



## Estimating parameter values of a socio-hydrological flood model

Marlies Holkje Barendrecht<sup>1</sup>, Alberto Viglione<sup>1</sup>, Heidi Kreibich<sup>2</sup>, Sergiy Vorogushyn<sup>2</sup>, Bruno Merz<sup>2</sup>, and Günter Blöschl<sup>1</sup>

<sup>1</sup>Vienna University of Technology, Institute of Hydraulic Engineering and Water Resources Management, Vienna, Austria

<sup>2</sup>GeoForschungsZentrum Potsdam (GFZ), Section Engineering Hydrology, Telegrafenberg, Potsdam, Germany

**Correspondence:** Marlies Holkje Barendrecht (barendrecht@hydro.tuwien.ac.at)

Received: 31 December 2017 – Accepted: 10 February 2018 – Published: 5 June 2018

**Abstract.** Socio-hydrological modelling studies that have been published so far show that dynamic coupled human-flood models are a promising tool to represent the phenomena and the feedbacks in human-flood systems. So far these models are mostly generic and have not been developed and calibrated to represent specific case studies. We believe that applying and calibrating these type of models to real world case studies can help us to further develop our understanding about the phenomena that occur in these systems. In this paper we propose a method to estimate the parameter values of a socio-hydrological model and we test it by applying it to an artificial case study. We postulate a model that describes the feedbacks between floods, awareness and preparedness. After simulating hypothetical time series with a given combination of parameters, we sample few data points for our variables and try to estimate the parameters given these data points using Bayesian Inference. The results show that, if we are able to collect data for our case study, we would, in theory, be able to estimate the parameter values for our socio-hydrological flood model.

### 1 Introduction

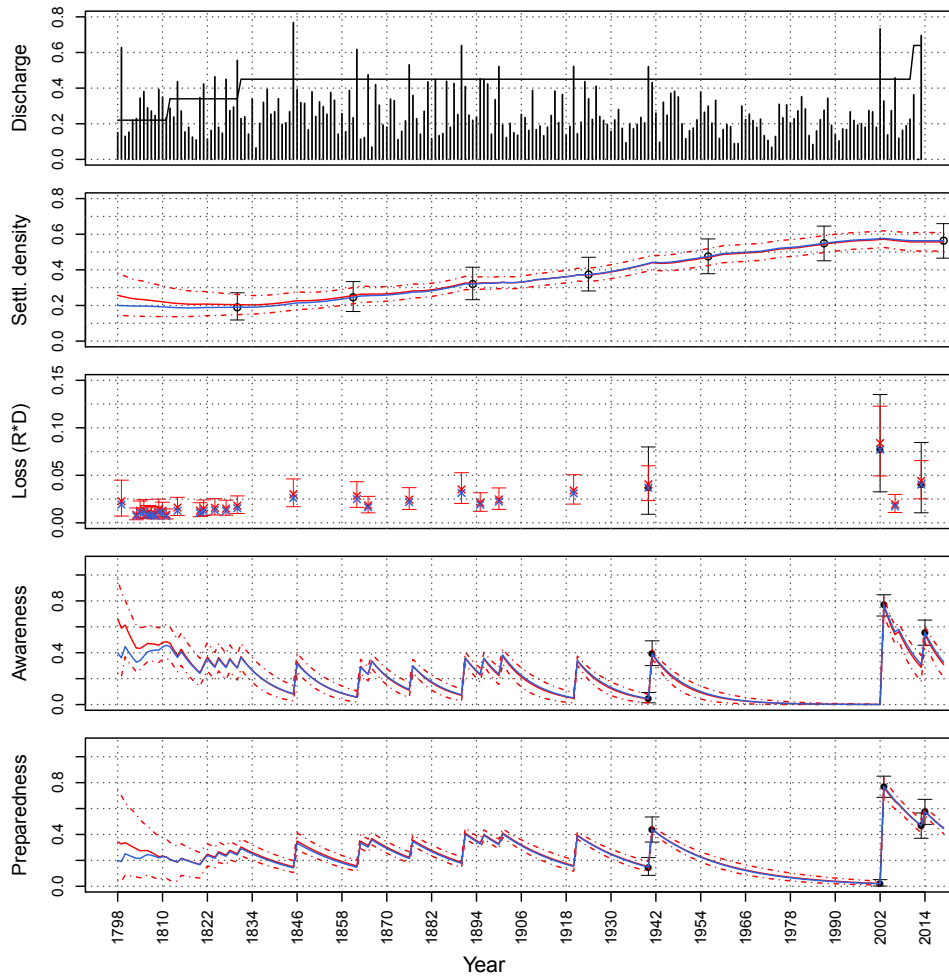
Socio-hydrology aims to study the long term feedbacks between humans and hydrology and tries to explain the phenomena that occur as a result of these feedbacks (Sivapalan and Blöschl, 2015; Sivapalan et al., 2012). So far, socio-hydrological flood models are mostly based on concepts derived from personal experience/observation of how the system works (Di Baldassarre et al., 2013, 2015, 2017; Grames et al., 2016; Grelot and Barreteau, 2012; Viglione et al., 2014). Most of them are general models that are not developed to reproduce and evaluate the dynamics of a specific case study and have not been compared to data. An exception to this is the work of Ciullo et al. (2017): they made a qualitative comparison between the model results of Di Baldassarre et al. (2015) and the human-flood system of the city of Rome and Bangladesh using data on population density, flood losses and levee heights. Chen et al. (2016) developed a socio-hydrological model to reflect the human-flood dynamics in the Kissimmee River Basin, Florida. Using flood

