Dynamics in the Flood Vulnerability of Companies

Lukas Schoppa

Cumulative dissertation for the degree

Doctor of Engineering
(Dr.-Ing.)

in Natural Hazards Research

submitted to the Faculty of Science
at the University of Potsdam

prepared at the Section Hydrology
of the German Research Centre for Geosciences (GFZ)

Date of submission: 23rd of November 2022

First Supervisor: Prof. Dr. Bruno Merz
Second Supervisor: Dr. habil. Heidi Kreibich

First Reviewer: Prof. Dr. Bruno Merz
Second Reviewer: Dr. habil. Heidi Kreibich
Independent Reviewer: Prof. Dr. Matthijs Kok

Examination Board Members:
Prof. Dr. Bruno Merz
Dr. habil. Heidi Kreibich
Prof. Dr. Matthijs Kok
Prof. Dr. Oliver Korup
Prof. Dr. Annegret Thieken
Prof. Dr. Axel Bronstert

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-59242
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-592424

Declaration of originality

Potsdam, 23rd of November 2022

I hereby declare that, to the best of my knowledge, this thesis does not bear resemblance to any other work in whole or in part and has been completed by myself. I did not use any other sources and means than specified. Furthermore, this work has not been previously submitted to any university. All sources have been referred to and this work gives adequate credit to others for their work.

Lukas Schoppa

Acknowledgements

During the past four years that I have spent working on this dissertation, I have been supported by many wonderful people. This thesis would not have been possible without the help of these individuals, and I would like to express my deep gratitude to them.

First, I would like to thank my supervisor Heidi Kreibich for her close scientific guidance and support at every step of my PhD. I am impressed by her commitment to her doctoral students, regardless of her busy schedule, and could not have wished for better supervision. Furthermore, I would like to thank Bruno Merz for the opportunity to pursue my PhD in the Hydrology section of GFZ. His helpful feedback at many occasions improved this research in many ways. I am also grateful to Gert Zöller for being my mentor and assisting me during the first phase of this project.

I would also like to thank Matthijs Kok very much for serving as a reviewer for this thesis.

A special thanks goes to Tobias Sieg, who worked closely with me on all three studies and was always available to discuss modeling problems or when I was looking for advice on life as a PhD-student. I also thank Marlies Barendrecht for the close collaboration and for teaching me so much about socio-hydrological modeling. In addition, I would like to thank Nivedita Sairam for being a great office mate and co-author and for proofreading this thesis. Finally, I thank Kristin Vogel and Dominik Paprotny for contributing to this work in the form of helpful comments, data, and reviewing the manuscripts.

It was a pleasure to do the PhD in the DFG research training group "NatRiskChange", which was perfectly managed by Annegret Thieken, Axel Bronstert and Theresia Petrow. It is thanks to them and my fellow PhD-students, some of whom became dear friends, that these four years were a fantastic experience, both inside and outside the office. My heartfelt thanks also go to all the members of the Hydrology section of GFZ who, despite the undeniable abundance of nice buildings in the science park, managed to make C4 the most beautiful place on the hill.

Finally, I would like to sincerely thank my family and longtime friends for their unconditional support. Without them, I would not have reached this point. Most importantly, I thank you, Sophie, for your support, understanding, and love during the ups and downs of this journey and beyond.

Zusammenfassung

Flussüberschwemmungen sind eine ständige Gefahr für die Gesellschaft und verursachen jedes Jahr weltweit wirtschaftliche Schäden in der Größenordnung von 100 Milliarden US-Dollar. Im Zuge des globalen Wandels erhöht sich die Konzentration von Menschen und Vermögenswerten in Überschwemmungsgebieten kontinuierlich, während der menschengemachte Klimawandel Hochwasserextreme verstärkt. Die Überlagerung dieser Prozesse führt zu einer Verschärfung des Hochwasserrisikos in vielen Weltregionen. Der Hochwasseranapassung kommt dabei eine Schlüsselrolle bei der Abschwächung von Schäden zu. Allerdings ist das Verständnis von Hochwasservulnerabilität (d.h., Anfälligkeit gegenüber Schäden) und damit verbundener Dynamiken noch sehr begrenzt, was die Risikoabschätzung und die Entwicklung von Anpassungsstrategien erschwert.

In dieser kumulativen Dissertation werden anhand von drei Studien neue Methoden zur Hochwasserrisikoabschätzung für den gewerblichen Sektor vorgestellt, der in der Vergangenheit wenig untersucht wurde. Die erste Studie präsentiert Hochwasserschadensmodelle die auf statistischen Methoden und maschinellem Lernen basieren und eine Vielzahl von Einflussfaktoren berücksichtigen. In Verbindung mit probabilistischen Vorhersagen führt dies zu einer Verbesserung der Modellgenauigkeit und -verlässlichkeit. Anschließend wird in einer Pilotstudie für Dresden, Deutschland, eines der neuen Schadensmodelle in ein ganzheitliches systemdynamisches Modell integriert, um Veränderungen in Hochwasservulnerabilität und -risiko kontinuierlich zu simulieren. Die Methode integriert zusätzliche Prozessdetails und Kalibrierungsdaten in das Modell und verbessert so die Simulationsleistung. Schließlich werden mit dem systemdynamischen Modell in der dritten Studie langfristige Projektionsläufe durchgeführt, um die Entwicklung des Hochwasserrisikos bis zum Ende des Jahrhunderts abzuschätzen. Die Ergebnisse der Studie unterstreichen das Potential von Hochwasseranpassung - insbesondere in Zeiten des Klimawandels - und demonstrieren die Fähigkeit ganzheitlicher Modellierungsansätze, ungünstige Entwicklungen des Risikos frühzeitig aufzudecken. Insgesamt verbessert diese Arbeit die Darstellung der Vulnerabilität in der Hochwasserrisikoabschätzung, indem sie Modellierungslösungen anbietet,

die der Komplexität der Wechselwirkungen zwischen Mensch und Hochwasser gerecht werden und Unsicherheiten konsequent quantifizieren.

Summary

River flooding is a constant peril for societies, causing direct economic losses in the order of $100 billion worldwide each year. Under global change, the prolonged concentration of people and assets in floodplains is accompanied by an emerging intensification of flood extremes due to anthropogenic global warming, ultimately exacerbating flood risk in many regions of the world. Flood adaptation plays a key role in the mitigation of impacts, but poor understanding of vulnerability and its dynamics limits the validity of predominant risk assessment methods and impedes effective adaptation strategies. Therefore, this thesis investigates new methods for flood risk assessment that embrace the complexity of flood vulnerability, using the understudied commercial sector as an application example.

Despite its importance for accurate risk evaluation, flood loss modeling has been based on univariable and deterministic stage-damage functions for a long time. However, such simplistic methods only insufficiently describe the large variation in damage processes, which initiated the development of multivariable and probabilistic loss estimation techniques. The first study of this thesis developed flood loss models for companies that are based on emerging statistical and machine learning approaches (i.e., random forest, Bayesian network, Bayesian regression). In a benchmarking experiment on basis of object-level loss survey data, the study showed that all proposed models reproduced the heterogeneity in damage processes and outperformed conventional stage-damage functions with respect to predictive accuracy. Another advantage of the novel methods is that they convey probabilistic information in predictions, which communicates the large remaining uncertainties transparently and, hence, supports well-informed risk assessment.

Flood risk assessment combines vulnerability assessment (e.g., loss estimation) with hazard and exposure analyses. Although all of the three risk drivers interact and change over time, such dependencies and dynamics are usually not explicitly included in flood risk models. Recently, systemic risk assessment that dissolves the isolated consideration of risk drivers has gained traction, but the move to holistic risk assessment comes with limited thoroughness in terms of loss estimation and

data limitations. In the second study, I augmented a socio-hydrological system dynamics model for companies in Dresden, Germany, with the multivariable Bayesian regression loss model from the first study. The additional process-detail and calibration data improved the loss estimation in the systemic risk assessment framework and contributed to more accurate and reliable simulations. The model uses Bayesian inference to quantify uncertainty and learn the model parameters from a combination of prior knowledge and diverse data.

The third study demonstrates the potential of the socio-hydrological flood risk model for continuous, long-term risk assessment and management. Using hydroclimatic ad socioeconomic forcing data, I projected a wide range of possible risk trajectories until the end of the century, taking into account the adaptive behavior of companies. The study results underline the necessity of increased adaptation efforts to counteract the expected intensification of flood risk due to climate change. A sensitivity analysis of the effectiveness of different adaptation measures and strategies revealed that optimized adaptation has the potential to mitigate flood risk by up to 60%, particularly when combining structural and non-structural measures. Additionally, the application shows that systemic risk assessment is capable of capturing adverse long-term feedbacks in the human-flood system such as the levee effect.

Overall, this thesis advances the representation of vulnerability in flood risk modeling by offering modeling solutions that embrace the complexity of human-flood interactions and quantify uncertainties consistently using probabilistic modeling. The studies show how scarce information in data and previous experiments can be integrated in the inference process to provide model predictions and simulations that are reliable and rich in information. Finally, the focus on the flood vulnerability of companies provides new insights into the heterogeneous damage processes and distinct flood coping of this sector.

Contents

List of Figures

List of Tables

Abbreviations

BEINF	Zero-and-One-Inflated Beta Distribution
BN	Bayesian Network
BR	Bayesian Regression
BUI	Building
CDF	Cumulative Distribution Function
CRPS	Continuous Ranked Probability Score
EAD	Expected Annual Damage
EQU	Equipment
EURO-CORDEX	European branch of the Coordinated Downscaling Experiment
EUROPOP2019	Eurostat population projections
GDP	Gross Domestic Product
GNS	Goods and Stock
HANZE	Historical Analysis of Natural Hazards in Europe
ID	Identification
IPCC	Intergovernmental Panel on Climate Change
LOO-CV	Leave-One-Out Cross-Validation
MAE	Mean Average Error
MBE	Mean Bias Error
MCMC	Markov Chain Monte Carlo
NACE	Nomenclature statistique des Activités économiques dans la Communauté Européenne (Statistical Classification of Economic Activities in the European Community)
NUTS	Nomenclature des Unités Territoriales Statistiques (Nomenclature of Territorial Units for Statistics)

RCP . . . . Representative Concentration Pathways
RF . . . . Random Forest
ROPE . . . . Region Of Practical Equivalence
SDF . . . . Stage-Damage Function
SDF-D . . . . Stage-Damage Function - Deterministic
SDF-P . . . . Stage-Damage Function - Probabilistic
SI . . . . Supporting Information
TVAR . . . . Tail Value At Risk
UN . . . . United Nations
VAR . . . . Value At Risk

1 Introduction

1.1 Flood risk under global change

The Central-European floods in July 2021 were a stark reminder of the devastating impacts of river floods. The widespread event heavily impacted Belgium, the Netherlands, and Germany, causing long-lasting disruptions in society in affected regions and claiming a total of 228 lives (Munich Re 2022a). Economically, the 2021 floods caused a record-breaking $46-54 billion of overall losses, which is the highest ever documented loss figure of a flood event globally and any natural disaster in Europe (Bevere and Remondi 2022; Munich Re 2022b; Munich Re 2022c). While events like the 2021 flood in Europe stand out due to their exceptional severity in spatial extent, magnitude, and impacts, statistical records underline that flooding is a constant threat to humans and their assets.

Globally, hydrological perils are the most frequent natural hazard with respect to the number of relevant loss events (Bevere and Remondi 2022; Munich Re 2018; World Meteorological Organization 2022). In an average year, river floods are estimated to affect 54-58 million people and cause direct economic damage of $75-163 billion worldwide (Tanoue et al. 2021; Ward et al. 2017; Alfieri et al. 2017; Dottori et al. 2018; Winsemius et al. 2016). Over the past decades, absolute flood losses have increased as a result of population growth and accumulating wealth in flood-prone areas (Tanoue et al. 2016; Paprotny et al. 2018b; Visser et al. 2014). The prospects of future flood impacts are alarming as the growth of wealth will continue (Hirabayashi et al. 2013; Jongman et al. 2012) and anthropogenic climate change is projected to intensify flood frequency and magnitude in most regions of the world (Arias et al. 2021). In its most recent assessment report, the Intergovernmental Panel on Climate Change (IPCC) concludes that an increase in both exposed population and direct economic flood damages has to be expected with high confidence, independent of the level of global warming (Caretta et al. 2022). Projections of global economic flood risk in the 21st century suggest a potential increase in annual average loss by a factor of 4-25 assuming the most severe global warming scenarios and different degrees of risk reduction (Tanoue et al. 2021; Ward et al. 2017; Alfieri et al. 2017; Dottori et al. 2018; Winsemius et al. 2016). In view of these figures, societies have to increase their efforts to counteract the exacerbating flood risk and manage adverse consequences. *Flood adaptation* - that is, the adjustment of the human system to current or expected risk in order to moderate the harm or exploit beneficial opportunities (Ara Begum et al. 2022) - is considered a cornerstone in mitigating the costs of upcoming flood disasters (Jongman 2018; Kreibich et al. 2017a; Winsemius et al. 2016; Kinoshita et al. 2018; Willner et al. 2018). For the development of effective adaptation strategies, it is necessary to understand which drivers control flood risk and how these drivers change over time.

Flood risk is the potential for adverse consequences for a system (e.g., human, environmental, infrastructure) due to flooding. Depending on the context and the scope of the analysis, this can comprise intangible impacts such as loss of life or ecosystem services or tangible impacts such as physical damage or financial losses due to business interruption (Merz et al. 2010b; Kreibich et al. 2014). Flood risk arises from the interaction of three drivers: *hazard*, *exposure*, and *vulnerability* (Kron 2005). In the context of river flooding, the *hazard* is the magnitude (e.g., river discharge or inundation extent) and occurrence probability of a flood event. The *exposure* comprises the population and inventory of assets that are located in the flood zone and could potentially be impacted by a flood. The *vulnerability* describes the susceptibility of these exposed elements to be harmed by the hazard (e.g., physical flood damage). Vulnerability is determined by physical, social, economic, and environmental factors and varies across societies, geographic regions, or sectors (UNDRR 2022). Moreover, each of the three risk drivers can change over time. Flood hazard has already begun to change as a result of anthropogenic global warming. The exposure in floodplains has risen due to population growth or urbanization (Merz et al. 2021). Temporal changes in flood vulnerability are more difficult to detect but there exists empirical evidence that societies adapt to recurring flood events or, conversely, forget about the risk in flood-poor periods (Kreibich et al. 2017a; Di Baldassarre et al. 2015; Fanta et al. 2019). Consequently, humans shape flood risk continuously (Di Baldassarre et al. 2013) and, hence, have the capacity to actively mitigate flood impacts through flood risk management (Jongman et al. 2015; Winsemius et al. 2016).

Flood risk management aims at the modification of the hazard, exposure, and vulnerability with the overarching objective of reducing risk (Kreibich et al. 2022; Merz et al. 2010a). This includes structural measures that control the flood hazard such as river training, reservoirs, or dykes and increasingly also actions that concentrate on the reduction of exposure and vulnerability; for instance, early warning and emergency management, land use planning, awareness raising, and private precaution. Finally, there exist instruments for risk-transfer that finance the residual risk (e.g., insurance) (Jongman 2018; Merz et al. 2021; Dottori et al. 2020). In order to identify effective and economically feasible risk reduction strategies (i.e. combinations of measures), decision makers rely on quantitative estimates of risk.

Flood risk assessment provides a formalized framework for determining risk and serves as the basis for risk-informed management of flooding. The approach starts with a flood frequency analysis and inundation modeling (hazard analysis), which is followed by an identification of objects that are exposed to the flooding (e.g., through analyses with geographical information systems). Afterwards, flood loss models estimate the economic damage at affected objects (i.e., vulnerability), taking into account the hazard intensity (e.g., inundation depth) and assets characteristics (e.g., building type). Repeating this process for several flood events with different magnitude and, hence, chance of occurrence (e.g., through simulation) results in a probability distribution of economic flood losses, that is, flood risk (Merz and Thieken 2009; de Moel et al. 2015; Falter et al. 2015).

This thesis advances methods for fluvial flood risk assessment in relation to vulnerability. The focus of the research lies on direct economic losses, that is, monetary damage to assets that occurs in affected areas as a direct consequence of a flood. The following section outlines the status quo, recent developments, and research gaps in this field and looks at the subject from two perspectives: (i) complexity in vulnerability and risk modeling and (ii) uncertainty and data analysis.

1.2 Flood risk assessment - recent advancements and research gaps

1.2.1 Complexity of flood vulnerability

Flood risk systems are complex due to the interactions between hydrological processes and human activity. People alter floodplains in multiple ways; for example, for economic benefit, settlement development, flood protection, or recreation. Often, the overall implications of decisions for such coupled systems are hard to evaluate and consequences of today's actions might only emerge after decades (Gober and Wheater 2015; Sivapalan and Blöschl 2015; Liu et al. 2007). Several factors characterize a coupled human and natural system; that is, non-linearity (e.g., threshold effects), interdependence and feedback loops, time lags, non-stationarity, and heterogeneity. These properties are also present in human-flood systems and the reason, why flood risk is difficult to understand and manage (Liu et al. 2007; Merz et al. 2015). Together with cognitive biases in human perception, complexity is considered a main source of surprising, catastrophic flood disasters (Merz et al. 2021; Merz et al. 2015). The characteristic factors of complexity appear at multiple scales in the flood risk system, which impedes the risk assessment and modeling.

Loss estimation

At the micro-scale, flood damage processes are highly complex. The damage grade of an object (i.e., buildings, contents, goods) in case of a flood is governed by a multitude of factors that can be classified as impact variables describing the flood intensity (e.g., inundation depth, flow velocity, inundation duration) and resistance variables that are characteristic for the exposed asset (e.g., building type and material, occupation, precaution, emergency measures) (Merz et al. 2010b; Thieken et al. 2005).

The intricate interplay of the physical damage processes is poorly understood and highly variable, so that risk analysts use empirical models to derive loss estimates. On the basis of damage data that is collected in the aftermath of flood events or synthetic data derived from what-if analyses so-called stage-damage functions (or vulnerability functions) are derived that relate the dominant flood impact variable - usually inundation depth - to the damage grade of the exposed asset (Merz et al. 2010b). Stage-damage functions are the most prevalent loss estimation tool in flood risk assessment due to low demands towards input data, simple model structure, and straightforward transferability. They are often developed individually for different regions, asset groups, or sectors (Scawthorn et al. 2006; Alfieri et al. 2016; Huizinga et al. 2017; Penning-Rowsell et al. 2005). However, such univariable loss models capture the complexity of damage processes insufficiently since they neglect relevant mechanisms and interactions between damage influencing parameters (Middelmann-Fernandes 2010; Kelman and Spence 2004; Gissing and Blong 2004).

An increasing body of literature proposes the use of multivariable loss models that include several predictor variables at once and their interdependency. Examples range from advanced multivariable regression models (Van Ootegem et al. 2015; Rözer et al. 2019), over rule-based models (Kreibich et al. 2010; Thieken et al. 2008), to machine learning methods such as Bayesian networks (Wagenaar et al. 2018; Vogel et al. 2014) or decision trees (Carisi et al. 2018; Merz et al. 2013). The inclusion of additional factors improved the predictive capacity of the loss models in the majority of cases (Thieken et al. 2008; Wagenaar et al. 2017; Rözer et al. 2019; Schröter et al. 2014; Merz et al. 2013; Kreibich et al. 2017b). Another advancement of these state-of-the-art loss models is their capability to capture complex patterns and non-linearity in the damage data. Loss observations are commonly highly dispersed with strongly skewed or bimodal distributions, which arises from threshold effects such as rounding of reported losses or insurance write-offs. Machine learning or mixture models (e.g., inflated Beta regression) allow for more flexibility in the functional model form and, hence, better accommodate the superposition of different processes in the data generating process (Vogel et al. 2014; Schröter et al. 2014; Van Ootegem et al. 2015; Rözer et al. 2019). Despite larger requirements towards input data than stage-damage functions, the practical value of complex loss models has been confirmed in model transferability and upscaling exercises (Steinhausen et al. 2022; Lüdtke et al. 2019; Wagenaar et al. 2018; Sieg et al. 2019b; Molinari et al. 2020).

However, a review of 47 flood loss models revealed that this new generation of loss models does not cover all sectors equally well (Gerl et al. 2016). Model development has focused on the residential sector for the most part since damage data for other sectors such as the commercial sector are more scarce and heterogeneous (Gissing and Blong 2004; Sieg et al. 2017; Merz et al. 2010b). Additionally, the growing number of diverse modeling approaches makes proper model validation, comparison, and benchmarking even more important in order to ensure model validity and facilitate an informed model selection.

Systems-level dynamics

The complexity of flood risk also manifests itself at the systems-scale (e.g., an entire floodplain). Feedbacks between the risk components can cause unforeseen, seemingly paradox risk dynamics such as the levee effect. The levee effect occurs when improvements in flood protection (e.g., levee heightening) reduce the frequency of floods, which eventually leads to a stimulation of urban development in the floodplain and a decline in risk awareness. The reduction of flood hazard through enhanced protection is outweighed by increasing exposure and vulnerability and causes an overall rise in flood risk (Montz and Tobin 2008). Another example is the adaptation effect, which describes the phenomenon that particularly flood-rich (i.e., hazardous) periods are often accompanied by decreasing vulnerability and flood losses. Experiencing repeated flood losses triggers a learning process in society, which becomes more aware of the flood risk and improves its preparedness. The levee and adaptation effect are both examples of feedbacks and non-linearity in flood risk systems and have been supported by empirical evidence at diverse study sites around the globe (Kreibich et al. 2017a; Di Baldassarre et al. 2015). Yet, conventional approaches to flood risk assessment do not account for these factors of complexity and, thus, fall short of replicating such phenomena (Di Baldassarre et al. 2016).

Commonly, flood risk models consider narrowly defined systems that put large emphasis on the representation of hazard processes; for example, by involving extensive hydrological and hydraulic computations. This also applies to risk dynamics, where flood hazard is often the only risk component that is considered non-stationary by accounting for climate change. Changes in the exposure are increasingly considered (e.g., population, GDP, asset values), but risk assessments that account for dynamics in vulnerability are still an exception (Metin et al. 2018). In some instances, risk simulations are conducted under the assumption of different vulnerability scenarios (Jongman et al. 2015; Winsemius et al. 2016; Metin et al. 2018; Steinhausen et al. 2022), but the majority of the existing flood risk models lacks the structural complexity to capture and replicate interactions between hazard, exposure, and vulnerability. That is, the three risk drivers are treated as isolated and - in the case of vulnerability - often stationary quantities that do not influence each other. Apart from that, there exists a mismatch between the temporal range (or legacy) of flood management decisions and the treatment of time in prevalent risk assessment tools. Processes in human-flood systems differ in their temporal scale from fast shocks (e.g. flood loss event) to slow transitions (e.g. climate change or degrading risk awareness in flood-poor periods), where the individual time scales influence each other (Thompson et al. 2013; Sivapalan and Blöschl 2015). However, dynamics and feedbacks are often neglected by assuming stationarity or independence between processes, particularly when looking at slow-onset, long-term processes (Merz et al. 2014a).

