RESEARCH ARTICLE - HYDROLOGY

Analysis of extreme flow uncertainty impact on size of flood hazard zones for the Wronki gauge station in the Warta river

Tomasz Dysarz[1] · Joanna Wicher-Dysarz[1] · Mariusz Sojka[2] · Joanna Jaskuła[2]

Abstract

In this paper, the impact of maximum flow uncertainty on flood hazard zone is analyzed. Two factors are taken into account: (1) the method for determination of maximum flows and (2) the limited length of the data series available for calculations. The importance of this problem is a consequence of the implementation of the EU Flood Directive in all EU member states. The factors mentioned seem to be among the most important elements responsible for potential uncertainty and inaccuracy of the developed flood hazard maps. Two methods are analyzed, namely the quantiles method and the maximum likelihood method. The maximum flows are estimated for the Wronki gauge station located in the reach of the Warta river. This simple river system is located in the central part of Poland. The length of the available data is 44 years. Hence, the series of the lengths 40, 30 and 20 years are tested and compared with reference calculations for 44 years. The hydrodynamic model HEC-RAS is used to calculate water surface profiles in steady state flow. The Python scripting language is applied for automation of HEC-RAS calculations and processing of final results in the form of inundation maps. The number of trials for each factor is not huge to keep the presented methodology useful in practice. The chosen measure of uncertainty is the range of variability for maximum flow values as well as inundation areas. The estimated values stressed the great importance of the factors analyzed for the uncertainty of the maximum flows as well as inundation areas. The impact of the data series length on the maximum flows is straightforward; a shorter data series gives a wider range of variability. However, the dependencies between other factors are more complex. Hence, the application of methodology based on the simulation and GIS data processing for assessment of this problem seems to be quite a good approach.

Keywords Flood hazard maps · Hydrological uncertainty · River flow simulation · Geo-processing · Python scripting

Introduction

The main focus in the present paper is uncertainty of flood hazard maps. The motivation for such research is implementation of the EU Flood Directive (European Parliament 2007) in EU member states. Although the concepts introduced in the EU Flood Directive could significantly reduce the risk related to flood damage, there are two general problems. These are (1) non-uniformity of the approaches applied in different EU member states and (2) uncertainty of

the developed maps. The first is reported by many researchers and consists in the differences at many steps of flood hazard map preparation. The most important are (a) different quality of topographic data (Van Alpen and Passchier 2007), (b) different methods applied for calculation of maximum flows with a specific return period (Van Alpen and Passchier 2007), (c) different return periods for representation in flood hazard maps (Van Alpen and Passchier 2007; Nones 2015) and (d) different hydraulic models applied (Van Alpen and Passchier 2007). Additionally, the lack of necessary data for flood damage analysis and related risk assessment is also reported (e.g., Albano et al. 2017) Although some researchers express the need for standardization over the EU member states (Nones 2017), legal regulations in this area have not been established yet.

The second problem seems to be more important, because it is related to our lack of knowledge. The uncertainty of the flood hazard maps should be investigated taking into

✉ Tomasz Dysarz
tomasz.dysarz@up.poznan.pl

1 Department of Hydraulic and Sanitary Engineering, Poznan University of Life Sciences, Poznan, Poland

2 Institute of Land Improvement, Environmental Development and Geodesy, Poznan University of Life Sciences, Poznan, Poland

account the multistage procedure of the maps' development including many different elements (e.g., Bates et al. 2014; Teng et al. 2017). Generally, the uncertainty in flood inundation modeling can be categorized into seven major types: (1) topographic data, (2) hydrologic data, (3) data for preliminary estimation of roughness, (4) method applied for final roughness calibration, (5) method applied for estimation of maximum flows, (6) structure of the applied model and (7) transition of the model results to the maps (e.g., Refsgaard and Storm 1990; Cook and Merwade 2009; Calenda et al. 2009; Liu and Merwade 2018). Some of the above-listed elements become less uncertain due to the development of the technology, e.g., increasing accuracy and resolution of LIDAR data for DEM elaboration (e.g., Gilles et al. 2012; Sampson et al. 2012; Walczak et al. 2013; Laks et al. 2017), involvement of satellite and remote sensing data (e.g., Jung et al. 2014; Arseni et al. 2017), elaboration of more detailed databases on flood events (e.g., Kundzewicz et al. 2017) and broader availability of more accurate hydraulic models (e.g., Szydłowski et al. 2013; Gąsiorowski 2013; Brunner 2016a; Kolerski 2018). A specific problem is the uncertainty related to the channel and floodplain roughness (e.g., Dimitriadis et al. 2016; Pappenberger et al. 2008; Engeland et al. 2016; Liu and Merwade 2018). The lack of proper data for the initial assessment of these factors may be corrected by robust application of the calibration method. On the other hand, the time-consuming procedure of model parameters' identification may be reduced by relatively good quality data describing land cover with cover of the channel bed.

In the area of uncertainty assessment, one of the still unsolved problems is the estimation of maximum flows, also called design floods. The magnitudes of such design floods depend on the assumed return period uniquely linked to probability of exceedance. In fact, the maximum flows are the basis of the whole procedure for elaboration of flood hazard maps after the model calibration. The estimation of their magnitudes consists of a number of steps (e.g., Calenda et al. 2009). Each of these steps may be prone to potential uncertainty sources (e.g., Merz and Thieken 2005; Griffis and Stedinger 2007; Laio et al. 2011). As it is reported by many researchers, the most important sources are (1) measurement errors and rating curve evaluation (Di Baldassarre and Montanari 2009; Di Baldassarre et al. 2010, 2012), (2) plotting position formula; (Hirsh and Stedinger 1987; Ewemoje and Ewemooje 2011), (3) assuming the randomness, stationarity, homogeneity, independence of the data (Serinaldi and Kilsby 2015; Serago and Vogel 2018), (4) choice of the observation period and data sampling (Calenda et al. 2009; Schendel and Thongwichian 2015, 2017), (5) probability distribution function (Calenda et al. 2009; Schendel and Thongwichian 2015, 2017; Yuan et al. 2017; Sun et al. 2017) and (6) method for estimation of parameters (Romanowicz and Beven 2003; Calenda et al. 2009; Beven

and Hall 2014; Schendel and Thongwichian 2015, 2017; Parkes and Demeritt 2016; Sun et al. 2017). In fact, some of the listed problems could be reduced with advances in the methodology applied. But two of them seem to be present independently of the methods developed. These are factors (4) and (6) in the above list. The first, choice of the observation period and data sampling, is related to the limited historical data available. This problem will not be overcome with new methods. The second, the method for estimation of parameters, is still a matter of subjective choice. Hence, the error made with such a decision is very important for the potential modeler.

In this study, we focus on the uncertainty related to the method applied for determination of design floods and the length of the data series. Of course, there are also other elements potentially important for the accuracy and reliability of flood hazard maps, e.g., digital elevation models (DEM), type of hydrodynamic simulation software. Taking into account achieved nowadays, quality of available data and applicability of the available methodologies, these two elements analyzed in our research seems to be the most crucial factors causing uncertainty in flood hazard assessment and management. The analysis was performed on the basis of the data collected in one selected gauge station. This is the Wronki gauge station on the Warta river. The station is located in the town of Wronki in the western part of Poland. The main purpose of the paper is to estimate the uncertainty of the maximum flow evaluation and its impact on the uncertainty of the flood hazard zone determination. The uncertainty of the maximum flow calculation is assessed with respect to factors mentioned earlier: the method for evaluation of the maximum flows and length of the data series. The uncertainty of the flood hazard zone is analyzed taking into account a very basic measure: inundation area.

Study site description

The Warta river is the third longest river in Poland. The length of the river is 808.2 km, and the total watershed area is 54,480 km^2. It is the greatest tributary of the Oder river (Fig. 1a). The total annual precipitation in the Warta watershed ranges from less than 500 mm in the central part to approximately 650 mm in upper part.

