Bayesian Data-Driven approach enhances synthetic flood loss models

Published Version:
Bayesian Data-Driven approach enhances synthetic flood loss models / Sairam N.; Schroter K.; Carisi F.; Wagenaar D.; Domeneghetti A.; Molinari D.; Brill F.; Priest S.; Viavattene C.; Merz B.; Kreibich H.. - In: ENVIRONMENTAL MODELLING & SOFTWARE. - ISSN 1364-8152. - ELETTRONICO. - 132:(2020), pp. 104798.1-104798.10. [10.1016/j.envsoft.2020.104798]

Availability:
This version is available at: https://hdl.handle.net/11585/778745 since: 2020-11-08

Published:
DOI: http://doi.org/10.1016/j.envsoft.2020.104798

Nivedita Sairam, Kai Schröter, Francesca Carisi, Dennis Wagenaar, Alessio Domeneghetti, Daniela Molinari, Fabio Brill, Sally Priest, Christophe Viavattene, Bruno Merz, Heidi Kreibich, Bayesian Data-Driven approach enhances synthetic flood loss models, Environmental Modelling & Software, Volume 132, 2020, 104798, ISSN 1364-8152

The final published version is available online at:

https://doi.org/10.1016/j.envsoft.2020.104798

Bayesian Data-Driven Approach Enhances Synthetic Flood Loss Models.

Nivedita Sairam[1,2], Kai Schröter[1], Francesca Carisi[3], Dennis Wagenaar[4], Alessio Domeneghetti[3], Daniela Molinari[5], Fabio Brill[1,7], Sally Priest[6], Christophe Viavattene[6], Bruno Merz[1,7], Heidi Kreibich[1]

[1]GFZ German Research Centre for Geosciences, Section 4.4. Hydrology, Potsdam, Germany.
[2]Geography Department, Humboldt University, Berlin, Germany.
[3]DICAM-Department of Civil, Chemical, Environmental and Materials Engineering, Alma Mater Studiorum - University of Bologna (Bologna, Italy)
[4]Deltares, Delft, The Netherlands
[5]Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano (Italy)
[6]Flood Hazard Research Centre, Middlesex University, The Burroughs, Hendon, London, UK.
[7]Institute for Environmental Sciences and Geography, University of Potsdam, Germany

Key Points

1. Bayesian Data-Driven approach integrates knowledge from the vast compendium of established synthetic models with empirical loss data.
2. This approach improves accuracy and quantifies reliability of synthetic flood loss models using local empirical data.
3. Continuous integration of empirical data from multiple flood events, using Bayesian Data-Driven approach improves loss predictions for a potential future event.

Abstract

Flood loss estimation models are developed using synthetic or empirical approaches. The synthetic approach consists of what-if scenarios developed by experts. The empirical models are based on statistical analysis of empirical loss data. In this study, we propose a novel Bayesian Data-Driven approach to enhance established synthetic models using available empirical data from recorded events. For five case studies in Western Europe, the resulting Bayesian Data-Driven Synthetic (BDDS) model enhances synthetic model predictions by reducing the prediction errors and quantifying the uncertainty and reliability of loss predictions for post-event scenarios and future events. The performance of the BDDS model for a potential future event is improved by integration of empirical data once a new flood event affects the region. The BDDS model, therefore, has high potential for combining established synthetic models with local empirical loss data to provide accurate and reliable flood loss predictions for quantifying future risk.

1. Introduction

Due to changing climate and increased settlements and assets in the flood plains, risk to life and property due to flooding is rising (Barredo 2009, Merz et al. 2012, Domeneghetti et al. 2015). Decisions concerning Flood Risk Management (FRM) focusing on new flood defense schemes and resilience initiatives are generally based on risk assessment encompassing of future hazard scenarios and the resulting damages. Models focusing on the hazard components (hydrology and hydraulics) are constantly being developed and improved by the research community, and are outside the scope of this paper; especially, the integration of physics-based models with Machine Learning algorithms have led to the development of high-resolution hazard maps (Teng et al. 2017, da Costa et al. 2019). In addition to flood hazard modelling, accounting for flood damage processes is crucial to predict losses. Flood damage processes are modelled using loss models, also called as vulnerability functions (Ward et al. 2019). Flood loss models are an essential component of the risk chain as they quantify flood risk in terms of economic losses (Merz et al., 2010). Flood loss models are generally developed using two approaches: 1. Synthetic or Engineering functions, 2. Empirical modelling. Synthetic models use expert opinions or engineering solutions that result in a set of What-If scenarios to estimate flood losses. They are not based on statistical analysis of observed data (Penning-Rowsell and Chatterton, 1977). One of the major advantages of synthetic loss models is their non-dependency on empirical data. However, the development of detailed damage scenarios covering all damage possibilities and building characteristics requires high effort (Smith, 1994). Since these models are synthesized based on a variety of data sources, such as expert knowledge and technical papers, the advantage is that these models are more generalized and lead to higher levels of standardization compared to empirical models and therefore are more suited to being used for actions that require accountability, such as investment decision-making (Smith, 1994; Merz et al. 2010; Amadio et al., 2019). For practical applications, the outputs from the synthetic models are required to capture the observed damage processes. However, except in very few models such as the INSYDE (Dottori et al. 2019), the empirical loss values do not constitute the model development.

Empirical models are developed based on real damage information observed from past events and hence, require large amounts of high-quality detailed data on flood damages and the damage-influencing factors, such as water depth (Merz et al. 2010, Smith, 1994). These models aim to represent the relationship between flood damage and its influencing factors using patterns that occurred in the past events. The empirical models may be based on data from a single event (localized model) or cumulative data from multiple events (generalized model). Flood loss models purely based on localized empirical datasets are unable to reliably predict building damages for other events (Wagenaar et al. 2018). In contrast, generalized models (e.g. Bayesian Network, multi-level parameterization) based on data from multiple events cover a wider range of damage processes and perform better for new events (Wagenaar et al. 2018, Sairam et al. 2019). As empirical models are based on real damage data, it is expected that they capture the observed damage processes and are less prone to surprises (Merz et al. 2015). However, an important disadvantage is their requirement for detailed damage surveys. These are often expensive and time consuming. Survey campaigns that are conducted after extreme events may result in a large sample of respondents that reported damage. However, in the case of surveys conducted after small localized events, the resulting datasets are often insufficient to model different damage processes.

Owing to lack of detailed object-level damage data, only a few studies have validated the flood loss models against observed loss estimates (Gerl et al. 2016; Amadio et al., 2019). An advantage of the empirical approach is the possibility to use a part of the empirical data for validation during model development. However, since synthetic models are generally developed when empirical data is unavailable, both calibration and validation of synthetic models remain a challenge. Both synthetic and empirical flood loss models may be deterministic or probabilistic. More than 95% of the state-of-the-art flood loss models are deterministic (Gerl et al. 2016).

Deterministic models result in one damage estimate based on the influencing factors. On the other hand, probabilistic models provide a distribution of losses. In reality, there exists variability in damage predictions given by the loss model based on the influencing factors. This may be due to the inherent stochastic nature of damage processes and other reasons such as uncertainty in empirical data, model structure and missing influencing factors in the model (Schröter et al. 2014, Winter et al. 2018). Decision makers and administrators are required to consider thoroughly the reliability of the flood loss models, in order to base FRM decisions and investments on the loss predictions. Hence, flood loss models should provide loss predictions along with an estimate of their uncertainty and reliability. A probabilistic flood loss model estimates the probability of occurrence of all possible loss scenarios for each object and results in a distribution of predicted losses. Probabilistic models potentially account for all sources of uncertainty in model parameters, structure and variability in the modelled processes based on observed data and assumptions concerning damage processes. Hence, there is an increasing interest in developing probabilistic approaches for flood loss modelling (Schröter et al. 2014, Wagenaar et al. 2018, Rözer et al. 2019, Lüdtke et al. 2019). In the presence of large detailed empirical datasets, advanced approaches for the development of probabilistic loss models are given by Wagenaar et al. (2018) and Rözer et al. (2019). Thus, another advantage of the empirical approach is the possibility to develop probabilistic models whose reliability can be determined. Since the synthetic models are not fitted to observed losses during development, they are commonly not calibrated. Hence, it is impossible to estimate the reliability of the synthetic model without validating the model against empirical loss data (Zischg et al. 2018).

We propose to combine the empirical and synthetic approaches to harness advantages of both concepts. To this end, we use relevant empirical loss data for enhancing the synthetic model predictions. The objective of this study is to propose and validate a Bayesian Data-Driven approach that calibrates the predictions of existing synthetic flood loss models using relevant empirical loss data at the object-level (residential buildings), within a probabilistic framework. The resulting flood loss estimation model is a Bayesian Data-Driven Synthetic (BDDS) Model. The BDDS model associates probability distributions with synthetic model outputs and can explain variability across households due to characteristics, which are not taken into account by the synthetic loss model. The BDDS model requires a synthetic model and local empirical data to calibrate the model for that region. The synthetic model can refer to any spatial scale (regional, national, continental). The BDDS model is aimed at enhancing the synthetic loss model by providing truly probabilistic loss predictions that are sharp (narrow width of distribution of predictions), calibrated and reliable for both central values and dispersion.

The BDDS model is tested for improvement in predictive capability compared to the standard national synthetic model, based on case studies from four countries in Western Europe – UK, Netherlands, Italy and Germany. We develop the BDDS model for residential buildings using the loss predictions from the synthetic flood loss models and empirical loss data from one or several (if available) flood events from the specific case study regions. Moreover, the BDDS model allows integrating synthetic model predictions with a continuous collection of empirical data after each flood event, in order to enhance prediction of flood losses due to potential flood events that may occur in the future.

The paper is organized as follows: Section 2 explains the Methods and Data including setting up the framework for BDDS model (2.1), BDDS model construction (2.2) and metrics for assessing model performances (2.3); explanation of case studies, object-level empirical data and the synthetic models used in the study (2.4). Results including damage prediction for post-event scenarios and future events are reported and discussed in Section 3. Section 4 includes concluding points focusing on implementation of the model, scope for future work and software availability.