With the recent establishment of socio-hydrology, an interdisciplinary field that studies the dynamics and co-evolution of coupled human-water systems, a new generation of holistic flood risk models emerged (Sivapalan et al. 2012; Sivapalan et al. 2014). Socio-hydrological flood risk models explicitly include human activity and the interaction between the risk components in their framework (e.g., through causal model links). Their systemic design allows for the establishment of temporal continuity in simulations and extends the temporal scope of the analyses (e.g., to centuries). Agent-based models and system dynamics models are the most prevalent socio-hydrological approaches used in flood risk assessment (Barendrecht et al. 2017; Ross and Chang 2020). Agent-based models operate on the micro-scale and resolve individual agents (e.g., a private household) that act according to a set of predefined decision rules on the basis of theories from social sciences (e.g. behavioural economics) (Blair and Buytaert 2016). The interactions and aggregated decisions of all agents eventually lead to emerging patterns on the macro-scale (i.e., flood risk dynamics) (Aerts et al. 2018). Agent-based models have already been used to study the adaptive behavior of households and governments (Haer et al. 2017; Haer et al. 2019), the levee-effect (Haer et al. 2020), flood-induced business interruption (Coates et al. 2014; Coates et al. 2019), flood risk communication (Haer et al. 2016), or early warning (Alonso

Vicario et al. 2020). In contrast, system dynamics models are lumped and directly explain the macro-scale system behavior through coupled differential equations that describe the governing system processes conceptually. Hence, the focus in system dynamics lies on the overall workings and behaviour of a system rather than on its sub-scale processes (e.g., physical, economic, psychological) (Blair and Buytaert 2016; Barendrecht et al. 2017). System dynamics models have been utilized successfully in simulating floodplain dynamics focusing on questions such as economic impacts and growth (Di Baldassarre et al. 2013; Di Baldassarre et al. 2015; Viglione et al. 2014; Grames et al. 2016), risk coping and management (Ciullo et al. 2017; Yu et al. 2017; Di Baldassarre et al. 2017), or flood perception and memory (Ridolfi et al. 2020; Ridolfi et al. 2021; Song et al. 2021).

Both approaches are united by the objective of integrating vulnerability and its dynamics in flood risk assessment frameworks. Studies showed that the negligence of vulnerability dynamics can lead to an overestimation of flood risk and biased decision-making (Haer et al. 2016; Aerts et al. 2018). The inclusion of vulnerability as a variable instead of a constant in systemic risk modeling is expected to advance simulations and the understanding of vulnerability change, which is essential for the development and implementation of effective flood adaptation strategies (UNDRR 2022; Mechler and Bouwer 2015). To date, the majority of systemic modeling solutions that account for vulnerability dynamics investigate stylized problems as the assembly of socio-hydrological model data is difficult (Barendrecht et al. 2017; Troy et al. 2015a). As with loss estimation, previous studies of flood risk dynamics have largely focused on private households so that knowledge on changes in vulnerability and risk for the commercial sector remains scarce.

1.2.2 Uncertainty and data analysis

Flood risk assessment is subject to uncertainty, which arises from a variety of sources and at different scales and steps of the analysis (Apel et al. 2004; de Moel et al. 2015). Uncertainty is classified into aleatory and epistemic uncertainty. While aleatory uncertainty describes the natural variability (i.e., randomness) of a process and cannot be controlled, epistemic uncertainty originates from a lack of knowledge and can, in principle, be reduced by gathering additional information. For instance, the timing and magnitude of flooding as a natural process is inherently random (i.e., aleatory), whereas measurement and model uncertainty emerge from erroneous observations or simplified mathematical descriptions of reality, which can potentially be improved (i.e., epistemic) (Merz and Thieken 2009; Merz et al. 2010a). Developments such as multivariable loss modeling and systemic risk assessment successfully diminish epistemic model uncertainty by including additional relevant variables and processes, but considerable uncertainty in the model outputs remains. Predictive errors of multivariable loss models are still substantial due to scarce damage data, heterogeneity across assets, and a restricted understanding of the causal mechanisms of flood damage (Meyer et al. 2013; Merz et al. 2010b). Similarly, not all processes in a floodplain that influence risk dynamics can be incorporated in socio-hydrological models because of lacking knowledge or computational constraints (Di Baldassarre et al. 2016; Blair and Buytaert 2016). Throughout the individual steps of risk assessment the uncertainty accumulates, necessitating proper uncertainty quantification. The disregard of relevant uncertainty sources could provide unreliable information to decision makers and lead to dysfunctional flood risk management (Apel et al. 2008; Hall and Solomatine 2008).

In order to account for uncertainty, modellers make use of model ensembles, sensitivity analysis, or forward uncertainty propagation (Figueiredo et al. 2018; Metin et al. 2018; Haer et al. 2017; de Moel and Aerts 2011; Merz et al. 2015). Additionally, inverse approaches that use probability theory to quantify uncertainty in a model and its parameters from data are becoming increasingly popular. In loss estimation, deterministic loss models are still prevalent, but probabilistic models that return loss distributions instead of point estimates are gaining traction (Gerl et al. 2016). The prevalence of machine learning and statistical methods such as random forests (Sieg et al. 2019b; Sieg et al. 2019a), Bayesian networks (Lüdtke et al. 2019; Vogel et al. 2014; Wagenaar et al. 2018; Paprotny et al. 2020), or Bayesian regression (Rözer et al. 2019; Sairam et al. 2019a; Mohor et al. 2021) boosts this trend as these approaches often inherently quantify uncertainty. In models for flood risk assessment, simulation-based approaches (e.g., Monte Carlo methods) are common, which propagate the uncertainty in the input variables through a risk modeling chain (Apel et al. 2008; Falter et al. 2015). Typically, such analyses are concentrated on the hazard component of the risk system, for instance, the stochasticity in flood discharge. Methods that estimate and combine the uncertainties in all determinants of flood risk in a homogeneous way are rare. Socio-hydrological flood risk models alleviate this shortcoming, as all relevant risk drivers are fully integrated in one model. For example, Barendrecht et al. (2019) demonstrated how Bayesian parameter estimation can be used to jointly quantify uncertainty across the variables of a systemic flood risk model (i.e., exposure, vulnerability, loss). Equally important, the model can combine uncertainties that originate from different sources (data, parameter, statistical) in a coherent probabilistic framework.

The combination of socio-hydrological modeling and Bayesian inference for systemic, long-term risk assessment could create synergies. The inclusion of additional dynamics and variables in flood loss and risk models commonly enhances the number of model parameters and the demands towards the amount of data which are required to inform the parameters. In particular, socio-hydrological models require extensive observational records (i.e., decades to centuries) and heterogeneous data in order to make inferences about slow feedback loops and diverse system processes. However, for certain variables data are only available for the most recent period and at discrete points in time (e.g., loss records) or not directly but in form of proxy variables (e.g., surveys for vulnerability) (Troy et al. 2015a; Blair and Buytaert 2016). Historical information - such as land use maps or loss reports - can be useful in estimating the state of a model variable decades ago, but observations that lie far in the past are commonly more uncertain than recently collected data (Barendrecht et al. 2019). The Bayesian framework is very flexible in incorporating all kinds of information such as findings from previous experiments (e.g., as priors) or data that were gathered with measurement uncertainty (Schoot et al. 2021). Yet the potential of Bayesian methods in systemic flood risk modeling requires further investigation.

Altogether, a selective and partial uncertainty analysis in loss and risk estimates is problematic and could lead to mismanagement and ill-informed decision making (Pappenberger and Beven 2006; Sayers et al. 2012; Hall and Solomatine 2008). The established treatment of uncertainty in flood risk assessment seeks to identify probable outcomes rather than all possible alternatives, leaving space for surprising, potentially devastating events (Merz et al. 2015; Di Baldassarre et al. 2016). While more complexity in flood modeling could widen the scope of risk estimates, the increased number of parameters and variables in socio-hydrological and loss modeling exacerbates the challenge of data scarcity. Consequently, there is a demand for modeling solutions that quantify and analyse uncertainty comprehensively, optimize the use of available information, and ideally also reduce the uncertainties in flood analyses.

1.3 Research objectives and outline

Flood loss and risk are the result of a complex interplay of the hydrological and the human domain. However, the established approaches in loss modeling and risk assessment neglect most of the complexity due to the lack of relevant damage-influencing factors, narrow definition of the risk system, and assumptions of stationarity. Although such oversimplifications introduce substantial bias in the resulting risk estimates, formalized approaches for a comprehensive exploration of uncertainty and possible system evolutions are lacking.

Therefore, this thesis aims to improve the representation of vulnerability in the modeling of economic flood loss and risk. In particular, the research investigates and compares methods for object-level loss estimation, seeks to advance the vulnerability assessment in socio-hydrological risk modeling, and explores the role of vulnerability change in flood risk dynamics. Since previous model development in the field predominantly concentrated on the residential sector, the focus of this work should lie on the commercial sector (i.e., small and medium sized companies). The **overarching research questions** of this thesis are:

1. How can flood loss and risk models better account for the complexity of vulnerability processes?
2. What is required to enhance the scope of uncertainty analysis in flood risk assessment and to optimize the use of available information?
3. Which new insights into company vulnerability does this model development provide?

As shown in Figure 1.1, this cumulative thesis is composed of five chapters. Chapter 1 explains the relevance of flood risk assessment and provides an overview of the scientific status quo and recent advancements in the field. Chapters 2-4 present original research on flood loss estimation and risk assessment in form of three manuscripts. Finally, Chapter 5 synthesises the results and findings of the three studies and provides recommendations for further research in the field. The three manuscripts build on each other methodologically, and all contribute with individual aspects to answering the research questions.

The first study develops multivariable, probabilistic flood loss models for companies on the object-level and evaluates their capacity to improve loss predictions. The new loss models consider various damage-influencing variables, take into account heterogeneity, and quantify predictive uncertainty.

The second study integrates one of the multivariable loss models into a socio-hydrological flood risk model for companies in Dresden, Germany. The model coupling augments the systemic risk model with process understanding and additional data sources and provides new insights on flood vulnerability dynamics in the past.

The third study uses the same augmented socio-hydrological flood risk model for the long-term projection of flood vulnerability and risk dynamics. It demonstrates how systemic risk assessment solutions can combine scenario-based and probabilistic uncertainty analysis to elicit the potential evolutions of future flood risk. The study also highlights the practical value of holistic risk assessment for effective and anticipatory flood adaptation.

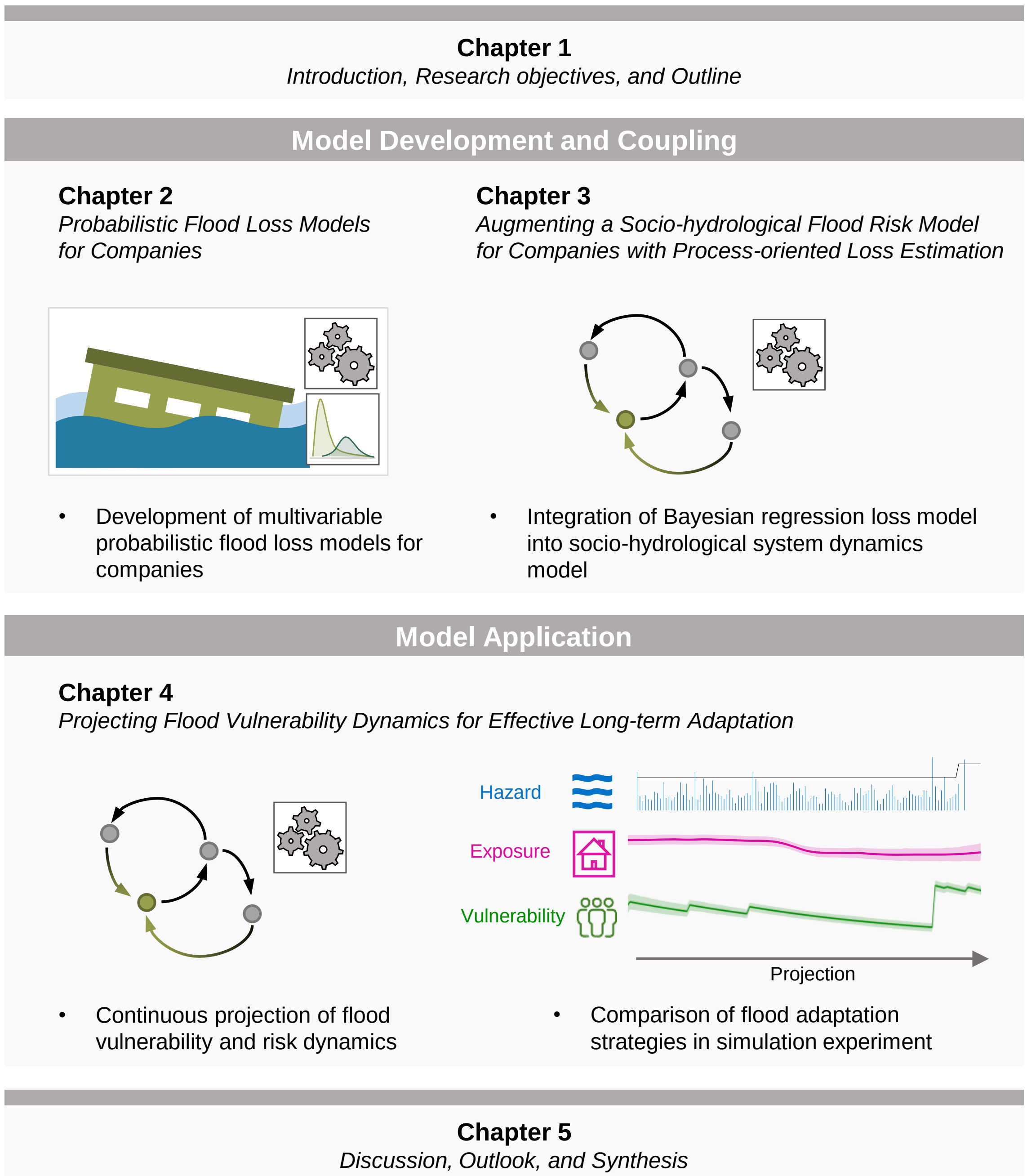

Figure 1.1: *Structure of this thesis.*

1.4 Author contributions

The main body of this thesis is composed of three manuscripts, that were published in peer-reviewed academic journals or are under review for publication. The majority of the presented work has been conducted by the author of this thesis (L.S.). The co-authors of the three studies contributed to the research in the form of discussions, comments, conceptualization, reviewing of manuscripts, or joint model development. Following the CRediT taxonomy (Brand et al. 2015), the author contributions for the individual manuscripts are as follows:

Chapter 2: Conceptualization: L.S., T.S., H.K.; Data curation: L.S., H.K.; Formal analysis: L.S.; Funding acquisition: H.K; Investigation: L.S., T.S., K.V., G.Z., H.K.; Methodology: L.S., T.S., K.V., H.K.; Project administration: H.K.; Software: L.S.; Supervision: G.Z., H.K.; Validation: L.S.; Visualization: L.S.; Writing – original draft: L.S.; Writing – review & editing: L.S., T.S., K.V., G.Z., H.K.

Chapter 3: Conceptualization: L.S., M.B., T.S., H.K. Data curation: L.S., M.B., H.K. Formal analysis: L.S. Funding acquisition: H.K. Investigation: L.S., M.B., T.S., N.S., H.K. Methodology: L.S., M.B., T.S., N.S., H.K.; Project administration: H.K.; Software: L.S., M.B.; Supervision: H.K.; Validation: L.S.; Visualization: L.S.; Writing – original draft: L.S.; Writing – review & editing: L.S., M.B., T.S., N.S., H.K.

Chapter 4: Conceptualization: L.S., H.K.; Data curation: L.S., D.P., H.K.; Formal analysis: L.S.; Funding acquisition: H.K.; Investigation: L.S., M.B., D.P., T.S., N.S., H.K.; Methodology: L.S., M.B., D.P., T.S., N.S., H.K.; Project administration: H.K.; Software: L.S., M.B., D.P.; Supervision: H.K.; Validation: L.S.; Visualization: L.S.; Writing – original draft: L.S.; Writing – review & editing: L.S., M.B., D.P., T.S., N.S., H.K.

Additionally, the author contributed to the following publications, which are not included in this thesis:

Schoppa, L., Disse, M., and Bachmair, S. (2020). Evaluating the performance of random forest for large-scale flood discharge simulation. Journal of Hydrology, 590, 125531. `https://doi.org/10.1016/j.jhydrol.2020.125531`

Schoppa L., Kreibich H., Sieg T., Vogel K. and Zöller G. (2021). Developing multivariable probabilistic flood loss models for companies [Paper presentation]. FLOODrisk 2020 - 4th European Conference on Flood Risk Management, Online-Conference. `https://doi.org/10.3311/floodrisk2020.11.12`

Berghäuser, L., **Schoppa, L.**, Ulrich, J., Dillenardt, L., Jurado, O. E., Passow, C., Mohor, G. S., Seleem, O., Petrow, T., Thieken, A. H. (2021). Starkregen in Berlin - Meteorologische Ereignisrekonstruktion und Betroffenenbefragung. Potsdam. `https://doi.org/10.25932/publishup-50056`

Caldas-Alvarez, A., Augenstein, M., Ayzel, G., Barfus, K., Cherian, R., Dillenardt, L., Fauer, F., Feldmann, H., Heistermann, M., Karwat, A., Kaspar, F., Kreibich, H., Lucio-Eceiza, E. E., Meredith, E. P., Mohr, S., Niermann, D., Pfahl, S., Ruff, F., Rust, H. W., **Schoppa, L.**, Schwitalla, T., Steidl, S., Thieken, A. H., Tradowsky, J. S., Wulfmeyer, V., Quaas, J. (2022). Meteorological, impact and climate perspectives of the 29 June 2017 heavy precipitation event in the Berlin metropolitan area. Natural Hazards., Earth System Sciences, 22(11), 3701-3724. `https://doi.org/10.5194/nhess-22-3701-2022`

2 Probabilistic Flood Loss Models for Companies

Authors:
Lukas Schoppa
Tobias Sieg
Kristin Vogel
Gert Zöller
Heidi Kreibich

Published as:
Schoppa, L., Sieg, T., Vogel, K., Zöller, G., and Kreibich, H., 2020. Probabilistic Flood Loss Models for Companies. Water Resources Research, 56 (9). https://doi.org/10.1029/2020WR027649.

Abstract

Flood loss modeling is a central component of flood risk analysis. Conventionally, this involves univariable and deterministic stage-damage functions. Recent advancements in the field promote the use of multivariable and probabilistic loss models, which consider variables beyond inundation depth and account for prediction uncertainty. Although companies contribute significantly to total loss figures, novel modeling approaches for companies are lacking. Scarce data and the heterogeneity among companies impede the development of company flood loss models. We present three multivariable flood loss models for companies from the manufacturing, commercial, financial, and service sector that intrinsically quantify prediction uncertainty. Based on object-level loss data (n = 1,306), we comparatively evaluate the predictive capacity of Bayesian networks, Bayesian regression, and random forest in relation to deterministic and probabilistic stage-damage functions, serving as benchmarks. The company loss data stem from four post event surveys in Germany between 2002 and 2013 and include information on flood intensity, company characteristics, emergency response, private precaution, and resulting loss to building, equipment, and goods and stock. We find that the multivariable probabilistic models successfully identify and reproduce essential relationships of flood damage processes in the data. The assessment of model skill focuses on the precision of the probabilistic predictions and reveals that the candidate models outperform the stage-damage functions, while differences among the proposed models are negligible. Although the combination of multivariable and probabilistic loss estimation improves predictive accuracy over the entire data set, wide predictive distributions stress the necessity for the quantification of uncertainty.

2.1 Introduction

Flooding poses immense risk to life and economic goods. Over the past four decades, 40% of globally recorded natural catastrophes were caused by pluvial or fluvial flooding and the share of hydrological events is rising (Munich Re 2018). Severe fluvial flooding such as the 2002 event (Ulbrich et al. 2003) or 2013 event (Merz et al. 2014b) in Germany can harm all components of society such as private households, infrastructure, or economy. Damage to companies constitutes a high share of total flood losses. For instance, businesses accounted for € 1.4 billion (15.9%) of the total direct flood loss of € 9.1 billion in 2002 (Mechler and Weichselgartner 2003). In the 2013 flood, companies suffered € 1.3 billion (19%) of the total € 6.7 billion damage (German Federal Ministry of the Interior 2013; Thieken et al. 2016). Despite the substantial contribution of companies to overall damage, previous flood loss research addressed residential damage for the most part (Gerl et al. 2016; Gissing and Blong 2004).

Flood risk assessment comprises the evaluation of flood hazard, exposure, and vulnerability (Merz et al. 2010a; Olsen et al. 2015). Vulnerability describes the susceptibility of exposed assets, such as buildings or contents, to sustain damage during a flood. The assessment of monetary loss through loss models represents a cornerstone in flood risk analysis and directly influences flood management practice, for instance in the cost-benefit analysis of flood management measures or in the calculation of insurance premiums (Merz et al. 2010b). Conventionally, flood loss estimation engages univariable stage-damage functions, which relate the hazard intensity at an asset, that is, inundation depth, to the damage grade or absolute damage (Alfieri et al. 2016; Grigg and Helweg 1975; Huizinga et al. 2017; White 1945). Most flood loss models feature a variety of distinct stage-damage functions differentiating between occupancy (e.g., residential, commercial, and industrial), asset type (e.g., building, contents, and equipment), and asset characteristics (e.g., building type, building material, and number of stories). Several models include explicit stage-damage functions for the commercial and industrial sector, for instance, the Multi-Coloured Manual (Penning-Rowsell et al. 2005), HAZUS-MH (Scawthorn et al. 2006), the stage-damage functions of the International Commission for the Protection of the Rhine (2016), or the global data set of stage-damage functions by Huizinga et al. (2017). Still, stage-damage functions often omit other damage influencing factors such as inundation duration or preparedness (Kelman and Spence 2004; Middelmann-Fernandes 2010; Thieken et al. 2005) and, more importantly, cannot account for interactions among the variables. As a result, stage-damage functions can only partially describe the damage processes (Gissing and Blong 2004; Merz et al. 2004; Rözer et al. 2019;

Schröter et al. 2014; Sieg et al. 2019b). The advance of machine learning and data mining promoted the development of multivariable flood loss models, which jointly consider a variety of damage influencing factors and their interdependency. The modeling community proposed ample methods for flood loss estimation including multivariate generalized regression (Rözer et al. 2019; Van Ootegem et al. 2015; Zhai et al. 2005), rule-based models (Elmer et al. 2010; Kreibich et al. 2010; Thieken et al. 2008), tree-based approaches (Carisi et al. 2018; Hasanzadeh Nafari et al. 2016b; Kreibich et al. 2017b; Merz et al. 2013; Sultana et al. 2018), and Bayesian networks (Lüdtke et al. 2019; Vogel et al. 2012; Vogel et al. 2014; Wagenaar et al. 2018).

Another advantage of such flood loss models is their ability to quantify the predictive uncertainty in their loss estimates. By returning predictive distributions instead of deterministic point estimates, probabilistic models inherently provide reliability information alongside their predictions (e.g., Lüdtke et al. 2019; Rözer et al. 2019; Sieg et al. 2019b; Wagenaar et al. 2018). Despite the evidently large uncertainties governing loss estimation, only a small number of existing models is probabilistic (Gerl et al. 2016). However, the explicit consideration of predictive uncertainty bears concrete value for flood risk management practice. For instance, flood loss estimates are central components of risk-based decision making in flood protection planning (Merz and Thieken 2009; Wagenaar et al. 2016). Decision-making frameworks such as expected utility theory or multicriteria analysis regard uncertainty information as integral for evaluating competing protection strategies (Brito and Evers 2016; Kreibich et al. 2014; Kunreuther et al. 2013). Probabilistic loss models inherently provide this uncertainty information and, hence, fit neatly into different decision support tools (Lüdtke et al. 2019). In this context, they represent an alternative to multimodel ensembles of deterministic flood models (see e.g., Figueiredo et al. 2018), where a sufficient number of models is lacking or the setup of an ensemble is too expensive. Furthermore, Sieg et al. (2019b) showed that probabilistic loss models can aid in bridging the gap between flood risk assessment at different scales, as they provide more accurate and informative loss estimates than deterministic models on the object level and are capable of propagating predictive uncertainty to aggregated levels, such as municipalities or states. Since both modelers and decision makers benefit from the transparent communication of uncertainty in damage estimates, further efforts should aim at the implementation of probabilistic loss models (Merz et al. 2010b; Meyer et al. 2013).