series and data about wetland area they made a qualitative assessment of the model's performance. While the socio-hydrological models and studies so far show that these models can yield valuable insights into human-flood dynamics, there is a lack of application to real world systems and none of the models so far have been calibrated to represent a specific case study. It is therefore of interest to assess whether calibrating socio-hydrological models is indeed feasible and understanding what are the data needed and their amount. Here we develop a socio-hydrological model and explore whether we are able to estimate the parameter values of this model using Bayesian Inference (Gelman et al., 2014) if we would have data available.

### 2 Methods

#### 2.1 Model structure

History has shown that repeated flood events may result in lower damages if flood events happen again, because peo-



**Figure 1.** Generated dynamics of system variables, sampled data points and estimated time series. The generated dynamics of the system variables is represented in blue. Sampled data points are shown in black, with their 90 % uncertainty bounds. Estimated time series are plotted in red, with the mean as a solid line and the 90 % credible bounds as dotted lines.

ple become aware of the flood risk and adapt. We want our model to capture these interactions between floods, awareness and preparedness. We model these interactions with the following variables: Floods ( $W$ ), Losses ( $L$ ), Relative Losses ( $R$ ), Settlement Density ( $D$ ), Awareness ( $A$ ) and Preparedness ( $P$ ). The behavior of these variables over time is described with the system of differential equations in Eq. (1).

$$L = RD \tag{1a}$$

$$R = \begin{cases} R_{\max} - \beta_R \exp\left(-\alpha_R (P_{\max} - P) \frac{W}{W_{\max}}\right), & \text{if } W > H \\ 0, & \text{otherwise} \end{cases} \tag{1b}$$

$$\frac{dD}{dt} = U(1 - \alpha_D A) D \left(1 - \frac{D}{D_{\max}}\right) \tag{1c}$$

$$\frac{dA}{dt} = \alpha_A L \left(1 - \frac{A}{A_{\max}}\right) - \mu_A A \tag{1d}$$

$$\frac{dP}{dt} = \alpha_P \frac{dA}{dt} \left(1 - \frac{P}{P_{\max}}\right) S(R) - \mu_P P \tag{1e}$$

$$S(R) = \tanh(1000R) \tag{1f}$$

High water levels ( $W$ ) (here called floods) and the protection level ( $H$ ) are external variables. For illustrative purposes, we use time series of the gauge at Dresden, Germany. If the water level ( $W$ ) is higher than the protection level ( $H$ ), this results in a relative loss ( $R$ ) as given in Eq. (1b).  $R$  increases exponentially to a maximum of one for a maximum flood. Together with the settlement density ( $D$ ), this results in an actual loss ( $L$ ), according to Eq. (1a). The actual loss is higher if the relative loss is higher or if the settlement density (and thus the exposure) is higher.