The mean discharge estimated for the whole period is 124.7 m^3/s. The selected discharges are presented in Fig. 2. The presented light brown diamonds, dark blue line and red dots denote the annual minimum, mean and maximum flows, respectively, in Fig. 2a. In this part, the greatest floods are also marked. As one can see, the most dramatic flood occurred in 1979, although other flood events are shown in Fig. 2a in the years 1975, 1981, 2010 and 2011, with dangerous flooding generally occurring every 2–5 years. The

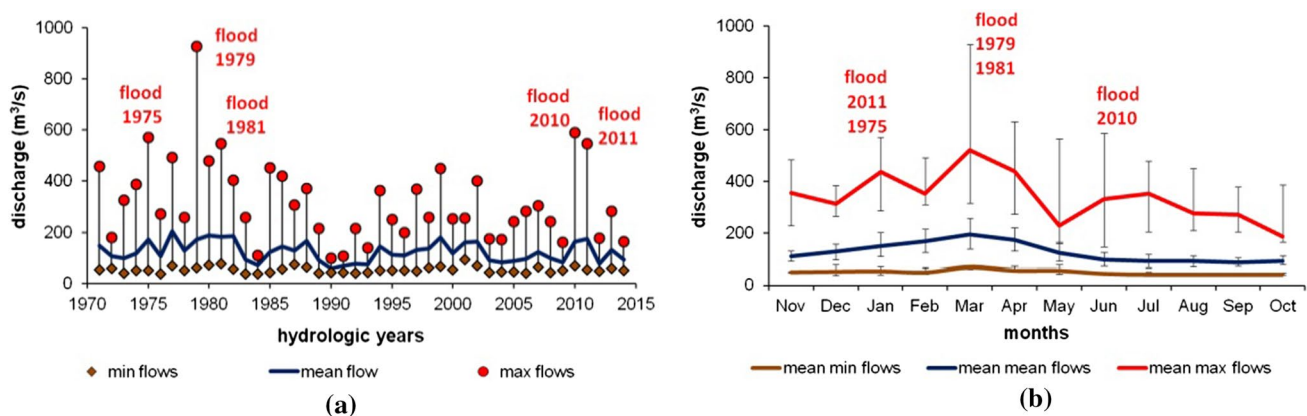**Fig. 1** Chosen case study: **a** location in Poland, **b** modeled river reach

Fig. 2 Variability of discharge in the Wronki gauge station for the period 1971–2014: **a** annual min, mean and max flows, **b** monthly min, mean and max flows

second part, Fig. 2b, presents characteristics of the monthly flows. The brown, blue and red lines denote average minimum, mean and maximum flows observed in each month of the hydrologic year. The error bars show the variability of the minimum, mean and maximum monthly flows. The same floods are marked. As we can see, the greatest flood events occur in March, during the season change from winter to spring. However, the floods were also observed in other months.

There is a high risk of flooding in the Warta river valley. The flood hazard maps were determined for the whole Warta river (Fig. 1a) during implementation of the EU Flood Directive (European Parliament 2007). It was done in the frame of the ISOK project, mentioned earlier. One of the most interesting reaches along the Warta is the channel near the town of Wronki (Fig. 1b). The reach seen in Fig. 1b is about 7.6 km. The average width of the channel is 70 m there. Along this short reach, the width changes slightly. The depth is about 1.5–1.6 m. The selected reach is located in the lower part, along the last 68 km of the river course,

where the slope is only 0.13‰ (www.mdwe70.pl). The flood threat along this reach is serious. In Fig. 1b, the range of the flood hazard zone determined for a 500-year flood was showed. The existing infrastructure, the town of Wronki and two bridges, is prone to flooding. Additionally, new road investments are planned there. The third bridge is going to be built upstream of the town. In the town of Wronki, the gauge station is located (Fig. 1b).

Materials and methods

Data collection

In this study, four data sets were used: (1) hydrologic data for the Wronki gauge station, (2) a digital elevation model (DEM) for the area surrounding modeled reach, (3) the measurements of the channel cross sections and structures, (4) additional GIS layers supporting preparation of the model and visualization of results.

The hydrologic data for the Wronki gauge station were obtained from the Institute of Meteorology and Water Management (IMGW—Polish: *Instytut Meteorologii i Gospodarki Wodnej*). The location of the gauge station is presented in Fig. 3a, b as a black dot. The data are daily observed water stages and estimated discharges for the period 1971–2014. The homogeneity of the hydrologic data was tested with the Mann–Kendall test recommended by Banasik et al. (2017). No significant trends were detected. The DEM applied in this research was obtained from the Head Office of Geodesy and Cartography (GUGiK—Polish *Główny Urząd Geodezji i Kartografii*). Its spatial resolution is 1 m × 1 m. The vertical accuracy equals 15 cm. The modeled reach of the Warta river is presented in Fig. 2b. The DEM represents the surrounding area and it is seen in Fig. 3b. It covers about 122 km². These data are stored as an ESRI GRID file. The memory size is about 467 MB. The cross-section measurements were obtained from the National Board for Water Management (KZGW—Polish: *Krajowy Zarząd Gospodarki Wodnej*). They are shown in Fig. 3a as green dots. There are 11 cross sections measured along the modeled reach. The average distance between measurements is 756 m. The obtained set of measurements also includes hydrostructures. In the modeled reach, two bridges are present. They are denoted as red squares in Fig. 3a. The measurements of cross sections near bridges are spaced closer. Their average distance apart is 52 m. On the other hand, the maximum distance between cross sections is 1460 m. The measurement of cross sections was done with a GPS device. The data have the form of a vector layer including points. The average number of measurement points in a single cross section is 136. The table of attributes includes position (*X*, *Y*), elevation (*Z*), code of point type and code of the land cover. The processing of the data enables interpolation of the bed, e.g., (Dysarz 2018a) and preparation

of the model (Dysarz et al. 2015). The computational cross sections are presented in Fig. 3b.

The additional GIS layers include vector layers as river centerline, river banks and lines of the computational cross sections as well as rasters such as orthophotomap, topographic map and OpenStreetMap (OSM). The listed vector layers were created during the process of model preparation. The river centerline is based on the Map of Hydrological Division of Poland (MPHP—Polish: *Mapa Podziału Hydrograficznego Polski*) available from KZGW. The two mentioned raster layers are applied as WMTS (Web Map Tile Service) servers linked to Geoportal 2—a web application prepared for management of spatial data sets. The OSM is a layer shared by OpenStreetMap Foundation (OSMF) on the Internet. All these layers are used here for better visualization of the results.

Estimation of maximum flows

For the period 1971–2014, annual maximum flow was determined. The annual maximum flows were the basis for calculating design floods with return periods of 10, 100 and 500 years. Two methods are applied for estimation of the maximum flows. These are: (1) the method of quantiles and (2) the maximum likelihood method. The methods are chosen due to their wide applicability in Polish conditions. Both of them are recommended in Banasik et al. (2017) as elements of flood hazard elaboration in Poland. In both cases, the probability distribution is modeled with the Pearson III function. In the first case, the quantiles of the distribution for cumulative probabilities 10%, 50% and 90% are estimated on the basis of the historical data. Then, the Pearson III curve was fitted in such a way that quantiles from historical data equal those calculated from the theoretical formulae. In the second method, the parameters of the distribution are set in such a way that the likelihood is maximized. The likelihood
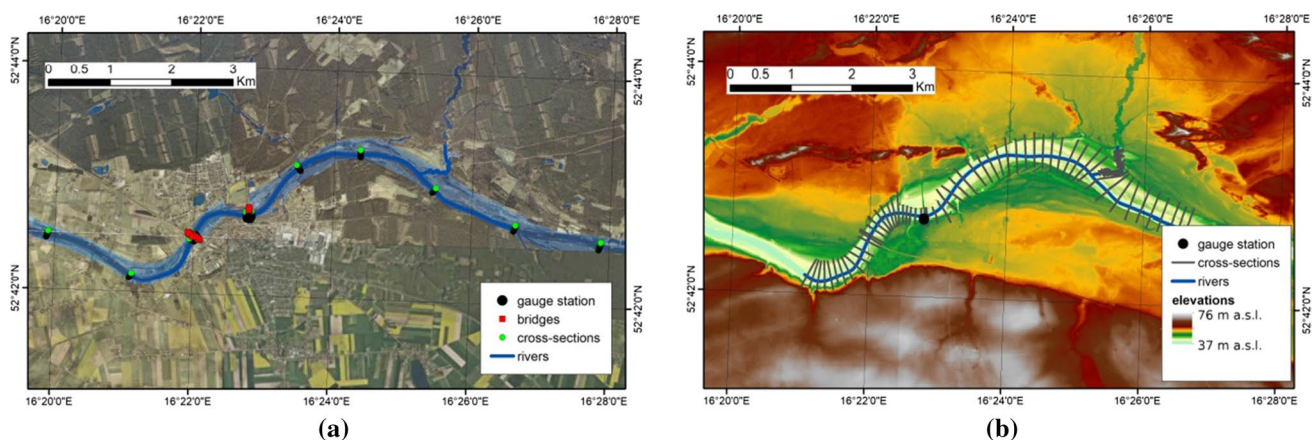

Fig. 3 Map of the chosen study case: **a** orthophotomap with inundation area for 100-year flood, **b** DEM and measured cross sections

is derived from the theoretical formula. Both methods are described in several hydrological books, e.g., Chow et al. (1988).

Modeling of river flow

The HEC-RAS is a well-known hydrodynamic model for rivers and water reservoirs. This program was designed at the Hydrologic Engineering Center (HEC). The second term in the name defines its application: River Analysis System (RAS). The concepts applied in the package are well described by Brunner (2016a). The HEC-RAS is applied for simulation of flow and transport processes in river networks, including floodplains and reservoirs. The modeled flow conditions include steady and unsteady longitudinal flow. The first is based on the simple Bernoulli equation. For description of the second model, the numerical solution of the St. Venant system of equations is implemented. There is also a module enabling the so-called quasi-unsteady flow simulation—simplified flow simulation in unsteady conditions on the basis of the Bernoulli equation. In the last versions of the package, the 2D flow module is also available. The HEC-RAS modules for simulation of transport processes include different transport of solutes, heat and sediments with deposition and erosion. The HEC-RAS package also includes several useful tools for data preparation and results processing. These tools include the module for GIS data processing called RAS Mapper (Brunner 2016b).