Company flood loss models that account for variable interactions and predictive uncertainty at the same time are still an exception. Several aspects impede the development of novel flood loss estimation techniques for companies. First, the damage processes of companies and residential buildings differ, which, in turn, requires the separate setup of company loss models (Merz et al. 2010b). Second, companies are more heterogeneous than private households, for instance, with respect to building type, size, or occupancy. Namely, company equipment ranges from heavy machinery over technical devices to office items depending on the business sector, whereas the composition of the contents varies less across private households, and the size of companies ranges from self-employed persons to production facilities with large numbers of employees, while household sizes range in the same order of magnitude. This heterogeneity reflects in the loss data as variance (Gissing and Blong 2004). Third, flood loss data are scarce and often inaccurate, especially for companies (Merz et al. 2010b; Molinari et al. 2014; Seifert et al. 2010; Sieg et al. 2017). Examples of multivariable flood loss models for companies that account for variable interactions are the empirical-synthetic FLFAcs model (Hasanzadeh Nafari et al. 2016a), the rule-based FLEMOcs model (Kreibich et al. 2010; Seifert et al. 2010), and the random forest model of Sultana et al. (2018). Sieg et al. (2017) and Sieg et al. (2019b) explored the capability of random forests to predict company flood loss for different economic sectors and spatial scales. Despite the necessity of proper model benchmarking (Gerl et al. 2016), an intercomparison of different multivariable probabilistic company flood loss models is still missing.

In this study, we present three multivariable probabilistic flood loss models for companies: Bayesian networks, Bayesian zero-and-one-inflated beta regression, and random forest. These models performed well in loss prediction exercises for the residential sector (Rözer et al. 2019; Schröter et al. 2014), where they outperformed other approaches such as rule-based models, probabilistic Gaussian regression models, or deterministic stage-damage functions; but except for random forest they have not been implemented for companies to date. The random forest model for companies of Sieg et al. (2017) and Sieg et al. (2019b) achieved promising performance scores but has not yet been tested against equally complex models. We aim at closing these gaps by implementing the models for the estimation of company flood loss and conducting a thorough comparison of their predictive capacity on basis of the same data. Since the three candidate models can deal with multidimensional, heterogeneously scaled model data and return predictive distributions of flood loss, they fulfill the requirements of modern flood loss estimation and match the highly variable company loss data. We benchmark the proposed models against a probabilistic and a deterministic stage-damage function, serving as standard reference models. We fit and validate all models separately for direct tangible loss to the company assets building (BUI), equipment (EQU), and goods and stock (GNS) on the basis of object-level company loss data (n = 1,306) collected in postevent surveys in Germany between 2002 and 2013. The multivariable candidate models use information on flood intensity, company characteristics, and private precaution to estimate the flood damage, whereas the stage-damage functions solely depend on water depth. The objective of this study is the comparative examination and assessment of

1. the predictive capacity of multivariable models against the established, univariable stage-damage functions and

2. differences in predictive power among the multivariable probabilistic candidate models

with a particularly focus on probabilistic forecasting. The results of this study offer new insights into flood damage processes of companies, the added value of complex modeling approaches, and the potential of probabilistic modeling in flood risk assessment.

2.2 Data and methods

2.2.1 Survey data

The empirical company flood loss data used in this study stem from four individual postevent surveys after major floods in Germany that occurred in the period from 2002 to 2013 (Kreibich et al. 2007; Thieken et al. 2016). The survey questionnaires remained consistent over all four surveys and gathered information on flood intensity, company characteristics, emergency and private precautionary measures, flood experience, and flood loss. Large flood events in the Danube and Elbe river catchments in 2002 and 2013 contribute the largest share (n = 1,014) to the total number of 1,346 completed company interviews. The remaining company loss data were collected in the aftermath of events in 2005, 2006, and 2010–2012 in the Danube, Elbe, Oder, and Rhine catchments. The data set is dominated by small- and medium-sized companies with less than 250 employees. For details on the survey data set and the collection methodology see Kreibich et al. (2007).

Table 2.1 lists a subset of variables from the survey data set, which we used for modeling in this study. The selection of the variable subset is primarily based on the studies of Kreibich et al. (2010) and Sieg et al. (2017), in which the authors quantitatively investigated variable importance with respect to relative loss on subsets of the same survey data. Furthermore, the composition of the predictor set was influenced by existing residential flood loss models (Elmer et al. 2010; Schröter et al. 2014; Wagenaar et al. 2018). In the following we motivate the predictor set and reference to studies, where the predictor was identified as influential or used

Table 2.1: *Predictor (n=8) and response (n=1) model variables.*

Variable	Abbreviation	Scale[a], unit, range
Predictors		
Flood intensity		
Water depth	wd	c: 0-960 cm above ground
Inundation duration	dur	c: 0-720 h
Return period	rp	c: 1-909 a
Company characteristics		
Size	size	c: 1-800 employees [-]
Business sector	sec	n: (1) manufacturing, (2) commercial, (3) financial, (4) service
Spatial situation	spat	n: (1) premises with several buildings, (2) one entire building, (3) one or more floors in shared building, (4) less than one floor in shared building
Experience and precaution		
Flood experience	exp	o: zero previous floods to five or more previous floods (6 classes)
Precaution ratio	pre	c: 0-1 [-]
Response		
Flood loss		
Relative loss to building	rloss	c: 0-1 [-]
Relative loss to equipment	rloss	c: 0-1 [-]
Relative loss to goods/stock	rloss	c: 0-1 [-]

Note. The rightmost column provides the observed ranges of each variable in the survey dataset.

[a] c: continuous, n: nominal, o: ordinal

in a loss model. Water depth (Kreibich et al. 2010; Penning-Rowsell et al. 2005; Scawthorn et al. 2006; Sieg et al. 2017) and inundation duration (Kreibich et al. 2010; Merz et al. 2013; Sieg et al. 2017; Vogel et al. 2014; Wagenaar et al. 2017) describe the intensity of the damaging flood event and are widely used in the prediction of flood loss. We augmented the surveyed flood intensity information on water depth and inundation duration by regionalized estimates of flood return periods (Elmer et al. 2010; Merz et al. 2013; Wagenaar et al. 2017; Wagenaar et al. 2018). Regional return period estimates provide additional insight on the general magnitude of the flood independent of spatially volatile inundation depths.

Table 2.2: *Precaution classes and measures. We estimate the degree of precaution for each company on the basis of the listed measures.*

Classification	Precautionary measure
Adaptation	Adapted use of flood-prone area
	Relocation of susceptible equipment
Mitigation	Improve flood resilience of building; e.g. basement waterproofing
	Installation of water barriers
Emergency	Saving equipment / saving goods and stock[a]
	Use of water pumps
	Shut-down of machinery and power
	Preventing contamination

[a] Ordered answer with four levels: from 0='nothing was saved' to 3='everything was saved'; this measure is possible for all companies

Moreover, return periods allow for implications on the flood experience of affected companies, since severe events might have an impact on infrequently inundated neighborhoods with low risk awareness (Elmer et al. 2010). The calculation of the return period estimates involved a statistical extreme value analysis of time series of annual maximum discharge at river gauges in affected regions and was carried out in analogy to Elmer et al. (2010). Company characteristics are included into the model through the business sector in which the company operates, the company size expressed by the number of employees, and the spatial situation of the premises at the affected site (Kreibich et al. 2010; Sieg et al. 2017). We assume that the flood experience of a company is tied to the number of previous floods that the company experienced (Kreibich et al. 2010; Merz et al. 2013; Schröter et al. 2014; Wagenaar et al. 2018). In that sense, companies that were flooded once or several times before the surveyed event exhibit higher flood experience than companies, which never encountered flooding before.

For the assessment of companies' flood precaution (Kreibich et al. 2007; Kreibich et al. 2010; Thieken et al. 2008; Vogel et al. 2018), we computed a ratio from a set of individual adaptation, mitigation, and emergency measures similar to Sieg et al. (2017). In contrast to Sieg et al., we combined adaption, mitigation, and emergency measures in one precaution ratio in order to reduce the number of predictor variables. The precaution ratio is defined as the number of precautionary measures that a specific company actually implemented prior to the damaging flood (nI) divided by the number of relevant measures that this company could have possibly implemented (nP)

$$pre = \frac{nI}{nP}. \tag{2.1}$$

Hence, company precaution is a ratio on the interval [0, 1], where well-prepared companies are assigned high ratios and poorly prepared companies are assigned low ratios. The observed values range from 4 to 10 for nP and from 0 to 10 for nI. The individual measures from which the precaution ratio was calculated are listed in Table 2.2. Except for the measure "saving equipment/saving goods and stock", which allowed for ordered answers depending on the amount of saved assets, all measures are treated as binary variables meaning that they were either implemented at the occurrence of the flood or not.

The damage to assets is expressed relative to their replacement value in order to facilitate the transferability of the derived models in space and time (Merz et al. 2010b). Consequently, losses to building, equipment, and goods and stock are a ratio on the interval [0, 1], where a relative loss of 0 corresponds to no damage and a relative loss of one corresponds to the total loss of the asset.

Furthermore, we excluded companies (n = 8) with extraordinary long-lasting inundation durations (>30 days) since we found evidence for erroneous survey answers in these cases. Prior to the model derivation, we removed companies with missing predictor values and subdivided the resulting data set (n = 1,306) into three asset-specific data sets (n_{bui} = 545, n_{equ} = 829, n_{gns} = 928).

Figure 2.1 shows the distributions of the predictor and response variables for the three asset-specific data sets in the form of violin plots (Hintze and Nelson 1998). The variable distributions are estimated through kernel density estimation (Silverman 2018). The response variable, relative loss, contains considerable shares of no (value: 0) and total (value: 1) loss cases for building (0: 32%, 1: 4%), equipment (0: 37%, 1: 17%), and goods and stock (0: 51%, 1: 20%). This results in bimodality of the relative loss distributions, which is particularly pronounced for equipment and goods and stock.

2.2.2 Development of probabilistic loss estimation models

Random forest

Random forest (RF) is a machine learning technique, which uses ensembles of decision trees for classification and regression problems (for details see Breiman 2001; Liaw and Wiener 2002). RFs are capable of handling high-dimensional, nonlinear data and offer large flexibility as they accept discrete and continuous predictors at the same time (James et al. 2013).

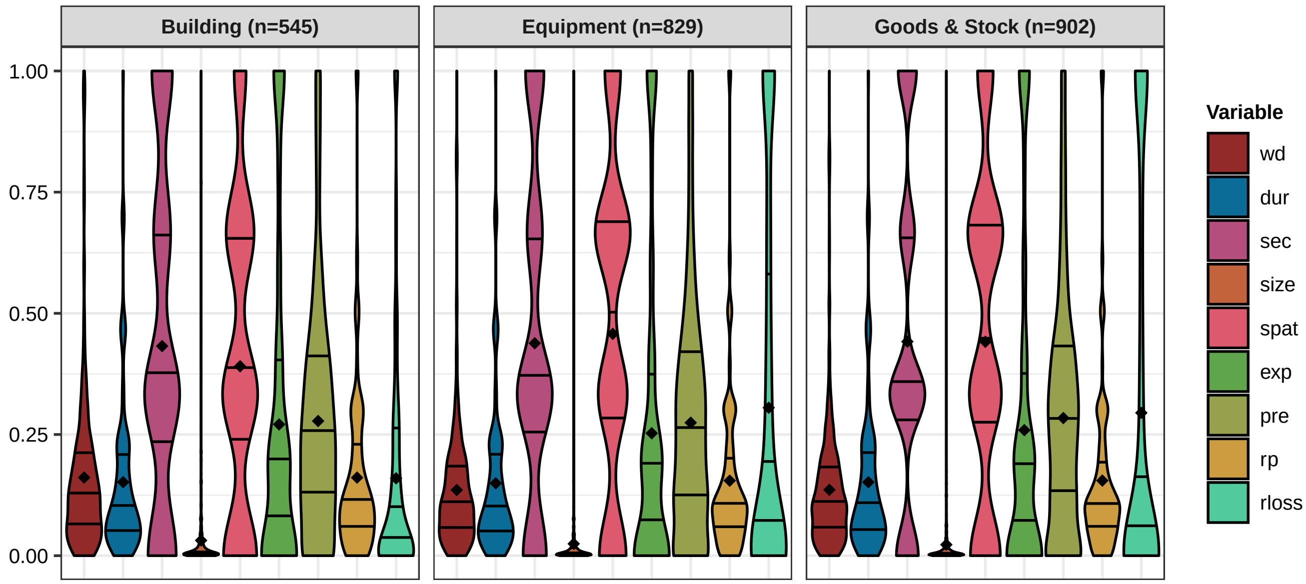

Figure 2.1: *Kernel density estimations of the model variable distributions for the three company assets building, equipment, and goods and stock. For this plot, we scaled all variables from zero to one. The lines in the violin plots indicate the quartiles while the dot represents the mean.*

RF is a supervised learning algorithm, which fits a large number of individual decision trees to data. The tree ensemble draws its predictive power from two techniques: bootstrap aggregation (bagging) and random feature selection. Bagging generates bootstrap samples of the original data before growing the trees and aggregates predictions of the individual trees afterward. During tree construction, random feature selection constrains the set of possible split variables at each splitting node, introducing additional randomness. The combination of bagging and random feature selection decreases the correlation among trees, which prevents overfitting and increases the prediction accuracy of the forest. Moreover, RFs inherently provide estimates of predictor importance. During tree construction, the algorithm randomly permutes each predictor and tracks the resulting mean decrease in prediction accuracy of the RF. A strong deterioration in prediction accuracy indicates that the respective predictor is more relevant for the predictive capacity of the RF.

The standard implementation of RF uses the classification and regression tree algorithm to construct the individual decision trees by recursively partitioning the training data into homogeneous subsets (Breiman et al. 1984). However, during recursive partitioning, this algorithm favors predictors with many possible splits (e.g., continuous variables) over predictors with few splits (e.g., categorical variables), leading to a variable selection bias (White and Liu 1994). Hothorn et al. (2006) developed a recursive partitioning routine based on permutation tests, termed conditional inference tree algorithm, which overcomes this bias. Since the company loss data set used in this study consists of continuous, ordinal, and nominal variables, we used the conditional inference tree algorithm. In addition, we obtained conditional response distributions of relative loss instead of mean values by employing the quantile regression forest methodology of Meinshausen (2006). The majority of previous studies on flood loss modeling used the conventional classification and regression tree algorithm (e.g., Carisi et al. 2018; Kreibich et al. 2017a; Merz et al. 2013; Schröter et al. 2014), but recent works increasingly applied the conditional inference tree algorithm (Sieg et al. 2017; Sultana et al. 2018) or a combination of conditional inference trees and quantile regression forests (Sieg et al. 2019b; Sieg et al. 2019a), which we also applied in this study.

Our RF model is controlled by two parameters, the number of trees n_{tree} and the number of randomly sampled predictors m_{try} during partitioning. We decided for a common parameter choice with $n_{tree} = 1,000$ and $m_{try} = 3$ (Hastie et al. 2009; Liaw and Wiener 2002). The supporting information (SI) to this paper provide information on the computational implementation of the RF model (Hothorn and Zeileis 2015).

Bayesian network

A Bayesian network (BN) is a probabilistic graphical model. It does not distinguish between predictor and response variable but represents the joint probability distribution of all variables in form of a directed acyclic graph (for details see Jensen and Nielsen 2007; Nagarajan et al. 2013; Pearl 2009). BNs encode the statistical dependence structure of the random variables into a set of nodes and arcs. Each variable is symbolized by a node, while the conditional dependence or independence of two variables is expressed by the presence or absence of a connecting arc between their corresponding nodes. This independence mapping of a BN facilitates efficient probabilistic computation as the global, joint distribution of the variable set can be factorized into a product of local, conditional probability distributions.

In theory, BNs are applicable to continuous and discrete variables. Yet in practice, continuous BNs are usually restricted to normally distributed variables in order to maintain closed-form expressions of the associated probability distributions (Scutari 2010). Since our flood loss data contain both discrete and continuous variables, which partly have skewed distributions, we implement discrete BNs in this study. The factorized formulation of the joint probability distribution for a discrete BN reads

$$P(X_1, ..., X_n) = \prod_{i=1}^{n} P(X_i \mid \Pi_{X_i}), \tag{2.2}$$

where X_i are all n variables of the BN and Π_{X_i} are the respective parent nodes of Xi in the directed acyclic graph, that is, nodes whose arcs point toward X_i. In a discrete BN, all probability distributions are multinomial, and the local distributions of the nodes are defined in conditional probability tables, which represent the parameters of the model (Nagarajan et al. 2013; Scutari and Denis 2014).

Consequently, the implementation of a BN requires (1) the definition of the graph structure and (2) the estimation of the conditional probability table values. We learned three separate network structures and their corresponding parameters to receive individual BN models for the company assets building, equipment, and goods and stock. For prediction we employed Bayesian inference.

The continuous variables in the survey data demanded for adjustments before we could use them for learning and prediction in a discrete BN. Therefore, we binned all continuous variables into intervals (Koller and Friedman 2009; Vogel et al. 2012; Vogel et al. 2014) by means of an equal-frequency discretization scheme (e.g., Wagenaar et al. 2018). This discretization routine calculates interval boundaries in a way that the resulting bins contain an equal amount of observations. In the interest of model accuracy, we assigned 10 bins to the presumably most influential predictors water depth and precaution ratio (Kreibich et al. 2010; Sieg et al. 2017) and to the target variable relative loss. The number of classes for the other continuous variables inundation duration, company size, and return period was set to 5. By definition, a discrete BN returns a probability mass function of the target variable. However, the other two candidate models of this study provide continuous predictive distributions on the interval [0,1] for the relative loss. For the purpose of comparability, we derive a continuous probability density for the binned BN, by fitting a distribution to data that we sampled from the observed relative loss cases with sampling weights according to the probability that the BN predicted. For further details on the BNs we refer to the SI (Højsgaard 2012).

Bayesian regression

In the Bayesian regression (BR) (for details see Gelman et al. 2013; McElreath 2018), we model relative loss with a zero-and-one-inflated beta distribution. The conventional beta distribution is a common choice for modeling fractional data, which range from 0 to 1 such as relative loss (Ferrari and Cribari-Neto 2004).

However, the beta distribution is not defined on those boundaries and, hence, cannot reproduce extreme cases of no (0) or total loss (1), which are abundant in the study data. The zero-and-one-inflated beta distribution (Ospina and Ferrari 2010) overcomes this limitation by combining the beta with the Bernoulli distribution, which accounts for the excess in zeros and ones in the model data. The resulting mixture distribution has the following cumulative distribution function (CDF):

$$BEINF\left(y \mid \lambda, \gamma, \mu, \phi\right) = \lambda \cdot F_{Bernoulli}\left(y \mid \gamma\right) + (1-\lambda) \cdot F_{Beta}\left(y \mid \mu, \phi\right) \tag{2.3}$$

where y is the response, relative loss, λ is the zero-and-one-inflation probability (i.e., the probability that the response is zero or one), $F_{Bernoulli}\left(\cdot \mid \gamma\right)$ is the cumulative distribution function of the Bernoulli distribution with parameter γ, which is the conditional one-inflation probability (i.e., the probability that the response is one rather than zero). $F_{Beta}\left(\cdot \mid \mu,\ \phi\right)$ is the cumulative distribution function of the reparameterized beta distribution with μ and ϕ as mean and precision parameter (Ferrari and Cribari-Neto 2004).

We configure the BR as a distributional model, which means that not only the mean μ of the beta distribution is predicted but also the remaining parameters λ, γ, and ϕ. We use different sets of predictor variables, $X_\lambda, X_\gamma, X_\mu, X_\phi$, for each parameter, receiving the following functions in the regression model

$$Y_i \sim BEINF\left(\lambda_i, \gamma_i, \mu_i, \phi_i\right) \tag{2.4}$$

$$logit\left(\mu_i\right) = \alpha_\mu + \beta_\mu X_{\mu,i} \tag{2.5}$$

$$\log\left(\phi_i\right) = \alpha_\phi + \beta_\phi X_{\phi,i} \tag{2.6}$$

$$logit\left(\lambda_i\right) = \alpha_\lambda + \beta_\lambda X_{\lambda,i} \tag{2.7}$$

$$logit\left(\gamma_i\right) = \alpha_\gamma + \beta_\gamma X_{\gamma,i} \tag{2.8}$$

where Y_i denotes the response variable for observation i (i.e., the relative loss of one company) and $X_{par,i}$ the respective values of the predictor variables for the corresponding parameter. α_{par} and β_{par} are the intercept and regression coefficients of the corresponding parameter in the combined regression model.

We estimate the mean μ of the beta distribution from all available predictors. In contrast, the inflation parameters λ and γ, and the precision parameter ϕ are predicted by a selection of the most influential predictor variables for the respective asset. In this way, we reduce the number of model parameters, which improves the model convergence during parameter estimation and accounts for differences in the flood damage processes across the assets. The analysis of Sieg et al. (2017), which has been conducted on a subset of the data used in this study, suggests that the spatial situation of a company is a major factor for flood loss

Table 2.3: *Predictor sets of the Zero-and-One-Inflated Beta regression. The predictors vary over the parameters and over the assets. Differences between the models for building and equipment/goods and stock are indicated in italics.*

	Building	Equipment/goods and stock
μ – beta mean	all predictors	all predictors
ϕ – beta precision	water depth, precaution	water depth, precaution
λ – zero-and-one-inflation	water depth, precaution, *spatial situation*	water depth, precaution, *sector*
γ – conditional one-inflation	water depth, precaution, *spatial situation*	water depth, precaution, *sector*

to buildings. Furthermore, the predictor importance measures for equipment, and goods and stock vary particularly strong across the economic sectors of the companies. Water depth and precaution exhibited high predictor importance across all assets. Table 2.3 shows which variables we used for predicting the zero-and-one-inflated beta parameters in the individual asset loss models. Before model fitting, we transformed continuous predictors by a Yeo-Johnson transformation in order to treat the pronounced skewness in the predictors variables (Yeo 2000). In addition, we centered and scaled continuous predictors. The regression terms contain individual coefficients for each level of the categorical predictors, sector and spatial situation (i.e., dummy encoding; McElreath 2018), and we model the ordinal variable flood experience as a monotonic effect (Bürkner 2017).

BR models require the definition of priors for model parameters as well as specifications for Markov chain Monte-Carlo (MCMC) sampling (Gelman et al. 2013; McElreath 2018). Details on the model implementation, including prior choice and MCMC-settings, are provided in the SI.

Comparison to stage-damage function

We compare the previously presented candidate models to univariable stage-damage functions (SDF). Stage-damage functions predict flood loss solely from water depth, *wd*, and represent the conventional approach to flood loss modeling (Merz et al. 2010b). Following similar studies (Rözer et al. 2019; Schröter et al. 2014; Sieg et al. 2019b; Wagenaar et al. 2017), we employ a square root stage-damage function as a baseline model for comparison. Square root stage-damage functions outperformed other functional forms (linear, polynomial) before (Elmer et al. 2010) and are arguably the most common instance of a SDF (Wagenaar et al. 2017). We implement a deterministic (SDF-D) and a probabilistic version (SDF-P) of the square root stage-damage function, in order to differentiate between the added value of multivariable and probabilistic prediction separately.

