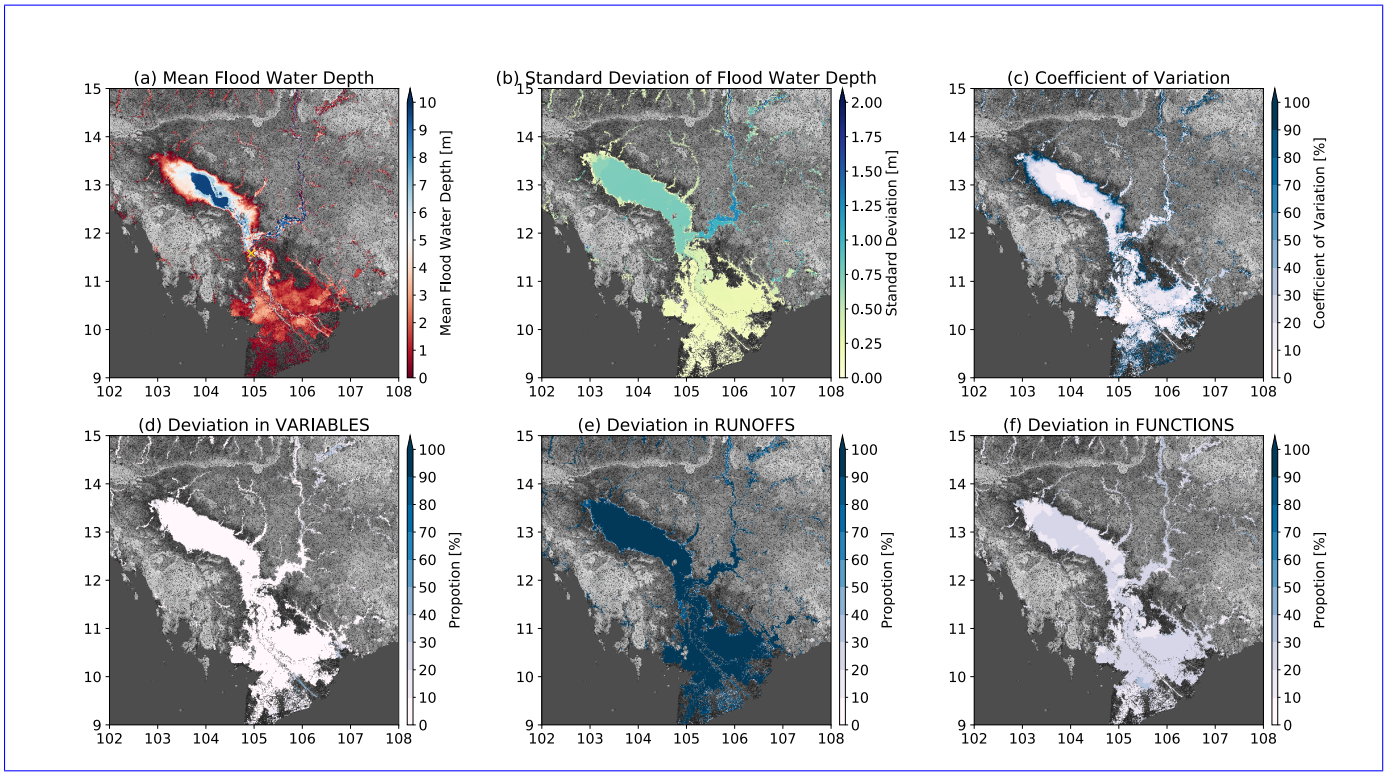
3.2.1 Regional flood water depth

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 ~~, e2o_cnrs (48.9%) provides the largest value for most of~~ 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 ~~land in North America, high latitudes in Europe and Asia, southeast Asia, central Africa and southeastern Australia. Then e2o_univk (24.9%) is the highest in the regions around the Mediterranean Sea and central Australia. e2o_ecmwf (6.2%) provides the highest floodplain water depth in the Amazon River basin. Regarding~~ 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 ~~ranking in the middle~~ (Figure 5-e) and from the fitting distributions (Figure ~~??-c) , the variety increases as no runoff input is~~ 5-f) and the variables (Figure 5-d). This is consistent with the conclusions from the ~~middle one for a large extent. The coverage of different runoff inputs ranges from 6.2% to 19.7% with a smaller variation than that for the lowest or highest estimates. In regional scale, the e2o_anu and e2o_cnrs are in the middle for most of the river channels in the Northern Hemisphere and south of the 10°S. In the low latitudes in the Northern Hemisphere and the tropical regions, the e2o_univu and e2o_ecmwf are probably being in the middle.~~ global analysis.

~~Therefore, for global-scale studies, there is no preference of runoff selection. Ensemble simulation is suggested to account for the different ability of the land surface or hydrological models in different climates or topographic conditions. While for regional studies, observations are recommended to validate the simulations. Ensemble simulations driven by all~~ 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~~are also preferable if the observation is lack. Otherwise, selecting the runoff inputs which provides the estimation of water level in the middle would be~~



~~The types of runoff inputs corresponding to (a) the lowest, (b) the highest and (c) middle estimates of the mean annual maximum river water depth. The coverage of each runoff corresponding to each map is shown in the right in unit of percentage.~~

~~Large variations are found among different runoff inputs in a different ranking. In Figure ??-a, e2o_anu and e2o_univu provide the lowest estimation of the maximum water depth in most regions in the world, except North America, the southeastern Asia and the Green Land (where e2o_ecmwf is the lowest). The regions with the two runoff account for 33.7% and 26.3% of the total continental grids, respectively. In Figure ??~~

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.

~~recommended to reduce the risk of large deviation in one single runoff input~~. The selected variables and fitting functions will not lead to large deviations compared to the runoff inputs.

4 ~~Regional uncertainties analysis~~

3.1 Flood water depth for multiple return periods

~~In this section, the uncertainty range in the water level and inundation due to the selection of investigated variables, fitting distributions and runoff inputs will be discussed. Different from the previous analysis, the results in this sectionare based on the original results on the FFA, rather than the results after normalization. The uncertainty analysis is concentrated on the lower Mekong region, where the delta is vulnerable to floods. Point analysis and analyses on regional maps are combined to better illustrate the uncertainties from various sources.~~

3.2 ~~Point analysis~~

~~A specific point (105.00°E, 11.54°N located one grid after the confluence of the main Mekong River and the outflow from Tonle Sap Lake, see the yellow cross point~~

3.1.1 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)~~was selected to analyze the uncertainties in the floodplain water depth.~~ , 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 ~~as well as~~ and the uncertainty range (~~doubled~~ standard deviation) ~~among~~ for different conditions are ~~shown~~ 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 ~~50% is 6.6~~1-in-2 year return period is 8.14 m and it is ~~7.9~~ 9.58 m for the ~~100-year return period flood(hereafter 100-yr flood)~~1-in-100 year flood. The overall ~~uncertainty range~~ standard deviation is large up to ~~1.3~~ 0.69 m and it is generally the same for different return ~~period frequency (from 0.1% to 99.9%)~~periods.