2. Methods and Data

2.1. <u>Setting up the framework for BDDS model:</u>

The BDDS model describes the relationship between empirical losses and their corresponding deterministic loss predictions from synthetic models by means of a full Bayesian approach. The parameters of the BDDS model are indicators pertaining to the deviation between the synthetic model predictions and empirical observations. Also, the full joint posterior probability distribution of the BDDS model parameters can be obtained along with the predictive distribution of flood losses given the synthetic model and empirical losses from events that occurred in the region. From the credibility intervals of the predictive distributions, it is possible to estimate the uncertainty in the flood loss predictions.

The BDDS model is based on the premise that the empirical losses and synthetic loss predictions may be seen as components of a statistical model, in which the synthetic loss predictions are considered as exogenous variables (one that is determined outside the model, and imposed on the model) that are used to determine the observed losses. The BDDS model estimates losses using a linear function with empirical loss as the dependent variable regressed against the synthetic loss prediction. We assume that the BDDS model is identifiable for households within a region: i.e., the damage processes that occur in households belonging to one region are the same. Hence, the BDDS model assumes a single set of parameters for each region.

In order to make the loss predictions comparable across the different case studies, we use relative loss to buildings, $rloss$, which is the ratio of absolute building loss to its total reconstruction value in the respective currencies, at the time of the event (Elmer et al., 2010). The rloss values are between 0 and 1, where 0 indicates no damage and 1 indicates complete damage, requiring reconstruction of the building. The BDDS model is given in 1.

$$\widetilde{rloss} | rloss\_syn \sim Beta(\alpha, \beta) \qquad \text{Equation - 1}$$
$$\alpha = \mu \times \varphi$$
$$\beta = (1 - \mu) \times \varphi$$
$$\mu = inv\ logit\ (\lambda \times rloss\_syn + \varepsilon)$$

In this model definition, the observed rloss is represented as $\widetilde{rloss}$ and the rloss predictions from synthetic model is represented as $rloss\_syn$. Since the observed losses are not included in the synthetic model development, the BDDS model definition uses a set of parameters to alter the synthetic model predictions to agree with the observations. $\widetilde{rloss}$ is modelled as a beta distribution with logit transformation, since, unbounded distributions might result in implausible values for $\widetilde{rloss}$ (Rözer et al. 2019). The beta distribution holds two parameters α and β which are algebraically determined using location parameter μ and variance parameter φ. μ is a function of the synthetic rloss predictions ($rloss\_syn$) with parameters slope (λ), intercept (ε). These parameters are estimated by modelling the deviations of the empirical loss data from the synthetic model predictions using Markov Chain Monte Carlo (MCMC) sampling implemented using STAN (Carpenter et al. 2017). We initially provide priors that describe our general belief about the distribution of the parameters. For example, φ is required to be positive and hence given a un-informative generic prior, $gamma(0.01{,}0.01)$. We provide un-informative generic priors to λ and ε to determine the parameterization of BDDS model based on the availability of evidence from empirical loss data. The MCMC sampling creates a large number of replications of these parameters explaining the data generation process of flood losses. This results in approximate posterior distributions of $\widetilde{rloss}$.

2.2. <u>BDDS model construction</u>

In reality, we are particularly interested in the capability of the BDDS model to estimate expected flood losses to buildings after an event (post-event scenarios) or predict expected losses for a potential future event. Therefore, we focus only on the temporal update of BDDS considering two scenarios:

1. Post-event: Comparison of a BDDS model developed using empirical data from one event against synthetic loss predictions, for the same event using 10-fold Cross Validation (local 10-fold CV). The empirical dataset from the event is split into 10 parts, a BDDS model is trained with 9 parts of the dataset and validated on the left-out data (10th part). This is repeated 10 times, i.e., until all of the dataset is validated. The model definition of the post-event scenario is given by Equation 2.

Future event: Comparison of a BDDS model developed using empirical data from one or more events against synthetic loss predictions, for a future event that occurs in the same region (Temporal one-step ahead Cross Validation; see Figure 1). Since flood damage processes are influenced by human-flood interactions such as preparedness and land use changes (Barendrecht et al. 2019), events occurring in the same region may show significant changes in terms of damage processes over time. Based on empirical evidence, it is expected that exposure and vulnerability show rather similar characteristics within one region than between regions (Schröter et al 2014, Sairam et al 2019).

A BDDS model (BDDS e_1) is developed using synthetic model and empirical flood loss data from the first event (e_1). This model provides calibrated probabilistic loss predictions for the future event, e_2. After the occurrence of the event e_2, a BDDS model (BDDS e_1, e_2) is developed using the same synthetic model and empirical loss data from both events e_1 and e_2. This model results in calibrated probabilistic loss predictions for the event e_3, which may potentially happen in the future. The BDDS model definition of the future event scenario is given by Equation 3.

Synthetic models are also sometimes updated to consider significant changes in damage processes over time. For example, in the UK, the MCM damage datasets have been incrementally updated and improved for over 40 years. Since the MCM online publication (https://www.mcm-online.co.uk/) in 2013, the MCM functions are updated considering available evidences on changes in building contents and structure as well as repair, drying and reconstruction costs and other socio-economic determinants. For predicting damages from potential future events, the recent models are preferable. Considering the available multi-event case studies, none of the corresponding synthetic models were updated between the events.

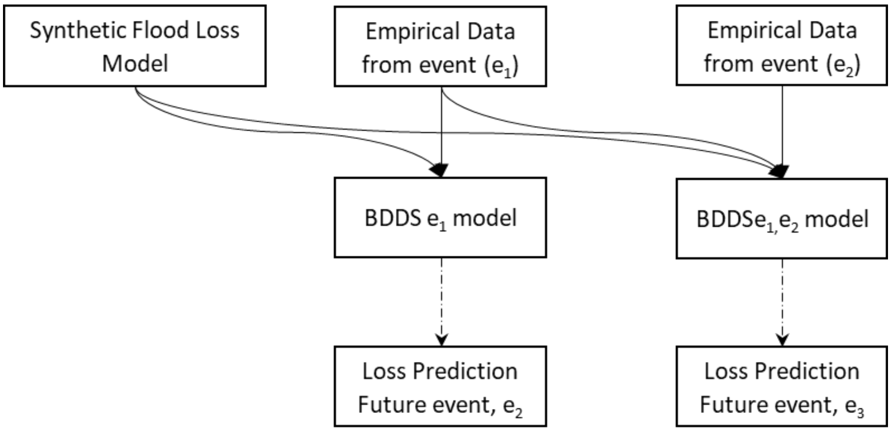


Figure 1: Framework for Temporal one-step ahead CV using a synthetic flood loss model and continuous collection of empirical flood loss data. The components involved in the development of BDDS model are shown with solid lines and the predictions are shown as dot-dash lines.

$$p\left(\widetilde{rloss}_{b'e} \middle| \widetilde{rloss}_{be}\right) = \int_{\Theta} p(\widetilde{rloss}_{b'e}|\theta)p(\theta|\widetilde{rloss}_{be})d\theta \qquad \text{Equation - 2}$$

$$p\left(\widetilde{rloss}_{b'e'} \middle| \widetilde{rloss}_{be}\right) = \int_{\Theta} p(\widetilde{rloss}_{b'e'}|\theta)p(\theta|\widetilde{rloss}_{be})d\theta \qquad \text{Equation - 3}$$

The BDDS model definition for the two scenarios of CV are given in equations 2 and 3, respectively. We are particularly interested in the posterior predictive distribution of the target variable $\widetilde{rloss}$ of residential buildings b' that are not included in training the BDDS model conditioned on the observed losses from the empirical dataset, $\widetilde{rloss}_{be}$ from buildings b and events e. For the post-event damage prediction, the posterior prediction consists of residential buildings that are from the same event e as the empirical data used in the BDDS model training/calibration (Equation 2). For the future event damage prediction, the posterior prediction of $\widetilde{rloss}$ are estimated for residential buildings from a future event e' that was not used in the BDDS model training/calibration. θ contains the beta model parameters (φ, λ and ε) as shown in Equation 1. Hence, after specifying a prior for θ, one finds the posterior distribution $p\left(\theta \middle| \widetilde{rloss}_{be}\right)$.

2.3. Metrics for assessing model performances

The influence of the BDDS model in enhancing synthetic flood loss models is quantified by comparing the predictive performance of the BDDS model against the synthetic model. The predictive performance is evaluated in terms of accuracy of the point estimate based on the median of the predictive distribution (50th percentile of the distribution), using the Mean Absolute Error (MAE) and Mean Bias Error (MBE); the reliability and uncertainty of the predictions are evaluated by means of the Hit rate (HR) and Interval Score (IS) metrics (Gneiting et al. 2007). The HR represents the percentage of predictions where the observed data falls into the 90% High Density Interval (HDI) of the prediction (HDI_{90}; values between the 5th and 95th percentiles of the distribution); the interval score (IS) penalizes the mean width of the 90% HDI, if the prediction lies outside the 90% HDI.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\widetilde{rloss}_i - rloss_i\right| \qquad \textit{Equation - 4}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}\widetilde{rloss}_i - rloss_i \qquad \textit{Equation – 5}$$

$$HR = \frac{1}{n}\sum_{i=1}^{n}h_i;\ h_i = 1\ if\ rloss_i \in HDI_{90i};\ 0\ ,otherwise \qquad \textit{Equation - 6}$$

$$IS = HDI_{90i} + \frac{1}{n}\sum_{i=1}^{n}\frac{2}{\beta}\left(\min(HDI_{90i}) - \widetilde{rloss}_i\right)\left|\left\{\widetilde{rloss}_i < \min\left(HDI_{90i}\right)\right\}\right. + \frac{2}{\beta}\left(\widetilde{rloss}_i - \max(HDI_{90i})\right)\left|\left\{\widetilde{rloss}_i > max(HDI_{90i})\right\}\right. \qquad \text{Equation - 7}$$

Where $\widetilde{rloss}$ is the observed rloss from empirical dataset, $rloss$ is the 50th percentile of the predictive distribution and β scales the score based on the considered HDI; $\beta = 1 - (0.95 - 0.05)$, for 90% HDI. Least MAE and least absolute value of MBE indicate the better performing model. High HR is characteristic of reliable estimates. A smaller IS indicates narrow 90% HDI, which may be potentially due to a larger coverage of empirical loss observations representing the damage processes. Thus, a smaller IS indicates a sharper distribution of the predictions with higher reliability. Most synthetic models considered in this study are deterministic and hence, do not provide a distribution of loss predictions. Thus, only MAE and MBE can be estimated for these synthetic models. However, if uncertainty due to stochastic processes or missing variables are considered by the synthetic model as it is the case for INSYDE (Dottori et al. 2016), the reliability of the synthetic and DDM models can be compared using IS and HR estimates.