The model applied here is prepared for the reach of the Warta river near the town of Wronki. The length of the reach is about 7.6 km. In the model, there are 26 cross sections and 2 bridges defined. The average distance between computational cross sections is about 130 m (Fig. 3b). However, this distance is much more shorter in the case of cross sections located upstream and downstream of the bridges. It is a little bit more than 60 m in such a case. There is also a tributary in the model. It is a short reach of the Smolnica river (Fig. 3b). The length of this reach is about 1050 m. The role of the Smolnica river is not huge, but in the valley and floodplains of this river, a backwater may occur during flooding in the Warta river.

The maximum flows estimated for the Wronki gauge station are the basis for the calibration process. The values of the discharges are 560 $m^3\ s^{-1}$, 928 $m^3\ s^{-1}$ and 1211 $m^3\ s^{-1}$ for 10-, 100- and 500-year floods, respectively. The flood hazard maps are used to read the expected water surface elevations determined for each flood along the reach. The water surface elevations in the reach outlet are used as downstream boundary conditions. Although the changes of the discharge would also induce the changes in the water stage in the downstream cross section, there is a lack of data precisely describing this relationship. It is worth to remember that the water level in particular cross section may be influenced by hydraulic conditions occurring downstream of this location, e.g., local weirs, flow contractions, etc. Because we focused on the selected reach, this problem is beyond the scope of our research. Hence, approach assuming consistency with other ISOK results is implemented. It is expected that the impact of the downstream boundary condition is not huge, though the problem was not carefully analyzed.

The rest of the water levels are applied as reference values in the process of model calibration. The implemented method of calibration is based on the standard trial-and-error procedure. However, the marching nature of the algorithm for water surface calculation is used. It let to search the optimal roughness values from downstream to upstream, what simplifies the whole process. This procedure is described by Dysarz et al. (2015). One of the relatively poorly known elements of HEC-RAS is HEC-RAS Controller (Goodell 2014). It enables access to the HEC-RAS elements, because it is compiled as the Component Object Module (COM). Hence, the HEC-RAS Controller is an application programming interface (API). Any program able to read the COM library may be used to control HEC-RAS computations. In this research, the HEC-RAS Controller is used to set the discharges for a simulation, to run a simulation, to read the results and store them in DBF tables.

Tools of spatial analysis

The ArcGIS 10.5.1. software developed by Esri company is applied in this research (e.g., Docan 2016; Law and Collins 2018). It enables quite easy processing of GIS data such as vector and raster layers. An integral part of Arc-GIS is the ArcToolbox. It is a module including the main external tools and methods. Some of them are concurrent to methods available in the basic ArcGIS interface. Others extend the capabilities of the standard interface. Extension of the ArcGIS is possible also with specific plug-ins installed as ArcGIS toolbars. One such plug-in is HEC-GeoRAS (Cameron and Ackerman 2012). It is a toolbar with methods designed to support preparation of the river flow model.

The most important are tools applied for the generation of flood hazard maps including range of inundation, maps of depths and maps of water surface elevations. These are mainly spatial interpolation and map algebra methods. The natural neighbor method (e.g., Sibson 1981; Van der Graaf 2016) is applied for modeling of water surface spatial distribution. The conditional method for raters and reclassification is applied to properly process the interpolation results. The data for interpolation are read from DBF tables, where the simulation results were stored before.

Automation with Python scripts

Python is one of the most popular programming languages today and it is still under development. In this paper, version 2.7.12 is used. Its usefulness in many areas has been reported. This scripting language is relatively simple for the beginner, but it is also very powerful if applied by an experienced coder. A description of this language may be found in many books or on internet websites, e.g., Downey et al. (2002), Python Software Foundation (2017a). The Python capabilities may be extended with specific modules dedicated to particular problems. The import of a module or single function is very simple. For the purposes of this research, two modules are used: (1) Pywin32 and (2) ArcPy. The methods available in the first module are applied to access COM (Component Object Module) objects and COM severs in Python code (Hammond and Robinson 2000; PythonCOM Documentation Index 2017). Brief descriptions of this package may be found in PythonCOM Documentation Index

(2017). The application of Python language with Pywin32 library is described in Dysarz (2018b). The second module, ArcPy, is fully integrated with ArcGIS (Zandbergen 2013). It enables access to all ArcToolbox methods. There are also mechanisms enabling access to objects of the vector layers such as points, lines and polygons.

Generation of flood hazard zones

The computations are configured in several steps shown in Fig. 4 and described below. The first is random choice of the annual discharge hydrographs for the estimation of maximum flows. The hydrographs are chosen from available data, namely years 1971–2014. The number of selected hydrographs depends on the length of the hydrological series assumed. There are four series assumed, namely 44 years, 40 years, 30 years and 20 years. Although the series shorter than 30 years are not recommended by Banasik et al. (2017), the shortest data scenarios are presented here to indicate what

Fig. 4 Scheme of the data and results processing
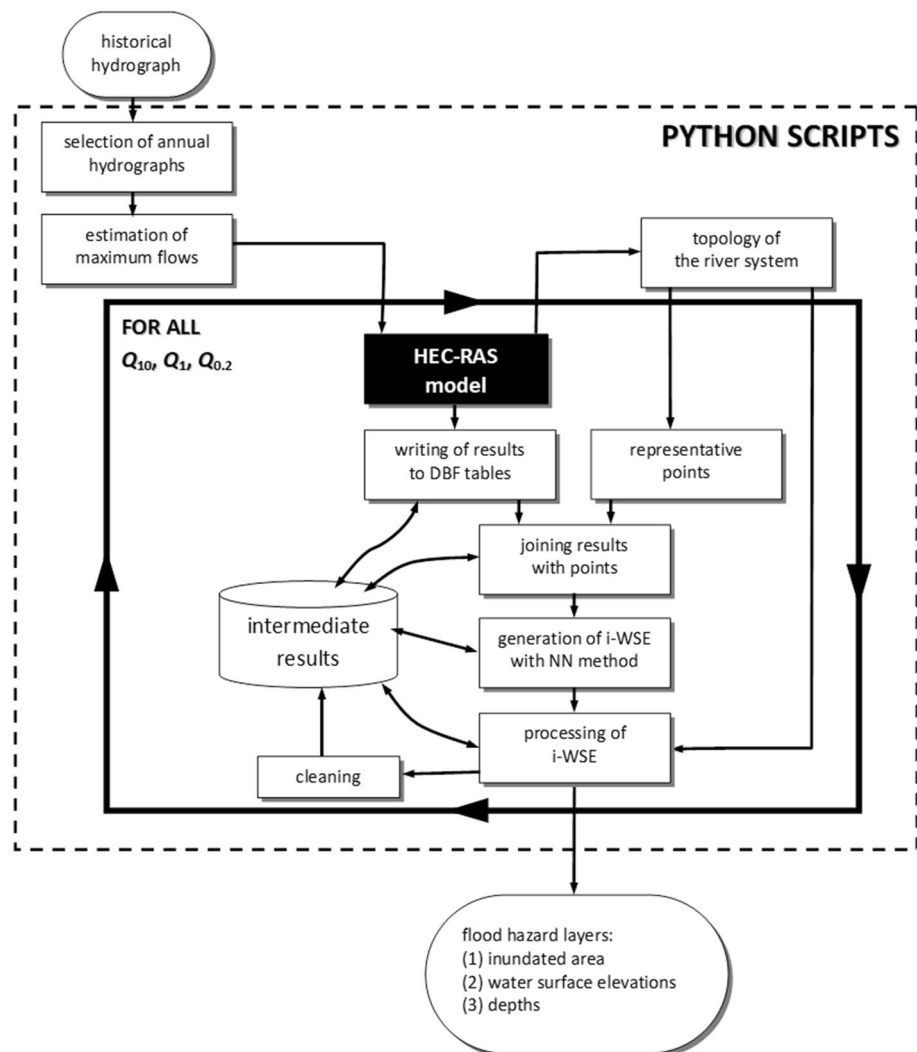
kind of uncertainty should be expected if this recommendation is ignored. Because the total number of years in this data set is 44, there is only one scenario tested for 44 years. The maximum flows estimated for this scenario are used as reference for other tests. For the rest of the tested series, the numbers of scenarios are given in Table 1. The numbers of tested scenarios were chosen arbitrarily taking into account the practical applicability of the presented methodology. In real cases, the number of simulations has to controlled due to the time limitations of project elaboration. As can be seen, the number of scenarios is inversely proportional to the length of the series. When the scenarios are selected, the maximum flows are estimated for each scenario. The quantiles and maximum likelihood methods described above are used for this purpose. The maximum flows used further are Q_{10}, Q_1 and $Q_{0.2}$ which correspond to 10-, 100- and 500-year floods.

The preliminary step of the simulations is reading of the river network topology from the model. Python scripting with the access to HEC-RAS Controller is applied for this purpose. The river centerline, the cross-section cut lines and bank lines are read and transformed into shapefiles. The last step is done with GIS functions available in the ArcPy module. From these layers, the points of results' location are generated. These are starting and ending vertexes of the cross sections as well as cutting points of cross sections with river centerline and banks. Finally, there are 5 representative points for each cross section.