**Table 1.** Model parameter description and units ( $n_h$  = number of households,  $n_m$  = number of precautionary measures)

Parameter	Definition	Unit
$\alpha_D$	Risk taking attitude	$[1/(n_h/n_h)]$
$\alpha_A$	Scaredness	$[1/(\text{€}/\text{€})]$
$\mu_A$	Forgetfulness	$[1/t]$
$\mu_P$	Decay rate of precautionary measures	$[1/t]$
$\alpha_P$	Activeness	$[(n_m/n_m)t/(n_h/n_h)]$
$\alpha_R$	Effectiveness of preparedness	$[1/(n_m/n_m)]$

According to the literature, awareness (Eq. 1d) increases when households experience damage (Bradford et al., 2012; Bubeck et al., 2013; Osberghaus, 2015; Owusu et al., 2015; Poussin et al., 2014; Wachinger et al., 2013), and decreases exponentially over time with rate  $\mu_A$  as households forget (ICPR, 2002; Kreibich et al., 2011). We assume the increase in awareness depends on the size of the awareness as well, i.e. if awareness is high already it will increase less than when it is close to zero.

An increase in awareness may result in an increase in the uptake of precautionary measures and thus preparedness. How big this increase in preparedness is depends on other factors as well, like coping appraisal (Bubeck et al., 2012), maladaptive coping responses (Bubeck et al., 2013) or worry (Miceli et al., 2008; Raaijmakers et al., 2008). The fact that not all households that are aware of the risk take precautionary measures is represented by the parameter  $\alpha_P$ , which determines how much of an increase in awareness results in an uptake of precautionary measures and thus an increase in preparedness. As described by Eq. (1e), preparedness increases after a flood event with an amount relative to the change in awareness and depending on how high the preparedness was before the flood event (if households have already implemented many measures, there is less room for an increase in preparedness than when they have not implemented any measures). Preparedness only increases when damage occurs, i.e. when the relative loss is higher than zero. We approximate this step function with a tangens hyperbolicus (Eq. 1f). Like the awareness, preparedness decreases exponentially over time with a rate  $\mu_P$  because households forget and measures deteriorate.

According to the literature, implementing precautionary measures can greatly reduce the losses because of a flood (Kreibich et al., 2015; Poussin et al., 2015). So if preparedness is higher the relative losses will be lower. How much lower the relative losses will be depends on the parameter  $\alpha_R$  (Eq. 1b). As explained the losses depend on the relative losses and on the settlement density. The settlement density grows with a general growth rate  $U$ , but if awareness is higher this will reduce the growth rate and thus settlement density will grow slower or may even decline as described in Eq. (1c).

**Table 2.** Parameter values. Real values used to generate our data points and the mean and standard deviation of the posterior distribution estimated with Bayesian inference

Parameter	Real value	Estimated
$\alpha_D$	3	mean = 3.30, sd = 0.64
$\alpha_A$	10	mean = 9.80, sd = 0.07
$\mu_A$	0.1	mean = 0.11, sd = 0.02
$\mu_P$	0.05	mean = 0.05, sd = 0.01
$\alpha_P$	1	mean = 1.00, sd = 0.09
$\alpha_R$	0.2	mean = 0.23, sd = 0.07
$D_0$	0.2	mean = 0.26, sd = 0.07
$A_0$	0.4	mean = 0.67, sd = 0.21
$P_0$	0.2	mean = 0.34, sd = 0.22