The deterministic SDF represents an established standard approach to flood loss estimation. The model is a simple, least square regression, which is defined as follows

$$Y_i = \alpha + \beta\sqrt{wd_i} + \varepsilon_i, \tag{2.9}$$

where Y_i is the observed relative loss, α and β are the intercept and regression coefficient, and ε_i is the error for observation i. During model fitting, the error sum of squares is minimized.

We implement the probabilistic SDF in a Bayesian framework in order to assure comparability with the probabilistic candidate models. Like in the BR model, we assume that relative loss follows a zero-and-one-inflated Beta distribution. The SDF model formulation reads

$$Y_i \sim BEINF\left(\lambda, \gamma, \mu_i, \phi\right) \tag{2.10}$$

$$logit\left(\mu_i\right) = \alpha + \beta\sqrt{wd_i}. \tag{2.11}$$

Other than in the BR model, we only predict the mean parameter μ of the beta distribution. The remaining distribution parameters, λ, γ, and ϕ are assumed to be constant across observations; that is, we estimate them during the inference but do not predict them. We estimated the parameters of SDF-P in analogy to the BR model via MCMC. The SI contains further information on the prior choice for the SDF-P model.

2.2.3 Model validation

We validate the predictive performance of BN, BR, RF, and SDF individually for the three assets building, equipment, and goods and stock. This results in 12 asset-model combinations. All candidate models return probabilistic predictions rather than deterministic loss estimates. However, the models do not provide analytical predictive distributions but simulated approximations in the form of samples. For each model, we sampled 1000 values from the conditional response distribution and evaluated this probabilistic response with respect to accuracy, sharpness and calibration. Within one asset dataset, we estimated the model test errors through repeated 10-fold cross-validation in order to receive robust estimates of true model performance. That is, we initiated 100 independent runs of 10-fold cross-validation with varying, random data partitioning. In each of the 10-fold cross-validation runs, every company is held out of the training set for prediction exactly once. We validate model performance for each cross-validation fold by means of three performance metrics:

1. The mean absolute error (MAE) for the mean of the predictive distribution. The MAE evaluates the accuracy of a point forecast and averages the absolute difference between the observed response and the predicted point estimate over the number of observations.

2. The mean bias error (MBE), which quantifies model over- and underestimation in the mean of the predictive distributions.

3. The continuous ranked probability score (CRPS), which is a proper scoring rule that evaluates the entire continuous distribution of a probabilistic forecast. It jointly assesses the sharpness and calibration of the predictive distribution and generalizes the absolute error (Gneiting and Katzfuss 2014; Matheson and Winkler 1976). Hence, the CRPS can be compared directly to the MAE. The CRPS for one observation y_i is defined as

$$CRPS_i\left(F_i, y_i\right) = \int_{-\infty}^{\infty} \left(F_i(x) - \mathbf{1}\left\{y_i \leq x\right\}\right)^2 dx, \qquad (2.12)$$

   where $F_i(x)$ is the cumulative distribution function (CDF) of the predictive distribution $f_i(x)$, and $\mathbf{1}\{\cdot\}$ is the indicator function. We compute the CRPS with an empirical CDF estimated from samples of $f_i(x)$. For details on the numerical implementation of the CRPS for simulated forecasts, we refer to the corresponding literature (Jordan et al. 2019; Krüger et al. 2016). For the proportional response variable, relative loss, the CRPS is defined on the interval [0,1] with the optimum at zero. Note that the CRPS is calculated individually for each observation. For the comparison with the MAE, we computed the mean CRPS value in each cross-validation fold.

2.3 Results and discussion

2.3.1 Variable importance in multivariable models

We compare the fitted multivariable models with respect to plausibility and consistency. First, the model fits should be in line with physical principles governing flood damage processes; for example, that loss increases with larger water depths. Secondly, the relative effect and influence of the predictors should be similar across models, since they are fit to the same training data.

Figure 2.2 compares the learned BN structures, the estimated BR regression parameters for the mean parameter of the beta distribution (μ), and the RF predictor importance measures for the three study assets. In the BN structures, variables with the strongest statistical dependence on relative loss are directly connected to its node. The relative magnitude and sign of the estimated BR regression coefficients yield information on the effect of the corresponding predictor on relative loss. The coefficients for the categorical predictors, spatial situation and sector, express the deviation in flood loss for each variable group individually and relative to the first group of the respective variable, which acts as a reference (see 'dummy encoding' in section 2.2.2). For example, companies that operate in the second sector group, 'commercial', experienced considerably higher building loss than companies belonging to the first sector group, i.e., 'manufacturing', since the respective coefficient 'sec[com]' is positive in the building model. Ultimately, RF expresses variable importance through the change of model accuracy that is induced by simulating the absence of a particular variable. The stronger the decrease in RF accuracy, the more relevant is the withheld predictor.

Water depth is a dominant influencing factor for flood loss to all three assets, as indicated by direct arc connections to relative loss in all BNs. This is confirmed by high absolute values of BR regression coefficients and RF predictor importance. The relevance of water depth deteriorates from building over equipment to goods and stock. The estimated signs of the BR regression coefficients show that water depth has a positive effect on relative loss. The high variable importance of water depth is in accordance with the majority of company flood loss models (e.g., Hasanzadeh Nafari et al. 2016a; Kreibich et al. 2010; Penning-Rowsell et al. 2005; Sieg et al. 2017), where water depth is the most influential predictor. The remaining flood intensity variables, return period and duration, predominantly drive relative loss as well; yet, to a lesser degree than water depth, confirming findings from similar studies for private households and companies (e.g., Kreibich et al. 2010; Merz et al. 2013; Sieg et al. 2017; Vogel et al. 2018).

Precaution is a likewise important predictor in the proposed models with direct arc connections to relative loss in all BNs. BR and RF reveal that the effect of precaution becomes more important for losses to equipment and, especially, goods and stock. Precaution was identified as an influential variable before (Kreibich et al. 2010; Sieg et al. 2017), but it exhibits striking relevance in the presented models, which might trace back to different approaches to estimating company precaution. The large negative impact of precaution on relative loss in the BR models implies that precautionary measures can reduce flood loss effectively.

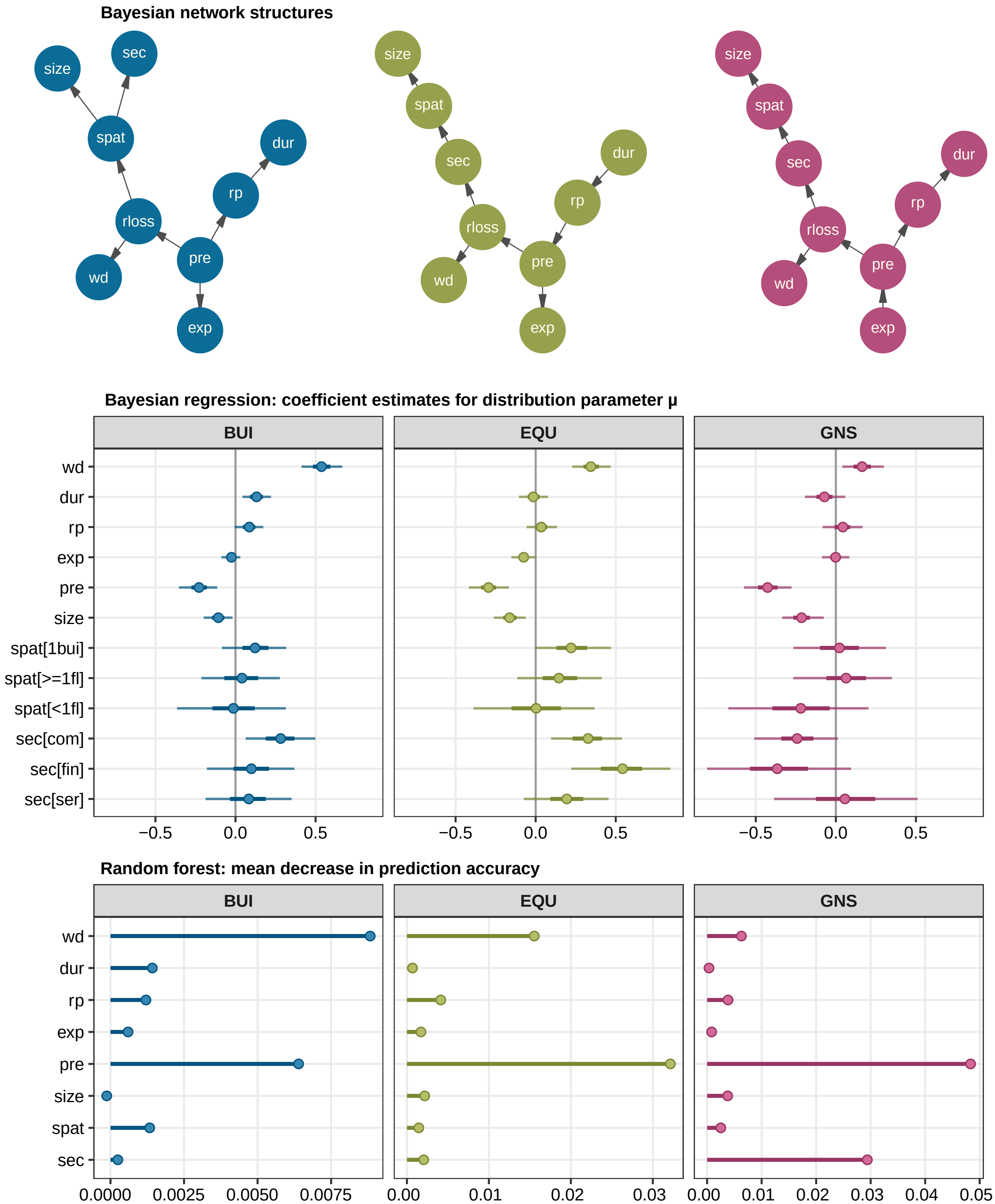

Figure 2.2: *Comparison of the fitted candidate models. From top to bottom the plots show Bayesian network structures, estimated regression coefficients for the mean parameter of the beta distribution in the Bayesian regression, and predictor importance measures of random forest. From left to right, the columns display the model fits for building (BUI), equipment (EQU), and goods and stock (GNS).*

The spatial situation is more significant for losses to building than to equipment, and goods and stock. In contrast, the economic sector exhibits higher explanatory power for equipment, and goods and stock as indicated by the direct arc connections from sector to relative loss in the respective BNs. The effect of the company size is negative with maximum magnitude for losses to goods and stock. Sieg et al. (2017) found similar patterns of predictor importance for the spatial situation and company size. Flood experience plays a minor role for all assets and seems to reduce relative loss as well. The BN graphs imply that the predictive power of company size and flood experience is covered by correlated variables in adjacent nodes that are closer to relative loss (spatial situation, precaution), which explains their inferior overall importance. Kreibich et al. (2010) also identified flood experience as a subordinate predictor.

The variation in the predictor effects across the assets suggests that damage processes differ for losses to building, equipment, and goods and stock. This was also reported by Sieg et al. (2017), who observed fluctuating predictor importance across asset types for a subset of the same survey data. In general, building loss is controlled by variables describing the hazard intensity, precaution, and the spatial situation. In contrast, variables that describe company characteristics (sector, size) and precaution bear most information for losses to equipment, and goods and stock and sometimes even exceed the effect of water depth. Considering the pronounced variable effect of the sector for these assets, it seems that the heterogeneity among companies mainly reflects in the damage processes for equipment, and goods and stock. For instance, company equipment ranges from heavy machinery over technical devices to office items depending on the business sector. Conversely, the low RF predictor importance of the sector in the building loss models suggests that the damage processes for buildings are more alike over different company types. These findings are in line with the results of Sieg et al. (2017), where damage processes across sectors diverged more for equipment, and goods and stock.

We conclude that the fitted candidate models satisfy the criteria of plausibility because the predictor effects agree with previous findings and match the physical understanding of damage processes. The dissimilarity in the model fits for different assets justifies the development of distinct loss models for building, equipment, and goods and stock and highlights the benefit of multivariable loss modeling approaches. Overall, the candidate models consistently identify the same predictors as most relevant (water depth, precaution, sector) and agree well within one asset. Minor discrepancies occur primarily for predictors with moderate to weak predictive power such as return period or flood experience.

2.3.2 Model performance

Model validation

Figure 2.3 shows the results of the repeated cross-validation runs for the three performance metrics MAE, MBE and mean CRPS. Each boxplot summarizes 100 repetitions of a 10-fold cross-validation for a specific asset (x-axis), metric (plot panels), and model (color-coded). Comparing MAE, we observe that all models achieve the lowest errors for building loss. The multivariable models (BN, BR, RF) perform similarly and exhibit median MAE values of 0.149, 0.158, and 0.150, respectively. The probabilistic and deterministic SDFs reach slightly higher MAEs of 0.174 and 0.165 in the median. MAE scores for equipment and goods and stock deteriorate in comparison to building loss and are in the same range across the multivariable models with values of approximately 0.27. For the SDFs, however, MAEs increase stronger for goods and stock (SDF-P: 0.355, SDF-D: 0.348) than for equipment (SDF-P: 0.329, SDF-D: 0.316). Among the multivariable models, BR shows slightly higher MAEs than BN and RF.

The cross-validated mean CRPS shows almost the same relative ranking of the models. Medians of mean CRPS values for the multivariable models are approximately 0.10 for building, and 0.16 for equipment and goods and stock. With mean CRPS values of 0.109 (BUI), 0.195 (EQU) and 0.200 (GNS), SDF-P is outperformed by the complex models, especially for equipment and goods and stock. RF reaches the best mean CRPS for all three assets, yet the difference to the other multivariable models is small. CRPS cannot be calculated for the deterministic stage-damage function, as it requires probabilistic predictions.

The boxplots of the MBE reveal that all models neither under- nor overestimate relative loss considerably in the median.

As described in section 2.2.3, the CRPS generalizes the MAE, which facilitates the direct comparison of deterministic and probabilistic forecasts. The larger values of MAE in comparison to mean CRPS suggest a loss of information about the observed response, when the predictive distribution is condensed to a single value, namely the mean. Moreover, in case of the probabilistic models, the MAE produces biased estimates of model skill, as the mean of the predictive distribution commonly deviates from the most probable loss. This could also explain why the deterministic SDF outperforms the probabilistic SDF when considering the MAE. Yet, computing the MAE on basis of the mode is likewise biased in this application since the predictive distributions are often bimodal (see Figure 2.4). We reason that scoring rules that evaluate entire predictive distributions rather than response means or modes, are more robust estimates of true model skill; at least for response distributions other than the normal.

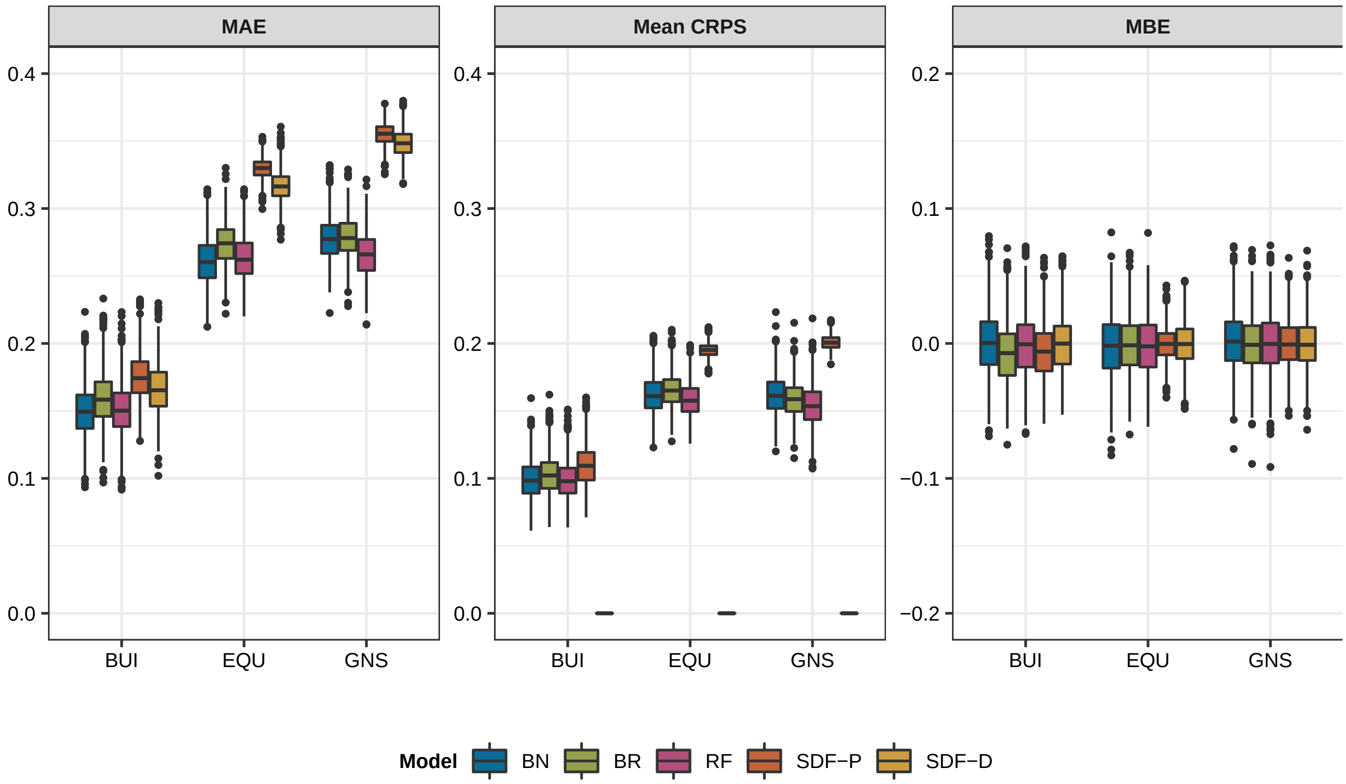

Figure 2.3: *Performance metrics mean average error (MAE), mean continuous ranked probability score (CRPS), and mean bias error (MBE) for the five models (color-coded) and assets (x-axis). Each boxplot summarizes 100 repetitions of a 10-fold cross-validation with varying data partitioning.*

Model performance for individual companies

Figure 2.4 compares the predictive distributions for building loss of the candidate and benchmark models for nine randomly selected companies. The predictive distributions are color-coded according to the models, and the actually observed relative loss is indicated by a vertical, black line. The yellow line shows the predicted loss of the deterministic stage-damage function (SDF-D). The predictive distributions of BN, BR, and RF are flexible and vary considerably from company to company. Conversely, the SDF-P distributions fluctuate less and their medians rarely exceed 0.10. The deterministic predictions of SDF-D vary the least across individual companies, as the model lacks a component that explicitly accounts for extreme losses at zero and one. In contrast to SDF-P, which predicts constant shares of zero and ones, the multivariable models inflate and deflate the modes of their predictive densities at zero and one dynamically, reflecting the actual observations of relative loss (see e.g., IDs: 165 and 136 in Figure 2.4). The invariant shape of the SDF-P predictive densities leads to overall higher errors of their predictions (Figure 2.3). In general, the prediction accuracy and sharpness is larger for companies with low loss magnitudes as compared to companies with more severe losses.

Figure 2.5 confirms that the differences in the example predictive distributions between the multivariable models and the SDF-P also apply to the entire dataset. Every point represents the CRPS error of the predictive distribution for one company, while the step-wise, black line indicates the mean CRPS in the corresponding interval of observed relative loss. The scatter plots show that the CRPS of the probabilistic predictions changes over different magnitudes of observed relative loss. The variation in the CRPS is stronger for the multivariable models than for the SDF-P, for which errors disperse less around the interval mean.

The steady predictive distributions of the SDF-P, and hence its errors, do not change significantly across observations. While this generalizing behavior of the SDF-P is favorable in principle, its mean CRPS values exceed the ones of the multivariable models. In addition, prediction errors tend to increase with larger values of relative loss. We encounter this trend for all models and assets and it is more pronounced for building loss than for losses to equipment, and goods and stock. The striking difference in the scatter point clouds between buildings on the one side and equipment, and goods and stock on the other side, traces back to stronger bimodality for the observed losses to equipment, and goods and stock (see rloss distributions in Figure 2.1).

2.3.3 Model comparison

Multivariable models and stage-damage functions

The cross-validated performance metrics in Figure 2.3 show that BN, BR, and RF are superior to the deterministic and probabilistic SDFs with respect to predictive capacity for all three study assets. In general, the prediction accuracy is higher for buildings than for equipment and goods and stock. We identify two reasons for the difference in prediction skill. First, Figure 2.6 shows that the relationship between water depth and relative loss is volatile for all assets and only insufficiently discriminates between severe and minor relative loss. BN, BR, and RF have access to information contained in predictors other than water depth, which fosters a more accurate determination of the loss magnitude. SDFs base their predictions solely on water depth and, thus, fail to explain irregular loss cases, for instance, when low water depth causes high relative loss. Secondly, the multivariable models exhibit higher structural complexity than the SDFs, which allows for

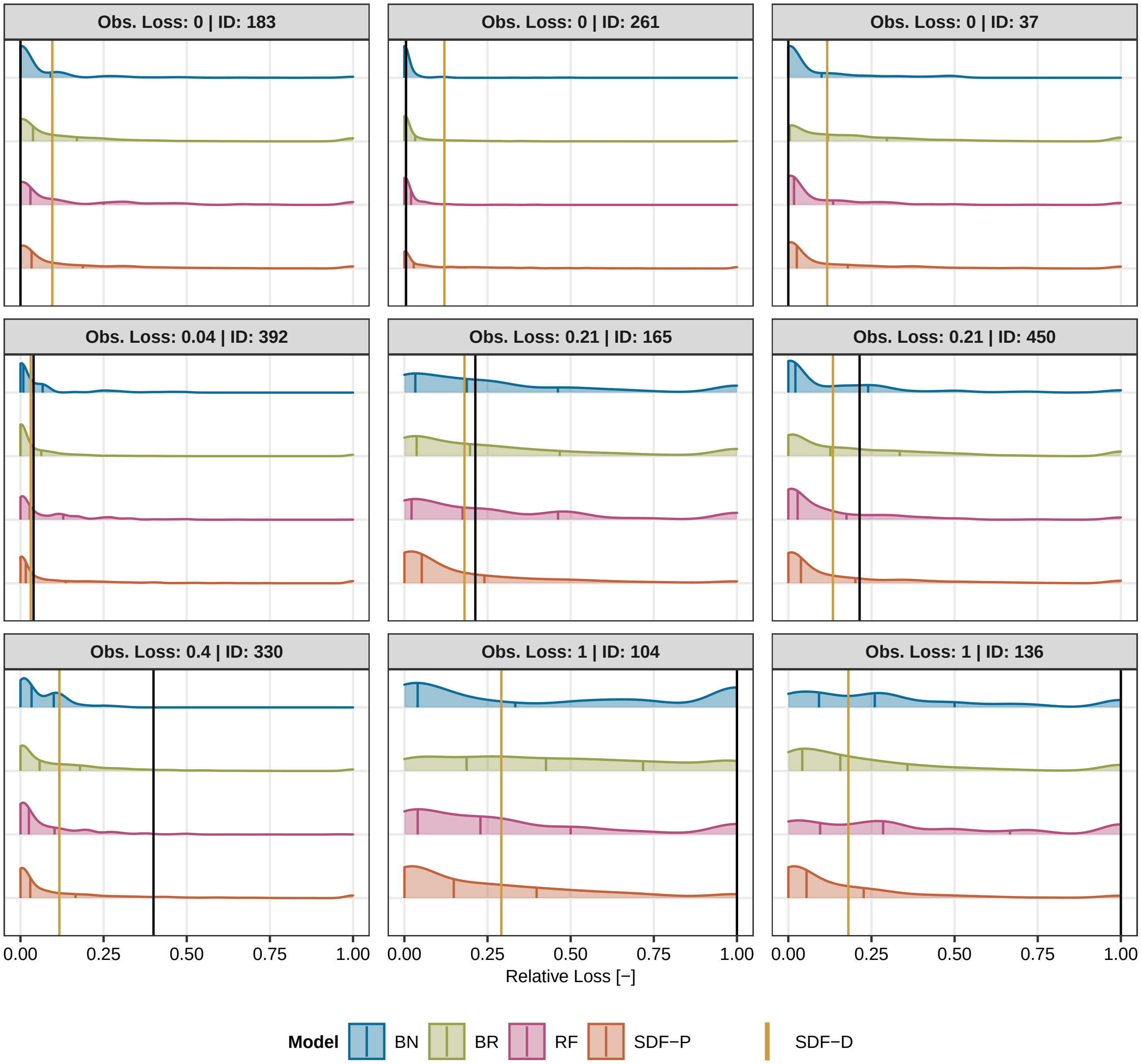

Figure 2.4: *Examples of predictive densities from the four probabilistic models (color-coded) for the building loss of nine randomly selected companies (identified by ID). Black and yellow lines display the observed loss and the predicted loss of the deterministic stage-damage function, respectively. The panels are sorted according to the observed loss which is indicated by the black, vertical lines. Colored lines beneath the distributions indicate the quartiles of the respective predictive density. The scaling of the four y-axes within each panel is consistent, ensuring the comparability of the predictive densities. The displayed densities originate from a cross-validation run.*

closer fits to the training data. For instance, SDF-P and BR both model relative loss with an inflated beta distribution. However, while SDF-P assumes constant inflation and precision parameters, BR predicts these parameters for each company individually. The increased number of parameters leads to higher flexibility in the predictive densities for the BR model. The capability to de- and inflate the modes at zero and one (see Figure 2.4) enables the complex models to capture both extremes of relative loss at the same time, while the SDFs have to find a balance between covering no and total loss cases. The boxplots in Figure 2.3 show that the larger shares of zero and ones in the data for equipment and goods and stock lead to larger performance difference between the models with complex (BN, BR, RF) and simple structure (SDFs).