After simulation, the results saved in DBF tables are joined with representative points generated earlier. A new point layer is generated for each simulation result. The points with water surface elevation values recorded are used for interpolation. The Natural Neighbor method from the Spatial Analyst extension of ArcGIS is used to generate the surface of spatially distributed water elevation. Then, the generated surfaces are processed for each simulation result. In the first step, the raster of differences between interpolated water surface elevations (i-WSE) and digital elevation model (DEM) is calculated. This is the raster of pseudo-depth including negative and positive values. The negative values are removed by simple reclassification, and the positive values are replaced with a single value, e.g., unity. Such a raster is transformed into polygons. In the next step, the polygonal objects are selected. The polygons intersecting with the river centerline are considered as inundated area. The inundated area is used as mask for the final processing of the water surface map and depth map generated after extraction of the previously computed difference between i-WSE and DEM. During this process, a number of

intermediate vector and raster layers are created. These are storage consuming processed data. At the end of the processing, the memory is cleaned. The final results for any simulation are three layers: (1) polygons of inundated area, (2) the raster of depths and (3) the raster of water surface elevations.

Applied methods for assessment of uncertainty

Two elements are of concern in the present research. These are the method applied for estimation of the maximum discharges and the length of the data series. The impact of these two elements on the obtained values of the discharges and the range of the inundation area are analyzed. The analysis is based on direct comparison of obtained results presented in the form of graphs or probabilistic maps.

The first results are analyzed as mean values obtained for particular series of data and deviations from these values. The values are maximum discharges and inundation areas for three flood events, namely 10-year, 100-year and 500-year floods. Such comparisons show the potential stability of the mean values. They also present the potential range of variability if the data series of a particular length is taken as the basis of the analysis.

The final results of uncertainty assessment are based on calculation of deviations from reference values. In any case, the reference values are those measures which are calculated for the total length of data series equaling 44 years. The deviations from reference are calculated for each of the analyzed flood events mentioned above. The final results are presented as the total range of variability, which seems to be the most suitable variable in the case of relatively small number of trials. If the number of trials is increased, the range is never smaller. It might be only greater. Hence, the range of variability in the case of a relatively small number of trials should be treated as the minimum, which may be only greater.

Results and discussion

Selected flood hazard maps

There are 25 data scenarios tested (Table 1). Taking into account the method and the scenario, the maximum flows are determined 39 times. The quantiles method was tested with all lengths of the data series. The maximum likelihood method was not tested with the 20-year scenario. The results of each maximum flow estimation are three discharges, namely the 10-year flood, 100-year flood and 500-year flood. Each discharge is used for hydraulic simulation, which gives 117 calculations of water surface profiles. For each profile, three layers are processed: the inundation zone, the water surface elevations and spatial distribution

Table 1 Numbers of scenarios tested for each series and method

Length of series (years)	44	40	30	20
Quantiles method	1	5	8	11
Maximum likelihood method	1	5	8	–

of depth. The first is a vector layer of polygon features. The last two are raster layers. The total number of processed layers is 351. An example is shown in Fig. 5. The presented maps cover only the area around the town of Wronki. These results were obtained with the quintile method and randomly chosen 20-year data series. Part (a) of Fig. 5 presents the inundation zone for the 10-year flood. The background is an orthophotomap of this area. The second map (Fig. 5b) shows the spatial distribution of the depth for the same flood. It is presented on the background of DEM. The gauge station is marked in both maps.

The obtained maps are then compared and analyzed. The results of these analyses are presented below.

Impact of observation period for each method

In Figs. 6 and 7, the impact of the data series length on the obtained discharges and inundation areas is presented. These analyses included three maximum flows: 10-year

(blue color), 100-year (red) and 500-year (green). In both figures, the continuous lines represent mean values obtained for particular series of tests characterized by length of the data series and method chosen for the estimation of maximum discharges. The dots denote particular results. The dashed lines mark the range of variability. In the graphs, the range of uncertainty is written as numbers expressed as a percentage of the mean value determined for a particular length of the data series.

In Fig. 6, the results for the quantiles method are presented. It may be seen that the uncertainty of discharge evaluation increases as the length of the data series is shortened (Fig. 6a). In the case of the shortest series, the uncertainty dramatically increases with the magnitude of the flow. Although the uncertainty also seems to increase with the magnitude of the flow in the cases of longer series, 30 and 40 years long, this effect is not so significant as it is in the case of 20-year-long series. In Fig. 6b, the uncertainty of the inundation area is presented. The dependence of the
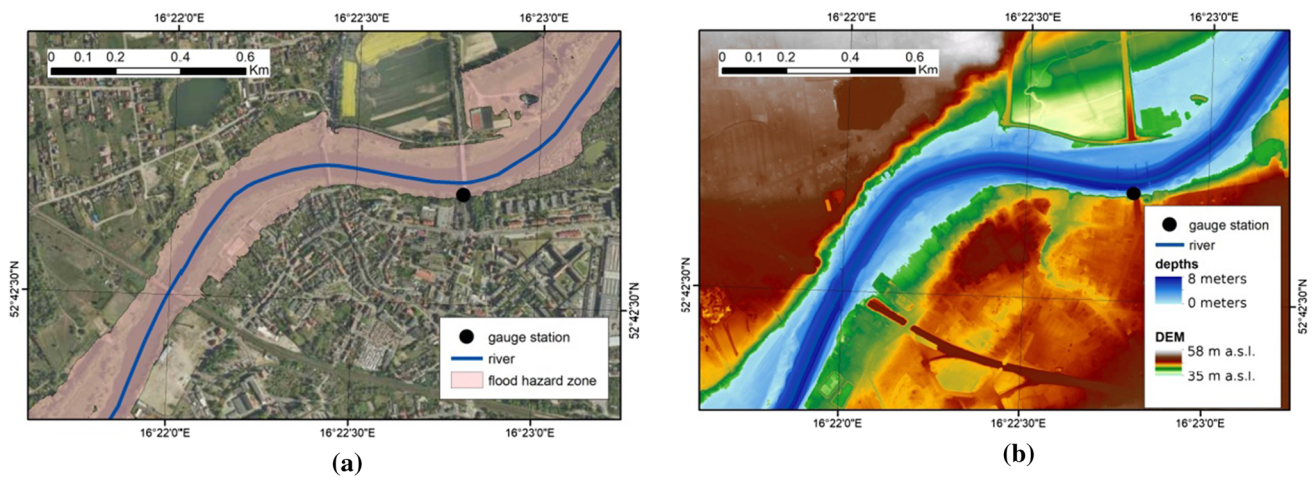

Fig. 5 Examples of flood hazard maps—quantiles method, 10-year flood: **a** inundation zone, **b** map of depths
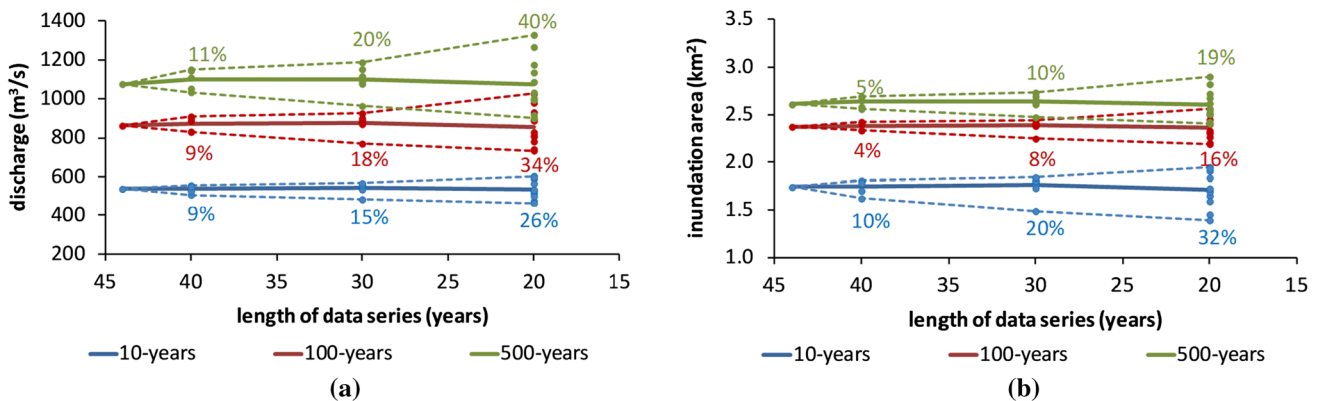