In Figure 6-b, the differences between mean floodplain water depth using river depth (V1~~_*rivdph*~~) and storage (~~V1_*sto*2*dph*~~V2) is very small. The uncertainty range is still as large as that in Figure 6-a~~, indicating~~ . This indicates that the uncertainty ~~is little contributed by~~ 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-~~c), indicating~~ -d). This indicates that the uncertainty that resulted from the selection of fitting distribution is still small. ~~Particularly~~In particular, the mean value for GUM function in the tails of the floods (more than ~~20-yr~~ 1-in-20 year flood) is higher ~~tha~~V2_n than results of other functions, indicating that GUM may provide a relatively deviated estimate of mean floodplain water depth for the extreme flood events. ~~The~~ This difference of GUM mainly happens because GUM only has two degrees of freedom. The uncertainty

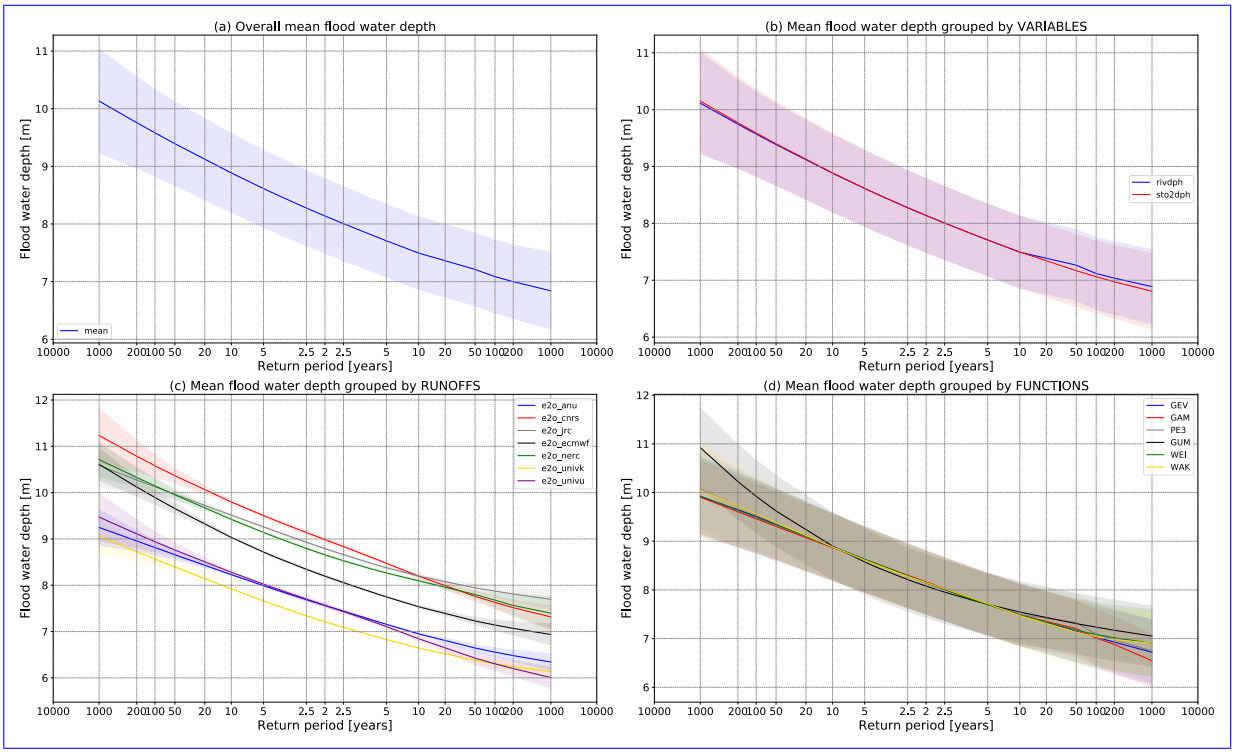


Figure 6. Uncertainties in the estimated floodplain water depth at Phnom Penh (104.9317°E, 11.5617°N) in the Mekong River Basin in different groups. a) the mean floodplain water depth and overall uncertainty; b) the mean and uncertainty in groups of different variables for FFA, the uncertainty is then 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.

ranges of other uncertainties ~~in GUM is still similarto the magnitude of uncertainties for other fitting distributions, indicating~~ except GUM are similar, which indicates that the uncertainty from experiments ~~other than~~ excluding the fitting distribution is still large.

Figure 6~~-d~~-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~~in Figure 6-d~~). For the ~~100-yr~~ 1-in-100 year flood, the mean water depth ranges from ~~6.9~~ 8.57 m in e2o_univk to ~~9.8~~ 10.58 m in e2o_cnrs (~~2.9~~ 2.01 m in difference). As for each of the ~~runoff~~runoffs, the uncertainty caused by other sources (variables and fitting distributions; the shaded area in Figure 6~~-d~~-c) is now very small, especially within the normal period ~~(5-yr flood and 5-yr drought)~~ covered by the modelled simulations (35 years in this study). ~~While~~ Meanwhile, the uncertainty range starts to increase for the extreme floods. The uncertainty range increases to 0.3-0.5 m for ~~100-yr~~ 1-in-100 year flood (on average ~~25~~5% of the total uncertainty) and ~~0.8-1.0 m for 200-yr flood(on average 33.3% of the total uncertainty). Though,~~ 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.

~~The above results~~

3.1.2 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).