2.4. Case studies: Synthetic models, event description and empirical data

2.4.1. **Cumbria, United Kingdom**

2.4.1.1. Synthetic model: Multi Coloured Manual (MCM)

The Multi-Coloured Manual (MCM) (Penning-Rowsell et al., 2013) was initiated in 1977 and incrementally improved thereafter and was developed for the purpose of benefit appraisal for flood investment. It aims to represent national economic losses in sterling. Adopting a deterministic approach, the MCM provides a range of synthetically-generated absolute depth-damage functions for residential and non-residential properties of different types which have been developed to provide national consistent values. The damage functions are generated for individual inventory items and building contents per social grade based on the best ownership and economic values available from market-based surveys and synthetically generated susceptibility curves. For residential properties, unique damage functions are provided according to the type and duration of flooding, warning lead time, building type, year of construction and social class; and estimates of damage are provided for the building fabric and contents and the costs of drying and cleaning. Weighted average damage function curves are then obtained for the different properties considering the national distribution of properties in flood prone areas. For comparability, we utilize MCM loss data to only the residential building fabric and divide by reconstruction cost to obtain an estimate of relative loss. Since empirical data concerning social class was not available, an initial MCM assessment for building fabric losses was performed utilizing different damage functions based on type of flooding, water depth, duration of inundation, warning lead time, building type and year of building construction.

2.4.1.2. <u>Event description and empirical data: Cumbria 2015</u>

The December 2015 flood event in Cumbria (Storm Desmond) was characterized by exceptionally high rainfall, temperature and soil moisture. This is the biggest recorded flooding in Cumbria in almost all the river basins. In comparison, the meteorological winter of 2015/2016 was the wettest on record across all of the UK. The December 2015 event with a return period of 800 to 1,000 years in some parts of Cumbria broke numerous climate records resulting in extreme flooding and strong winds. This event is estimated to have caused impacts between £520 and £662 Million (Szönyi et al. 2016). In most parts of Cumbria, the flooding occurred due to overtopping of the structural protection measures such as dikes and flood walls. In Cockermouth and Keswick, the improved flood protection reduced the impacts of the 2015 event. Further information on the event can be found in Szönyi et al. (2016) and Cumbria County Council (2018). The households reported up to 3 meters of inundation depth and the duration of inundation was between a few hours to almost 48 hours in many regions.

After the 2015 event, computer-aided telephone surveys were undertaken targeting the households that suffered damage during the 2015 flooding. A list of affected streets was obtained using the flood outlines published by the Environment Agency DEFRA (Environment Agency DEFRA, 2019) and the telephone numbers of households in these streets were obtained from public telephone directory. The survey locations were mainly spread over northern UK, mainly focused on the Cumbria region covering, Appleby, Keswick, Kendal, Carlisle and Cockermouth. The survey consisted of questions concerning the hazard (water depth, duration, velocity, contamination etc.), exposure (rebuilding cost and content value), vulnerability (building type, construction year, private precautionary measures, emergency measures, warning information etc.) and incurred damage to building structure and contents. The reconstruction costs for the houses were obtained from the Association of British Insurers (https://www.abi.org.uk/). The households that provided water depth and building loss information from the Cumbria region were selected for this analysis. This resulted in a dataset with 33 residential buildings. All of these households provided information pertaining to the initial appraisal of the MCM. The summary statistics of the responses from the households are provided in Table 1.

2.4.2. **Meuse, Netherlands**

2.4.2.1. <u>Synthetic model: SSM</u>

SSM is a flood loss model developed for the Dutch national government (De Bruijn et al., 2014). It is \]the standard model applied in all Dutch flood risk management studies for the national government. It is an update of an earlier model called Standard Damage and Fatality assessment model (HIS-SSM) (Kok et al., 2005). The damage function applied in this paper, for residential structural damage was first proposed in Duiser (1982). This damage function is based on a combination of information synthesized from empirical observations concerning flood damages from three events: the coastal floods in Zeeland in 1953, the Wieringermeer flood of 1945 from a large lake and a flood in Tuindorp-

Oostzaan in 1960 (canal dike breach), interviews from experts and damage functions from Penning-Rowsell et al. (1977).

2.4.2.2. Event description and empirical data: Meuse 1993

This dataset is based on the 1993 flood of the Meuse River in the Dutch province of Limburg. It has been described in WL Delft (1994), Wind et al. (1999) and Wagenaar et al. (2017). The 1993 Meuse discharge was 3,120 m3/s, the highest recorded up to that point. 8% of the province was flooded causing about 180 Million Euro damage (price level 2016) (Wagenaar et al., 2017). Unlike most of the rest of Dutch rivers, in 1993 the Meuse River didn't have dikes yet in Limburg.

The data was collected to compensate affected households. Every flooded building was visited, resulting in a complete data set of 5,780 records. The data collection was carried out by insurance experts who visited the affected buildings weeks after the flood, often before restoration activities were completed. The experts also recorded the water depth in the buildings but this wasn't their primary objective and was sometimes difficult to assess because the flood had happened weeks prior. In Wagenaar et al. (2018) the recorded flood losses have been transferred to relative losses. The summary statistics of the survey responses are given in Table 1.

2.4.3. **Adda, Caldogno and Secchia, Northern Italy**

2.4.3.1. Synthetic model: INSYDE (Dottori et al, 2016)

INSYDE is an expert-based synthetic model, developed for the Italian context. The model is based on a what-if analysis, consisting in a virtual step-by-step inundation of a residential building and in the evaluation of the corresponding physical and monetary damage as a function of hazard and building characteristics. A mathematical function describes the damage mechanisms for each building subcomponent (walls, doors, etc.), and the associated cost for reparation, removal, and replacement; when the influence of hazard and building variables cannot be determined a priori, damage mechanisms are modelled using a probabilistic approach. In total, INSYDE adopts 23 input variables, six describing the flood event and 17 referring to building features. However, the model can be also applied when the available knowledge of the flood event and building characteristics is incomplete, given the possibility of automatically considering default values for unknown parameters and of expressing some of the variables as functions of other ones. The model supplies damage in absolute terms but an estimation of relative damage can be obtained.

2.4.3.2. Event descriptions and empirical data: Adda 2002, Caldogno 2010, Secchia 2014

In this case study three flood events in the Po valley in Northern Italy are considered. The first one happened in November 2002 in the town of Lodi. The flood resulted from a most critical combination of events for the lower part of the Adda river, namely the simultaneous increase of the discharges from the Como lake and of the Brembo river, that is the largest tributary of the Adda upstream of Lodi. Between the 25th and 26th of November, the Adda reached the hydrometric height of 3.43 m above the reference level (68.28 m a.s.l.), corresponding to a discharge between 1,800 and 2,000 m^3/s. The return period has been estimated as 100-200 years. Large portions of the town were flooded with water levels above 2 m in some neighbourhoods. The second flood event happened in the Veneto region, where from the 31st of October to the 2nd of November 2010, persistent rainfall affected the pre-Alpine and foothill areas, with peaks of more than 500 mm in some locations (ARPAV, 2010). Consequently, about 140 km^2 of land was inundated, involving 130 municipalities, some of which were particularly negatively affected. The situation of Bacchiglione River and its tributaries was especially critical, where hydrometric levels overcame historical records (water velocities in the river higher than 330m^3/s were registered; see Belcaro et al., 2011), causing the opening of a breach on the right levee of the river on the morning of the 1st of November. The countryside and the settlements of Caldogno, Cresole and Rettorgole were flooded with an average water depth of 0.5 m (ARPAV, 2010) for about 48 hours. The total damage, including residential properties, economic activities, agriculture and public infrastructures, was estimated to be about EUR 26 million, of which EUR 7.5 million relate to the residential sector (Scorzini and Frank, 2017). Finally, the last event happened in January 2014 in the central area of the Emilia–Romagna region (Modena province), where in the early morning of the 19th of January the water started to overtop the right levee of the Secchia River, flooding the countryside. The breach was not caused by an extreme river discharge (the return period of the event was estimated around 5 years), but by the collapse of the river embankment, weakened by animal burrows (D'Alpaos et al., 2014). Seven municipalities were affected with an inundated area of around 52 km^2 with the small towns of Bastiglia and Bomporto suffering the largest impacts remaining flooded for more than 48 h. The total volume of overflowing water was estimated about $36x10^6$ m^3, with an average water depth of 1 m (D'Alpaos et al., 2014). The economic cost inflicted on residential properties, according to damage declaration, amounted to EUR 36 million.

After the three floods, public funding was made available by the national Civil Protection Authority. In order to be reimbursed, with similar procedures for all inundation events, citizens were requested to fill in pre-filled claim forms; the latter were then mostly collected by the affected municipalities and, in a small part, by the Regional Authorities. In total, our dataset includes 1,158 buildings in the flooded areas (Amadio et al. 2019). They include information on the owner, the address of the flooded building, its typology (e.g. apartment, single house), the number of affected floors, a description of the physical damage and its translation into monetary terms (distinguishing for the different rooms among damage to walls, windows and doors, floor and content). More information about the individual flood events, their hydrodynamic simulations and the data collection campaigns were published in Scorzini et al. (2018), Molinari et al. (2020), Scorzini and Frank (2017), Carisi et al (2018), Amadio et al. (2019).