Fig. 6 Impact of the observation period for quantiles method: **a** variability of extreme flows, **b** variability of inundation area
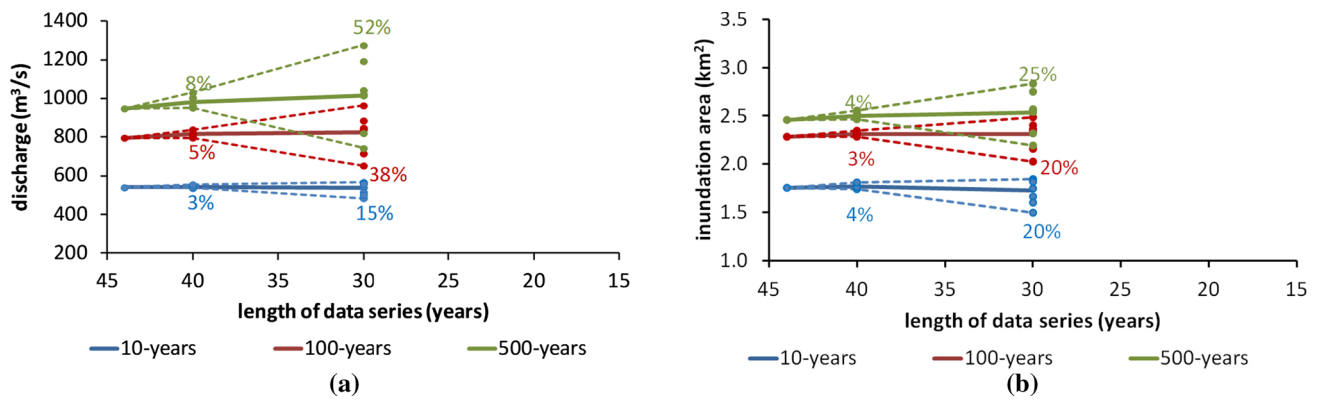
Fig. 7 Impact of the observation period for maximum likelihood method: **a** variability of extreme flows, **b** variability of inundation area

uncertainty on length of the data series is still visible. However, the relationship between the magnitude of the flow and the length of the series is changed. The greatest uncertainties are observed in the case of the lowest flows, 10-year floods. The impact of the terrain topography on the inundation area is the most important in the case of lower discharges.

In Fig. 7, the results obtained with the maximum likelihood method are presented. In this case, only the 40 and 30 year-series are tested, because the maximum flows were not determined for shorter series. It may be seen that this method gives smaller uncertainty of discharges for 40-year-long series (Fig. 7a). However, the uncertainty "jumps" to greater values than was observed previously as the length of the series is shortened to 30 years. The uncertainty of the inundation area is also smaller for longer series. For shorter series, the results are not so unique. It gives a greater range of variability for higher flows, 100-year flood and 500-year flood. At the same time the uncertainty for the lowers one, the 10-year flood, is smaller. A decrease in uncertainty with the increase in the flow magnitude is not seen in these results.

In Fig. 8, the comparison of the methods for the 100-year flood is presented. In part (a) we may see the changes of maximum flows, while part (b) shows variability of the inundation area. Both methods are presented as seen in Figs. 6 and 7, but colored polygons representing ranges of variability are added to indicate the convergence and discrepancy areas. The horizontal shift of the polygons is caused by different values of maximum flows obtained from each method. The comparison presents well the previous findings. The uncertainty of the maximum likelihood methods is smaller for longer series, but it dramatically increases as the series is shortened.

Propagation of uncertainty

The next graphs presented in Fig. 9 show the dependence of inundation area uncertainty on the discharge uncertainty. The axes represent values of difference between the results of the particular computations and results obtained for the reference scenario of 44 years long. These differences are expressed as the percentage of change with respect to the
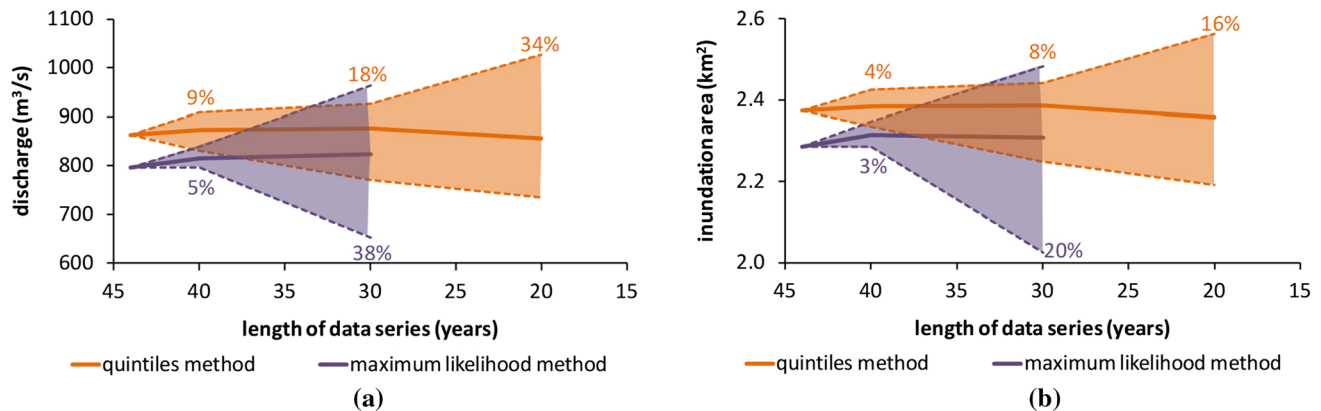

Fig. 8 Comparison of methods for 100-year flood: **a** variability of extreme flows, **b** variability of inundation area

reference results. The changes are positive and negative, which means that the obtained discharges and inundation areas could be greater or smaller than reference values. The gray dashed lines crossing at point (0, 0) represent the reference computations. The colored dots represent previous results. Figure 9a is composed from the results of the quantiles method, where Fig. 9b is prepared with the results of the maximum likelihood method.

The ranges of variability are marked with colored belts and proper values. In the case of the quantiles method (Fig. 9a), the 34% uncertainty of maximum flows obtained with the 20-year-long series gives 16% uncertainty in inundation areas. For the 30-year-long series, this relationships look as follows: 18% uncertainty of maximum flows gives 8% uncertainty of the inundation area. In the case of the 40-year-long series, the 9% uncertainty of maximum flows is related to 4% uncertainty of inundation areas. For the maximum likelihood method (Fig. 9b), the relations are as follows: 39% to 20% for 30-year-long series and 5% to 3% for 40-year-long series.

The results of analyzed tests are presented in Tables 2 and 3 in a slightly different form. In both tables, there are shown ranges of variability calculated as the percentage of

the reference values obtained for 44-year-long series of data. The first one presents the results of the quantiles method. In the second, the results obtained with the maximum likelihood method are shown. The tables include the ranges of variability of maximum flow calculation as well as the ranges of variability in determination of related inundation areas. All measures are estimated for each tested maximum flow, namely the 10-year flood, 100-year flood and 500-year flood. In the case of the quantiles method, the lengths of the data series are 40, 30 and 20 years. In the case of the maximum likelihood method, the tested lengths of data series are 40 and 30 years. The results of analysis of tendencies presented in Tables 2 and 3 are averaged over flood events as well as over the length of the data series. The first averages are shown in the last rows denoted as "EVENT AVERAGE." The processed results of the second kind are placed in the last columns of each table denoted as "SERIES AVERAGE."

The results presented in Tables 2, 3 show that the uncertainty measured in this way is a monotonically decreasing function of the length of the data series. For the case of maximum flows determined with quantiles method (Table 2), the range of variability increases from about 10% for a 40-year series to over 30% for a 20-year series. The range
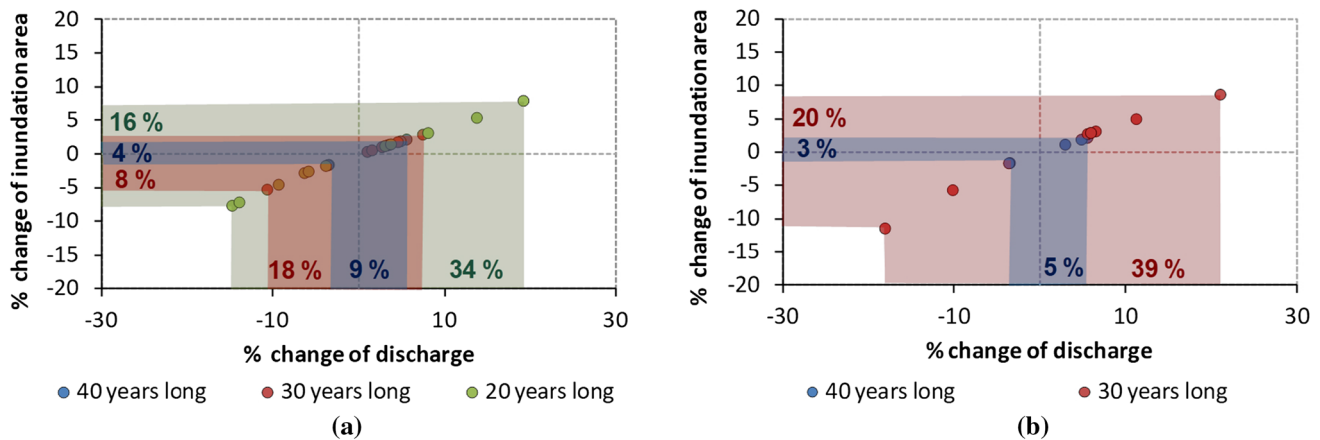