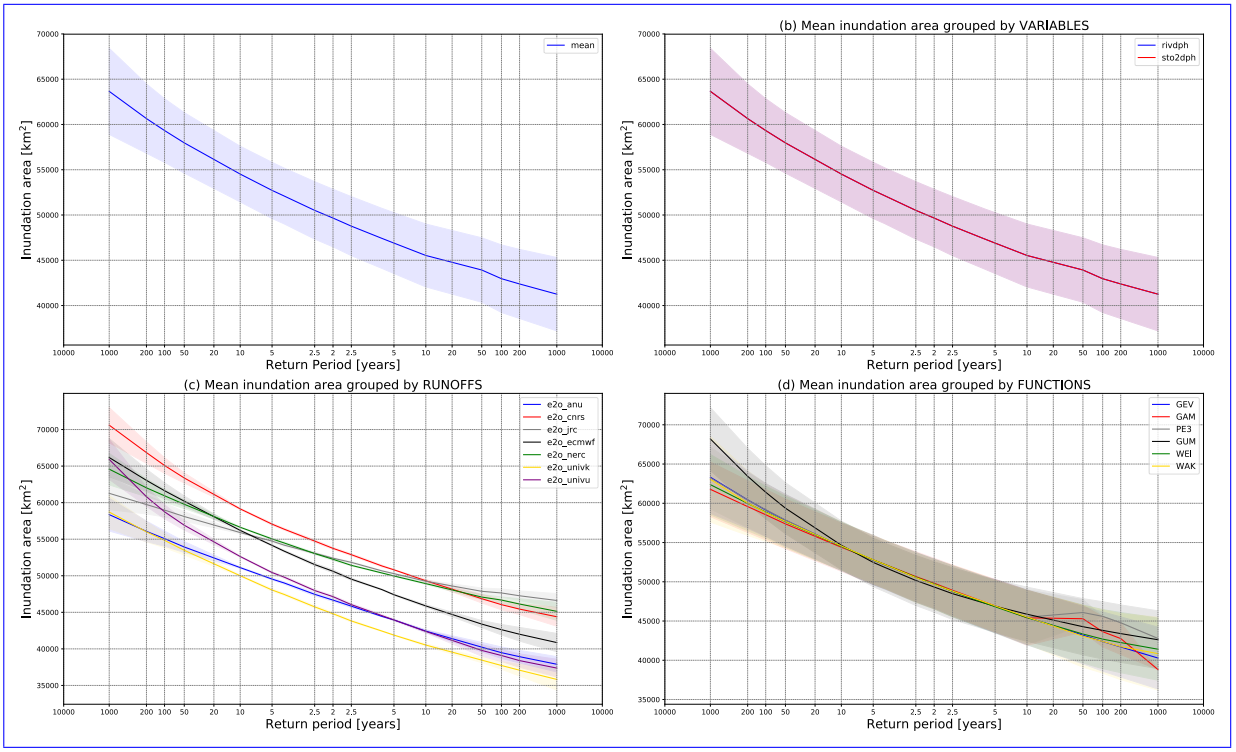


Figure 7. The uncertainties in the estimated inundation area for the study area: a) shows the mean inundation area and the overall uncertainty, (b)-(d) show the mean and uncertainty in different groups by variables, runoff inputs and the fitting functions, respectively.

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 ~~to the~~ in river water depth simulation. ~~The~~ 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). ~~In the tails that~~ That extrapolation is applied to FFA ~~,~~ in the tails, where the uncertainty range ~~gets increasing~~ is increased, mainly due to the different tail shape of various fitting distributions. ~~But~~ 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. ~~Then attention can~~ 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.

~~Uncertainties in the estimated floodplain water depth at a specific point (105.00°E, 11.54°N) 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 fitting distributions, d) the mean and uncertainty in groups of different runoff inputs.~~

3.1.3 Population and economic exposure to floods

3.2 ~~Regional analysis – floodplain water depth~~

~~The floodplain water depth at a 100-yr flood was first downscaled to high-resolution map (90 m) to show~~ 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. ~~details with topography and the mean floodplain water depth for the lower Mekong is shown as Figure 5-a. The largest water depth (>10.0 m) is found in the centre of Tonle Sap Lake and the main channel of the Mekong River~~

In total, the inundation area for floods at 1-in-100 year return period reaches $13 \times 10^6 km^2$, accounting for 9.1% of the global area (excluding Antarctica). The ratio for different runoff inputs ranges in 8.1–10.3%. ~~Large extent in the lower Mekong delta is suffering relatively low inundation water depth (in dark red). The low water depth is also occurring along the boundaries of lakes and main channels. The river tributaries are also with low average water depth among all the experiments. In summary~~

(Table ??), Regarding the population exposure, the inundation area (water depth >0.01 m) of the study area during 100-yr flood is 68809.1 km^2. Among which 22.2% of the area is with high water depth (>5.0 m, 15129.5 km^2). 33.1% of the inundation area is with water depth less than 1.0 m and 8.8% with water depth under 0.1 m.

(a) Ensemble mean water depth among all the experiments for 100-yr flood and (b) the overall uncertaint (standard deviation) for all experiments. (c) and (d) illustrate the average uncertainty for fitting distributions and runoff inputs, respectively. Area with floodplain water depth less than 0.01 m are masked out.

Inundation area with the floodplain water depth corresponding to 100-yr floods (Figure 5-a) with different water depth categories. all 0.01-0.1 0.1-0.3 0.3-0.5 0.5-1.0 1.0-5.0 >5.0 all gross number is 1.17 billion, accounting for 13.4% [km^2 12.1%–15%] 68809.16078.3 5030.6 3707.9 7976.1 30886.7 15129.5 percentage to all 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%] 100.0 8.8 7.3 5.4 11.6 44.922.0 of the total values.

Figure 5-b shows the uncertainties resulted from different experiments listed in Table 1 except for fitting distributions of GAM and GUM because of their poor fitting performance. In general, 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 uncertainty range is higher where the estimated water depth is higher (Figure 5-a) as the lowest uncertainties are found in the lower Mekong delta and largest uncertainties in the main channel of Mekong with magnitude higher than 2.0 m. The uncertainty in the Tonle Sap Lake is homogeneous with a magnitude around 1.0 m.

The overall uncertainties mainly result from the fitting distributions (Figure 5-c)and floods. Although these values are not the runoff inputs (Figure 5-d) . Whilst the uncertainties from runoff inputs contributed the most because the magnitude in Figure 5-d is very similar to the overall uncertainties (Figure 5-b). This is consistent with the conclusion from point analysis in the previous subsection. It further strengthens the point analysis and makes it valid over the entire region. The uncertainties of fitting distributions are small in the lower deltas, but it gets larger when the water depth increases. The largest uncertainty is approaching 1.0 m in the upper Mekong reaches of the study area. Though, the increases in water depth will not lead to significantly increase inundation area.

3.2 Regional analysis – inundation agreement

Despite the uncertainties with estimated mean floodplain water depth, final flood damages, the agreement on the prediction of inundation among different FFA settings might be more important because the inundation will cause damages regardless of the water depth. High inundation agreement will provide the confidence of adapting corresponding actions, for example, evacuation in the most serious condition. Figure ?? illustrates the inundation agreement among all the experiments in the study area. In general, the agreement is high for the lakes, river channels and the lower Mekong deltas (coloured in dark blue, ~100%), indicating that if suffering a 100-yr flood, all these regions will be inundated regardless of different runoff inputs and used fitting distributions. Lower agreement is generally with lower estimated water depth around the boundaries of lakes and other inundation areas. Particularly, the large area of the Krong Prey Veng (white square in Figure ??)is with inundation agreement

around 50%, indicating that the resulting inundation is not highly consistent, which means selection of different runoff input and fitting distributions will lead to big differences (inundation or non-inundation).

Inundation agreement of the estimated inundation (>0.01 m) for the ensemble 100-yr flood. It is calculated as the ratio of the number of experiments which predict the inundation to the total number of experiments. The white square represents the region of Krong Prey Veng.

Figure **??** shows the inundation area in different categories of mean floodplain water depth. Meanwhile, the inundation area is separated according to the level of inundation agreement, with $\geq$50%as high agreement and <50% as low agreement. For the regions with high predicted inundated water depth (>1.0 m), different experiments are highly consistent as all the inundation area is with agreement larger than 50%. In other words, more than 50% of the experiments predict inundation at this location. On the contrary, the agreement is low for inundation area with mean floodplain water depth less than 0.1 m, as 83.8% of the area has a agreement less than 50%. The percentage of area with low inundation agreement decreases from 83.8%for low mean water depth 0.01 – 0.1 mto 5.1% for the area with water depth 0.5 – 1.0 mand to 0% for the area with water depth larger than 1.0 m.