The areas flooded in the three cases are characterized by similar exposure characteristics and economic well-being (Amadio et al. 2019). Previous studies compared the same cases and the findings sustain the opportunity to merge the dataset (Amadio et al. 2019). Hence, the three events are combined into one case study. The summary of empirical data from this case study is provided in Table 1.

2.4.4. **Danube, Germany**

2.4.4.1. Synthetic model: Rhine Atlas Model (RAM) (ICPR, 2001)

The Rhine Atlas Model (RAM) was developed in 2001 in order to determine the regions with high flood risk in the Rhine catchment based on the 1995 floods and develop risk management strategies (ICPR, 2001). Since, the RAM is intended for the Rhine catchment, an inherent transfer scenario exists when the RAM is generalized to the other catchments within Germany. However, given that a number of studies consider RAM as a standard synthetic flood loss model (Jongman et al. 2012), we use the model as the standard synthetic flood loss model for Germany. The RAM is mostly based on expert judgment as well as some information based on the HOWAS empirical flood damage data (Buck & Merkel. 1999). It is a stage-damage function using water depth as the only predictor. The RAM loss prediction is based on the resolution of land-use classes similar to that of the CORINE land use data (Jongman et al. 2012). We apply the stage-damage function corresponding to losses to building structure in the residential land-use class to estimate flood loss for each residential building.

2.4.4.2. Event descriptions and empirical data: Danube 2002-2013

In this case study, three flood events that occurred between 2002 and 2013 in the Danube catchment is considered. Among the events, the 2013 flood was quite extreme with return period up to greater than 1000 years in some parts of the catchment. These were summer floods caused due to heavy rainfall resulting in surface water flooding and flash floods (Vogel et al. 2018). The 2013 floods were characterized by high antecedent soil moisture combined with heavy precipitation resulting in large spatial extent of flood peaks with high magnitudes resulting in the most severe flooding in Germany over the past 6 decades (Merz et al., 2014, Schröter et al. 2015). Another distinguishing feature is the occurrence of dike breaches during the Danube 2013 event. Many properties were affected after dike breaches (e.g. at Deggendorf).

After these events, computer-aided cross-sectional telephone surveys of private households that had suffered from losses were undertaken using a standardized questionnaire. A list of affected streets was obtained using the flood masks derived from satellite data, (DLR, Center for Satellite Based Crisis information, https://www.zki.dlr.de/), and the telephone numbers of households in these streets were obtained from public telephone directory. The survey campaigns always focused on a single event. Depth of water within the house is determined using the reported water level in the highest affected storey by applying corrections based on the presence of a basement and height of the ground floor. Building reconstruction costs are adjusted for inflation to values as of 2013 using the building price index (DESTATIS, 2013). We consider all datasets which refer to households with basement (for unbiased measurements of water depth) and for which information on water depth and relative building loss were provided. Hence, the empirical data used in this study consists of 408 buildings from three events in the Danube catchment, that have a considerable number of completed surveys (sample size>25). The summary of empirical data from this case study is provided in Table 1.

2.4.5. **Elbe, Germany**

2.4.5.1. Synthetic model: Rhine Atlas Model (RAM) (ICPR, 2001)

The Rhine Atlas Model (RAM), described in section 2.4.4.1 is implemented for estimating losses in the Elbe catchment.

2.4.5.2. Event descriptions and empirical data: Elbe 2002-2013

In the Elbe catchment, the 2002 and 2013 events were extreme with return periods greater than 100 years. These events affected a large number of households. The 2002 event was characterized by a large number of dike breaches affecting households with low preparedness. However, after the 2002 event, preparedness increased among households via implementation of private precautionary measures and emergency measures. Hence, a reduction in average losses is observed after the 2002 event in the Elbe catchment. The other flood events (2006 and 2011) were smaller with return periods less than 50 years. They were caused due to rain-on-snow after the winter periods (Vogel et al. 2018).

Empirical damage data was collected from the affected households in the Elbe catchment during the same survey campaigns, explained in section 2.4.4.2. The study uses four events comprising of a total of 1,110 households, that provided information on water depth and relative building loss and have a considerable number of completed surveys (sample size>25). The summary of empirical data from this case study is provided in Table 1. More information about the individual flood events in the Elbe and Danube, the surveys and their results were published in Thieken et al. (2007), Kreibich et al. (2011, 2017), Kienzler et al. (2015) and Vogel et al. (2018).

In this study, the Danube and Elbe catchments are considered as different case studies due to their strikingly different socio-economic and exposure characteristics which affect flood damage processes (Thieken et al. 2007). These regional differences have historical roots since the Danube catchment belonged to former West Germany and the Elbe catchment to the former East.

Table 1: Sample size, the summary (average) of water depth (wd) in meters, exposed building value (bv in EUR) , absolute and relative losses to residential buildings (bloss in EUR, rloss) for the five case studies.

Case study	Event	Sample size	wd	bv[1]	bloss[1]	rloss
Cumbria, United Kingdom (UK)	Cumbria 2015	33	0.6	390,320[2]	32,640[2]	0.08
Meuse, Netherlands (NL)	Meuse 1993	5780	0.4	138,000	4,307	0.03
Northern Italy (IT)	Adda 2002	270	0.9	197,356	10,592	0.05
	Caldogno 2010	294	0.4	268,175	18,398	0.07
	Secchia 2014	594	1.0	229,670	22,832	0.10
Danube, Germany (DE)	Danube 2002	225	1.7	360,107	6,352	0.02
	Danube 2005	104	2.0	412,102	7,992	0.02
	Danube 2013	79	3.0	580,109	45,675	0.08
Elbe, Germany (DE)	Elbe 2002	518	3.5	306,535	44,462	0.14
	Elbe 2006	42	2.9	312,417	7,066	0.02
	Elbe 2011	58	2.7	482,588	9,277	0.02
	Elbe 2013	492	2.7	434,095	23,599	0.05
Total		8489				

Note: [1] Values in € adjusted for inflation to values as of 2015; [2] Values in £ converted to € using conversion rate 1€ = 0.73£.

3. Results and Discussion - Comparison of predictions from synthetic loss models and BDDS models

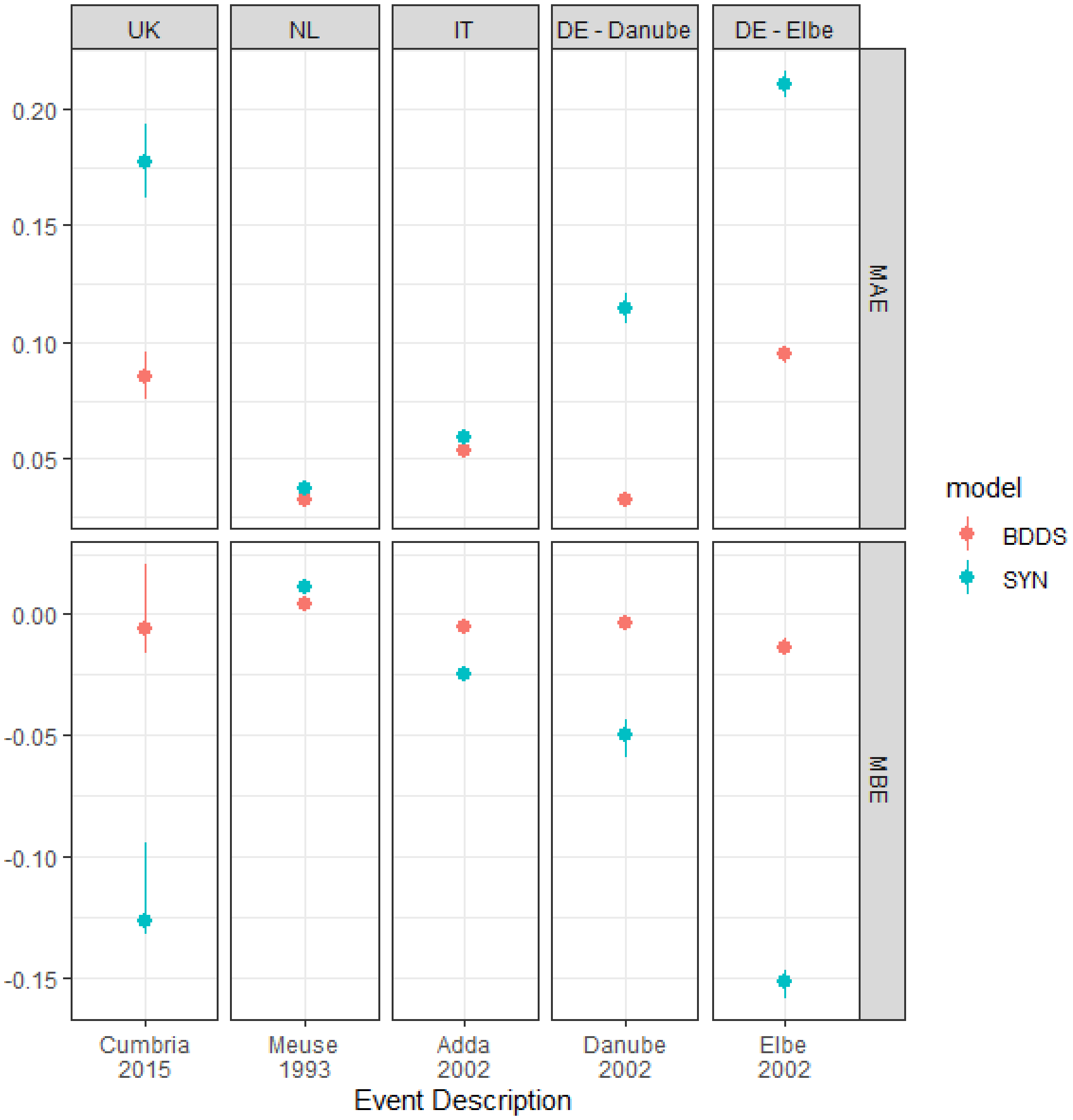
The performance of the BDDS model is compared with the synthetic models from the respective regions. Since the development of BDDS models requires empirical data, the model is independently trained for each of the local 10-fold CV as well as temporal one-step-ahead CV and is validated on the left-out dataset. During both validation scenarios, there are no variations in definition and parameterization of the synthetic models. Point estimates are assessed via MAE and MBE and prediction uncertainty and reliability via IS and HR (section 2.3). Reliability and uncertainty of loss predictions are provided by all BDDS models, representing an enhancement over the deterministic synthetic models (4 out of 5 models). Among the synthetic models, INSYDE is the only synthetic model that provides distribution of loss estimates from which IS and HR can be determined. The model validation is performed by bootstrap sampling of the synthetic and BDDS model predictions with 1,000 iterations with replacement, while preserving the sample size of the empirical data during each iteration.