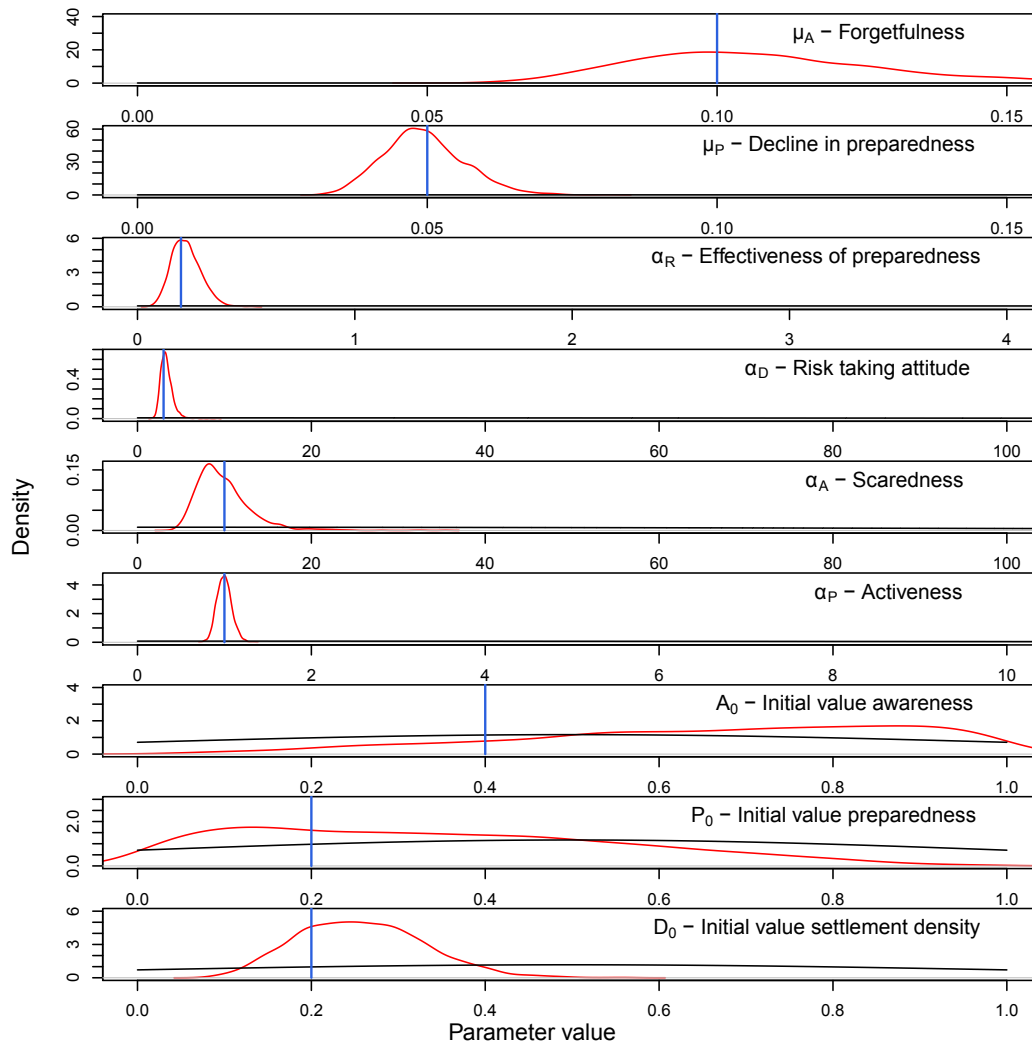
## 2.2 Bayesian inference

We want to estimate the values of the model parameters from data. The variables are scaled between zero and one. We assume  $\beta_R$  is one and try to estimate the values of the six parameters that do not have a value of one. These parameters are described in Table 1. In addition we need to estimate the starting values of the variables settlement density, awareness and preparedness, because we describe the evolution of these variables with a differential equation.

Since socio-hydrology studies both the human and the hydrological system, we have to deal with various types of data with different uncertainties. Bayesian inference allows us to include all of these different types of data. Bayes' Theorem (Eq. 2) tells us that the posterior distribution of our parameters depends on the likelihood of the data and our prior estimation of the parameter distributions:

$$p(\theta|y) = \frac{p(y|\theta) p(\theta)}{\int p(y|\theta) p(\theta) d\theta} \tag{2}$$

$p(\theta|y)$  is the estimated posterior distribution of our parameter values.  $p(\theta)$  represents our prior belief about the distribution of the parameters. If we have any information about our parameters beforehand we can specify an informative distribution, if we do not have any information we use non-informative distributions, which is what we will do here.  $p(y|\theta)$  gives us the likelihood of our data given the parameter values. Here we can incorporate our data and specify the un-