Fig. 9 Relation between uncertainty of discharge assessment and uncertainty of inundation area: **a** quantiles method, **b** maximum likelihood method

Table 2 Quantiles method: variability of max flows and related inundation areas with respect to length of the series

Length of series (years)	Range of variability (%)							
	10-Year flood		100-Year flood		500-Year flood		SERIES AVERAGE	
	Max flow	Inundation area	Max flow	Inundation area	Max flow	Inundation area	Max flow	Inundation area
40	8.92	10.47	9.15	3.82	10.84	5.13	9.64	6.47
30	15.53	20.13	18.11	8.18	20.81	9.97	18.15	12.76
20	26.03	31.80	33.89	15.60	39.83	18.99	33.25	22.13
EVENT AVERAGE	16.83	20.80	20.38	9.20	23.83	11.37		

Table 3 Maximum likelihood method: variability of max flows and related inundation areas with respect to length of the series

Length of series (years)	Range of variability (%)							
	10-Year flood		100-Year flood		500-Year flood		SERIES AVERAGE	
	Max flow	Inundation area	Max flow	Inundation area	Max flow	Inundation area	Max flow	Inundation area
40	3.39	3.76	5.43	2.75	9.03	4.14	5.95	3.55
30	15.37	19.89	39.10	20.06	56.17	26.05	36.88	22.00
EVENT AVERAGE	9.38	11.83	22.27	11.40	32.60	15.09		

of variability of the inundation area for the same method increases from almost 7% to over 20%. The results obtained with the second method, maximum likelihood (Table 3), show that the range of variability for maximum discharges increases from about 6% for the 40-year series to over 35% for the 30-year series. The range of the inundation areas for the same method varies from 4% to over 20%.

There are some interesting results reported in the scientific literature, which could be compared with the present research. Sun and co-authors (Sun et al. 2017) mentioned earlier analyzed only the uncertainty in determination of the maximum discharges. The study was conducted on two Chinese rivers with available observation periods of 70 years long. This study showed that the great increase in uncertainty is related to the increase in return period, which means the decrease in flood frequency. Sun and co-authors (Sun et al., 2017) analyzed six return periods ranging from 20 to 1000 years. In the first gauge station, Cuntan, they obtained uncertainty increasing with return period from 15 to 58%. In the second station, Pingshan, the uncertainty is also increasing with return period and varies in the range 11–45%.

In our example, this factor seems not to be so important. The results presented in Table 2, calculated for the quantiles method, show that the average uncertainty of maximum flow for 10-, 100- and 500-year events increases from 17 to 24%. In the case of the maximum likelihood method (Table 3), this uncertainty changes from 9 to 33%. In both cases, the uncertainty of discharges increases with the return period. If we focus on inundation area, the increase in uncertainty is not so obvious. In the cases of both methods, we can see that the uncertainty of a 100-year event may be lower than the uncertainty of a 10-year event. The only reason may be the special features of the local topography, e.g., local dikes, storage areas, etc. Such elements are present in the analyzed region.

Sun and co-authors (Sun et al. 2017) also noted that the method chosen for estimation of probability distribution is responsible for 3–8% of uncertainty. In the cases analyzed in this paper, the dependence of flow uncertainty on the method chosen is not so simple, because it also depends on the length of the data series. We may see that this uncertainty for the 40-year-long series is greater in the case of the quantiles method (Table 2). It changes from 9 to 11%. In the case of the likelihood method (Table 3) and the same length of the data series, these values are in the range 3–9%. However, this tendency is inverted if the length of the data series is shortened. For the 30-year series, the quantiles method gives 15–21%, when the second method produces results with 15–57% uncertainty. This variation is lower if we compare the inundation area. For the 40-year series, we obtained uncertainty of 8–20% for the quantiles method (Table 2) and 20–26% for the maximum likelihood method (Table 3).

Merz and Thieken (2005) also presented interesting results for the Rhine River obtained with a 120-year observation series. They observed that the uncertainty of discharges calculated for the 100-year return period is about 11%. In our case, this uncertainty depends on the length of the series. For the quantiles method (Table 2), it varies from 9 to 34% for a series of 40–20 years. For the maximum likelihood method (Table 3), it varies from 5 to 39% for a series 40–30 years. What is interesting, the uncertainty of the inundation area is smaller. In the first example it varies in the range 4–16%, while in the second the range of variability is 3–20%.

As can be seen, the present results are compatible with those presented in the literature, but not exactly the same. The specific features of the local river system play an important role. Additionally, the transition of the discharge uncertainty to uncertainty of the inundation area is not straightforward. The local topography could disturb or even change the tendencies of uncertainty growth or decrease.

Examples of probabilistic maps

Selected probabilistic maps of inundation area are shown in Figs. 10 and 11. The maps are composed for tests with 100-year flood flow. The first figure is composed for the quantiles method, the second for the maximum likelihood method. In both cases, the total zone is presented on the left and the selected piece of the area is zoomed and shown on the right. The colors applied indicate the relative frequency of inundation. The blue and green are rarely flooded cells, while the yellow and red are frequently flooded areas. The
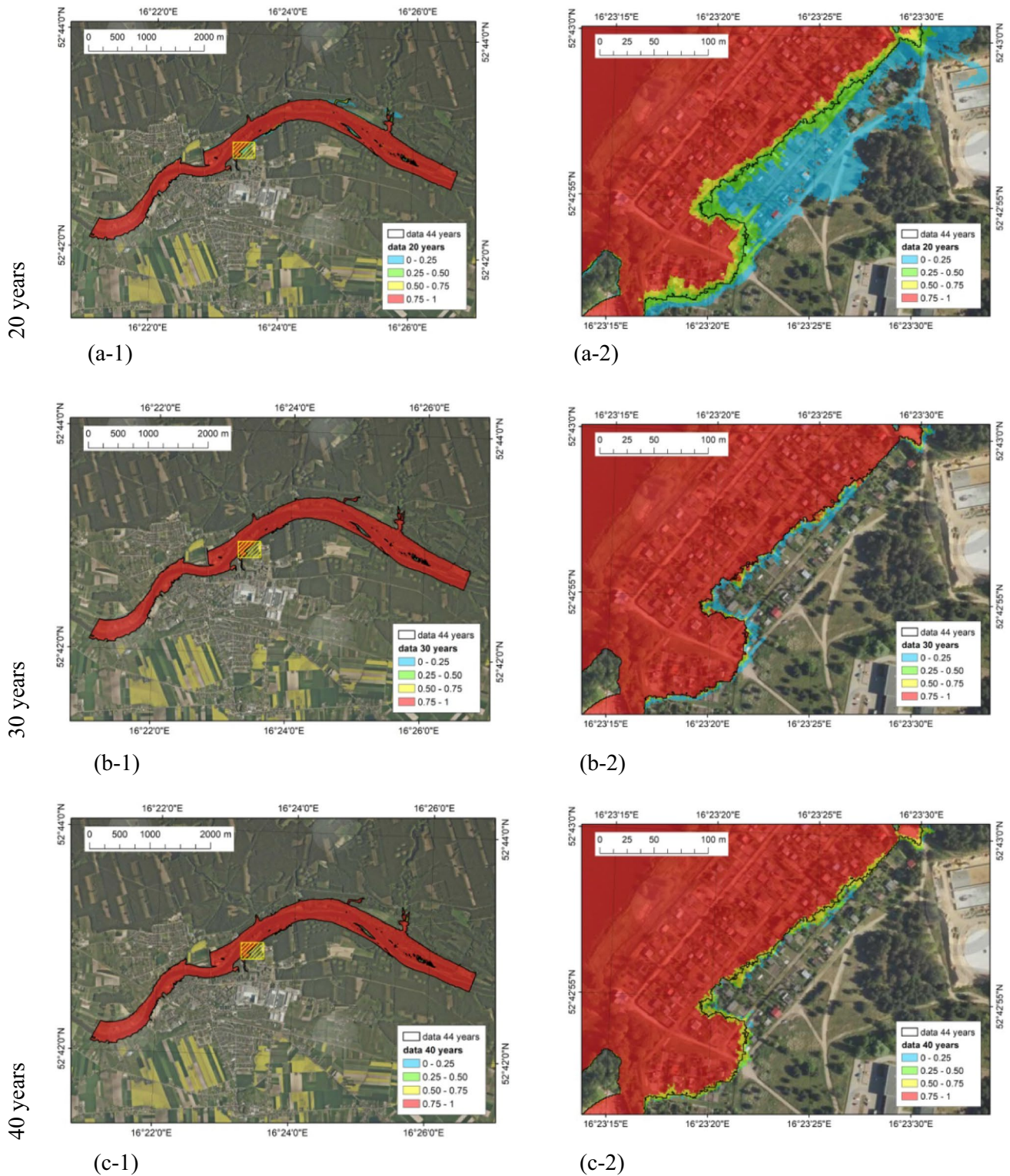
Fig. 10 Relative frequency of flood inundation for 100-year flood and quantiles method

background map is an orthophotomap, which allows us to see the difference of inundation in relation to the town of Wronki infrastructure, e.g., building and roads.