Area (km^2) and proportions (%) of the inundation with different uncertainty in different categories of the ensemble mean floodplain water depth for the ensemble 100-yr flood. For each water depth category, low agreement and high agreement are divided by criterion of 50%. The labels for each bar is the inundation area (km^2, up) and the percentage of the area (%, low) in each category of floodplain water depth.

To conclude, for a potential 100-yr flood, 13.4% of the predicted inundation area (9239.1 km^2) is with low model agreement less than 50% as half of the experiments/FFA settings predict non-inundation for this location. Selection of the appropriate experiments will become more important for flood prediction and risk analysis. For regions with high model agreement (>50%), the adaptions to the predicted floods have to be taken in high confidence. However, the required actions can be different for different flood water depth.

3.2 Regional analysis – inundation area

In addition to the 100-yr flood, the predicted inundation area, as well as the uncertainty for all the return periods, are investigated in this subsection. The mean inundation area averaged over the estimated inundation area in each experiment and potential impact of the uncertainty of the inundation area are plotted as Figure 7. The mean inundation area increases from a normal flood (return period as 50%, 52135.2 km^2) to 62234.8 km^2 corresponding to a 100-yr flood (Figure 7) . However, it is notable that the inundation area for 100-yr flood is 68809.1 km^2 (10.6% higher) in Table **??** if inundation is calculated according to the mean floodplain water depth averaged over different experiments. This difference is mainly caused by the different ways of estimations, as in this subsection, the inundation area is estimated first for each experiment and then they are averaged to reach the mean value. While for the previous estimates in Table **??**, the inundation area is calculated by the averaged floodplain water depth over multiple experiments. In that case, one single experiment with very high floodplain water depth can lead to high mean water depth (>0.01 m) even if the other experiments do not predict inundation in the same location. This is why, in

the Figure ??, the model agreement is very low especially for the area with the ensemble mean floodplain water depth less than 0.1 m.

Similar to the features of floodplain water depth at the selected point (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 are also mainly resulted from the deviation of means values in different runoff inputs (Figure 7-b). The predicted inundation area for a 100-yr flood ranges from 56000 km^2 to 70000 km^2 in different experiments, indicating a 20% difference to the largest extent.

The uncertainties in the estimated inundation area for the study area. a) shows the mean inundation area and the overall uncertainty, b) shows the mean and uncertainty related to runoff inputs.

3.2 Validation with other results

In this subsection, the inundation map is compared to two flood hazard maps from different sources. The first source of the flood hazard map is from GAR (Global Assessment Report on Disaster Risk Reduction, GAR, 2015). This dataset was observation-based for large rivers. Quantiles of the river discharge were estimated based on the collected stream-flow or proxy data from homogeneous regions. The calculated quantiles were then introduced to river sections with topographic data (SRTM) and a simplified approach based on Manning's equation to model the water levels (Rudari et al., 2015). The second source is JRC (Joint Research Centre Data Catalogue) data based on streamflow data from the European and Global Flood Awareness System (EFAS and GloFAS) and computed using two-dimensional hydrodynamic models-CA2D (Alfieri et al., 2014; Dottori et al., 2016) .Though, it was already mentioned in the references that there are limitations in the model and the maps might differ from official flood hazard maps. The two maps are plotted as Figure ??-a and ??-b.

Comparison of inundation map with other flood hazard map for 100-yr flood (a) GAR, (b) JRC, (c) CaMa-Flood. The three maps are all at 30 arcsec spatial resolution. Pakse is one of the GRDC gauges just upstream of the study area of the lower Mekong River.

The spatial resolution of the GAR dataset and JRC dataset is 30 arcsec. It is recommended that the comparisons are conducted on the same spatial resolution. We therefore downscaled the original CaMa-Flood results to 30 arcsec (Figure ??-c) . Large differences of the inundation area in the tributaries can be seen by comparing the three maps in the lower Mekong River, as almost all the tributaries are inundated in GAR dataset, while very few tributaries are in high risk in JRC dataset. The risk of inundation in tributaries is in the middle for CaMa-Flood result. Flood extent simulation in tributaries is affected by many factors including model's spatial resolutions, model parameters such as channel cross-section, and also flood frequency analysis in observation based product. The other difference is flood in Asia (AS) is the floodplain water depth: both the two model-based results (JRC and CaMa-Flood) predict similar but much higher water depth (>10 m) along the main river channels.While the water depth in the GAR data is only around 5 m.The water depth in the Tonle Sap lake is probably not considered as flood hazard in GAR and JRC. The third difference is related to the topography. Stripes are found in the lower Mekong delta regions in JRC, which is caused by the biases of the topography (this has been well explained and addressed by the MERIT dataset, see reference of (Yamazaki et al., 2017)) . In GAR, the spatial distribution of the water depth over the delta regions are also not

vary realistic. The accuracy of the topography (DEM) is of vital importance for inundation calculation especially for the lower flat regions. The flood extent with water depth between 0.01m-1.0m is the most sensitive to the topography (grey areain Figure **??**). Corresponding extent in GAR is mainly over the Tonle Sap lake, where the underwater bathymetry is not accessible and the backwater effect is difficult to accurately modelled; the corresponding extent in the CaMa-Flood is mainly distributed in the coastal deltas because of the bifurcation channels (Yamazaki et al., 2014) and around the boundaries of the flood extent which seems more realistic. Despite of the differences, CaMa-Flood hazard map (Figure **??**-c) is reasonable since its extent is almost within range of existing hazard maps (Figure **??**-a,b) even though the baseline topography or methodology are different. The floodplain water depth in CaMa-Flood also exhibits similar spatial patterns especially with the JRC flood hazard map, despite of the values in the Tonle Sap lakehighest. 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.

A summary of the total inundation area based on the three maps is provided as Figure**??**-a, with two different threshold for inundation as 0.01 m (left panel)and 1.0 m (right panel), respectively. For either of the criterion, the sum of the inundation area in CaMa-Flood (averaged of different experiments) is within the value range of GAR and JRC. The flood extent with 0.01-1.0 m water depth accounts for around 25%for the three products and contributes the most to the overall deviation. While for larger flood with flood water depth larger than 1.0 m, the three products provide similar total inundation area . The remaining differences can be explained by the runoff uncertainty because the deviation of results driven by different runoff inputs exceeds the difference among the mean value of the three products (right panel of Figure **??**-a).

The variation of the inundation area is closely related to the variation of the predicted discharge at Pakse which represents most the river discharge to the lower Mekong from upstream (Figure **??**-c). The discharge driven by R2 (e2o_cnrs) is the highest and R6 (e2o_univk) is the lowest in both the inundation area and the discharge (Figure **??**-b), while the real discharge (GRDC) is among the spread ranges of all the seven different runoff inputs. It is also notable that the inundation area for CaMa-Flood at 3 arcsec is 68809.1 km^2 and 46016.2 km^2 (Table **??**) for area with water depth larger than 0.01 m and 1.0 m, respectively. The total inundation area is around ~20%lower than the results based on 30 arcsec resolution shown in (Figure **??**-a, 79791.9 km^2 and 60073.7km^2, respectively). The difference at the two spatial resolution is resulted from the ability to describe the heterogeneity in topography with the high-resolution topography data. This large difference indicates that the current assessment on the flood risk/impact could have been overestimated and it requires a necessity to apply a much higher-resolution topography to the current methodologies.