3.1. Local 10-fold CV

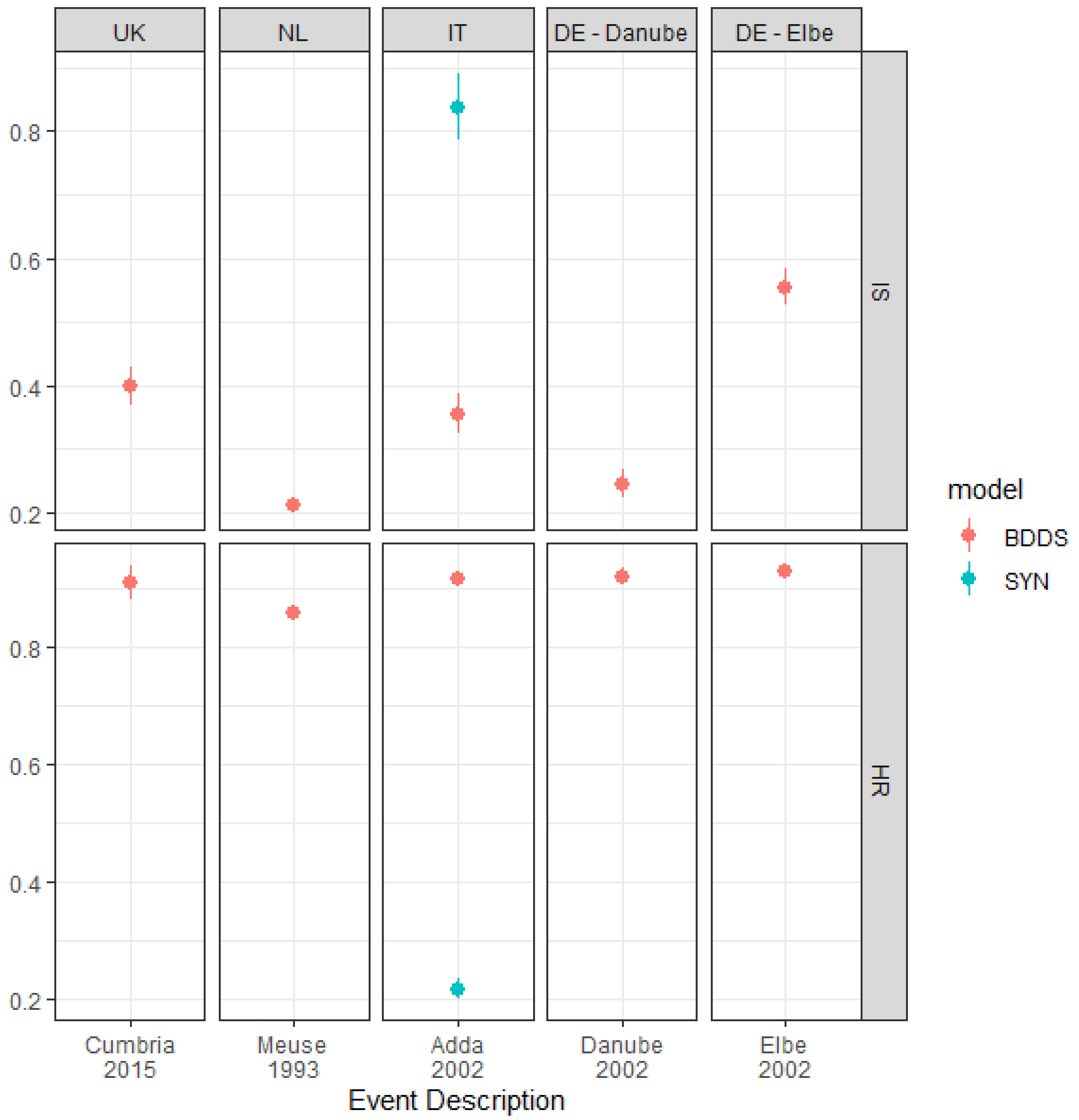
We perform a local 10-fold CV in order to validate the BDDS model predictions against the synthetic model predictions for the post-event scenario. The case studies with no empirical data from the region prior to the event are used for local 10-fold CV. This scenario (Equation 2) is applicable for the Cumbria 2015, Meuse 1993, Adda 2002, Danube 2002 and Elbe 2002 flood events. These events are either the only available empirical data from the respective regions or the first event of the continuous empirical data collection campaigns. All synthetic models, except SSM, result in a negative MBE which indicates that on average, all these synthetic models over-estimate the building losses (see Figure 2a).

The prediction performance of the BDDS model with one event is compared against the performance of the synthetic models from the corresponding countries (Figure 2a). The BDDS model performs better than the synthetic model in terms of point estimates. As described in Equation 6, during the local 10-fold CV, the model is iteratively validated on residential buildings that are not used in the model development. Thus, the local 10-fold CV evaluates out-of-sample model performance of the BDDS model. The BDDS model with RAM and empirical data from the Elbe 2002 event results in the highest improvement in predictive performance in terms of MAE and MBE. Small improvement in predictive performance is exhibited by the BDDS models - SSM and empirical data from Meuse 1993 event and INSYDE with empirical data from the Adda 2002 event. However, among the tested synthetic models, the INSYDE and SSM models result in the smallest errors in the 10-fold CV. Among the two catchments in Germany, the RAM results in larger errors predicting losses for the Elbe 2002 event compared to the Danube 2002 event. The BDDS model consistently improves the predictions for the 2002 event in both catchments.

The uncertainty and reliability of the loss predictions is quantified using the IS and HR metrics. For the Adda 2002 event, the IS (HR) of the predictions from the INSYDE model is high (low) compared to the corresponding BDDS model. Hence, integrating empirical data with the INSYDE model using BDDS model reduces uncertainty and improves the reliability. The predictions from BDDS model with SSM and empirical data from the Meuse 1993 event have the least IS which represents a narrow prediction interval/HDI_{90}. The predictions from BDDS model with RAM and empirical data from Elbe 2002 event results in the highest HR with approximately 93% of the empirical loss data lying within the HDI_{90} of the predictions, representing high model reliability. However, the IS of these predictions is also high suggesting a large uncertainty. The predictions from BDDS model with empirical data from Danube 2002 event show low IS and high HR representing a good balance between reliability and uncertainty. The HDI_{90} is narrow for these predictions and also a large percentage (92%) of the observed losses is captured within the HDI_{90} of the predictions.


UK
NL
IT
DE - Danube
DE - Elbe
MAE
MBE
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
model
BDDS
SYN
Cumbria 2015
Meuse 1993
Adda 2002
Danube 2002
Elbe 2002
Event Description

(a)



(b)

Figure 2 Model performances for local 10-fold CV using events and their corresponding synthetic loss models (shown in brackets) -– Cumbria 2015 (MCM), Meuse 1993 (SSM), Adda 2002 (INSYDE), Danube 2002 (RAM) and Elbe 2002 (RAM). (a) MAE and MBE of flood loss predictions using synthetic models and BDDS models (b) IS and HR of loss predictions using BDDS models.

Among the tested synthetic models, the SSM and INSYDE models result in the least errors (see, Figure 2a). These models were developed after the occurrence of the respective events and may potentially capture flood damage processes based on recent events, which are comparable with the tested events. This may explain the better fit compared to the other models. Another plausible reason for the small errors from the SSM model is that the Meuse 1993 event resulted in small damage values (Table 1). This may lead to smaller errors in terms of MAE and MBE (Wagenaar et al. 2018). From the bootstrap iterations of MAE and MBE, the spread of the errors from the Cumbria 2015 event is the largest. This can be attributed to the low coverage (small sample) of empirical data from the Cumbria 2015 event. However, despite the limited availability of empirical data, the BDDS model enhances loss predictions from the MCM as well. The BDDS model reduces errors and provides predictive distributions indicating uncertainty and reliability of the predictions. In the case of Elbe 2002, the hit rate of the BDDS model is high and comparable with the performance of other BDDS models. However, the high IS indicates that the loss distributions are not sharp. This high uncertainty may be attributed to variability in damage processes that are not adequately captured by the variables in the RAM (i.e. water depth only). This quantification of uncertainty and reliability from BDDS model is an enhancement over the established synthetic models, which is crucial for risk-based decision making (Polasky et al. 2011).