It is clearly seen as the area of probable inundation changes with the length of the time series. The biggest differences are seen for the 20-year-long series for the
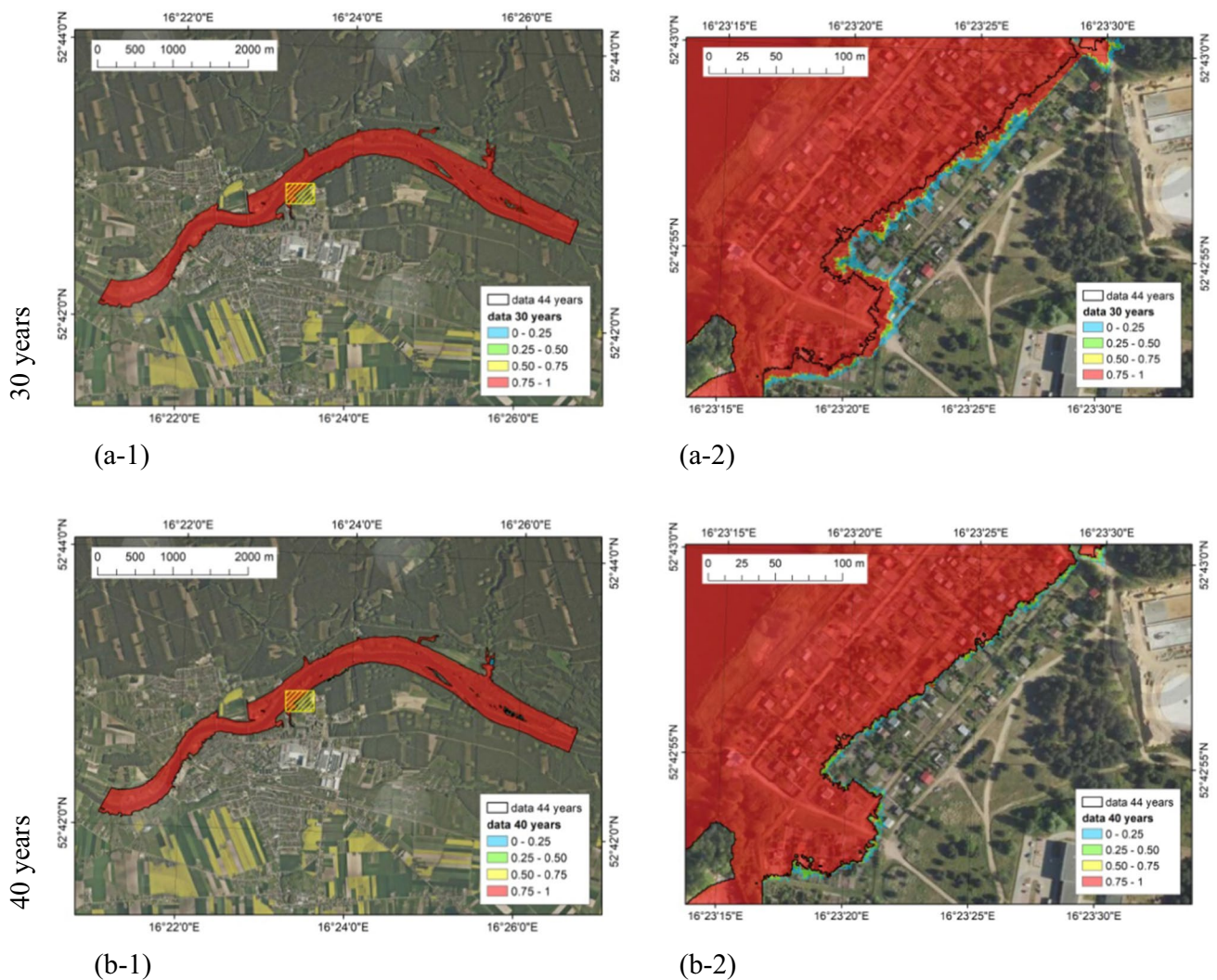
30 years
(a-1)   (a-2)

40 years
(b-1)   (b-2)

Fig. 11 Relative frequency of flood inundation for 100-year flood and maximum likelihood method

quantiles method (Fig. 10a-2). In this case, the inundation area may extend beyond the reference zone several dozens of meters in the shown region. As can be seen, the region is quite populated and such a difference significantly affects the risk of flooding. In the cases of longer data series (Fig. 10b-2, c-2), the differences in inundation area are smaller but their values vary from a few to over a dozen meters, which may be important in such urbanized areas as those shown in the quoted map.

In the case of the maximum likelihood method, the spread of the inundation area seems to be similar (Fig. 11a-2, b-2). The differences presented in Figs. 8 and 9 are not visible on this scale. However, it is clearly visible that the application of the maximum likelihood method does not protect against the uncertainty related to the length of the data series.

Conclusions

In this paper, the uncertainty assessment of the flood hazard analysis is presented. The chosen case study is the selected reach of the Warta river near the Wronki gauge station. The assessment is performed taking into account two main sources of uncertainty, namely the length of the data series and the method for evaluation of extreme flows. Hydraulic modeling and GIS processing is used to transform the calculated maximum flows into flood inundation zones, water surface and depth maps. The sophisticated scripting tools enabled effective processing of data and results. The obtained results are presented in informative graphs and maps.

The conducted research allowed us to formulate the following conclusions:

- The length of the annual maxima series has the greatest impact on the estimation of design flow. In the case of the quantiles method, the range of variability averaged over 10-, 100- and 500-year flood events changes from about 10% to over 33% with the length of the data series changing from 40 to 20 years. In the second case, the maximum likelihood method, this range varies from about 6% to slightly less than 37% with the tested length of data series 40 and 30 years.
- The impact of the method for estimation of distribution parameters is less important, but still significant. In the case of the quantiles method, the range of variability averaged over the length of the data series changes from about 17% to 24% with the return periods 10-, 100- and 500-years. The results obtained with the maximum likelihood method show variability from about 9% to less than 33% with the same return periods.
- Uncertainty associated with the calculation of design flows propagates into the size of flood hazard zones, but the main tendencies of variability increase or decrease are not kept during this transition. In both methods, the variability in determination of inundation area averaged over the length of the data series decreases with the return period when analyzed from the 10-year flood to the 100-year flood. But this tendency increases if analyzed from the 100-year flood to the 500-year flood.

In general, the presented examples indicate the role and importance of uncertainty in flood hazard assessment. The results proved that the inundation area may be overestimated or underestimated with insufficient data series. The choice of the proper method for evaluation of extreme flows is also crucial. Besides the problem of statistical parameters fitting in first-order regression, the resistance to uncertainties to the insufficient data series is also important. Although the main results present the inundation areas, it is supposed that the impact of these two factors analyzed is similar for the spatial distribution of the depths.

The challenge of the uncertainty assessment in determination of flood hazard maps is the number of necessary trials. In this research, only the minimum number of scenarios was tested. However, even this limited approach indicated a strong need for automation of the river flow modeling and spatial data processing. The response for this requirement is application of the scripting language for management of the whole computational process.

Taking into account the number of applied approaches, methods and data used in the implementation of the Flood Directive, it should be noted that flood hazard analyses in the different EU states are inconsistent. As the European integration covers new areas of economic and environmental management, the need for one uniform methodology in this area of technical activity becomes more and more

crucial. The need for development of robust methodology for uncertainty assessment is obvious. Although the number of uncertainty sources is huge in the case of flood modeling, the length of the data series and the method for evaluation of extreme flows are among the most important.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Albano R, Crăciun I, Mancusi L, Sole A, Ozunu A (2017) Flood damage assessment and uncertainty analysis: the case study of 2006 flood in Ilisua basin in Romania. Carpathian J Earth Environ Sci 12(2):335–346

Arseni M, Roşu A, Bocăneală C, Constantin D-E, Georgescu LP (2017) Flood hazard monitoring using GIS and remote sensing observations. Carpathian J Earth Environ Sci 12(2):329–334

Banasik K, Wałęga A, Weglarczyk S, Więzik B (2017) Update of methodology for calculation of maximum discharges and maximum precipitation with determined exceedance probability for controlled and uncontrolled watersheds and calibration of rainfall—runoff models (in Polish). Association of Polish Hydrologists, Warsaw. http://www.kzgw.gov.pl/files/zam-pub/20170402-przeglad-i-aktualizacja-map/zal-2-do-OPZ-ujednolicony-1.pdf. Accessed 22nd Dec 2018

Bates PD, Pappenberger F, Romanowicz RJ (2014) Uncertainty in flood inundation modelling. In: Beven K, Hall J (eds) Applied uncertainty analysis for flood risk management. Imperial Collage Press, London, pp 232–269

Beven KJ, Hall J (eds) (2014) Applied uncertainty analysis for flood risk management. Imperial College Press, London

Brunner GW (2016a) HEC-RAS river analysis system hydraulic reference manual. US Army Corps of Engineers. Report No. CPD-69; Hydrologic Engineering Center (HEC), Davis

Brunner GW (2016b) HEC-RAS river analysis system user's manual version 5.0. US Army Corps of Engineers. Report No. CPD-68; Hydrologic Engineering Center (HEC), Davis