However, the flexibility in the predictive distributions of the multivariable models propagates to the CRPS, resulting in considerable variance in the errors for individual companies (Figure 2.5). This observation reflects the bias-variance tradeoff, a typical phenomenon in predictive statistical modeling (see e.g., James et al. 2013). It describes that complex, multivariable models, such as BN, BR, and RF, incur lower bias than models with fewer parameters, such as the SDFs, at the cost of larger variance in their predictions and errors. While overly flexible models are at risk of undesirably capturing random noise in the data (i.e., overfitting), inflexible models might be unable to reproduce essential features of the data generating process (i.e., underfitting). The required degree of model complexity depends on the data and the question under consideration. We assume that the heterogeneity of companies and damage processes demands for a fair amount of model complexity. Given the results of the validation experiment, we conclude that it is the combination of multivariable and probabilistic modeling, which causes the candidate models to outperform the benchmark models, albeit the large variation in CRPS error. Schröter et al. (2014), who developed and validated multivariable probabilistic models for the residential sector, also observed that complex models perform better than simpler modeling approaches. The ability of the proposed models to account for variable interactions and to capture complicated data-generating processes (i.e., zero-one-inflation) might even be more useful for modeling company loss data, where heterogeneity across company types leads to particularly noisy relationships between predictors and loss.

Further, the notable difference in the values of the mean CRPS and the MAE within the same models in Figure 2.3 shows that the predictions of the probabilistic models are more informative than the loss forecasts of the deterministic stage-damage function. This gain in information can be employed for practical applications in risk analysis or decision making, where estimates of prediction reliability provide additional decision support.

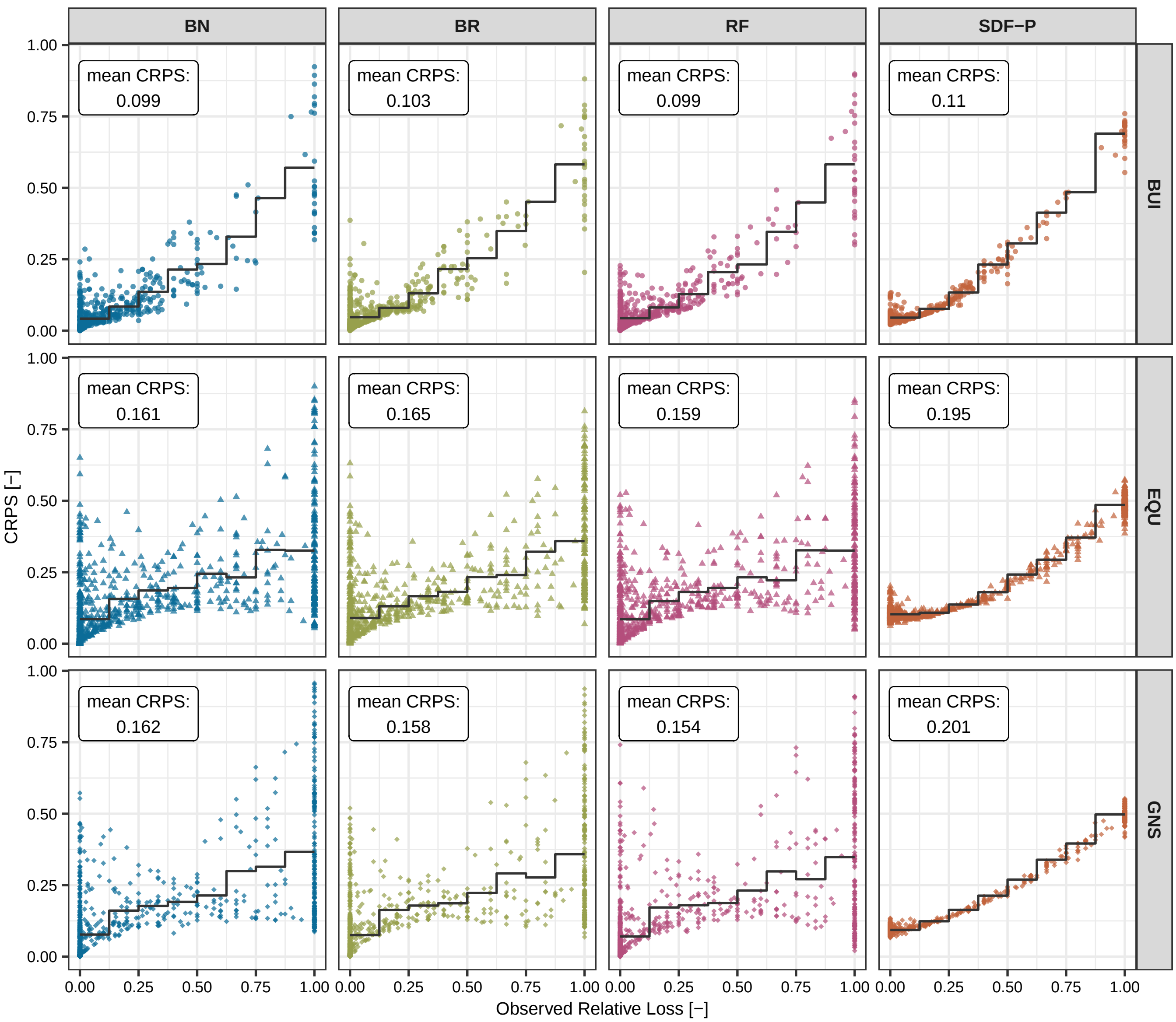

Figure 2.5: *Scatter plots of observed relative loss versus cross-validated continuous ranked probability scores (CRPS) for all combinations of assets (rows; symbol-coded) and probabilistic models (columns, color-coded). Each symbol represents the prediction error incurred by the respective model for one company. The black, step-wise lines show the average CRPS in different intervals of observed relative loss. The labels in the top-left corner of each panel contain the mean CRPS over all predictions (=symbols) of the respective asset-model combination.*

Intercomparison of multivariable models

The intercomparison of the different multivariable models does not reveal clear performance differences. BN and RF outperform BR slightly. Yet, the magnitude of the performance differences is small. The high agreement on the aggregated and company-level performance metrics of BN, BR, and RF implies that the predictive capacity of the multivariable approaches is rather constrained by the information content in the training data than by model specific characteristics. It remains an open question, whether limited knowledge about flood damage processes hinders the composition of more meaningful predictor sets, or whether the inherent variation in the flood damage processes restricts the forecasting capacity of existing models at a certain threshold. Either way, the model choice should be guided by the study task and data availability. BNs allow for intuitive inference on the flood damage processes through the graphical dependency structure and have advantages in the treatment of missing data. The strength of BR lies in the flexibility of the Bayesian framework, where multilevel modeling and the definition of strong priors facilitate predictions even with few loss data. RF provides accurate predictions with relatively small implementation effort and is tolerant with respect to differently scaled model variables. However, modelers have less influence on the internal model structure, and the interpretation of the RF functionality is difficult.

BN, BR, and RF outperform other multivariable company loss models which have been validated on subsets of the same survey data. For instance, Seifert et al. (2010) reported MAE values of 0.23 (BUI), 0.30 (EQU), and 0.30 (GNS) for their FLEMOcs+ model. The random forest model of Sieg et al. (2017) achieved MAE values of 0.18 (BUI), 0.31 (EQU), and 0.37 (GNS). We assume that the performance advantages of the presented models are a joint result of different model configurations, changes in the predictor variables, and a larger data basis in this study.

Although the multivariable probabilistic models improve the accuracy and sharpness of the loss estimations over the entire dataset, they incur considerable CRPS errors for severe losses. This is problematic, since large relative loss cases can have a strong effect on the total estimates of post-event loss in a flooded area. The poor performance for severe losses arises from the imbalance between frequent but small, and infrequent but major damages, which is common in natural disaster loss datasets (Pisarenko and Rodkin 2010). The number of high losses provides too few training samples for the algorithms to reliably identify whether a company experiences severe flood loss or not. The problem of undersampled extremes might eventually resolve when observational periods become long enough to contain a sufficient number of severe losses. Yet, in the

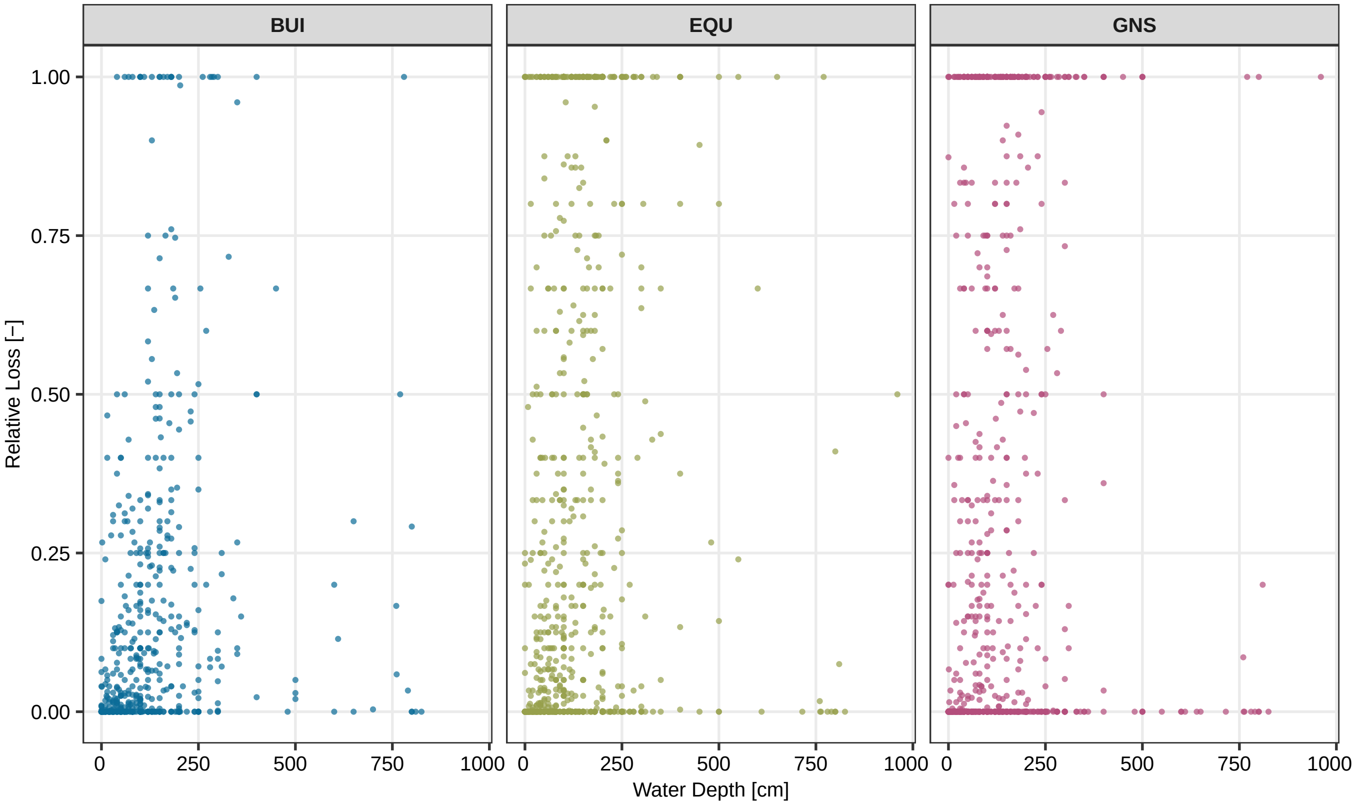

__Figure 2.6:__ Scatter plots of water depth versus relative loss. Each plot panel is color coded according to the three study assets building (BUI), equipment (EQU), and goods and stock (GNS).

analysis of natural hazards the required time horizons quickly exceed decades (Zöller 2013). Here, the enrichment of loss datasets with severe loss cases from other regions, as practiced in the modeling of extreme earthquakes (i.e., method of analogs; see e.g., Holschneider et al. 2011; Wheeler 2009), could represent an alternative. Additionally, if available in the data, further refinements of the predictor set could improve the predictive power of the models; for example, by including variables that describe the structural characteristic of buildings more accurately (see e.g., Hasanzadeh Nafari et al. 2016a; Scawthorn et al. 2006). In general, the predictive distributions of the multivariable models are relatively wide, especially for companies which experienced large relative loss and for the assets equipment, and goods and stock. Hence, further analysis of the distinct uncertainty sources and the potential to reduce their contribution to the overall variance in the loss estimates could improve the reliability of the proposed models.

2.4 Conclusions

This study presents three multivariable flood loss models for companies which return probabilistic loss predictions. Referring to the objectives of this study:

1. Bayesian networks, Bayesian regression and random forest outperform the deterministic and probabilistic stage-damage functions due to additional information contained in predictors other than water depth and larger flexibility in their predictive densities.

2. The predictive capabilities across the multivariable models are very similar and constrained by the explanatory power of the predictor set rather than by model choice.

Although the cross-validated performance metrics for the multivariable models confirm higher predictive skill in comparison to existing company flood loss models, our analysis identified substantial uncertainty in the predictive distributions and deteriorating predictive power for large losses.

Since we have to accept the inherent complexity of flood damage processes and poor coverage of severe losses in the data, we advocate the proper treatment of the resulting uncertainties. Probabilistic modeling explicitly quantifies the associated uncertainties and, hence, provides more honest loss estimates than deterministic approaches. Moreover, the additional uncertainty information could directly contribute to flood risk management practice. For instance, by providing the probabilistic foundation for an informed decision making, where the attractiveness of a certain flood protection measure not only depends on the expected reduction in damage but also on the confidence in the predicted damage reduction; or by facilitating the seamless propagation of predictive uncertainty across different exposure aggregation levels. Therefore, in our opinion, probabilistic models should become the standard approach in flood loss estimation. Further, this study underlines that the demand for probabilistic loss estimation is particularly strong for companies, given the large variation of loss influencing variables across individual companies and their exposed assets. In conclusion, the combination of multivariable and probabilistic modeling advances the representation of company vulnerability in flood risk assessment through improved loss estimations and transparent communication of their reliability.

3 Augmenting a Socio-hydrological Flood Risk Model for Companies with Process-oriented Loss Estimation

Authors:
Lukas Schoppa
Marlies H. Barendrecht
Tobias Sieg
Nivedita Sairam
Heidi Kreibich

Published as:
Schoppa, L., Barendrecht, M., Sieg, T., Sairam, N., and Kreibich, H., 2022. Augmenting a socio-hydrological flood risk model for companies with process-oriented loss estimation. Hydrological Sciences Journal, 00 (00), 1–17. https://doi.org/10.1080/02626667.2022.2095207

Abstract

Socio-hydrological flood risk models describe the temporal co-evolution of coupled human–flood systems. However, most models oversimplify the flood loss processes and do not consider companies' substantial contribution to total losses. This work presents a socio-hydrological flood risk model for companies that focuses on changes in vulnerability. In addition, we augment the socio-hydrological model with a process- oriented, sector-specific loss model in order to capture damage processes more realistically. In a case study, we simulate the historical flood risk dynamics of companies in the floodplain of Dresden, Germany, over the course of 120 years. Our analysis suggests that the companies in Dresden increase their exposure more cautiously than private households and decrease their vulnerability more actively through private precaution. The augmentation, consisting of informative predictors, a refined probabilistic model, and the incorporation of additional data, improves the accuracy and reliability of the flood loss estimates and reduces their uncertainty.

3.1 Introduction

Flood risk is determined by hazard, exposure, and vulnerability, which change and interact over time, resulting in nonlinear risk dynamics such as the adaptation effect (Merz et al. 2010a; Di Baldassarre et al. 2015; Merz et al. 2015). The adaptation effect describes how societies decrease their vulnerability after repeatedly being affected by damaging flood events, eventually diminishing overall losses (Di Baldassarre et al. 2015; Kreibich et al. 2017a). Traditional scenario-based approaches in flood risk assessment can fall short of capturing such risk dynamics as they do not account for feedbacks between the hydrological and socioeconomic domain (Di Baldassarre et al. 2013; Barendrecht et al. 2017; Srinivasan et al. 2017). The negligence of interactions can produce biased estimates of future flood risk and, hence, affect risk management negatively. The interplay between society and floods has been studied with different approaches such as hydro-social theory (Marks 2019; Devkota et al. 2020; Haeffner and Hellman 2020), socio-ecological systems (Ishtiaque et al. 2017), coupled human and natural systems (O'Connell and O'Donnell 2014; Abebe et al. 2019), and socio-hydrology. Socio-hydrology (Sivapalan et al. 2012; Sivapalan et al. 2014) focuses on quantitative methods and employs a rich collection of modeling techniques (Blair and Buytaert 2016; Ross and Chang 2020). The objective of socio-hydrological flood risk assessment is a more realistic exploration of the possible pathways that a human-flood system might traverse in the future (Di Baldassarre et al. 2015; Merz et al. 2015; Barendrecht et al. 2017).

Stylized, conceptual models are a prevalent type of socio-hydrological models and describe the interactions between selected state variables through a set of coupled differential equations, each representing a system process (Blair and Buytaert 2016). They are commonly lumped and explain the macroscale behavior of the human-flood system, which promotes the model interpretability. Socio-hydrological models focus on the understanding of the system, and they are better suited for the strategic guidance of long-term decision making than for specific management problems (Sivapalan and Blöschl 2015; Barendrecht et al. 2017).

Di Baldassarre et al. (2013) and Di Baldassarre et al. (2015) introduced a conceptual model that explains how societies and rivers co-evolve within floodplains and is capable of capturing flood risk dynamics such as the adaptation effect. The model has been reproduced and refined widely in subsequent works that explore risk coping cultures (Viglione et al. 2014), flood memory (Ridolfi et al. 2021; Song et al. 2021), flood control and management (Di Baldassarre et al. 2017), risk perception (Ridolfi et al. 2020), resilience (Ciullo et al. 2017; Yu et al. 2017), or the relationship between flooding and economic growth (Grames et al. 2016).

In an effort to develop a fully quantitative parameter estimation procedure for socio-hydrological models, Barendrecht et al. (2019) used empirical data from private households to study the human-flood system in Dresden, Germany. The study of Barendrecht et al. (2019) is a first step towards more rigorous socio-hydrological models that explore specific case studies and could provide useful results for practical decision support. As a consequent next step, the informed inclusion of process-oriented modeling approaches has the potential to improve socio-hydrological flood risk assessment.

First, the mathematical representation of the flood loss processes in socio-hydrological models is oversimplified. For example, the model by Di Baldassarre et al. (2013) and its successor models derive monetary flood loss directly from the maximum flood discharge by parameterizing the topographic characteristics of the floodplain. However, from a loss modeling perspective, estimating flood loss from inundation depth would capture the physical loss processes with more detail, and empirical analyses confirmed the strong explanatory power of inundation depth as a predictor variable (Merz et al. 2010b; Hasanzadeh Nafari et al. 2016b; Wagenaar et al. 2017; Vogel et al. 2018). In addition, other characteristics of flood loss data such as bimodality (i.e., disproportionally high shares of zero and total building loss) adds to the complexity (Wing et al. 2020). Barendrecht et al. (2019) used a prevalent probabilistic beta model for loss estimation, which resulted in the overestimation of minor and the underestimation of major loss events. Dedicated probabilistic models that account for the frequent overdispersion in loss data could improve the accuracy of the estimates and reduce associated uncertainties (Rözer et al. 2019). Ignoring these advances in loss modeling biases the loss estimates and consequently the socio-hydrological model as a whole.

Secondly, the heterogeneity within society is a crucial but often neglected process detail in socio-hydrological models. The majority of conceptual models treat societies in the floodplain as homogenous entities (Viglione et al. 2014; Ciullo et al. 2017; Ridolfi et al. 2020). Yet individual societal groups (i.e., households, companies, institutions, government) follow their own motives, which influence the relevant damage processes or their decisions regarding flood protection (Haer et al. 2017; Haer et al. 2019; Bubeck et al. 2018). Song et al. (2021) investigated collective flood memory with a model that distinguishes between urban and rural societies in China and found differences in the accumulation of flood memory between the two groups. The variables that determine flood vulnerability also differ between private households and companies (Merz et al. 2010b) and even across economic sectors (Kreibich et al. 2007; Sieg et al. 2017). Therefore, flood loss models are either developed for specific sectors (e.g., residential, manufacturing, services) or they include the sector as a predictor in the model (Kreibich et al. 2010; Sieg et al. 2017; Paprotny et al. 2020; Schoppa et al. 2020). To date, methods for a sector-specific loss estimation in socio-hydrological modeling are lacking.

Thirdly, previous efforts in model development concentrated on private households (Haer et al. 2017; Haer et al. 2020; Barendrecht et al. 2019). Companies have not been addressed extensively in the socio-hydrological literature before, even though they usually account for large shares of total flood losses (Paprotny et al. 2020). Earlier works shed light on specific aspects of company flood risk; for instance, on flood impacts, adaptive behavior, or recovery (Wedawatta and Ingirige 2012; Wedawatta et al. 2014; Li and Coates 2016; Jehmlich et al. 2020). (Coates et al. 2014; Coates et al. 2019) coupled an agent-based model to a hydrodynamic model to examine the behavior of individual companies in the aftermath of a flood event. Nevertheless, there are no studies that explore the long-term dynamics of company flood risk, including feedbacks between the determinants of risk. In summary, the explicit consideration of new sectors and inter-sectorial differences could not only improve loss estimation in socio-hydrological models but also uncover variations in the decisions and behavior within societies.