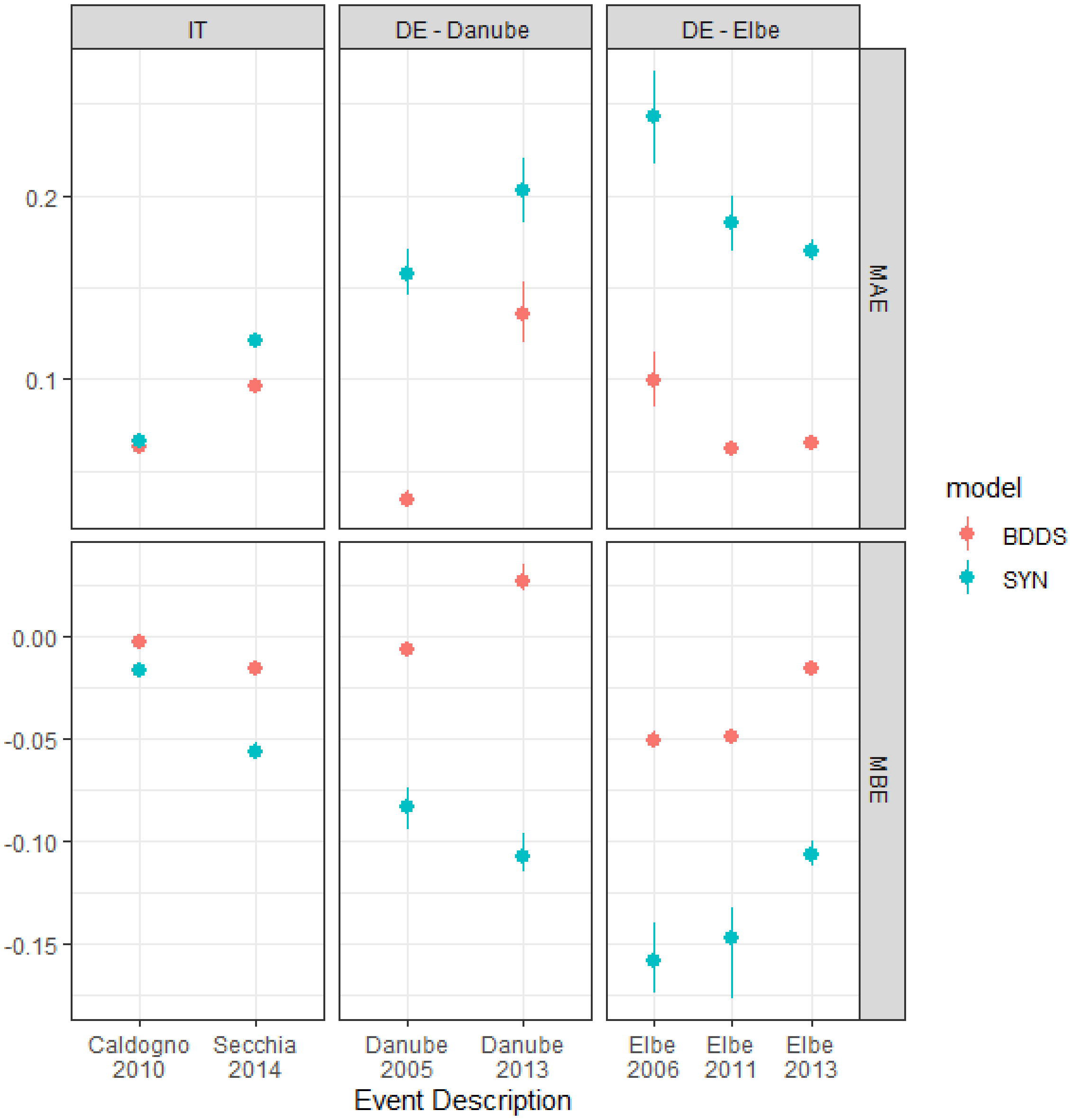
3.2. Temporal One-step ahead CV

In regions where, continuous empirical flood damage data is available, the predictions from synthetic models and BDDS models are compared using temporal one-step ahead CV. The losses suffered by residential buildings due to an event in the future is predicted from a BDDS model developed using the synthetic model and all available empirical data from the past events (Figure 1 and Equation 3). From our case studies, empirical damage data from northern Italy and Germany can be used to implement temporal one-step ahead CV.

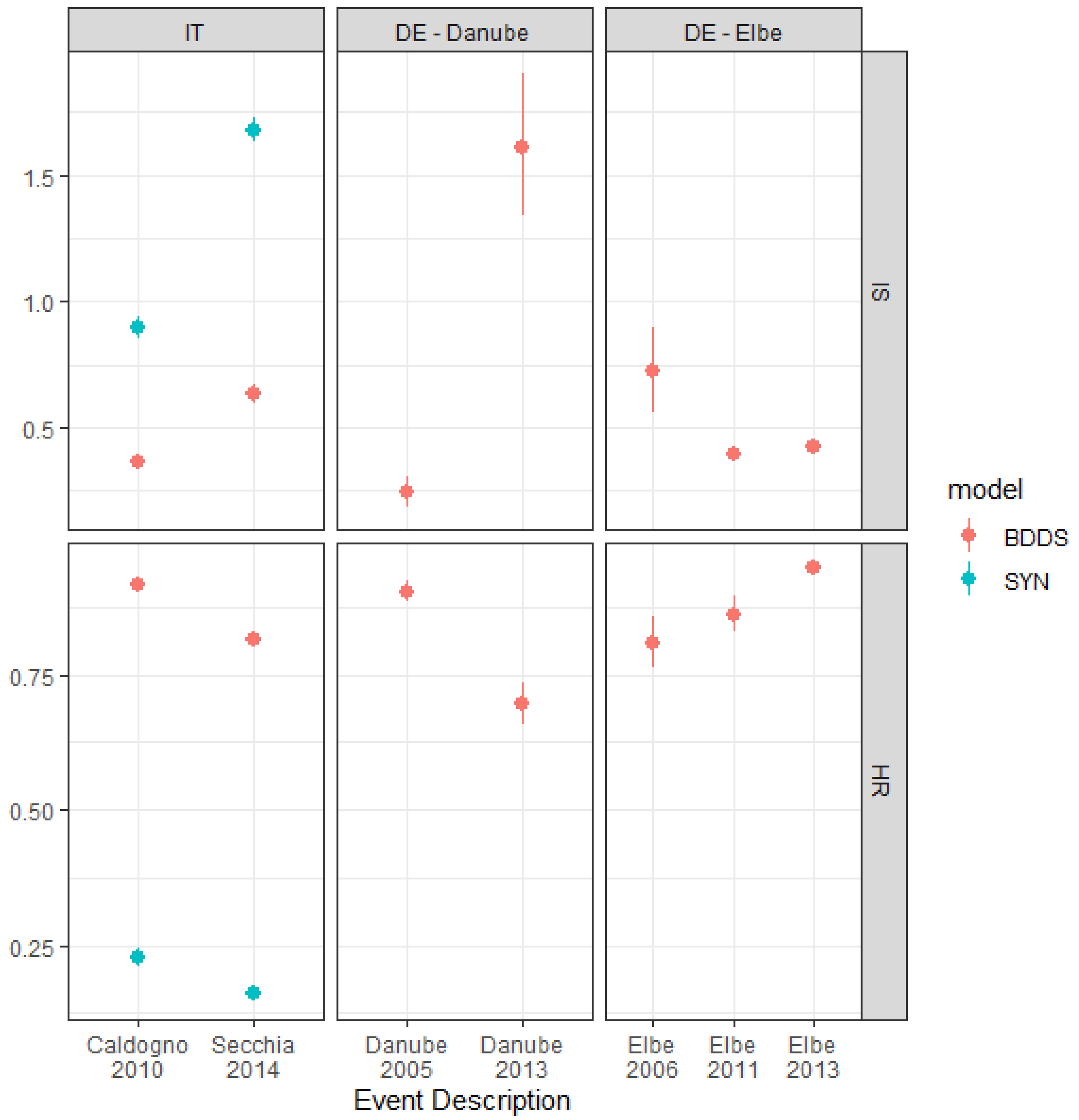
Since we have empirical data from three events from Northern Italy, two BDDS models are developed, i.e. to predict losses from Caldogno 2010, the BDDS model is developed using INSYDE model and empirical data from Adda 2002, and to predict losses from Secchia 2014, the BDDS model is based on INSYDE model and empirical data from Adda 2002 and Caldogno 2010. Five BDDS models are developed for Germany using the RAM and empirical data from the past events to predict future losses. In the Danube catchment, to predict losses from the 2005 (2013) event, a BDDS model is developed using RAM and empirical data from 2002 (2002 and 2005). In the Elbe catchment, to predict losses from the 2006 (2011 / 2013) event, a BDDS model is developed using RAM and empirical data from 2002 (2002 and 2006/ 2002, 2006 and 2011).

The results of the temporal one-step ahead CV are provided in Figure 3a. For all the case studies, the errors (MAE and MBE) from the BDDS model temporal one-step ahead prediction are smaller than the errors from the corresponding synthetic models. The results show that compared to the INSYDE model, the performance of the INSYDE model continuously integrated with empirical data from more events is higher. For the Elbe catchment, the BDDS model's improvement in predictive performance is observed for all future event predictions when integrated with a continuous collection of empirical data. These results suggest that, in these two regions, parameterizing the BDDS model with empirical data from events in the recent past improves the damage prediction for following events.

In the Danube catchment in Germany, the BDDS model outperforms the RAM for temporal one-step ahead predictions. However, the BDDS model shows a lower performance when data from an additional event is integrated. We also notice a change from negative to positive bias. This suggests that in the case of Danube 2013 event, the BDDS model developed by integrating RAM with empirical data from 2002 and 2005 events under-estimates the losses. The uncertainty and reliability estimates, i.e. IS and HR, from BDDS model one-step ahead temporal predictions are shown in Figure 3b. The two BDDS models developed for the case study in Northern Italy result in better HR and IS estimates compared with the INSYDE model. The BDDS model shows best reliability and least uncertainty for the Elbe 2013 event with a HR close to 100% and a relatively small IS, suggesting small uncertainty. On the other hand, loss predictions for the 2013 event in the Danube catchment from the BDDS model performs the worst with the least HR of 70% and a high IS, suggesting low reliability and large uncertainty.


IT
DE - Danube
DE - Elbe
MAE
MBE
0.2
0.1
0.00
-0.05
-0.10
-0.15
Caldogno 2010
Secchia 2014
Danube 2005
Danube 2013
Elbe 2006
Elbe 2011
Elbe 2013
Event Description
model
BDDS
SYN

(a)



(b)

Figure 3: Model performances for temporal one-step ahead CV of events using empirical data from past events and their corresponding synthetic loss models (shown in brackets) -– Caldogno 2010 (Adda 2002; INSYDE), Secchia 2014(Adda 2002, Caldogno 2010; INSYDE), Danube 2005 (Danube 2002; RAM), Danube 2013 (Danube 2002, 2005; RAM), Elbe 2006 (Elbe 2002; RAM), Elbe 2011 (Elbe 2002, 2006; RAM), Elbe 2011 (Elbe 2002, 2006, 2011; RAM). (a) MAE and MBE of flood loss predictions using synthetic models (SYN) and BDDS models (b) IS and HR of loss predictions using BDDS models.