Calenda G, Mancini CP, Volpi E (2009) Selection of the probabilistic model of extreme floods: the case of the River Tiber in Rome. J Hydrol 371(1–4):1–11

Cameron T, Ackerman PE (2012) HEC-GeoRAS GIS tools for support of HEC-RAS using ArcGIS user's manual. US Army Corps of Engineers, Institute for Water Resources, Hydroogic Engineering Center (HEC). http://www.hec.usace.army.mil/software/hec-georas/documentation/HEC-GeoRAS_43_Users_Manual.pdf. Accessed 7th July 2018

Chow VT, Maidment DR, Mays LW (1988) Applied hydrology. McGraw-Hill Book Company, New Year

Cook A, Merwade V (2009) Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping. J Hydrol 377(1–2):131–142

Di Baldassarre G, Montanari A (2009) Uncertainty in river discharge observations: a quantitative analysis. Hydrol Earth Syst Sci 13(6):913–921

Di Baldassarre G, Montanari A, Lins H, Koutsoyiannis D, Brandimarte L, Blöschl G (2010) Flood fatalities in Africa: from diagnosis to mitigation. Geophys Res Lett 37(22):L22402

Di Baldassarre G, Laio F, Montanari A (2012) Effect of observation errors on the uncertainty of design floods. Phys Chem Earth Parts A/B/C 42:85–90

Dimitriadis P, Tegos A, Oikonomou A, Pagana V, Koukouvinos A, Mamassis N, Efstratiadis A (2016) Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping. J Hydrol 534:478–492

Docan DC (2016) Learning ArcGIS for desktop. Packt Publishing. https://www.packtpub.com/application-development/learning-arcgis-desktop. Accessed 7th July 2018

Downey A, Elkner J, Meyers Ch (2002) How to think like a computer scientist. Learning with Python. Green Tea Press, Wellesley. http://www.greenteapress.com/thinkpython/thinkCSpy.pdf. Accessed 8th Nov 2017

Dysarz T (2018a) Development of RiverBox—an ArcGIS toolbox for river bathymetry reconstruction. Water 10(9):1266

Dysarz T (2018b) Application of Python scripting techniques for control and automation of HEC-RAS simulations. Water 10(10):1382

Dysarz T, Wicher-Dysarz J, Sojka M (2015) Assessment of the impact of new investments on flood hazard-study case: the bridge on the Warta River near Wronki. Water 7:5752–5767. http://www.mdpi.com/2073-4441/7/10/5752. Accessed 23rd July 2018

Engeland K, Steinsland I, Johansen SS, Petersen-Øverleir A, Kolberg S (2016) Effects of uncertainties in hydrological modelling. A case study of a mountainous catchment in Southern Norway. J Hydrol 536:147–160

European Parliament (2007) Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. http://data.europa.eu/eli/dir/2007/60/oj

Ewemoje TA, Ewemooje OS (2011) Best distribution and plotting positions of daily maximum flood estimation at Ona River in Ogun-oshun river basin, Nigeria. Agric Eng Int CIGR J 13(3), Manuscript No. 1380

Gąsiorowski D (2013) Analysis of floodplain inundation using 2D nonlinear diffusive wave equation solved with splitting technique. Acta Geophys 61(3):668–689

Gilles D, Young N, Schroeder H, Piotrowski J, Chang YJ (2012) Inundation mapping initiatives of the Iowa Flood Center: statewide coverage and detailed urban flooding analysis. Water 4(1):85–106

Goodell Ch (2014) Breaking HEC-RAS code. A user's guide to automating HEC-RA. h2ls, Portland

Griffis VW, Stedinger JR (2007) The use of GLS regression in regional hydrologic analyses. J Hydrol 344(1–2):82–95

Hammond M, Robinson A (2000) Python programming on Win32. O'Reilly Media Inc, Sebastopol

Hirsh RM, Stedinger JR (1987) Plotting position for historical floods and their precision. Water Resour Res 23(4):715–727

Jung Y, Kim D, Kim D, Kim M, Lee SO (2014) Simplified flood inundation mapping based on flood elevation-discharge rating curves using satellite images in gauged watersheds. Water 6:1280–1299

Kolerski T (2018) Mathematical modeling of ice dynamics as a decision support tool in river engineering. Water 10(9):1241

Kundzewicz ZW, Stoffel M, Wyżga B, Ruiz-Villanueva V, Niedźwiedź T, Kaczka R, Ballesteros-Cánovas JA, Pińskwar I, Łupikasza E, Zawiejska J, Mikuś P, Choryński A, Hajdukiewicz H, Spyt B,

Janecka K (2017) Changes of flood risk on the northern foothills of the Tatra Mountains. Acta Geophys 65:799–807

Laio F, Ganora D, Claps P, Galeati G (2011) Spatially smooth regional estimation of the flood frequency curve (with uncertainty). J Hydrol 408(1–2):67–77

Laks I, Sojka M, Walczak Z, Wróżyński R (2017) Possibilities of using low quality digital elevation models of floodplains in hydraulic numerical models. Water 9(4):283

Law M, Collins A (2018) Getting to know ArcGIS desktop, 5th edn. Esri Press, Redlands

Liu Z, Merwade V (2018) Accounting for model structure, parameter and input forcing uncertainty in flood inundation modeling using Bayesian model averaging. J Hydrol 565:138–149

Merz B, Thieken AH (2005) Separating natural and epistemic uncertainty in flood frequency analysis. J Hydrol 309(1–4):114–132

Nones M (2015) Implementation of the floods directive in selected EU member states. Water Environ J 29(3):412–418

Nones M (2017) Flood hazard maps in the European context. Water Int. 42(3):324–332

Pappenberger F, Beven KJ, Ratto M, Matgen P (2008) Multi-method global sensitivity analysis of flood inundation models. Adv Water Resour 31(1):1–14

Parkes B, Demeritt D (2016) Defining the hundred year flood: a Bayesian approach for using historic data to reduce uncertainty in flood frequency estimates. J Hydrol 540:1189–1208

Python Software Foundation (2017a) Python 2.7.14 documentation. https://docs.python.org/2/index.html. Accessed 8th Nov 2017

PythonCOM Documentation Index (2017) Python and COM. Blowing the rest away! http://docs.activestate.com/activepython/2.4/pywin32/html/com/win32com/HTML/docindex.html. Accessed 8 Nov 2017

Refsgaard JC, Storm B (1990) Construction, calibration and validation of hydrological models. Distributed hydrological modeling. Springer, Dordrecht, pp 41–54

Romanowicz R, Beven K (2003) Estimation of flood inundation probabilities as conditioned on event inundation maps. Water Resour Res 39(3):1073

Sampson CC, Fewtrell TJ, Duncan A, Shaad K, Horritt MS, Bates PD (2012) Use of terrestrial laser scanning data to drive decimetric resolution urban inundation models. Adv Water Resour 41:1–17

Schendel T, Thongwichian R (2015) Flood frequency analysis: confidence interval estimation by test inversion bootstrapping. Adv Water Resour 83:1–9

Schendel T, Thongwichian R (2017) Considering historical flood events in flood frequency analysis: is it worth the effort? Adv Water Resour 105:144–153

Serago JM, Vogel RM (2018) Parsimonious nonstationary flood frequency analysis. Adv Water Resour 112:1–16

Serinaldi F, Kilsby CG (2015) Stationarity is undead: uncertainty dominates the distribution of extremes. Adv Water Resour 77:17–36

Sibson R (1981) A brief description of nearest neighbor interpolation. Interpolating multivariate data. Wiley, New York

Sun H, Jiang T, Jing C, Su B, Wang G (2017) Uncertainty analysis of hydrological return period estimation, taking the upper Yangtze River as an example. Hydrol Earth Syst Sci Discuss 1–26. https://doi.org/10.5194/hess-2016-566

Szydłowski M, Szpakowski W, Zima P (2013) Numerical simulation of catastrophic flood: the case study of hypothetical failure of the Bielkowo hydro-power plant reservoir. Acta Geophys 61(5):1229–1245

Teng J, Jakeman AJ, Vaze J, Croke BF, Dutta D, Kim S (2017) Flood inundation modelling: a review of methods, recent advances and uncertainty analysis. Environ Model Softw 90:201–216

Van Alpen J, Passchier R (2007) Atlas of flood maps, examples from 19 European countries, USA and Japan. Ministry of Transport, Public Works and Water Management, The Hague

Van der Graaf SC (2016) Natural Neighbour Kriging and its potential for quality mapping and grid design, M.Sc. thesis, Delft University of Technology. https://repository.tudelft.nl/. Accessed 7th July 2018

Walczak Z, Sojka M, Laks I (2013) Assessment of mapping of embankments and control structure on digital elevation model based upon Majdany polder. Rocz Ochr Śr 15:2711–2724

Yuan F, Zhao C, Jiang Y, Ren L, Shan H, Zhang L, Shen H (2017) Evaluation on uncertainty sources in projecting hydrological changes over the Xijiang River basin in South China. J Hydrol 554:434–450

Zandbergen PA (2013) Python scripting for ArcGIS. Esri Press, Redlands