**Figure 2.** Prior (black) and posterior (red) distributions of the parameters. The “real” parameters to be estimated are represented by the blue vertical lines (see Table 2 for their values).

certainties. In most cases it is not possible to compute the integral in the denominator, therefore we approximate this using Markov chain Monte-Carlo (MCMC) methods (Gelman et al., 2014). In this case we use the software Stan (Carpenter et al., 2017) to perform the estimation. Stan uses Hamiltonian Monte Carlo sampling, which uses the gradient of the log probability to speed up convergence and parameter exploration (Stan Development Team, 2017).

### 3 Results

We assume that the parameters have the values given in Table 2 and simulate time series of all state variables (blue lines in Fig. 1). The question we ask is: if we could observe, with uncertainty, the variables only in a limited number of time points, would we be able to infer the temporal evolution of the variables or, which is analogous, would we be able to es-

timate the parameters of the model? To test whether this is possible, we sample some data points for each of the variables: Settlement Density ( $D$ ), Loss ( $L = R \cdot D$ ), Awareness ( $A$ ) and Preparedness ( $P$ ). These data points and their uncertainties are plotted in black in Fig. 1. Using these data points and uninformative priors for our parameters, we perform the Bayesian Inference using MCMC. Figure 1 shows the mean (solid red line) and 90 % credible bounds (dashed red lines) of the time series as estimated with the inference.

The top graph gives normalized discharges and protection level time series at the annual time scales. If the discharge is higher than the protection level this results in a loss as shown in the third graph. The size of this loss depends on the relative loss and the settlement density. If losses occur, this causes an increase in awareness (second graph from the bottom) and subsequently an increase in preparedness (bottom graph). Both the awareness and preparedness decrease grad-

ually over time when no flood occurs. The awareness does slightly influence the growth of the settlement density since every time a flood occurs the settlement density grows more slowly in the following years.

The credible bounds of the simulation are quite narrow, which means that the posterior distributions of our parameter values are narrow as well and that we are confident that we almost got the parameter values right. Figure 2 shows the real parameter values (used to generate our data points) in blue, the prior distributions that we used for the Bayesian inference in black and the posterior distributions estimated with the inference in red. The posterior mean estimates are very close to the real parameter values (see also Table 2). The estimates of the initial values are bit further off from the real values, which is not surprising given that we did not pick any data points at the start of our time series and therefore we do not have any information available to estimate the variables at the start. We used flat uninformative prior distributions and we are able to determine the posterior distributions quite well. The posterior distributions are not much influenced by the prior distributions and mainly depend on the data.

#### 4 Conclusions

We have conducted an artificial experiment where a Socio-Hydrological model is assumed whose “true parameters” are known to us, since we impose them, and observations are sampled from variables generated with the model. The aim of this paper is to assess whether the model can be calibrated to these observed data. This is not trivial because the model is highly nonlinear and it is not clear what amount of data is needed for calibration and, more importantly, whether the amount we can imagine to find in a real world case study can be enough. We demonstrate that, if the assumptions underlying the model are valid, we are able to estimate quite accurately the parameter values from relatively few data, which could be available in real case studies, using Bayesian Inference. Other estimation techniques, such as those used in rainfall-runoff modelling, could be used instead but the fact that Bayesian Inference is a flexible tool that can incorporate different types of information makes it ideal for the application to socio-hydrological models, where data are highly uncertain and prior information may be available to constrain model parameters.

The next step will be to gather data for a specific real world case study, develop and apply a model to it. This will pose a bigger challenge than it is in the case presented here, since the model will not represent correctly and fully the real world. Also, another challenge will be to map the information gathered on variables such as awareness and preparedness to the zero-to-one space adopted here. However it is promising to know that, based on what we show in this paper, if we can rely on the assumptions of the model so that the uncertainty is mainly due to not knowing its parameters,

we can actually calibrate a socio-hydrological flood model to the few data available.