During temporal one-step ahead CV, the BDDS model shows an overall improvement over the synthetic models. In the case of Danube 2013, integrating the RAM with Danube 2002 and 2005 events result in high IS and low HR (Figure 3b). This effect is also in agreement with the inferences from MBE for Danube 2013 estimated from the same model (Figure 3a). For all temporal one-step ahead CV cases, the synthetic models over-estimate the losses. However, when enhanced with empirical data from past events using BDDS model, the MBE is shifted towards zero. In the case of Danube 2013, the empirical data from past events reduces the overall bias, but leads to an underestimation of losses. This effect may result from some characteristics of the Danube 2013 event that differ from the other Danube events. For example, dike breaches that occurred during the Danube 2013 event inundated properties that were located away from the river with high water depths. These households had low flood experience and were not prepared for flooding. Hence, high intensity flooding combined with low preparedness resulted in large damages (e.g. oil contamination from heating systems). Such effects are not sufficiently captured either by the uni-variable RAM or the empirical data from past events. Hence, it is important to evaluate if the empirical data is representative of the target event's damage processes. One example is the implementation of ensemble models based on the individual model characteristics and target case study (Figueiredo et al. 2018). A potential approach to capture the difference in damage processes between events is to introduce a multi-level model that allows both shared and separate parameters representing the similarities and differences between the damage processes exhibited by the different events (Sairam et al. 2019). The criteria for similarities in damage processes used by these studies were established on the basis of expert knowledge. To reduce the subjectivity in choice of models and relevance of empirical data, standardization of data for flood loss estimation along with a rigorous benchmarking of the loss models are important next steps.

In order to interpret the importance of local empirical data, we discuss the performances of the BDDS model that is built with empirical data from the same event (local 10-fold CV) and past events (temporal one-step ahead CV). Local empirical data from the same event improves the overall reliability of the BDDS model and also results in low uncertainty, i.e. reduces IS and increases HR (Figures 2b and 3b). Hence, the use of empirical data from the same event is useful for post-event risk analysis and damage estimation. For risk-based decision making for future scenarios, we need accurate and reliable models, which can only be validated using empirical data from past events. Therefore, the IS and HR estimates obtained from the temporal one-step ahead loss predictions are more relevant. These metrics can be considered by decision makers and flood risk managers as the estimates of uncertainty and reliability of the damage model for future flood risk portfolios. In general, the BDDS model enhances synthetic models using local empirical data.

4. Conclusions

Synthetic models are based on what-if analyses and are hardly validated and compared with observations. Models purely developed using empirical data require large samples of detailed object-level damage data, preferably from various events. By the presented approach it becomes possible to use the vast compendium of established synthetic damage functions in a harmonized probabilistic framework in order to improve damage estimation and quantify the reliability of the model predictions. We calibrate the synthetic models with local empirical damage data, for which not as many observations are necessary as for the development of empirical damage models.

We have performed 10-fold and temporal one-step ahead Cross Validation (CV) for assessing the model performances for post-event and future event scenarios, respectively. Some empirical damage data from the event is used in model training for 10-fold CV. Whereas, only empirical damage data from past events are used for model training for temporal one-step ahead CV. Our validation results show that empirical loss data from past events are valuable for enhancing the synthetic models to predict damage more accurately. From the tested case studies, on average, a reduction of 50% (51%) and 88% (74%) in mean absolute error and mean bias error were achieved by BDDS model for the post(future)-event scenarios, respectively. In respect to reliability, average hit rates of 90% and 85% were achieved for post and future event scenarios, respectively. Hence, for improving estimates of future risk, empirical data collection campaigns after flood events are crucial. However, the loss predictions from the post-event scenario show higher reliability compared to the future risk predictions. This suggests that flood damage processes vary across events and therefore dynamic damage models are required to capture this variability. Within the scope of this study, the models are not tested for regional (cross-country) transferability. This is considered as a follow-up research work for the future.

An important feature of the presented approach is the uncertainty quantification of the damage estimate, since this provides valuable information for improved decision making. In order to train a BDDS model for a new case study, availability of empirical damage data from past event(s) and ability to run the national standard synthetic loss model for the same event(s) are required. From the modelling perspective, knowledge concerning formulating regression equations in R (R Core Team, 2019), interpretation of regression coefficients and understating probability distributions may help in customizing the presented model structure and parameter definitions, if needed. With respect to model application, no special skills are needed to use a trained BDDS model. The input data required to run the BDDS model are the same as that of the national standard synthetic model. The running time of the BDDS model is comparable to the national standard synthetic models for the samples in the tested case studies. Thus, the Bayesian Data-Driven approach is valuable for flood risk managers.

Acknowledgements

This research has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement 676027 MSCA ETN System-Risk, as well as from the project DECIDER (BMBF, 01LZ1703G).

Data and Software Availability

The Multi Coloured Manual (MCM) database handbook is published by the Flood Hazard Research Centre (FHRC) at MiddleSex Univerisity, London, UK. The functions are proprietary and not publicly accessible. The SSM model is available from de Bruijn et al. (2014). INSYDE functions are available for download as R open source code, currently hosted on GitHub (https://github.com/ruipcfig/insyde/). The Rhine Atlas Model (RAM) is available from ICPR (2001).

The data implemented in the Cumbria 2015 case study is currently not publicly accessible. The dataset may be obtained upon request. The data used in the Meuse 1993 case study is available from Wagenaar et al. 2017. The dataset used in the Northern Italy case study are not publicly accessible. The first reason behind this is that some data come from private sources (i.e., businesses, utilities companies) that agreed on sharing their data only for research objectives, including sensitive information. The dataset may be obtained upon request. For the Danube and Elbe case studies, flood damage data of the 2005, 2006, 2010, 2011, and 2013 events along with instructions on how to access the data are available via the German flood damage database, HOWAS21 (http://howas21.gfz-potsdam.de/howas21/). Flood damage data of the 2002 event was partly funded by the reinsurance company Deutsche Rückversicherung (www. deutscherueck.de) and may be obtained upon request. The surveys were supported by the German Research Network Natural Disasters (German Ministry of Education and Research (BMBF), 01SFR9969/5), the MEDIS project (BMBF; 0330688) the project “Hochwasser 2013” (BMBF; 13N13017), and by a joint venture between the German Research Centre for Geosciences GFZ, the University of Potsdam, and the Deutsche Ruckversicherung AG, Dusseldorf.

The models presented in this paper are implemented in the stan modeling language (Carpenter et al., 2017) using the brms package version 3.3.2 (Bürkner, 2017) in R (R Core Team, 2019).

References

1. Amadio, M., Scorzini, A. R., Carisi, F., Essenfelder, A. H., Domeneghetti, A., Mysiak, J., & Castellarin, A. (2019). Testing empirical and synthetic flood damage models: the case of Italy. *Natural Hazards and Earth System Sciences*, *19*(3), 661-678.

2. ARPAV: Scheda Evento “Idro” 31 Ottobre–5 Novembre 2010, available at: http://www.regione.veneto.it/ (last access: 21 March 2019), 2010.
3. Barendrecht, M. H., Viglione, A., Kreibich, H., Merz, B., Vorogushyn, S., & Blöschl, G. (2019). The value of empirical data for estimating the parameters of a sociohydrological flood risk model. *Water Resources Research*, 55, 1312– 1336. https://doi.org/10.1029/2018WR024128
4. Barredo JI (2009) Normalised flood losses in Europe: 1970–2006. Nat Hazards Earth Syst Sci 9:97–104. doi: 10.5194/nhess-9-97-2009
5. Belcaro, P., Gasparini, D., and Baldessari, M.: 31 ottobre–2 novembre 2010: l’alluvione dei Santi, Regione Veneto, available at: http://statistica.regione.veneto.it/ (last access: 21 March 2019), 2011.
6. Buck, W., & Merkel, U. (1999). *Auswertung der HOWAS-Schadendatenbank* (No. HY98/15). Institut für Wasserwirtschaft und Kulturtechnik: Universität Karlsruhe.
7. Bürkner, P.-C. (2017), ‘brms: An R Package for Bayesian Multilevel Models Using Stan’, Journal of Statistical Software 80(1), 1–28.
8. Carisi, F., Schröter, K., Domeneghetti, A., Kreibich, H., and Castellarin, A.: Development and assessment of uni- and multivariable flood loss models for Emilia-Romagna (Italy), Nat. Hazards Earth Syst. Sci., 18, 2057–2079, https://doi.org/10.5194/nhess-18-2057-2018, 2018.
9. Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*(1).
10. Cumbria County Council. (2018). Flooding in Cumbria, December 2015 Impact Assessment, Carlisle, UK: Cumbria County Council – Performance and Intelligence Team. URL: https://www.cumbria.gov.uk/eLibrary/Content/Internet/536/671/4674/17217/17225/43312152830.pdf (last accessed 9 May 2019).
11. D’Alpaos, L., Brath, A., and Fioravante, V.: Relazione tecnicoscientifica sulle cause del collasso dell’ argine del fiume Secchia avvenuto il giorno 19 gennaio 2014 presso la frazione San Matteo, available at: http://www.comune.bastiglia.mo.it/ (last access: 21 March 2019), 2014.
12. da Costa, R. T., Manfreda, S., Luzzi, V., Samela, C., Mazzoli, P., Castellarin, A., & Bagli, S. (2019). A web application for hydrogeomorphic flood hazard mapping. Environmental Modelling & Software, 118, 172-186.
13. De Bruijn, K. M., Wagenaar, D. J., Slager, K., De Bel, M., and Burzel, A.: Proposal for the new flood damage assessment method: SSM2015, Deltares, 2014
14. DESTATIS (Federal Statistical Office): Statistisches Jahrbuch – Deutschland und Internationales (2013). Available at: https://www.destatis.de/DE/Publikationen/StatistischesJahrbuch/StatistischesJahrbuch2013.pdf (last access: 20 May 2018), 2013 (in German).
15. Domeneghetti, A., Carisi, F., Castellarin, A., Brath, A. Evolution of flood risk over large areas: quantitative assessment for the Po River. Journal of Hydrology 527 (2015) 809–823, 2015; doi:10.1016/j.jhydrol.2015.05.043
16. Dottori, F., Figueiredo, R., Martina, M. L., Molinari, D., & Scorzini, A. (2016). INSYDE: a synthetic, probabilistic flood damage model based on explicit cost analysis.
17. Duiser, J. A.: Een verkennend onderzoek naar methoden ter bepaling van inundatieschade bij dijkdoorbraak, 82-0644 TNO, Delft, the Netherlands, 1982
18. Elmer, F., Thieken, A. H., Pech, I., & Kreibich, H. (2010). Influence of flood frequency on residential building losses. *Natural Hazards and Earth System Sciences, 10*(10), 2145–2159.
19. Environment Agency DEFRA. (2019): Recorded Flood Outlines. URL: https://environment.data.gov.uk/dataset/8c75e700-d465-11e4-8b5b-f0def148f590