4 Overview of uncertainties for different floods at a global scale

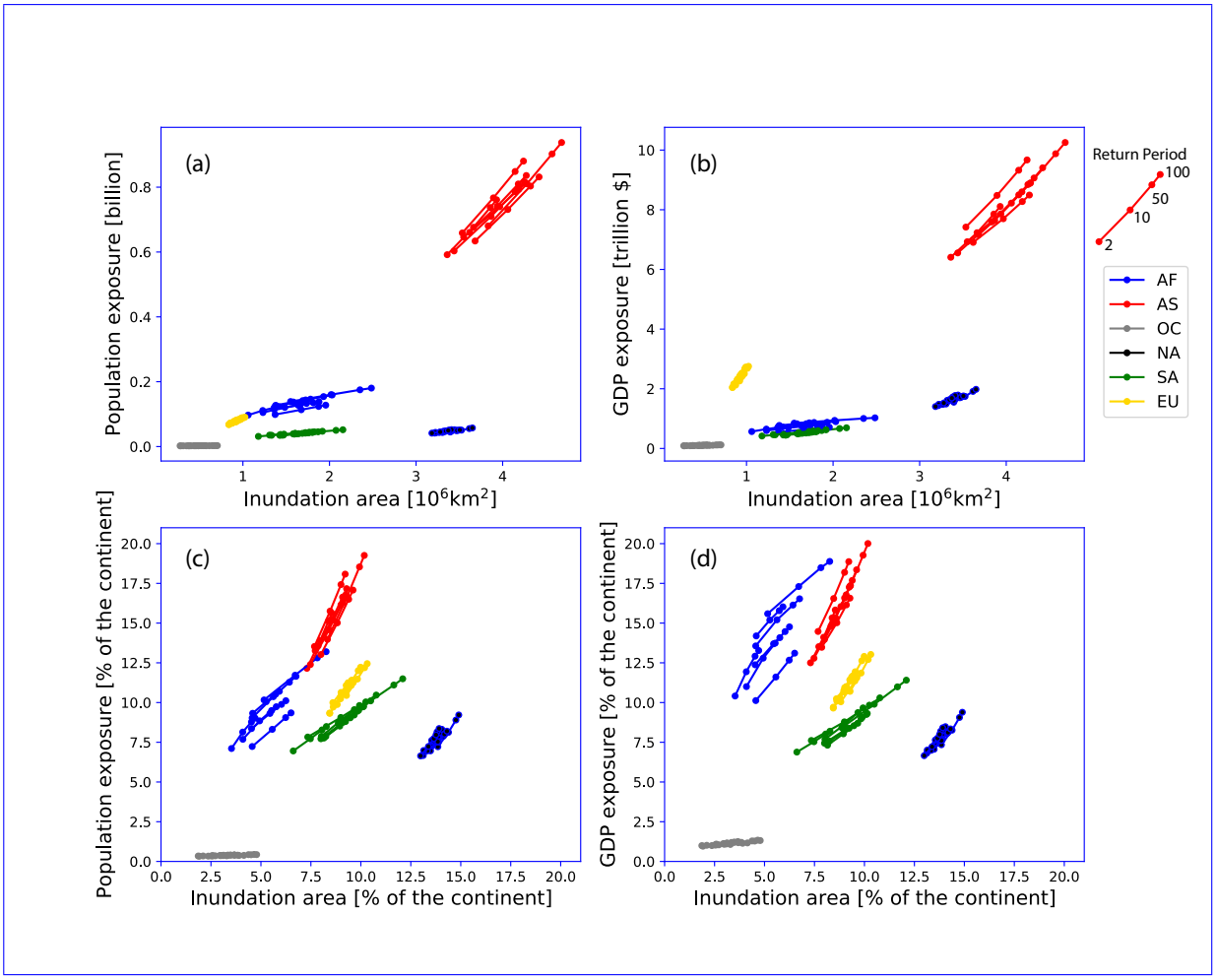


Figure 8. ~~(a) shows~~ Population and economic exposure to floods in different return periods and the uncertainties ~~of CaMa-Flood total inundation area driven by different~~ due to runoff~~(boxplots~~. The first row shows the absolute values, ~~from R1~~ while the second shows the proportion to ~~R7, Table 1) and~~ total values in the ~~comparisons of average area with other two maps (blue dashed and red dashed line)~~specific continent. The ~~sum of~~ left-hand column shows the ~~inundation area is counted by area with floodplain water depth higher than 0.01 m (left panel) and 1.0 m (right panel)~~population, ~~respectively~~while the right-hand column shows the Gross Domestic Product. ~~(b) Comparison of~~ Different colors represent different continents and the ~~estimated annual maximum discharge at Pakse gauge by CaMa-Flood driven by~~ values for four different ~~runoff inputs. Pakse gauge~~ return periods (~~105.8°E~~i.e., ~~15.1167°N~~1-in-2, 1-in-10, 1-in-50 and 1-in-100 year) ~~is located right up to~~ are marked. Different curves in the ~~edge of~~ same color represent the ~~upper reaches~~ results for ~~the study area (see Figure ??-c)~~various runoff inputs. The ~~plot is only for the annual maximum daily discharge covering 1980-1993 determined by the available period~~ names of ~~GRDC~~different runoff inputs are not specified.

Figure 4 summaries the mean floodplain water depth and the related uncertainties over the globe in order to have a broader view of FFA above regional results. Results corresponding to floods in two different return period (i.e., 100 year and 20 year) were compared. 20-yr flood is within the computation period (35 years in this study) while the water depth for 100-yr floods has to be obtained through extrapolation.

For the mean values, the floodplain water depth will exceed 10 m in most of the main channels of large rivers if suffering a 100-yr flood (Figure 4-a). The risk is especially high 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 the Amazon River, Congo River, large rivers in southern China, southeastern 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 rivers in Siberia. The spatial patterns of the floodplain water depth for 20-yr flood are very similar to the 100-yr flood but with lower magnitudes (Figure 4-b) . The total uncertainties, including different fitting distributions and multiple runoff inputs are shown as Figure 4-c,d for 100-yr and 20-yr flood, respectively. In general, high uncertainties are along with high predicted water depths. What needs to be noticed is the higher uncertainties tend to occur in mountainous regions rather than the flat regions, typically in the rivers originating from the Tibetan Plateau and Siberia. The uncertainties in the lower Yangtze, Mekong, Salween rivers are much lower than estimated uncertainties in their upstreams while the mean water depths are higher in the lower reaches. High uncertainties are also found in the Congo River. Although it is not surrounded by high mountains, its main channel is relatively short compared to its drainage area. The fluctuation in the river discharge will lead to high gradients in the water level which could be associated with its high uncertainties shown in (Figure 4-c and dflood-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 uncertainties in the Amazon is relatively low compared to the high mean water depths, this could result from the higher consistency of modelled runoff or fitting distributions and its relatively flat topography. The uncertainties in 20-yr show relatively lower values than that in the 100-yr floodinundation 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 mean floodplain water depth and uncertainties over the globe for 100-yr flood (left column) and 20-yr flood (right column). The first row is the mean floodplain water depth. The other three rows are the total uncertainties and uncertainties from fitting distribution (with e2o_ecmwf runoff) and runoff inputs (with GEV fitting function).