In this study, we aim to improve the currently available socio-hydrological flood risk models by addressing these shortcomings (i.e., oversimplified loss estimation, lack of heterogeneity, scarcity of models for companies). We integrate a process-oriented, sector-specific regression for loss estimation into a socio-hydrological model. Additionally, we study the risk dynamics of companies by transferring the socio-hydrological flood risk model for the residential sector by Barendrecht et al. (2019) to companies in the city of Dresden, Germany, where recurring flood events induced the society to reduce its vulnerability (i.e., adaptation effect) (Kreibich et al. 2005; Thieken et al. 2007; Kreibich and Thieken 2009; Jehmlich et al. 2020). The research questions of this study are:

1. What is the added value of augmenting the socio-hydrological model with a process-oriented loss estimation and differentiating between economic sectors?

2. Can the socio-hydrological flood risk model for companies reproduce the observed adaptation effect, and do companies behave differently in respect of flood risk than private households?

In a modeling experiment, we assess the benefits of the process-oriented loss estimation and the sector differentiation. Figure 3.1(a) displays the study area, the city of Dresden, Germany, which is located on the banks of the Elbe River.

3.2 Methods and data

On the basis of the model by Barendrecht et al. (2019) for the residential sector, we developed a socio-hydrological flood risk model for small and medium-sized companies from the manufacturing and service sector. Subsequently, we augmented the new company model by a process-oriented loss estimation and a sector differentiation. In the following, we introduce four model versions with increasing complexity, which we used in the modeling experiment. Afterwards, we present the socio-hydrological model and the two augmentations in detail.

3.2.1 Model versions

For the systematic examination of the added value of the process-oriented loss estimation and the sector differentiation, we configured four model versions, incrementally adding one augmentation option or both to the company model. We refer to the four model versions as follows

- *Parsimonious model ("pars")*: the adaptation of the socio-hydrological model by Barendrecht et al. (2019) for companies, which acts as the benchmark. It pools economic sectors and uses a simplistic loss estimation (Equations 3.1 and 3.2).

- *Intermediate model with sector differentiation ("int_sd")*: distinguishes between economic sectors but uses the simplistic loss estimation.

- *Intermediate model with process-oriented loss estimation ("int_lm"):* includes the process-oriented loss estimation but does not differentiate between economic sectors.

- *Fully augmented model ("aug"):* the most complex model as it differentiates between economic sectors and features the process-oriented loss estimation.

The four model versions enable the isolated and joint assessment of the effect of the two augmentation options (process-oriented loss estimation and sector differentiation) on the socio-hydrological simulation. Figure 1(b) presents the four candidate models in the form of causal loop diagrams including all model variables, their interrelation, and feedbacks. Additionally, the diagrams highlight which system processes are affected by the respective augmentation option. First, we evaluated the fit of the candidate models to the observed socio-hydrological data - in particular, the accuracy and uncertainty of the loss estimates. Second, we conducted a leave-one-out cross-validation experiment to test the predictive capacity of the models for flood loss events out of the training sample. We quantify the predictive capacity of the models with the continuous ranked probability score (Matheson and Winkler 1976; Gneiting and Katzfuss 2014; Krüger et al. 2016; Jordan et al. 2019), a proper scoring rule that indicates the distance between a probabilistic forecast and an observation (see Supplementary material, Text B.8).

3.2.2 Socio-hydrological flood risk model for companies

The socio-hydrological model considers the three determinants that affect the flood risk - *hazard*, *vulnerability*, and *exposure* (Kron 2005) - and focuses on the adaptive behavior of the companies. We explain the socio-hydrological system using the example of the parsimonious model (pars) in Figure 3.1(b).

The flood *hazard* is represented by an annual maxima series of the discharge W of the Elbe at the Dresden gauge. The public structural flood protection in Dresden, such as levees, is encoded as a protection level and expressed in the form of a design discharge H. Since the implementation of public structural flood protection lies within the authority of the federal state and its institutions, we consider the protection level exogenous to the socio-hydrological system. Flooding occurs in the model once the annual maximum discharge W exceeds the current protection level H. We assume that flooding impacts the company buildings in the floodplain, which is quantified by the monetary flood loss L.

After a damaging event, the flood risk awareness A of the companies increases. An increase in the awareness leads to higher flood preparedness P. In this context, the term "preparedness" comprises the implementation of private precautionary measures by the companies themselves, such as the flood proofing of buildings. The awareness and preparedness describe the current *vulnerability* of the companies and rise instantaneously after a flood event. The degree of the increase depends on the total flood loss suffered by companies in Dresden (for awareness) and the resulting increment in the awareness (for preparedness). At times where the flood protection withstands the annual maximum discharge, the awareness and preparedness decay since companies forget about the flood risk and precautionary measures deteriorate.

The *exposure* dynamics in the floodplain are captured by the economic density D, which is the share of the floodplain area that is covered by company premises. On the one hand, the economic density is driven by the economic growth rate U, which is also an exogenous forcing variable. When the economic growth rate is positive, more companies settle in the floodplain. On the other hand, high flood risk awareness motivates companies to move out of the floodplain and settle in safer places. The causal loop of the socio-hydrological system is closed since the economic density and the preparedness feed back into the total loss caused by an event. The area share of companies in the floodplain determines whether and how

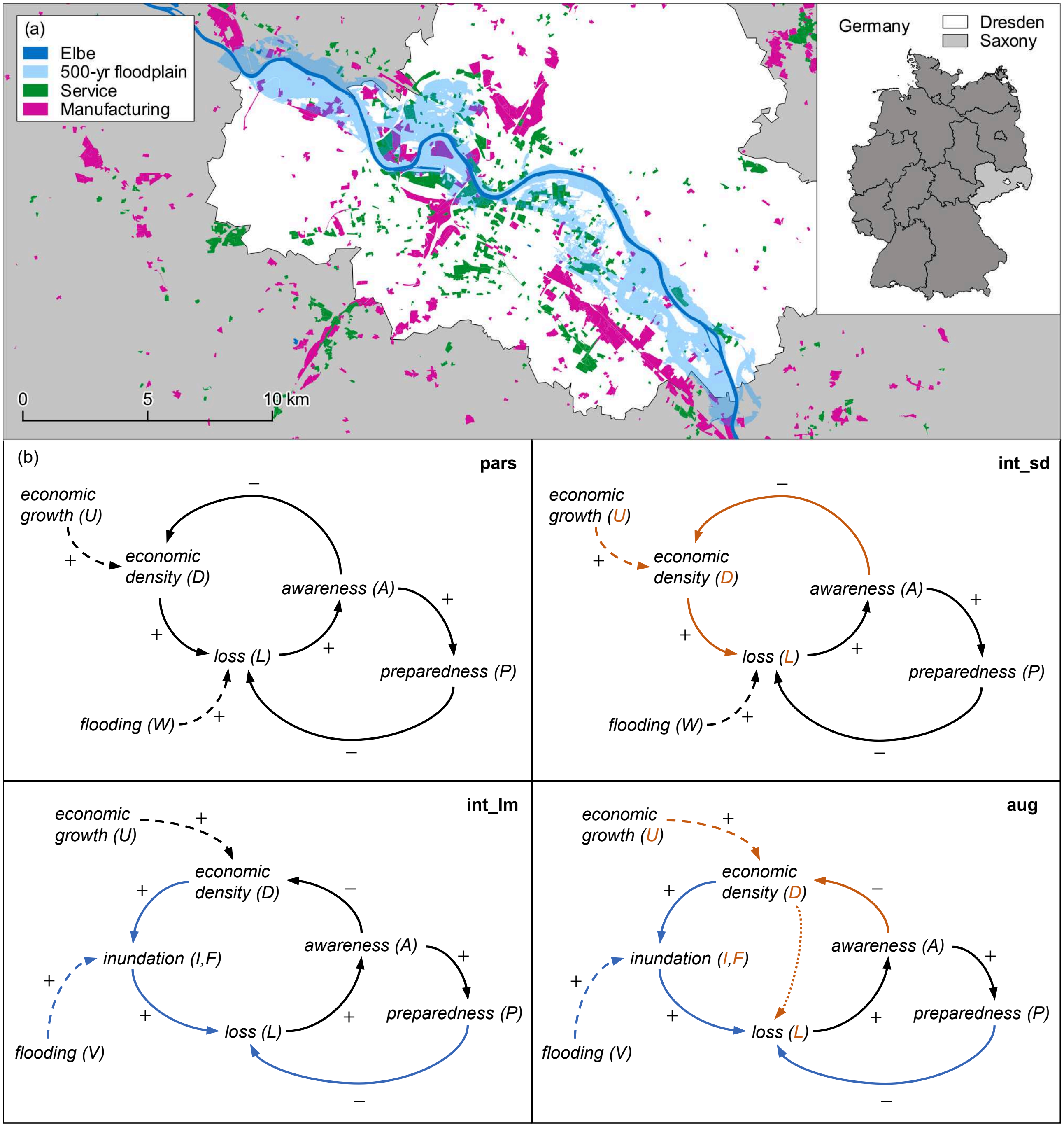

Figure 3.1: *(a) The 500-year floodplain of Dresden in Saxony, Germany, with manufacturing and service company premises (as in 2009). (b) Causal loop diagrams of the four socio-hydrological candidate models (model abbreviation in the top right of each box). System variables are represented by words and letters, while internal (solid) and exogenous (dashed) processes are indicated by arcs. Variables and processes that are augmented by the process-oriented loss estimation (blue) and the sector differentiation (orange) are highlighted in colour. The dotted orange arc between D and L in the "aug" model does not have a sign as it visualizes the weighting of sector-specific losses according to occupied floodplain area (see section 3.2.3).*

many companies are exposed to flooding and can actually incur damages. The level of preparedness influences the susceptibility of the companies to flood loss and, hence, the loss magnitude. Consequently, the flood loss and, thus, the flood risk is the product of the economic density (i.e., exposure) and the relative loss R, which depends on the flood discharge W (i.e., hazard) and the preparedness P (i.e., vulnerability). In this context, the relative loss R is the flood loss per unit area (i.e., €$/m^2$).

The socio-hydrological processes are described mathematically by three differential equations, which we split up into five equations for readability:

$$L = RD \qquad [€/€] \tag{3.1}$$

$$R = \begin{cases} R_{max} - \beta_R \exp\left(-\alpha_R \left(P_{max} - P\right) \frac{W}{W_{max}}\right), & W > H \\ 0, & W \leq H \end{cases} \qquad [\tfrac{€/m^2}{€/m^2}] \tag{3.2}$$

$$\frac{dA}{dt} = \alpha_P L \left(1 - \frac{A}{A_{max}}\right) - \mu_A A \qquad [n_c/n_c] \tag{3.3}$$

$$\frac{dP}{dt} = \begin{cases} \alpha_P \frac{dA}{dt} \left(1 - \frac{P}{P_{max}}\right) - \mu_P P, & R > 0 \\ -\mu_P P, & R = 0 \end{cases} \qquad [n_m/n_m] \tag{3.4}$$

$$\frac{dD}{dt} = U\left(1 - \alpha_D A\right) D \left(1 - \frac{D}{D_{max}}\right) \qquad [m^2/m^2] \tag{3.5}$$

Model variables (capital letters) vary over time t, which we omit in the notation for brevity. The equations contain a set of model parameters (Greek symbols) that control the strength of the variable interactions and their decay rate. The model is spatially lumped so that the parameters and variables describe the average characteristics and state of the companies in Dresden. These characteristics control the companies' behavior and, in turn, the entire dynamic of the coupled human-flood system.

Table 3.1 provides an overview of all model variables and parameters including descriptions. We chose a non-dimensional model formulation by scaling all socio-hydrological variables (i.e., W, A, P, D, R) from 0 to 1, which reduces the number of free parameters (Viglione et al. 2014). As a result, the variables $W_{\max}$, $A_{\max}$, $P_{\max}$, $D_{\max}$, and $R_{\max}$ take a value of 1. Since the awareness, preparedness, and economic density evolve over time according to the three differential equations, they require the definition of initial values A_0, P_0, and D_0. We simulated the evolution of the socio-hydrological system with a time step dt of one year, which is a reasonable time scale for the property-level adaptation through private precautionary measures of households or companies (Kreibich et al. 2007; Kienzler et al. 2015; Bubeck et al. 2020). For a more elaborate explanation of the

Table 3.1: *Variables and parameters of the socio-hydrological model. The units n_c and n_m refer to the number of companies in the floodplain and the number of implemented precautionary measures, respectively.*

External variable	Internal variable	Parameter	Description	Unit
W			Annual maximum discharge	$[(m^3/s)/(m^3/s)]$
$V^{(*)}$			Return period of annual maximum discharge	$[a/a]$
H			Protection level	$[(m^3/s)/(m^3/s)]$
U			Economic growth rate	$[1/a]$
	A		Awareness	$[n_c/n_c]$
	P		Preparedness	$[n_m/n_m]$
	D		Economic density	$[m^2/m^2]$
	$R^{(\dagger)}$		Relative loss	$[(€/m^2)/(€/m^2)]$
	L		Loss	$[€/€]$
	$I^{(*)}$		Inundation depth	$[cm]$
	$F^{(*)}$		Flooded area	$[m^2/m^2]$
		$W_{\max}$	Maximum flood discharge	$[(m^3/s)/(m^3/s)]$
		$A_{\max}$	Maximum awareness	$[n_c/n_c]$
		α_A	Anxiousness	$[1/(€/€)]$
		μ_A	Forgetfulness	$[1/a]$
		$P_{\max}$	Maximum preparedness	$[n_m/n_m]$
		α_P	Activeness	$[(n_m/n_m)/(n_c/n_c)]$
		μ_P	Deterioration rate of precautionary measures	$[1/a]$
		$D_{\max}$	Maximum economic density	$[m^2/m^2]$
		α_D	Risk-taking attitude	$[1/(n_c/n_c)]$
		$R_{\max}$	Maximum relative loss	$[(€/m^2)/(€/m^2)]$
		α_R	Effectiveness of preparedness	$[1/(n_m/n_m)]$
		β_R	Discharge to loss relationship	$[(€/m^2)/(€/m^2)]$

(*) only in models with process-oriented loss estimation; (†) only in models with simplistic loss estimation parameter interpretations and the motivation for the individual equations, refer to Barendrecht et al. (2019).

3.2.3 Model augmentation

Process-oriented loss estimation

The simplistic loss estimation in the parsimonious model infers the flood loss to buildings directly from the river discharge although monetary flood loss is commonly estimated from the inundation depth at the building, e.g., through depth-damage functions (Merz et al. 2010b; Gerl et al. 2016). Further, the simplistic loss estimation only considers the absolute flood discharge in the loss computation, although the magnitude by which the protection level is exceeded might also influence the loss severity (i.e., the difference between W and H). Apart from structural flood protection, the inundation is controlled by the topographic conditions and the location of the companies in the floodplain. In the parsimonious model, this inundation process is not modelled explicitly but rather captured by one parameter, the discharge to loss relationship β_R. Here, we substituted this simplistic loss estimation with dedicated regression models that describe the inundation and loss processes in the floodplain with more detail. As in the conceptual socio-hydrological model, these regression models are lumped and describe the average inundation and loss of companies in the floodplain. As indicated by the blue arcs in Figure 3.1(b), we fully integrated these regression models into the overarching conceptual socio-hydrological model as sub-models.

For each event, the inundation regression predicts the share of the total commercial area F that is flooded and the mean inundation depth I in these areas. The sub-model uses the event return period V and the economic density D in the floodplain at the time of the flood as predictors. This assumes that the economic density in the floodplain influences where new companies can settle. For instance, companies might have to move closer to the river as safer locations in the floodplain are already occupied. Previous socio-hydrological studies modelled this aspect similarly by simulating the distance of settlements to the river (Di Baldassarre et al. 2013; Viglione et al. 2014; Ridolfi et al. 2021). The inundation and loss regression, which we present in the following Equations 3.6-3.9, substitute for Equations (3.1) and (3.2) from the parsimonious socio-hydrological model. Since the return period, which is derived from the annual maxima series of flood discharge, and the economic density determine the flood loss via the inundation, the feedback loop of the socio-hydrological system is maintained (see Figure 3.1(b)).

Given that the observed share of flooded area in the floodplain F_{obs} can only take values between 0 and 1, we modelled it with a beta distribution (Ferrari and Cribari-Neto 2004). The observed inundation depth I_{obs} is constrained to positive values, which is why we modelled it with a gamma distribution (see e.g., Sieg et al. 2019b). The two linear regression terms of the inundation model read as follows:

$$\begin{aligned} F_{obs} &\sim Beta\,(F,\ \phi_F) \\ logit(F) &= \alpha_F + \beta_{F,D} D + \beta_{F,V} V, \end{aligned} \tag{3.6}$$

$$\begin{aligned} I_{obs} &\sim Gamma\,(I,\ \phi_I) \\ \log(I) &= \alpha_I + \beta_{I,D} D + \beta_{I,V} V, \end{aligned} \tag{3.7}$$

with intercepts α, predictor coefficients β, gamma shape parameter ϕ_F, and beta precision parameter ϕ_I. The variables F and I are the location parameters of the beta and gamma distribution respectively. The logarithm and the logit function act as link functions that guarantee plausible parameter values (e.g., I can only take positive values). Subsequently, F and I are used in the loss regression.

The loss regression is based on the Bayesian regression model by Schoppa et al. (2020). Here, we adopted a reduced version of this model considering only the two predictors that exhibited the highest explanatory power with respect to flood loss: inundation depth and preparedness (termed "precaution" in Schoppa et al. (2020)). With the predicted mean inundation depth I from the inundation regression and the preparedness P of the companies from the differential Equation (3.4), the socio-hydrological model provides two corresponding variables that can be used as predictors in the loss regression. Flood loss is commonly expressed relative to the replacement value of the building (Merz et al. 2010b) and, thus, ranges from 0 to 1. Therefore, the loss sub-model assumes that the observed building loss to companies in the floodplain L_{obs} follows a zero-and-one-inflated beta distribution ($BEINF$) (Ospina and Ferrari 2010), which is supported on the entire interval [0, 1]. This distribution mixes a beta distribution with a Bernoulli distribution and has four distribution parameters, three of which we predicted with linear predictor terms as follows:

$$\begin{aligned} L_{obs} &\sim BEINF\,(\mu_L, \phi_L, \lambda, \gamma) \\ logit\,(\mu_L) &= \alpha_{\mu_L} + \beta_{\mu_L,I} I + \beta_{\mu_L,P} P \\ logit(\lambda) &= \alpha_\lambda + \beta_{\lambda,I} I + \beta_{\lambda,P} P \\ logit(\gamma) &= \alpha_\gamma + \beta_{\gamma,I} I + \beta_{\gamma,P} P, \end{aligned} \tag{3.8}$$

where μ_L is the location parameter of the beta distribution, λ is the zero-and-one-inflation probability (i.e., the probability that the loss is 0 or 1), γ is the conditional one-inflation probability (i.e., the probability that the loss is 1 rather than 0), and ϕ_L is the precision of the beta distribution, which was not predicted. The regression intercepts and predictor coefficients are denoted by α and β. In contrast to the loss estimation in the parsimonious model, this approach differentiates between areas in the floodplain that are flooded and those that are not. The loss of the companies in the floodplain is the product of the share of flooded commercial area F, which we obtain from the inundation regression, and the mean of the zero-and-one-inflated beta distribution, which is the weighted mean of the beta and Bernoulli components of the mixture distribution (term in parentheses):

$$L = F\,(\lambda\gamma + (1 - \lambda\mu_L)). \qquad [€/€] \quad (3.9)$$

As the predicted flood loss L is expressed in relative terms, the object-level loss model can be used to approximate the aggregated flood loss to all companies in the floodplain. That is, the loss prediction is the absolute building loss of the inundated companies divided by the sum of all company building values in the floodplain. Consequently, the lumped socio-hydrological model treats the companies in the floodplain as one collective, average entity.

Sector differentiation

The second model augmentation accounts for the heterogeneity among the companies. In this way, we consider differences in the vulnerability (e.g., damage processes) and exposure (e.g., economic growth) between economic sectors. We applied a coarse sector split between companies in the industrial and manufacturing sector and the service sector, in accordance with the "NACE Rev. 2" statistical classification of economic activities of the European Union (Eurostat 2008). For instance, the manufacturing sector comprises handicraft, construction, and fabrication companies (NACE codes: B-F), while the service sector includes enterprises from commerce, finance, education, or accommodation (NACE codes: G-U). This split was primarily motivated by the thematic resolution of the available data, which did not allow for a more detailed sector differentiation. Moreover, previous findings on sectorial differences in the damage processes of building values suggest that this is a reasonable separation (Sieg et al. 2017; Schoppa et al. 2020).

We adapted the previously presented sub-models (socio-hydrological, inundation, and loss) so that they capture the differences between the two sectors. As highlighted with orange color in Figure 3.1(b), the models that include sector differentiation produce sector-specific estimations of the inundation I/F and flood loss L and allow the economic density D in the floodplain to develop separately for manufacturing and service companies. In the sector-differentiating models (int_sd, aug), the overall flood loss L is the weighted sum of the sector specific loss estimates, where the weights correspond to the contribution of each sector to the total commercial area (represented by the dotted orange arc in the "aug" model in Figure 3.1(b)). Introducing the sector differentiation required adjustments to the model structure. Firstly, the models are sector-specific for the parameters risk-taking attitude, effectiveness of preparedness, and initial economic density (α_D, α_R, and D_0). For example, the risk-taking attitude became a parameter vector $\boldsymbol{\alpha_D}$ instead of a scalar, with one entry for each sector (i.e., manufacturing and service). Secondly, we added the economic sector as a discrete predictor variable in the inundation and loss regressions (in Equations 3.6-3.8), similar to Schoppa et al. (2020). Finally, we reparametrized the probabilistic model to account for the presence of multiple sectors.

Limited detail in the historical data for awareness, preparedness, and loss hindered the creation of a full sector-specific model configuration across all variables. We had to lump the parameters that control the awareness and preparedness so that the simulations were constrained to the same value for these variables. Similarly, the loss regression could not be calibrated with sector-specific loss reports since disaggregated estimates were only available for the 2002 flood. However, the economic density and loss estimation allowed for disparity between the sectors, which could propagate through the coupled socio-hydrological system and reveal distinct risk dynamics. The model equations for the sector differentiating models can be obtained from the Supplementary material (Text B.6).