20. Figueiredo, R., Schröter, K., Weiss-Motz, A., Martina, M. L. V., and Kreibich, H.: Multi-model ensembles for assessment of flood losses and associated uncertainty, Nat. Hazards Earth Syst. Sci., 18, 1297–1314, https://doi.org/10.5194/nhess-18-1297-2018, 2018.
21. Gerl, T., Kreibich, H., Franco, G., Marechal, D., & Schröter, K. (2016). A review of flood loss models as basis for harmonization and benchmarking. *PloS one*, *11*(7), e0159791.
22. Gneiting, T. and Raftery, A. E.: Strictly Proper Scoring Rules, Prediction, and Estimation, Journal of the American Statistical Association, 102(477), 359–378, doi:10.1198/016214506000001437, 2007.
23. ICPR: Atlas of flood danger and potential damage due to extreme floods of the Rhine, International Commission for the Protection of the Rhine, Koblenz, 2001.
24. Jongman, B., Kreibich, H., Apel, H., Barredo, JI, Bates, PD, Feyen, L., ... & Ward, PJ (2012). Comparative flood damage model assessment: towards a European approach.
25. Kienzler, S., Pech, I., Kreibich, H., Müller, M., & Thieken, A. H. (2015). After the extreme flood in 2002: Changes in preparedness, response and recovery of flood-affected residents in Germany between 2005 and 2011. *Natural Hazards and Earth System Sciences*, *15*(3), 505–526.
26. Kok, M., Huizinga, H. J., Vrouwenvelder, A. C. W. M., and van den Braak, W. E. W.: Standaardmethode 2005, Schade en Slachtoffers als gevolg van overstroming, HKV, TNObouw, Rijkswaterstaat DWW, 2005
27. Kreibich, H., Di Baldassarre, G., Vorogushyn, S., Aerts, J. C., Apel, H., Aronica, G. T., et al. (2017). Adaptation to flood risk: Results of international paired flood event studies. *Earth's Future*, *5*, 953–965. https://doi.org/10.1002/2017EF000606
28. Kreibich, H., Seifert, I., Thieken, A. H., Lindquist, E., Wagner, K., & Merz, B. (2011). Recent changes in flood preparedness of private households and businesses in Germany. *Regional Environmental Change*, *11*(1), 59–71.
29. Lüdtke, S., Schröter, K., Steinhausen, M., Weise, L., Figueiredo, R., & Kreibich, H. (2019). A consistent approach for probabilistic residential flood loss modeling in Europe. *Water Resources Research*, 55. https://doi.org/10.1029/2019WR026213
30. Merz B, Elmer F, Kunz M, Mühr B, Schröter K, Uhlemann-Elmer S. The extreme flood in June 2013 in Germany. La Houille Blanche. 2014; 1(June 2013):5–10. https://doi.org/10.1051/lhb/2014001
31. Merz B, Kundzewicz ZW, Delgado J, Hundecha Y, Kreibich H (2012) Detection and attribution of changes in flood hazard and risk. In: Kundzewicz ZW (ed) Changes in flood risk in Europe. IAHS Special Publication 10:435–458
32. Merz, B., Kreibich, H., Schwarze, R., Thieken, A. (2010): Review article 'Assessment of economic flood damage'. - Natural Hazards and Earth System Sciences (NHESS), 10, 8, pp. 1697-1724. doi: http://doi.org/10.5194/nhess-10-1697-2010
33. Merz, B., Vorogushyn, S., Lall, U., Viglione, A., Blöschl, G. (2015): Charting unknown waters - On the role of surprise in flood risk assessment and management. - Water Resources Research, 51, 8, 6399-6416. https://doi.org/10.1002/2015WR017464
34. Molinari, D., Scorzini, A. R., Arrighi, C., Carisi, F., Castelli, F., Domeneghetti, A., Gallazzi, A., Galliani, M., Grelot, F., Kellermann, P., Kreibich, H., Mohor, G. S., Mosimann, M., Natho, S., Richert, C., Schroeter, K., Thieken, A. H., Zischg, A. P., and Ballio, F.: Are flood damage models converging to reality? Lessons learnt from a blind test, Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2020-40, in review, 2020.
35. Penning-Rowsell, E. C. and Chatterton, J. B.: The benefits of flood alleviation: A manual of assessment techniques, Gower Technical Press, Aldershot, 1977
36. Penning-Rowsell, Edmund & Priest, Sally & Parker, Dennis & Morris, Joe & Tunstall, Sylvia & Viavattene, Christophe & Chatterton, John & Owen, Damon. (2013). Flood and Coastal Erosion Risk Management A Manual for Economic Appraisal.

37. Polasky, S., Carpenter, S. R., Folke, C. and Keeler, B.: Decision-making under great uncertainty: environmental management in an era of global change, Trends in Ecology & Evolution, 26(8), 398–404, doi:10.1016/j.tree.2011.04.007, 2011.
38. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
39. Rözer, V., Kreibich, H., Schröter, K., Müller, M., Sairam, N., Doss-Gollin, J., ... & Merz, B. (2019). Probabilistic Models Significantly Reduce Uncertainty in Hurricane Harvey Pluvial Flood Loss Estimates. *Earth's Future*, *7*(4), 384-394.
40. Sairam, N., Schröter, K., Rözer, V., Merz, B., & Kreibich, H. (2019). Hierarchical Bayesian approach for modeling spatiotemporal variability in flood damage processes. Water Resources Research, 55. https://doi.org/10.1029/2019WR025068
41. Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F., & Merz, B. (2014). How useful are complex flood damage models?. *Water Resources Research*, *50*(4), 3378-3395.
42. Schröter, K., Kunz, M., Elmer, F., Mühr, B., & Merz, B. (2015). What made the June 2013 flood in Germany an exceptional event? A hydrometeorological evaluation. *Hydrology and Earth System Sciences*, *19*(1), 309–327. http://doi.org/10.5194/hess-19-309-2015
43. Scorzini A.R., Radice A., Molinari D., A New Tool to Estimate Inundation Depths by Spatial Interpolation (RAPIDE): Design, Application and Impact on Quantitative Assessment of Flood Damage, Water 2018, 10, 1805; doi:10.3390/w10121805
44. Scorzini, A. R. and Frank, E.: Flood damage curves: new insights from the 2010 flood in Veneto, Italy, J. Flood Risk Manage., 10, 381–392, https://doi.org/10.1111/jfr3.12163, 2017.
45. Smith, D. I.: Flood damage estimation – A review of urban stagedamage curves and loss functions, Water SA, 20(3), 231–238, 1994.
46. Szönyi, M., May, P. and Lamb, R. (2016). Flooding after Storm Desmond. UK: Zurich Insurance Group. https://www.jbatrust.org/wp-content/uploads/2016/08/flooding-after-storm-desmond-PUBLISHED-24-August-2016.pdf (last accessed 9 May 2019).
47. Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environmental modelling & software, 90, 201-216.
48. Thieken, A. H., Kreibich, H., Muller, M., & Merz, B. (2007). Coping with floods: Preparedness, response and recovery of flood-affected residents in Germany in 2002. *Hydrological Sciences Journal*, *52*(5), 1016–1037.
49. Vogel, K., Weise, L., Schröter, K., & Thieken, A. H. (2018). Identifying driving factors in flood-damaging processes using graphical models. *Water Resources Research*, *54*, 8864–8889. https://doi.org/10.1029/2018WR022858
50. Wagenaar, D., de Jong, J., and Bouwer, L. M.: Multi-variable flood damage modelling with limited data using supervised learning approaches, Nat. Hazards Earth Syst. Sci., 17, 1683–1696, https://doi.org/10.5194/nhess-17-1683-2017, 2017.
51. Wagenaar, D., Lüdtke, S., Schröter, K., Bouwer, L. M., & Kreibich, H. (2018). Regional and temporal transferability of multivariable flood damage models. Water Resources Research, 54, 3688– 3703. https://doi.org/10.1029/ 2017WR022233
52. Ward, P. J., Blauhut, V., Bloemendaal, N., Daniell, J. E., de Ruiter, M. C., Duncan, M., Emberson, R., Jenkins, S. F., Kirschbaum, D., Kunz, M., Mohr, S., Muis, S., Riddell, G., Schäfer, A., Stanley, T., Veldkamp, T. I. E., and Winsemius, H. C.: Review article: Natural hazard risk assessments at the global scale, Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2019-403, in review, 2019.
53. Wind, H. G., Nierop, T. M., de Blois, C. J., and de Kok, J. L.: Analysis of flood damages from the 1993 and 1995 Meuse floods, Water Resour. Res., 35, 3459–3466, 1999.

54. Winter, B., Schneeberger, K., Huttenlau, M. et al. Nat Hazards (2018) 91: 431. https://doi.org/10.1007/s11069-017-3135-5
55. WL Delft: Onderzoek watersnood Maas, Deelrapport 1: Wateroverlast December, Deelrapport 9: Schade, WL Delft, Delft, Appendix A, 1994 (in Dutch).
56. Zischg, A.P., Felder, G., Mosimann, M., Röthlisberger, V., Weingartner, R., Extending coupled hydrological-hydraulic model chains with a surrogate model for the estimation of flood losses, *Environmental Modelling and Software* (2018), doi: 10.1016/j.envsoft.2018.08.009