As stated in the regional studies, the uncertainties of predicted water depth are mainly contributed by the deviations 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 runoff inputs. At a global scale, the same results are found that the uncertaintiesresulted from fitting distributions contribute much less than that from the runoff. Figure 4-e and f show the uncertainties due to selections of the fitting distributions in the experiments of group driven by e2o_ecmwf. Very small values are found in the same scale of colours for the total uncertainties. On the contrary,

the uncertainties due to runoff inputs (within experiments group fitted by GEV distribution as an example, Figure 4-g,h) have similar magnitude of the total uncertainties, indicating that most of the uncertainties can be attributed to the deviations in runoff inputs.

The other finding is for the flood within the period of simulations (20-yr flood in the case of 35-years' simulation in this study), the uncertainties due to fitting distribution is nearly zero for the globe (Figure 4-f) . While, higher uncertainties can be visualized for the 100-yr flood(Figure 4-e) especially in the Amazon, Indus and rivers in the southeast Asia, southern China and the Siberia. The differences tell that the uncertainties due to fitting distribution are mainly because of the extrapolation out of the period with simulated or observed data. While within the period of available data, uncertainties due to fitting distributions will be efficiently constrained.

Figure ?? shows the changes in contribution of uncertainties in fitting distribution (and runoff inputs) to the total uncertainties from 20-yr flood to 100-yr flood. Positive values show that the uncertainty contribution is higher in 100-yr flood than that in 20-yr flood. Figure ??-a indicates that for almost all the global grids, the contribution of the uncertainties due to fitting distribution increases. This is mainly 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 extrapolation of fitting and larger uncertainties occur in the fitting tails. Although the uncertainties in runoff inputs are still the dominant, we need to pay more attention to the selection of fitting distribution for the rarer floods . 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.

Changes in the contribution of uncertainties to the full uncertainties for 20-yr flood and 100-yr flood. a) represents the contribution of uncertainties among fitting distributions and b) represents the contribution of uncertainties among runoff inputs. The contribution is calculated as the ratio of standard deviation to the total uncertainties of multiple experiments, corresponding to the third (and fourth) row and the second row, respectively.

Figure ??-b indicates the changes in contribution of uncertainties from runoff inputs. Obvious deviations are found between the mountainous/dry regions (e.g., The Rocky Mountains, Sahel, Central Australia, Central Asia) and floodplain/wet regions (e.g., Amazon, Congo, Ganges, Indonesia). The change means, for the rare floods (100-yr flood), the contribution of uncertainties from runoff inputs increases for the wet regions. Selection of the appropriate runoff becomes more important as well in this case. Note that the total uncertainty is not equal to the sum of the uncertainties from runoff and fitting distributions, the increases of both contributions indicate higher necessity of accounting for the uncertainties especially in the wet regions.

4 Discussion Discussions and Conclusionconclusion

4.1 DiscussionDiscussions

This study assesses the flood risk based on pure assessed the FHM based on simulations with a global hydrodynamic model (CaMa-Flood). Due to The analysis of flood hazards can be uncertain because of the multiple choices of runoff inputs, fitting

distributions for ~~flood frequency analysis (FFA)~~ FFA and the variables for FFA~~, the analysis of flood risk can be uncertain. The performance of the FFA is also varying in experiments with various combination of the above conditions. We conclude from the analysis with a performance metric *aic* that the river water depth (V1_*rivdph*) is more suitable for the FFA than converting the river storage to river depth (V2_*sto2dph*). Applying F2-GAM and F4-GUM is not suitable for the FFA because it leads to low fitting performance at a global scale through a ranking approach. F6-WAK performs the best mainly because it has five parameters while other 3-parameters fitting distributions (F1-GEV, F3-PE3 and F5-WEI) are in the middle. The fundamental limitation of this study is that the conclusion is based on uncalibrated simulations rather than observations and conclusions can differ if different routing method (rather than CaMa-Flood) is used. However, the attempts to assess the flood risk at a large scale provide reference information and routines for similar analysis. Moreover, FFA at the global scale based on observations is not feasible because of the lack of long-term observations on the water level or water storage and the difficulty to reach all the existing data. Model simulation has its advantage as it covers a larger areaand longer period because the forcing variables are much better in temporal and spatial coverage (Jones and Kay, 2007). Simulations can also estimate the high floods while the *in-situ* measurement for the large discharge can be with high uncertainty as well (Di Baldassarre and Montanari, 2009). Thus at the current stage model estimations are still important and can be relied on as a good reference data to conduct the flood frequency analysis if the uncertainties are properly treated~~. 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 ~~uncertainties in the estimated floodplain water level can be derived from multiple sources.~~ Trigg et al. ~~(~~2016~~);~~ Bernhofen et al. ~~(~~2018~~)~~ ~~compared the flood hazard maps (inundation area) over Africa with multiple products from six different institute. Large diversity is found in the inundation area among different model products, however, they have difficulties to attribute the variations or explain the behind reasons because different products use different forcing input, different model, different topography and different frequency analysis. They suggested rather than the product-level comparisons, component-level comparisons with limited variables could be better to attribute the uncertainties. In this study, by fixing the hydrodynamic model (CaMa-Flood), other uncertainties can be more easily quantified and attributed. Runoff inputs are regarded as the largest contributor to the final uncertainty. Because the runoff inputs are driven by the same WFDEI forcing, the differences in the output therefore explicit the difference of land surface models or hydrological models (**?**). Moreover,~~ FHM and FFA only uses the annual maximum water level (or water storage)~~;~~ therefore, the variety ~~in the FFA only demonstrate the performance~~ only demonstrates the performance of rainfall-runoff models in reproducing the ~~discharge peaks~~peak discharge. Separation of surface runoff and subsurface runoff, and the evaporation rate during the extreme raining events can lead to the ~~differences~~ deviations in total runoff and the ~~hydrodynamic processes during routing. Among the different runoff inputs, the e2o_ecmwf by HTESSEL stands in the middle according to the point analysis in Figure 6-d. However, the runoff input providing the middle estimation of water depth varies for the world (Figure **??**). This shows no runoff input is preferable~~ hydrography after 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~~for estimating the high water states, nor the land surface or hydrological models. Either the estimation needs validation with observations (~~Lin et al., 2019~~), or ensemble simulation is needed at a large scale (~~Warszawski et al., 2014~~; ~~?~~). For specific regions, Figure ~~??~~-c can be used as a reference for selecting a runoff input in the middle to reduce the risk of large uncertainty in floods if conducting all the runoff inputs consumes a large amount of computation.~~