3.2.4 Bayesian parameter estimation using empirical data

We estimated the parameters of the four socio-hydrological model versions from empirical data by means of Bayesian inference (Gelman et al. 2013; McElreath 2018; Schoot et al. 2021). The data that inform the models are composed of hydrological time series, inundation maps, telephone surveys, historical land-use maps, and economic data. Table 3.2 provides an overview of the model data. We confined the socio-hydrological system spatially by the area that a flood with a return period of 500 years would inundate (see Figure 1(a)). Accordingly, the data describe the average of the model variables within this maximum floodplain area. While data for the forcing variables (W, V, H, U) are available for the entire study period, observations for the socio-hydrological system variables (A, P, D, L) are only available in certain years. The model simulations estimate the state of these variables in years without data coverage. The introduced model augmentations enhance the amount of data that is available for parameter learning by time-invariant observations. The process-oriented loss estimation is informed by object-level loss data from telephone surveys (n=597) and inundation data (int_lm:

Table 3.2: *Data used for parameter estimation. The temporal coverage indicates for which years of the simulation period the respective data are available.*

Model variable	Data	Temporal coverage	References
W/V - flood discharge & return period	Annual maxima discharge series	1900-2017	German Federal Institute of Hydrology (BfG) (2021)
H - protection level	Reconstruction from previous studies, historical and authority reports	1900-2019	Barendrecht et al. (2019), Federal Dam Operation Authority of Saxony (2013), Kreibich and Thieken (2009), Pohl (2004), Weikinn (2000), and Weikinn (2002)
U - economic growth rate	Gross domestic product growth rate in Dresden	1900-2019	Paprotny et al. (2018a)
A - awareness	Telephone surveys	2002, 2003, 2007, 2014	GFZ (2021), Kreibich et al. (2007), and Thieken et al. (2016)
P - preparedness	Telephone surveys	2002, 2003, 2006, 2007, 2013, 2014	GFZ (2021), Kreibich et al. (2007), and Thieken et al. (2016)
D - economic density	Historical land use maps	1900, 1940, 1953, 1968, 1986, 1998, 2009	Gruner (2012) and Rosina et al. (2020)
L - loss	Data and report on flood relief, telephone surveys, spatial data on asset values	2002, 2006, 2013	GFZ (2021), Kreibich et al. (2007), Paprotny et al. (2018a), Saxonian Relief Bank (2007), and Thieken et al. (2016)
I/F - inundation depth & flooded area	Inundation maps, historical land use maps	Assumed to be time-invariant	Gruner (2012), Rosina et al. (2020), and Saxonian Environmental Agency (2012)

n=26; aug: n=56), while the sector differentiation doubles the economic density data (i.e., one set of observations per sector) in comparison to the aggregated approach. Bayesian parameter estimation inherently quantifies uncertainties in the model, parameters, and observations. For the socio-hydrological system variables, we assessed the observational uncertainty based on the dataset size or domain knowledge. The Supplementary material (Texts B.3-B.5) provides further information on data processing and uncertainty (Hosking 1990; Ferrari and Cribari-Neto 2004; Maier 2014; Delignette-Muller and Dutang 2015; Sennhenn-Reulen 2018).

Bayesian inference allows for the incorporation of information from previous experiments into the parameter estimation through priors. Here, we adopted the posterior parameter estimates from the residential model from Barendrecht et al. (2019) as priors for the socio-hydrological parameters in the new company models. In doing so, we assumed that the adaptive behavior of companies in Dresden is to some degree related to the actions of residential households. To ensure that the adopted, informative priors do not bias the inference, we conducted prior predictive checks (i.e., checking the plausibility of the prior through simulation) and tested different priors. Details on the prior distributions, the prior checking, and the computational implementation of the Bayesian models in the probabilistic software Stan are contained in the Supplementary material (Texts B.2 and B.7, and Tables B.1-B.3) (Hoffman and Gelman 2014; Bürkner 2017; Carpenter et al. 2017; Gelman et al. 2020).

3.3 Results and discussion

3.3.1 Socio-hydrological simulation

Temporal dynamics of flood risk

Using the four candidate models and empirical data, we estimated the model parameters and simulated the co-evolution of the socio-hydrological flood system for companies in Dresden over the period 1900-2019. In the following, we evaluate whether the models reproduce the observed adaptation effect in Dresden successfully. Figure 3.2 shows the fit of the four candidate models to the socio-hydrological observations. The simulated means of the model variables are shown with 95% credible intervals against the observations. The candidate models with the sector differentiation predict the development of the economic density separately for the manufacturing and service sectors.

The models agree on the evolution of the economic density in the floodplain,

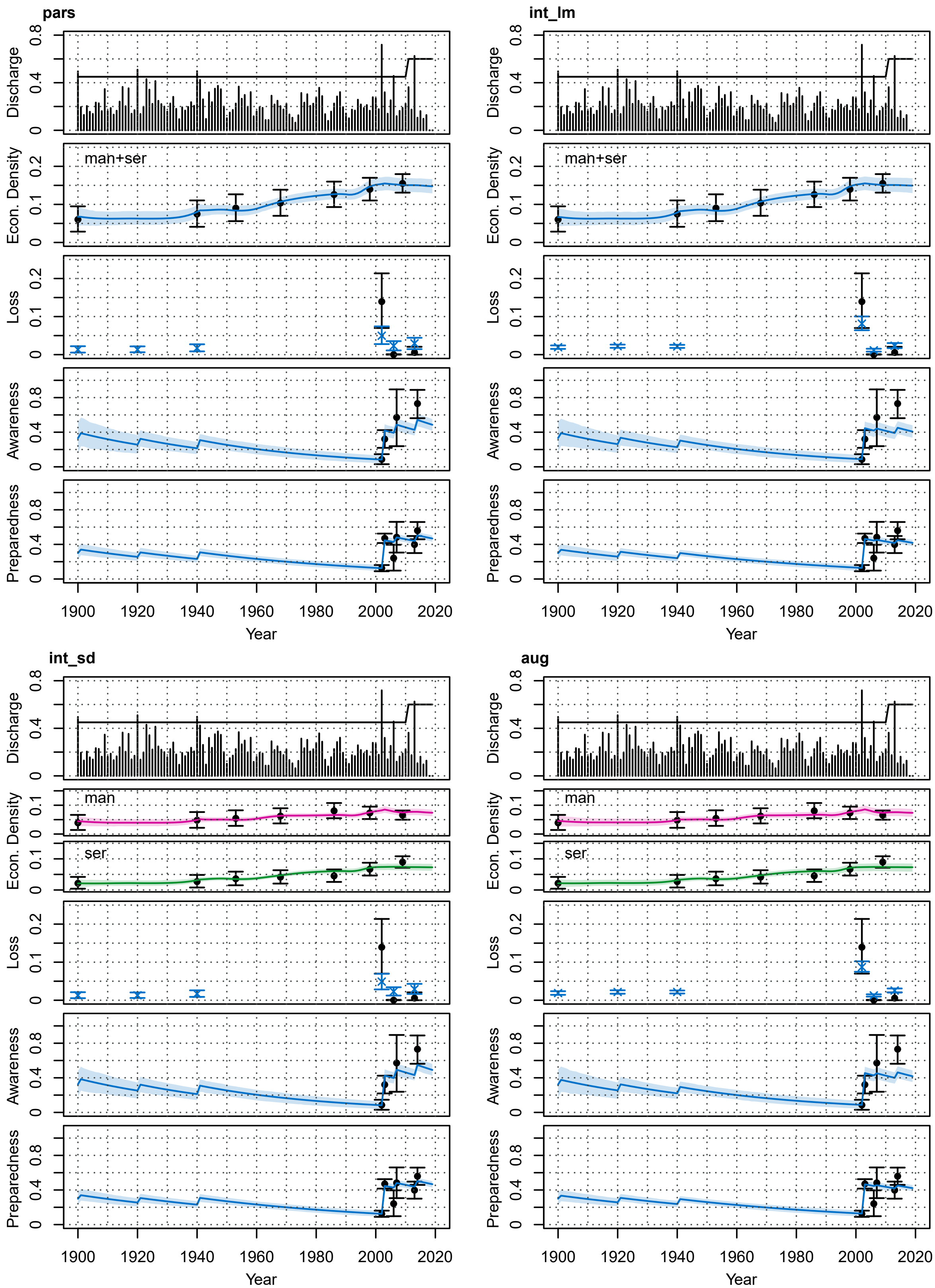

Figure 3.2: *Fit of the candidate models to data. Each plot panel shows one candidate model: pars, parsimonious; int_lm, intermediate with process-oriented loss estimation; int_sd, intermediate with sector differentiation; aug, fully augmented.*

and the simulations are generally within the credible intervals of the observations. In contrast, the candidate models show larger variation in the estimations of flood loss and the predictions match the reported losses worse than they do for the economic density. The models with the process-oriented loss estimation (int_lm, aug) predict larger losses for the 2002 event and lower losses for the 2006 and 2013 events than the models with the simplistic loss estimation (pars, int_sd). We discuss the performance of the individual models in the loss estimation in more detail in section 3.3.2. As the awareness directly depends on the loss magnitude, the awareness time series of the candidate models diverge after the severe 2002 flood. The models with the simplistic loss estimation reproduce the awareness data better, but at the cost of overestimating the 2006 and 2013 flood losses. Model differences in the preparedness time series are less pronounced since the preparedness only indirectly depends on the flood loss via the awareness. Overall, the preparedness simulations agree with the observations.

The adaptation of the companies after the severe 2002 flood is captured accurately. The increase in awareness and preparedness was also reported in comparable empirical analyses of the flood event (Kreibich et al. 2007; Jehmlich et al. 2020). The models do not suggest that damaging flood events substantially affect the settling or abandonment of the floodplain by the companies. Instead, other motives such as economic growth seem to govern the development of the economic density in the Elbe floodplain. Jehmlich et al. (2020) conducted qualitative interviews with companies in Dresden and reported that emotional attachment, tradition, and continued benefits of a location in the floodplain (e.g., proximity to customers) also induce companies to stay.

The uncertainty in the simulations of the economic density, awareness, and preparedness is largest in 1900 and decreases towards the present. Overall, the confidence is particularly low in the case of the awareness, compared to the other variables. The uncertainties reflect the availability of historical data and the information content in the prior for the respective variable. Specifically, a comparably large number of observations is available for the economic density, and the prior for the initial preparedness P_0 is comparably strong (see Figure 3.3). In contrast, the prior on the initial awareness A_0 is relatively weak, and the awareness data are most uncertain and smallest in number. In addition, awareness takes a pivotal position in the socio-hydrological system with connections to three other random variables (see Figure 3.1(b)). This allows for strong variable interaction and leads to an accumulation of uncertainties in the awareness simulations.

In summary, all tested model structures are capable of reproducing the essential dynamics of the coupled human-flood system, especially the adaptation effect. Variations across model simulations mainly affect the loss and awareness estimates and arise from the difference in the loss estimation.

Insights on the adaptive behavior of companies

The parameter estimates of the candidate models describe the adaptive behavior of companies in Dresden with respect to the flood risk. Figure 3.3 shows the marginal prior and posterior distributions of the socio-hydrological parameters in the four candidate models. Model parameters with subscripts (i.e., "man" and "ser") refer to sector-specific parameters that are included in the candidate models with the differentiation.

The posteriors reveal whether companies behave differently than private households, as they can be compared to the model fit of Barendrecht et al. (2019) for the residential sector in Dresden. Unless otherwise noted, we adopted the posteriors of this residential model as priors for our company models so that differences are directly visible in Figure 3.3. The estimated risk-taking attitude (α_D) of the companies is larger in the median than the adopted a priori parameter value. This indicates that companies in Dresden are less risk-taking than private households with respect to populating the floodplain. That is, commercially used areas grow more slowly and disintegrate more rapidly than residential areas. The fits suggest a slight difference between the economic sectors in the risk-taking attitude, but its magnitude is small given the level of uncertainty. For the anxiousness (α_A), we assigned a prior that is smaller than the residential posterior because the estimate for the private households proved to be implausibly high for the company model. With median values around 4.7 (int_lm, aug) and 7.6 (pars, int_sd), the posterior company anxiousness is lower than the reported anxiousness of private households (median: 11). The parameter directly depends on the magnitude of flood loss and, hence, the estimates differ relatively strongly between the company models with and without the process-oriented loss estimation. The candidate models agree on the activeness (α_P) and suggest that, given the same level of awareness, companies implement more precautionary measures than private households because the posterior estimates exceed the prior. For the effectiveness of the precautionary measures (α_R), we chose a prior that allowed for larger parameter values and was less informative than the posterior from the residential model. The comparison of the posteriors points towards larger effectiveness of the preparedness for companies (α_R: 0.61, $\alpha_{R,man}$: 0.65, $\alpha_{R,ser}$: 0.53) than for private households in the median (0.16). The forgetfulness (μ_A) and the decay rate of precautionary measures (μ_P) are lower than for the private households, which can be interpreted more intuitively when expressed as half times (i.e., the time until the awareness and preparedness are halved). Depending on the candidate model, the median half time of the awareness lies between 32 and 35 years, which is substantially longer than the half time for private households (21 years). The median half time of precautionary measures varies between 46 and 50 years across the company

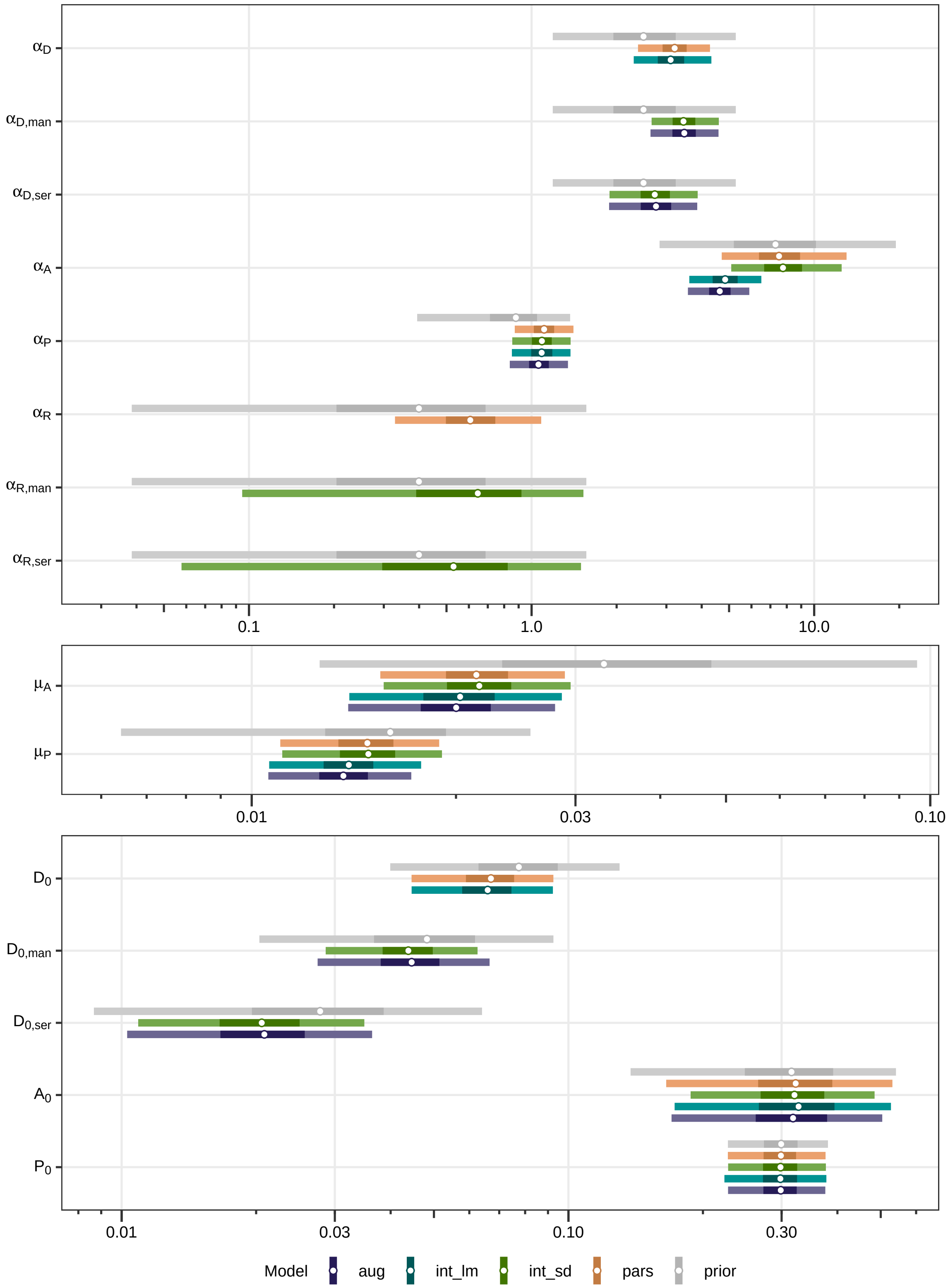

Figure 3.3: *Marginal posterior distributions (log-scale) of the socio-hydrological parameters in the four candidate models: pars, parsimonious; int_lm, intermediate with process-oriented loss estimation; int_sd, intermediate with sector differentiation; aug, fully augmented. The marginal prior distributions are the adopted posterior distributions from Barendrecht et al.'s (2019) model for the residential sector. The points show the median, while the bars correspond to 50% and 95% credible intervals.*

models, which is only slightly larger than the value for the residential sector (43 years). The initial values of the economic density (D_0) cannot be compared to the settlement density of private households since the variables describe distinct quantities. The variation in the simulated awareness time series also reflects in the initial awareness (A_0), which varies comparably strongly across the candidate models. The initial preparedness (P_0), however, is similar for the four company models. The posteriors indicate that the company awareness and preparedness in the year 1900 was similar to that of the private households, yet the initial values of these two variables are relatively uncertain parameters.

In summary, the companies in Dresden are not as anxious as private households, but they are less risk-taking and less forgetful, and more actively undertake precautionary measures. The posteriors of the sector differentiating parameters imply minor differences in behavior between the manufacturing and the service sector. However, these deviations are small in comparison to the associated uncertainties and do not allow for robust statements. Overall, the parameter estimates and the simulated time series (section 3.3.1) show that companies reduce their vulnerability through private precautions, rather than reducing their exposure through resettling. This is in line with the qualitative interviews of Jehmlich et al. (2020), where a considerably larger share of companies decided to undertake precautionary measures instead of dissolving or moving away.

Information content in priors and data

The contraction of the posterior relative to the prior densities in Figure 3.3 shows that, for most parameters, the data convey additional information that reduces the a priori parameter uncertainty. The sector specific effectiveness of preparedness and the initial values of the awareness and preparedness ($\alpha_{R,man}$, $\alpha_{R,ser}$, A_0, P_0), however, are informed less by the data and, in turn, depend more strongly on their priors.

The plot also highlights the benefit of using informative priors, especially for socio-hydrological models where datasets are usually small. The majority of the posterior parameters in the company models and, thus, the simulated variable time series exhibit considerably lower uncertainty than the posteriors of the private model, which act as priors in the company models. Yet the number of socio-hydrological data points for the inference were comparable in the two studies. For the residential model, however, no informative a priori knowledge from previous studies was available, resulting in larger a posteriori parameter uncertainty. The prior predictive checking during model setup indicated that the informative priors did not bias the inference (e.g., through underfitting or underestimating uncertainty) but rather increased the numerical stability of the models.

Correlations between model parameters or model overparameterization can inflate the associated uncertainties. For instance, the intermediate model with the sector differentiation (int_sd) resolves differences in the effectiveness of preparedness across sectors ($\alpha_{R,man}$, $\alpha_{R,ser}$) although no loss data for individual sectors is available. As a result, the parameters can only be identified indirectly via the sector-specific economic density data, leading to comparably large parameter uncertainty. The fully augmented model (aug), which also differentiates between sectors, does not suffer from this problem as the object-level survey data carry the necessary information on the inter-sectorial differences of damage processes.

Consequently, the socio-hydrological system processes that are resolved in the model require sources of information for parameter identification, either directly or indirectly through connected system variables. Our results show that the use of informative prior distributions, obtained from previous works, can complement the information provided by data, ultimately reducing uncertainty. In general, a deliberated prior choice in consideration of established practices such as prior predictive checking (Gabry et al. 2019; Gelman et al. 2020) promotes meaningful socio-hydrological inference.

3.3.2 Flood loss estimation

Predictive accuracy and uncertainty

This work aims at improving the loss estimation in socio-hydrological flood risk models. Based on the accuracy and the uncertainty of the loss predictions, we assess the skill of the simplistic and process-oriented loss estimation. Figure 3.4(a) compares the estimated flood loss distributions of the four candidate models and the observed loss. The predictive error of each probabilistic loss estimate is quantified by the continuous ranked probability score (CRPS), where a perfect fit is indicated by a value of 0. In each plot panel, the best CRPS value is underlined. The loss estimates differ particularly between the models that feature the process-oriented loss estimation (int_lm, aug) and those that rely on the simplistic loss estimation (int_sd, pars). The process-oriented loss estimation predicts all three loss events more accurately as indicated by consistently lower CRPS values, which are up to twice (i.e., 2006) as high for the simplistic loss estimation. In general, the predictions of the process-oriented loss estimation better capture the range in observed loss magnitudes between the individual events - from the minor 2006 to the severe 2002 loss. Moreover, the loss distributions of the process-oriented loss estimation are associated with considerably lower uncertainties than the predictions of the simplistic loss estimation. The parsimonious model (pars)

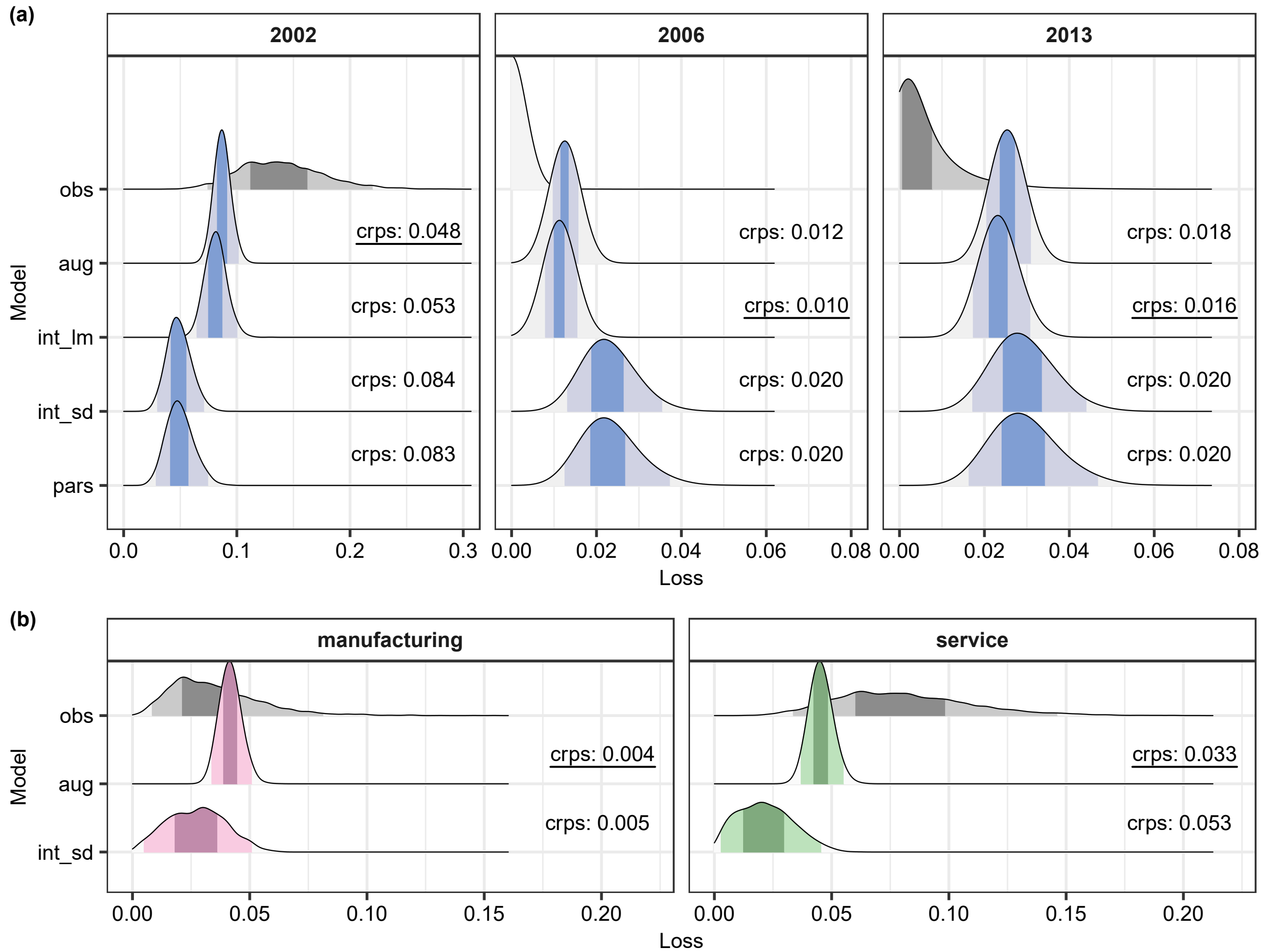

Figure 3.4: *Comparison of modelled and reported flood losses. (a) Aggregated losses for the three observed flood events; (b) sector-specific losses of the 2002 flood. The shaded areas under the curves show 50% and 95% credible intervals. The continuous ranked probability score (crps) quantifies the error of the loss predictions; the best fit is underlined. Model codes: obs, observation; pars, parsimonious; int_lm, intermediate with process-oriented loss estimation; int_sd, intermediate with sector differentiation; aug, fully augmented.*

yields the widest predictive distributions across the three observed events whereas the fully augmented model (aug) produces the narrowest predictive distributions, with 95% credible intervals up to four times smaller.