**Data availability.** The analyses in the paper are based on generated data, with the exception of the time series of water levels of the river Elbe at Dresden, which is the result of the research project “Integration von historischen und hydraulisch/hydrologischen Analysen zur Verbesserung der regionalen Gefährdungsabschätzung und zur Erhöhung des Hochwasserbewusstseins 2005–2007” (BMBF-Projekt, 2007) (Uwe Grünwald and Sabine Schümborg, Hydrology and Water Resource Management, Brandenburg University of Technology BTU Cottbus, Germany).

**Competing interests.** The authors declare that they have no conflict of interest.

**Special issue statement.** This article is part of the special issue “Innovative water resources management – understanding and balancing interactions between humankind and nature”. It is a result of the 8th International Water Resources Management Conference of ICWRS, Beijing, China, 13–15 June 2018.

**Acknowledgements.** We would like to acknowledge funding from the Marie Skłodowska-Curie Innovative Training Network “A Large-Scale Systems Approach to Flood Risk Assessment and Management – SYSTEM-RISK” (Grant Agreement number: 676027) and the project “SozioHydroKopplung” funded by the German Bundesanstalt für Gewässerkunde. We would like to thank Hongrui Wang, Dingzhi Peng and one anonymous reviewer for their useful comments to the original version of the paper.

Edited by: Dingzhi Peng

Reviewed by: Hongrui Wang and one anonymous referee

#### References

- BMBF-Projekt: “Integration von historischen und hydrologisch/hydraulischen Analysen zur Verbesserung der regionalen Gefährdungsabschätzung und zur Erhöhung des Hochwasserbewusstseins”, BTU Cottbus, Cottbus, available at: [http://www-docs.tu-cottbus.de/hydrologie/public/files/Schlussbericht\\_BMBF0330686.pdf](http://www-docs.tu-cottbus.de/hydrologie/public/files/Schlussbericht_BMBF0330686.pdf), 2007.
- Bradford, R. A., O’Sullivan, J. J., van der Craats, I. M., Krywkow, J., Rotko, P., Aaltonen, J., Bonaiuto, M., De Dominicis, S., Waylen, K., and Schelfaut, K.: Risk perception – issues for flood management in Europe, *Nat. Hazards Earth Syst. Sci.*, 12, 2299–2309, <https://doi.org/10.5194/nhess-12-2299-2012>, 2012.
- Bubeck, P., Botzen, W. J. W., and Aerts, J. C. J. H.: A review of risk perceptions and other factors that influence flood mitigation behavior, *Risk Anal.*, 32, 1481–1495, 2012.
- Bubeck, P., Botzen, W. J. W., Kreibich, H., and Aerts, J.: Detailed insights into the influence of flood-coping appraisals on mitigation behaviour, *Global Environ. Chang.*, 23, 1327–1338, 2013.

- Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., and Riddell, A.: Stan: A Probabilistic Programming Language, *J. Stat. Softw.*, 76, 1–32, <https://doi.org/10.18637/jss.v076.i01>, 2017.
- Chen, X., Wang, D., Tian, F., and Sivapalan, M.: From channelization to restoration: Sociohydrologic modeling with changing community preferences in the Kissimmee River Basin, Florida, *Water Resour. Res.*, 52, 1227–1244, <https://doi.org/10.1002/2015WR018194>, 2016.
- Ciullo, A., Viglione, A., Castellarin, A., Crisci, M., and Di Baldassarre, G.: Socio-hydrological modelling of flood-risk dynamics: comparing the resilience of green and technological systems, *Hydrolog. Sci. J.*, 62, 880–891, <https://doi.org/10.1080/02626667.2016.1273527>, 2017.
- Di Baldassarre, G., Martinez, F., Kalantari, Z., and Viglione, A.: Drought and flood in the Anthropocene: feedback mechanisms in reservoir operation, *Earth Syst. Dynam.*, 8, 225–233, <https://doi.org/10.5194/esd-8-225-2017>, 2017.
- Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Salinas, J. L., and Blöschl, G.: Socio-hydrology: conceptualising human-flood interactions, *Hydrol. Earth Syst. Sci.*, 17, 3295–3303, <https://doi.org/10.5194/hess-17-3295-2013>, 2013.
- Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Yan, K., Brandimarte, L., and Blöschl, G.: Debates – Perspectives on socio-hydrology: Capturing feedbacks between physical and social processes, *Water Resour. Res.*, 51, 4770–4781, <https://doi.org/10.1002/2014WR016416>, 2015.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B.: Bayesian data analysis (Vol. 2), CRC press Boca Raton, FL, 2014.
- Grames, J., Prskawetz, A., Grass, D., Viglione, A., and Blöschl, G.: Modeling the interaction between flooding events and economic growth, *Ecol. Econ.*, 129, 193–209, <https://doi.org/10.1016/j.ecolecon.2016.06.014>, 2016.
- Grelot, F. and Barreteau, O.: Simulation of Resilience of an Insurance System to Flood Risk, International Environmental Modelling and Software Society (iEMSs) 2012 International Congress on Environmental Modelling and Software Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty, Sixth Biennial Meeting, Leipzig, Germany, 2012.
- ICPR (International Commission for the Protection of the Rhine): Non structural flood plain management – measures and their effectiveness, ICPR, Koblenz, 2002.
- Kreibich, H., Seifert, I., Thielen, A.H., Lindquist, E., Wagner, K., and Merz, B.: Recent changes in flood preparedness of private households and businesses in Germany, *Reg. Environ. Change*, 11, 59–71, <https://doi.org/10.1007/s10113-010-0119-3>, 2011.
- Kreibich, H., Bubeck, P., Van Vliet, M., and De Moel, H.: A review of damage-reducing measures to manage fluvial flood risks in a changing climate, *Mitigation and Adaptation Strategies for Global Change*, 20, 967–989, 2015.
- Miceli, R., Sotgiu, I., and Settanni, M.: Disaster preparedness and perception of flood risk: A study in an alpine valley in Italy, *J. Environ. Psychol.*, 28, 164–173, 2008.
- Osberghaus, D.: The determinants of private flood mitigation measures in Germany – Evidence from a nationwide survey, *Ecol. Econ.*, 110, 36–50, 2015.
- Owusu, S., Wright, G., and Arthur, S.: Public attitudes towards flooding and property-level flood protection measures, *Nat. Hazards*, 77, 1963–1978, 2015.
- Poussin, J. K., Botzen, W. J. W., and Aerts, J. C. J. H.: Factors of influence on flood damage mitigation behaviour by households, *Environ. Sci. Policy*, 40, 69–77, 2014.
- Poussin, J. K., Botzen, W. J. W., and Aerts, J. C. J. H.: Effectiveness of flood damage mitigation measures: Empirical evidence from French flood disasters, *Global Environmental Change*, 31, 74–84, 2015.
- Raaijmakers, R., Krywkow, J., and van der Veen, A.: Flood risk perceptions and spatial multi-criteria analysis: an exploratory research for hazard mitigation, *Natural Hazards*, 46, 307–322, 2008.
- Sivapalan, M. and Blöschl, G.: Time scale interactions and the co-evolution of humans and water, *Water Resour. Res.*, 51, 6988–7022, <https://doi.org/10.1002/2015WR017896>, 2015.
- Sivapalan, M., Savenije, H. H. G., and Blöschl, G.: Socio-hydrology: A new science of people and water, *Hydrol. Process.*, 26, 1270–1276, <https://doi.org/10.1002/hyp.8426>, 2012.
- Stan Development Team: Stan Modeling Language Users Guide and Reference Manual, Version 2.17.0, <http://mc-stan.org>, 2017.
- Viglione, A., Di Baldassarre, G., Brandimarte, L., Kuil, L., Carr, G., Salinas, J. L., and Bloeschl, G.: Insights from socio-hydrology modelling on dealing with flood risk – Roles of collective memory, risk-taking attitude and trust, *J. Hydrol.*, 518, 71–82, <https://doi.org/10.1016/j.jhydrol.2014.01.018>, 2014.
- Wachinger, G., Renn, O., Begg, C., and Kuhlicke, C.: The risk perception paradox – implications for governance and communication of natural hazards, *Risk Analysis*, 33, 1049–1065, 2013.