~~The agreement on the inundation prediction is assessed for the lower Mekong basin. In general, the agreement is high (>50%) for most of the inundation area (86.6%), and approaching 100%if the predicted mean floodplain water depth is larger than 0.5 m. Despite the highly certain inundation area, attention should be also paid to regions where the inundation is predicted while the model agreement is relatively low (<50%) because the uncertainty from multiple sources will cause different consequences (inundation or non-inundation), especially the low human habits as shown in Figure ~~??~~.~~ Although ~~?~~ 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, ~~both the regions with different inundation agreement are in high risk if suffering large floods, different reactions should be taken in different priority according to the water depth and agreement on inundation as well as the local conditions regarding population and property. Inundation area is also calculated. We would like to note that, because of the different ways of estimating the inundation area, there is a 10%difference in the total inundation area for a 100-yr flood. The difference is mainly due to the low model agreement over the regions with the low predicted floodplain water level.The variations of the inundation area by different experiments can be as large as 20%~~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. ~~The ratio can be higher if excluding permanent water bodies such as the lakes and river channels. This necessitates large effort to decreases the uncertainties related to the FFA. Assimilation of the altimetry data (~~Revel et al., 2018~~) or inundation area(~~Hostache et al., 2018~~) from satellites can be a solution to evaluate model results at a large scale especially in regions where ground observations are not available.~~

~~Two other sources of flood hazard maps are utilized to validate the results from CaMa-Flood. However, because the two maps are also generated by models to some degree, we cannot quantitatively evaluate the CaMa-Flood performance from the comparison. The differences are mainly in inundation area with shallow water depth and in the tributaries. While for larger floods (water depth >1.0 m),~~ 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 differences among three products are much smaller and the variation of CaMa-Flood results from the variations of runoff inputs. The variation is related to the discharge estimation at the upper gauge. However, the proper validation of the inundation at a theoretical level (100-yr flood) is still not feasible because of the lack of spatial observations for the equivalent flood. Comparison with other products only increases the credibility of estimations but helps to identify the discrepancies among different products.To some degree, CaMa-Flood is superior in describing the spatial patterns and is more flexible in the selection of different spatial resolution. The variation between results driven by different runoff inputs requires an improvement of runoff estimation through bias correction or data assimilation.

During the validation, we found a ~20% inundation to roughness (Pappenberger et al., 2008; **?**) and spatial resolution (**??**) , although their results needs validation over a global scale. **?** 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 inundation area between two different spatial resolution (3 arcsec and 30 arcsec). Similar result is found in Hinkel et al. (2014) as they evaluated the coastal flood damages by using two kinds of topography data GLOBE (30 arcsec) and SRTM (1 arcsec). The exposed area, population and assets were lower by 50% to 70% in assessment with high-resolution topography (SRTM) than the low-resolution GLOBE. Therefore studies with 30 arcsec (1km) (Jongman et al., 2012; Ward et al., 2013; Jongman et al., 2015) could provide overestimated results for this end. This difference requires us the ability to have higher resolution topography and corresponding technologies to obtain the results with more spatial details.

The floodplain water depth and its uncertainties are investigated at a global scale . The conclusions are similar to regional studies as that mean values, they cannot reflect real sensitivity if using entirely different forcing. With improvement of the computation capability, the major contribution to the final uncertainties resulted from the deviations of runoff inputs. Although the uncertainties in the inundation area is yet investigated, the results will be in the same direction since the inundation area is highly associated with the floodplain water depth. Comparisons of the contribution of uncertainties from fitting distribution and runoff inputs to different floods 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., 100-yr, 20-yr)indicates that uncertainties from fitting distributions resulted from the extrapolation out of the period with data. Having longer-term modelled or observed data will greatly reduce the uncertainties. Investigation of historical floods will also benefit for the improvement of FFA (Payrastre et al., 2011). The deviating results about the uncertainties contribution from runoff inputs for wet and dry (or mountainous and flat regions) also requires a differentiating treatment to different kinds of floods in different regions. The behind reasons for the differences are associated with the topography or model performance in different regions, while they remain to be investigated.

The water depth , flood risk and damages are sensitive to the flood protection adaptions (Ward et al., 2013). Dam regulation (Wang et al., 2017) and river levees (Berning et al., 2001) are effective ways to mitigate the potential risks of floods. However,

levees, dams), which will lead to overestimation of the ~~database is not accessible at a large scale and the flood protection is not applied in the current model. Attempts on the improvement of~~ 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 ~~floodplain water depth after flood frequency analysis. Uncertainties can result from the selection of~~ 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 ~~uncertainty sources. Uncertainties from the~~ 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 ~~suggest~~ suggests the importance of rainfall-runoff model calibration (or runoff bias correction) if gauge discharge observation is available. ~~No preferable runoff inputs are available at the global scale, but the fitting performance implies that directly using the river water depth for FFA is better than using converted water depth from water storage. The fitting distribution WAK is the best among the various fitting distributions. The results of model agreement for inundation estimation is expectable as high agreement is found for inundation regionswith high predicted floodplain water depth. But additional information of model agreement will be helpful for the decision-makers during the flood protection. Inundation area related to the water level also shows large uncertainties, which will increase the difficulty of assessing flood risk and flood damages. The variation of contribution of uncertainties from fitting distribution and runoff inputs for two different level of floods (100-yr flood and 20-yr flood)shows that uncertainty from fitting distribution is due to the extrapolation out of the period with data, and increases with the flood magnitude. While uncertainties from runoff are spatial varied and the contribution from runoff can be higher for larger floods in wet regions~~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) ~~as well as the~~ and also model improvement by taking into account ~~the~~ human interventions are needed to reduce the various uncertainties.

~~The global hydrodynamic model CaMa-Flood is available from . The topography data MERIT is available from . The JRC flood hazard map is available from and the GAR flood hazard map is available from . 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 .~~

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.

~~The authors declare no competing interests.~~

Competing interests. The authors declare no competing interests.

Acknowledgements. This study was supported by "KAKENHI" 20H02251 & "~~by MEXT Japan~~ 20K22428" by JSPS and also by "LaRC-Flood project" by MS&AD Holdings. The computation is run at the server in ~~the~~ 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.

Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., and Feyen, L.: Advances in pan-European flood hazard mapping, Hydrological Processes, 28, 4067–4077, https://doi.org/10.1002/hyp.9947, 2014.

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.

Berning, C., Du Plessis, L. A., and Viljoen, M. F.: Loss functions for structural flood mitigation measures, Water SA, 27, 35–38, https://doi.org/10.4314/wsa.v27i1.5007, 2001.

Beven, K. and Hall, J.: Applied Uncertainty Analysis for Flood Risk Management, IMPERIAL COLLEGE PRESS, https://doi.org/10.1142/p588, https://www.worldscientific.com/doi/abs/10.1142/p588, 2014.

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.

Blazkova, S. and Beven, K.: Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty), Water Resources Research, 38, 1–14, https://doi.org/10.1029/2001wr000500, 2002.

Bras, R., Moughamian, M., and McLaughlin, D.: Estimation of flood frequency: a comparison of physically based procedures, in: US China Bilateral Symposium on the analysis of extraordinary flood events, Nanjing, 1985.

Di Baldassarre, G. and Montanari, A.: Uncertainty in river discharge observations: A quantitative analysis, Hydrology and Earth System Sciences, 13, 913–921, https://doi.org/10.5194/hess-13-913-2009, 2009.

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.

Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A., and Feyen, L.: Development and evaluation of a framework for global flood hazard mapping, Advances in Water Resources, 94, 87–102, https://doi.org/10.1016/j.advwatres.2016.05.002, http://dx.doi.org/10.1016/j.advwatres.2016.05.002, 2016.

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.

GAR: GAR Global Risk Assessment: Data, Methodology, Sources and Usage, Tech. rep., UNISDR, 2015.

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

Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S., Marzeion, B., Fettweis, X., Ionescu, C., and Levermann, A.: Coastal flood damage and adaptation costs under 21st century sea-level rise, Proceedings of the National Academy of Sciences of the United States of America, 111, 3292–3297, https://doi.org/10.1073/pnas.1222469111, 2014.

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.

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, https://doi.org/10.1029/2017WR022205, 2018.

Jones, D. A. and Kay, A. L.: Uncertainty analysis for estimating flood frequencies for ungauged catchments using rainfall-runoff models, Advances in Water Resources, 30, 1190–1204, https://doi.org/10.1016/j.advwatres.2006.10.009, 2007.

Jongman, B., Ward, P. J., and Aerts, J. C.: Global exposure to river and coastal flooding: Long term trends and changes, Global Environmental Change, 22, 823–835, https://doi.org/10.1016/j.gloenvcha.2012.07.004, http://dx.doi.org/10.1016/j.gloenvcha.2012.07.004, 2012.

Jongman, B., Winsemius, H. C., Aerts, J. C., Coughlan De Perez, E., Van Aalst, M. K., Kron, W., and Ward, P. J.: Declining vulnerability to river floods and the global benefits of adaptation, Proceedings of the National Academy of Sciences of the United States of America, 112, E2271–E2280, https://doi.org/10.1073/pnas.1414439112, 2015.

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.

Kjeldsen, T. R., Lamb, R., and Blazkova, S. D.: Uncertainty in flood frequency analysis, Applied Uncertainty Analysis for Flood Risk Management, pp. 153–197, 2014.

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., David, C. H., Durand, M., Pavelsky, T. M., Allen, G. H., Gleason, C. J., and Wood, E. F.: Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches, Water Resources Research, 55, 6499–6516, https://doi.org/10.1029/2019WR025287, 2019.

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.

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.

Merz, B. and Thieken, A. H.: Flood risk curves and uncertainty bounds, Natural Hazards, 51, 437–458, https://doi.org/10.1007/s11069-009-9452-6, 2009.

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.

Odunuga, S. and Raji, S. A.: Flood Frequency Analysis and Inundation Mapping of Lower Ogun River Basin, Journal of Water Resource and Hydraulic Engineering, 3, 48–59, file:///C:/Users/Becks/Downloads/JWRHE10055-20141218-140824-9711-45082.pdf, 2014.

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.

Payrastre, O., Gaume, E., and Andrieu, H.: Usefulness of historical information for flood frequency analyses: Developments based on a case study, Water Resources Research, 47, 1–15, https://doi.org/10.1029/2010WR009812, 2011.

Qi, S., Brown, D. G., Tian, Q., Jiang, L., Zhao, T., and Bergen, K. M.: Inundation extent and flood frequency mapping using LANDSAT imagery and digital elevation models, GIScience and Remote Sensing, 46, 101–127, https://doi.org/10.2747/1548-1603.46.1.101, 2009.

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.

Revel, M., Yamazaki, D., and Kanae, S.: Experimental Investigation of Sediment Trap Efficiency in Reservoirs, Journal of Japan Society of Civil Engineers, 74, 307–312, https://doi.org/10.4038/engineer.v47i2.6863, 2018.

Rudari, R., Silvestro, F., Campo, L., Rebora, N., and Boni, G.: Improvement of the Global Flood Model for the GAR 2015, Tech. rep., The United Nations Office for Disaster Risk Reduction, 2015.

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

Salinas, J. L., Laaha, G., Rogger, M., Parajka, J., Viglione, A., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins-Part 2: Flood and low flow studies, Hydrology and Earth System Sciences, 17, 2637–2652, https://doi.org/10.5194/hess-17-2637-2013, 2013.

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.

Sarhadi, A., Soltani, S., and Modarres, R.: Probabilistic flood inundation mapping of ungauged rivers: Linking GIS techniques and frequency analysis, Journal of Hydrology, 458-459, 68–86, https://doi.org/10.1016/j.jhydrol.2012.06.039, http://dx.doi.org/10.1016/j.jhydrol.2012.06.039, 2012.

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.

Smith, A., Sampson, C., and Bates, P.: Regional flood frequency analysis at the global scale, Water Resources Research, 51, 539–553, https://doi.org/10.1002/2014WR015814, 2015.

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.

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.

Wang, W., Lu, H., Ruby Leung, L., Li, H. Y., Zhao, J., Tian, F., Yang, K., and Sothea, K.: Dam Construction in Lancang-Mekong River Basin Could Mitigate Future Flood Risk From Warming-Induced Intensified Rainfall, Geophysical Research Letters, 44, 10,378–10,386, https://doi.org/10.1002/2017GL075037, 2017.

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.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J.: The inter-sectoral impact model intercomparison project (ISI-MIP): Project framework, Proceedings of the National Academy of Sciences of the United States of America, 111, 3228–3232, https://doi.org/10.1073/pnas.1312330110, 2014.

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.

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.

Proof-Reading-Service.com
PhD theses, journal papers, books and other professional documents

Proof-Reading-Service.com Ltd, Devonshire Business Centre, Works Road, Letchworth Garden City, Hertfordshire, SG6 1GJ, United Kingdom
Office phone: +44(0)20 31 500 431
E-mail: enquiries@proof-reading-service.com
Internet: http://www.proof-reading-service.com
VAT registration number: 911 4788 21
Company registration number: 8391405

19 January 2021

To whom it may concern,

RE: Proof-Reading-Service.com Editorial Certification

This is to confirm that the document described below has been submitted to Proof-Reading-Service.com for editing and proofreading.

We certify that the editor has corrected the document, ensured consistency of the spelling, grammar and punctuation, and checked the format of the sub-headings, bibliographical references, tables, figures etc. The editor has further checked that the document is formatted according to the style guide supplied by the author. If no style guide was supplied, the editor has corrected the references in accordance with the style that appeared to be prevalent in the document and imposed internal consistency, at least, on the format.

It is up to the author to accept, reject or respond to any changes, corrections, suggestions and recommendations made by the editor. This often involves the need to add or complete bibliographical references and respond to any comments made by the editor, in particular regarding clarification of the text or the need for further information or explanation.

We are one of the largest proofreading and editing services worldwide for research documents, covering all academic areas including Engineering, Medicine, Physical and Biological Sciences, Social Sciences, Economics, Law, Management and the Humanities. All our editors are native English speakers and educated at least to Master's degree level (many hold a PhD) with extensive university and scientific editorial experience.

Document title: **Uncertainty in flood frequency analysis of hydrodynamic model simulations**

Author(s): **Xudong Zhou et al.**

Format: **British English**

Style guide: **http://www.natural-hazards-and-earth-system-sciences.net/submission/general_terms.html**