The advantages of the process-oriented loss estimation arise from three aspects: increased detail in the representation of the damage process, greater flexibility of the probabilistic model, and additional data. First, the simplistic loss estimation is based on the diffuse relationship between flood discharge and loss. In contrast, the process-oriented model estimates the flooded area and the inundation depth, allowing for a predictor set with higher explanatory power. Secondly, the loss model augmentation addresses the common overdispersion of loss data with the dedicated inflation parameters of the zero-and-one-inflated beta distribution (λ and γ in Equation 3.9). The 2006 event underlines the benefit of this inflation, where the flood protection level was exceeded and caused a flood, but the resulting loss was nearly zero due to the small margin between the discharge and the protection level and the efficacy of the preparedness. The simplistic structure of the standard loss estimation is not capable of reproducing such threshold effects. Thirdly, the inundation and loss regression models are jointly informed by the socio-hydrological loss observations and the survey loss data. Although this complex loss model estimation comprises more parameters than the standard loss estimation, it has access to a far larger data pool for parameter inference (n=656 vs. n=3).

Yet even the fully augmented model (aug) underestimates the variation in the reported loss values. In the case of the 2002 flood, the underestimation can be explained by the spatial domain of the model, which only covers the Elbe floodplain. In this event, however, considerable parts of the city were inundated by the Elbe tributaries Weißeritz and the Lockwitzbach, which also flow through Dresden (Kreibich and Thieken 2009). As it is difficult to allocate the contribution to the overall loss in Dresden to the different rivers, we adopted the reported 2002 loss for the entire city. Under these circumstances, we can conclude that the loss estimates for 2002 are better than suggested by the figures, since the loss that is caused by the river Elbe must have been lower than the overall loss. While the confinement of the model domain to the main river is necessary to maintain a manageable socio-hydrological system, this oversimplification can cause biased loss estimates in the occurrence of compound events as in 2002.

The variation in the loss distributions due to the sector differentiation is small compared to the variation between models with different loss estimation approaches. Candidate models that share the same loss estimation routine (simplistic: pars, int_sd; process-oriented: int_lm, aug) exhibit similar CRPS values independently of how they treat the economic sectors (aggregated or differentiated). Small differences in CRPS (up to 0.005) occur between the fully augmented (aug) and the intermediate model with the process-oriented loss estimation (int_lm), with an advantage of the former (aug) for major and of the latter (int_lm) for minor loss events. Figure 3.4(b) displays loss predictions of the sector-differentiating models (int_sd, aug) for the manufacturing and service sectors for the 2002 flood, the only event for which sector-specific loss reports are available. Again, the model with the process-oriented loss estimation outperforms the model with the simplistic loss estimation for both sectors. Both models predict the loss of the manufacturing sector more accurately than for that of the service sector.

Reliability of loss estimation

The previously presented loss estimates reflect the training performance of the models and, hence, overestimate the true predictive capacity of the loss estimations for unseen data. Therefore, we conducted an LOO-CV experiment, in which we recursively fitted the models to the data, each time leaving out one of the three observed loss events in Dresden. The goodness of fit to the held-out loss events provides insight on the models' capacities to assess the flood loss of new events and has implications for the reliability of the candidate models.

Figure 3.5 summarizes the results of the LOO-CV experiment. Again, the columns of the plot show the estimated and observed company flood loss for the three reported flood events (2002, 2006, 2013). In each row, another loss event was held-out of the training dataset. This means that the panels on the diagonal (background shading) are of special importance because they express the predictive skill of the models for new data. The loss estimates in the LOO-CV experiment are indistinguishable from the estimates of the model calibration runs (Figure 3.4). The CRPS metrics show that the candidate models with the process-oriented loss estimation assess the three held-out loss events more precisely and with less uncertainty than the models with the simplistic loss estimation. In addition, the increase from training to validation error for the simplistic loss estimation (pars, int_sd; up to 75% increase in CRPS) is larger than that for the process-oriented loss estimation (int_lm, aug; up to 9% increase in CRPS).

More importantly, the plot reveals that the process-oriented loss estimation provides more robust predictions than the simplistic loss estimation. When considering the plot panels within one column, we see that the loss distributions and predictive errors (i.e., CRPS) of the models with the augmented loss estimation (int_lm, aug) fluctuate less across the different training datasets than the distributions of the simplistic loss estimation (pars, int_sd). This implies that the simplistic loss estimation relies more strongly on the available loss data, which can lead to systematic underestimation when the training dataset does not contain observations of rare, high magnitude loss events. Since the process-oriented loss estimation combines the aggregated, large-scale losses from the socio-hydrological data with the vulnerability information from the object-level flood loss data, it is capable of extrapolating more reliably to unseen flood magnitudes. This is of particular advantage in socio-hydrological studies since historical flood loss reports are commonly scarce and short discharge records might not contain extreme floods.

Overall, the scarcity of loss reports for historical floods only allows for an evaluation of the predictive model performance for three events. Nevertheless, the training and validation errors coherently indicate that the process-oriented

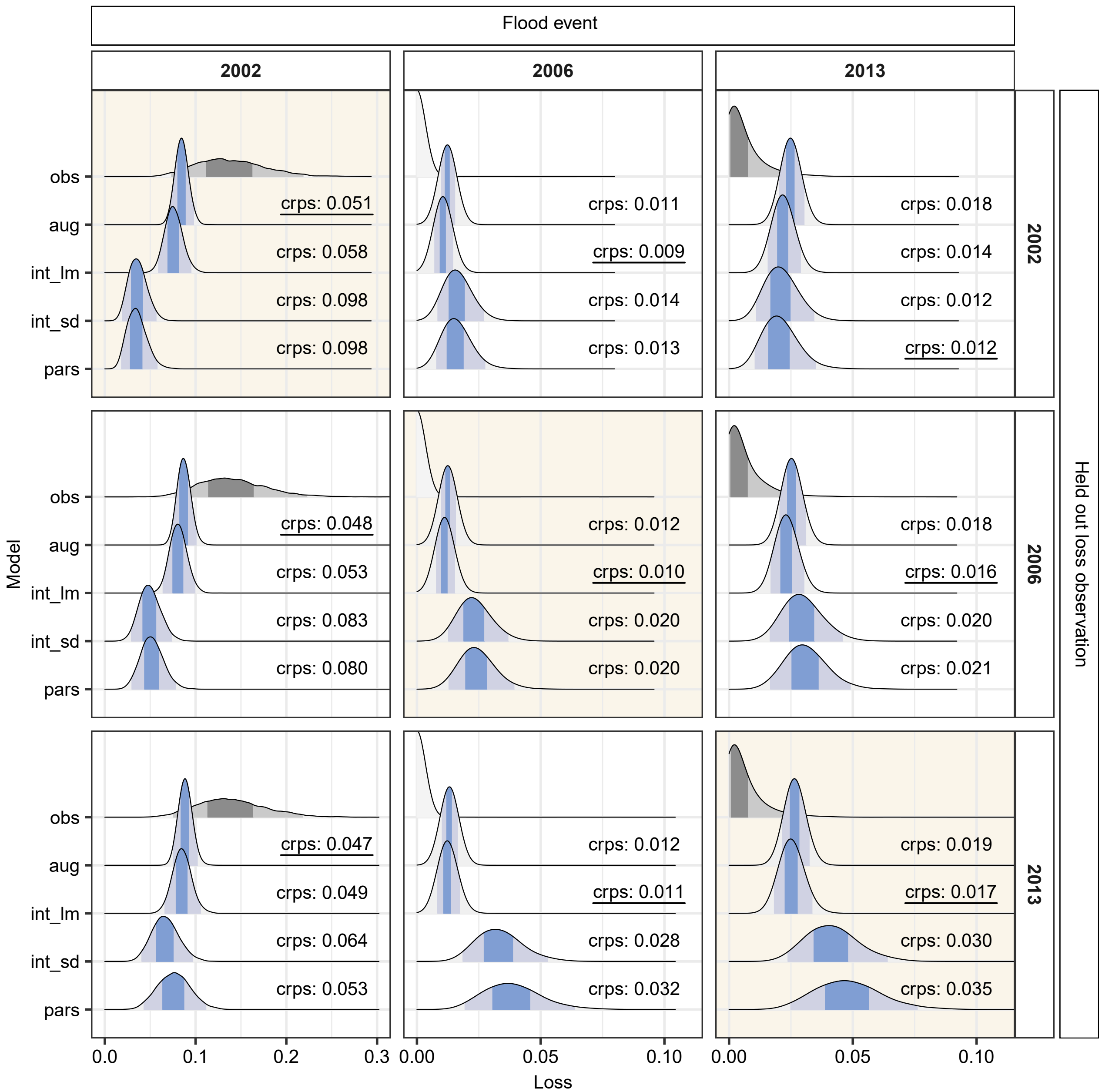

Figure 3.5: *Modelled and observed losses (as in Fig. 3.4(a)) for the leave-one-out cross-validation experiment. Panel rows indicate which flood event was held-out during model training, while in each column the same loss event is displayed. Plot panels with background shading highlight the predictions for unseen data. Model codes: obs, observation; pars, parsimonious; int_lm, intermediate with process-oriented loss estimation; int_sd, intermediate with sector differentiation; aug, fully augmented.*

loss regression model (int_lm, aug) outperforms the simplistic loss estimation (pars, int_sd). On the contrary, the sector-specific modeling has a minor influence on the loss estimates and, given the level of uncertainty, we cannot assess the performance differences between the aggregated (par, int_lm) and sector specific candidate models (int_sd, aug) confidently. Possibly, performance differences might emerge when additional sector-specific loss reports become available for the validation of the loss estimates.

3.3.3 Potential of augmentations in socio-hydrological modeling

Our results show that the presented augmentations increase the accuracy, confidence, and reliability in the loss estimates of the socio-hydrological flood risk model. The loss estimation benefits from the inclusion of the inundation and loss regression, which resemble the physical reality of the damage processes more closely and feature a refined probabilistic model. The sector differentiation did not improve the loss estimation conclusively. Since we lumped the awareness, preparedness, and loss across sectors due to data constraints, more distinctive risk dynamics between the sectors might have been attenuated. Conceivably, the influence of the sector differentiation on the loss prediction and the entire socio-hydrological system could be larger if these variables, conditional on sector-specific observations, were also allowed to develop individually for each sector or if the society under consideration involved more distinct actors - for example, in a model that considers private households and companies. The augmentations add further complexity to the socio-hydrological flood risk model, and, yet, the substantial increase in training data outweighs the increase in the number of parameters, ultimately reducing uncertainty.

As flood loss represents a central component in the coupled human-flood system (see Figure 3.1(b)), the effect of the improved loss estimation enhances the validity of the entire socio-hydrological flood risk model. A biased loss estimation could propagate through the entire socio-hydrological system, leading to unrealistic system evolutions and misguided conclusions about the behavior of society. The LOO-CV experiment shows that the process-oriented loss estimation provides more reliable loss estimates even in the absence of numerous reported loss events. This characteristic promotes the prospective transferability of the socio-hydrological flood risk model in space and time. Thus, the object-level flood loss data, which stem from various regions in Germany, facilitate the model application at other study sites with comparable socioeconomic conditions (e.g., building codes). In addition, credible loss estimates are a prerequisite for sound projections of the socio-hydrological flood system in Dresden into the future.

While this study focused on the improvement of one specific process in a socio-hydrological flood risk model (i.e., loss estimation), the notion of process augmentation could be extended to other components of the human-flood system. Socio-hydrological models are modular frameworks that stipulate how the considered system variables interact and co-evolve. Depending on the required degree of process detail, we could selectively replace one or several simplistic mathematical process representations with more informed estimation techniques, conditional on domain knowledge and additional data. As the targeted enhancement of socio-hydrological processes increases model complexity, it is only advisable when suitable and sufficient data are available to inform the additional parameters. Similarly, model augmentations might hinder the spatial transfer to other case studies where these additional data requirements cannot be satisfied. Particularly for variables that map the individuals or entities of society, like awareness or preparedness, data collection is intricate and expensive because it commonly relies on interviews or surveys (Barendrecht et al. 2019).

Returning to the human-flood system, next steps could aim at improving the representation of how households and companies become aware of the flood risk and what drives them to take action to protect themselves. Protection motivation theory provides a conceptual basis and models that could be added the socio-hydrological flood risk model in addition to the process-oriented loss model (Grothmann and Reusswig 2006; Bubeck et al. 2018). In the end, model development remains an iterative process, where recursive updates of the employed data streams or the model structure can improve the capacity of existing models to reproduce human-water dynamics and reduce the simulation uncertainty (Thompson et al. 2013; Hipsey et al. 2015; Sivapalan and Blöschl 2015).

3.4 Conclusions

All versions of the developed socio-hydrological flood risk model are capable of reproducing the adaptation effect for companies in Dresden that was observed over the past 20 years. The model augmentation, mainly in the form of process-oriented loss estimation, improves the accuracy and reliability of the loss estimates and reduces their predictive uncertainty (research question 1). The simulations suggest that companies settle more cautiously in exposed locations in the floodplain and prepare themselves more actively against flooding than private households do (research question 2).

Consequently, the augmented socio-hydrological flood risk model provides higher reliability for further analyses than the parsimonious model; for example, for projecting the evolution of the coupled human-flood system in Dresden into the future. In general, the informed augmentation of socio-hydrological models of all kinds (e.g., for drought or water resources management) by process-oriented model components facilitates the model transfer in space (i.e., to other study sites) and time (i.e., projections). After the integration of empirical data, the inclusion of validated, empirical models that reflect current process understanding represents the next step towards more precise and credible socio-hydrological modeling.

4 Projecting Flood Vulnerability Dynamics for Effective Long-term Adaptation

Abstract

Authors:
Lukas Schoppa
Marlies H. Barendrecht
Dominik Paprotny
Nivedita Sairam
Tobias Sieg
Heidi Kreibich

Under review:
Schoppa, L., Barendrecht, M.H., Paprotny D., Sairam, N., Sieg, T. and Kreibich, H. Projecting Flood Vulnerability Dynamics for Effective Long-term Adaptation. Earth's Future.

Flood losses have steadily increased in the past and are expected to grow even further owing to climate and socioeconomic change. The reduction of flood vulnerability, for example through adaptation, plays a key role in the mitigation of future flood risk. However, lacking knowledge about vulnerability dynamics, which arise from the interaction between floods and the ensuing response by society, limits the scope of current risk projections. We present a socio-hydrological method for flood risk assessment that simulates the interaction between society and flooding continuously, including changes in vulnerability through collective (structural) and private (non structural) measures. Our probabilistic approach quantifies uncertainties and exploits empirical data to chart risk dynamics including how society copes with flooding. In a case study for the commercial sector in Dresden, Germany, we show that increased adaptation is necessary to counteract the expected four-fold growth in flood risk due to transient hydroclimatic and socioeconomic boundary conditions. We further use our holistic approach to identify solutions for effective long-term adaptation, demonstrating that integrated adaptation strategies (i.e., combined structural and non structural measures) can reduce the average risk by up to 60% at the study site. Ultimately, our case study highlights the benefit of the model for robust flood risk assessment as it can capture unintended, adverse feedbacks of adaptation measures such as the levee effect. Consequently, our socio-hydrological method contributes to a more systemic and reliable flood risk assessment that can inform adaptation planning by exploring the possible system evolutions comprehensively including unlikely futures.

4.1 Introduction

Global change has sparked an increase in economic river flood losses over the past decades (Bevere and Remondi 2022; Tanoue et al. 2016; Barthel and Neumayer 2012). The historic rise in flood losses has mostly been attributed to demographic and economic growth and a concomitant accumulation of exposure in floodplains (Visser et al. 2014; Kundzewicz et al. 2014; Paprotny et al. 2018b; Jongman et al. 2012). Anthropogenic global warming has not been a dominant control of flood risk change in the past, but its influence might grow in the future due to emerging shifts in flood hazard (Merz et al. 2021; Neumayer and Barthel 2011; Bouwer 2011). Continued exposure growth and climate change are going to propel flood risk even further in most regions of the world (Merz et al. 2021; Hirabayashi et al. 2013; Jongman et al. 2012). As a result, global average annual flood loss could grow by up to a factor of ten until the end of the century compared to today under the assumption of constant flood vulnerability (Dottori et al. 2018; Winsemius et al. 2016; Alfieri et al. 2018). The reduction of vulnerability through flood adaptation has proven effective in the past (Tanoue et al. 2016; Jongman et al. 2015; Hudson et al. 2014; de Moel et al. 2013; Sairam et al. 2019b; Poussin et al. 2015) and, hence, is a key element in the effort to offset the expected intensification of impacts (Dottori et al. 2018; Winsemius et al. 2016; Hirabayashi et al. 2013; Jongman 2018; Kinoshita et al. 2018).

Changes in flood vulnerability are difficult to trace as it is a multidimensional quantity that is determined by physical, economic, institutional, and social factors (UNDRR 2022; Merz et al. 2010a), which impedes the collection of continuous and extensive data. Similarly, adaptation measures differ in type (e.g., structural, nature-based), scale (country, object-level) and the implementing actor (government, individual households) (Jongman 2018; Dottori et al. 2020). Therefore, the dynamics of vulnerability and their effect on flood risk are understood less in comparison to hazard and exposure (Kreibich et al. 2017a). This also reflects in prevalent flood risk assessment, where vulnerability changes are usually not considered (Metin et al. 2018). Some studies attempted to bridge this gap by running risk simulations assuming different levels of adaptation (i.e., discrete and constant vulnerability scenarios) (Metin et al. 2018; Jongman et al. 2015; Steinhausen et al. 2022). While this approach goes further than most previous risk analyses, it still cannot capture the continuity of vulnerability change which arises from the constant interplay between flood events and society at different time scales (Sivapalan and Blöschl 2015). For instance, damaging floods demonstrably trigger adaptation response by the affected society in the aftermath of the event (Di Baldassarre et al. 2015; Kreibich et al. 2017a) and, conversely, flood-poor periods might lead to a decay in a societies' risk awareness (Fanta et al. 2019; Viglione et al. 2014). Altogether, limited understanding of the causal factors of vulnerability change and narrowly defined model boundaries do not embrace the complex reality of floodplains (Merz et al. 2015; Jongman et al. 2015), which becomes even more relevant in an increasingly globalized and rapidly changing world. Such knowledge is essential for the reliable projection of future adaptation and the quantification of its risk reduction potential (Aerts et al. 2018; Dottori et al. 2018).

Following the call for more systems-thinking in flood risk analysis (UNDRR 2022; Schröter et al. 2021; Barendrecht et al. 2020; Merz et al. 2015; Di Baldassarre et al. 2016), recent studies integrated vulnerability as an inherent component in applied flood risk assessment. Haer et al. (2017) developed an agent based model that incorporates the dynamic flood adaptation by households. It was then extended to governments to study the future evolution of vulnerability and risk under different behavioral scenarios (Haer et al. 2019; Haer et al. 2020).

Socio-hydrological system dynamics models are a parsimonious alternative to agent based models as they only resolve the most essential components and processes on the systems-level (e.g., a floodplain) and focus on the overall co-evolution of human-flood systems (Blair and Buytaert 2016). Having originated from studies that investigated hypothetical systems (Di Baldassarre et al. 2013; Viglione et al. 2014), these models have recently progressed towards a data-informed solution for quantitative tracing of vulnerability and risk dynamics (Barendrecht et al. 2019; Schoppa et al. 2022). In such models, small-scale variations are treated probabilistically (e.g., via Bayesian methods) rather than being resolved explicitly, which facilitates the efficient exploration of future flood risk projections while including the numerous sources of uncertainty.

Here, we present an efficient socio-hydrological method for continuous flood risk assessment that expands the conventional focus on hazard and exposure changes by explicitly including vulnerability dynamics. The probabilistic approach is calibrated on empirical data and quantifies uncertainties by means of Bayesian inference. In a pilot application for the commercial sector, we (1) project flood risk trajectories until the end of the 21st century accounting for the transient nature of all three risk drivers (hazard, exposure, vulnerability) and (2) assess the effectivity and robustness of adaptation strategies against the background of exacerbating hydroclimatic and socioeconomic boundary conditions. To our knowledge, this is the first time that a socio-hydrological system dynamics model is used in a quantitative projection study of flood risk. Our method helps to link the expected large-scale patterns of change (e.g., in climate, demographics, economy) to the still largely unknown local response of human-flood systems (Jongman et al. 2015) and could contribute to unfolding the full potential of societal flood adaptation.

4.2 Methods and data

4.2.1 Socio-hydrological flood risk projection

At the core of the method is a socio-hydrological model (Barendrecht et al. 2019; Schoppa et al. 2022) that captures the temporal interactions between hazard, exposure, and vulnerability in a coupled human-flood system of a given floodplain (Figure 4.1(a)). The model is forced by a time series of annual flood maxima and a socioeconomic growth indicator such as gross domestic product or population growth rate. The socioeconomic forcing accounts for the fact that the floodplain development is also shaped by factors other than flooding, such as social, economic, or political interests. A flood event occurs once the protection level of the public flood protection infrastructure (e.g., dykes) is exceeded, which causes monetary damage to the assets in the floodplain and triggers a cascade of reactions by the resident society. Experiencing losses increases the society's flood awareness which, in turn, enhances its preparedness through private precaution (i.e., reducing vulnerability) or withdraws from the floodplain and settles in safer locations (i.e., reducing settlement density and, hence, exposure). These choices affect the exposure and vulnerability in subsequent time steps so that the temporal dependency between flood events and the actions of society are incorporated. Eventually, the model continuously traces the evolution of the settlement density (physical exposure), awareness and preparedness (vulnerability), and flood losses (risk) conditional on the hydroclimatological flood signal (hazard) over the long term.

The reaction of the society to flood events is described by socio-hydrological model parameters (Table 4.1) that characterize the flood coping strategy of the resident society. We assume the parameters to be constant or change at much slower rates than the *socio-hydrological dynamics* (i.e., change in the variables floodplain exposure, loss, awareness, preparedness) which they control. The parameters are calibrated on historic observations of settlement density, awareness, preparedness, and flood loss. We compiled these socio-hydrological training data from heterogeneous data sources such as historic land use maps, surveys, published loss reports, and economic and population statistics. The motivation for the structure of the socio-hydrological model and the calibration process are explained elaborately in Barendrecht et al. (2019).

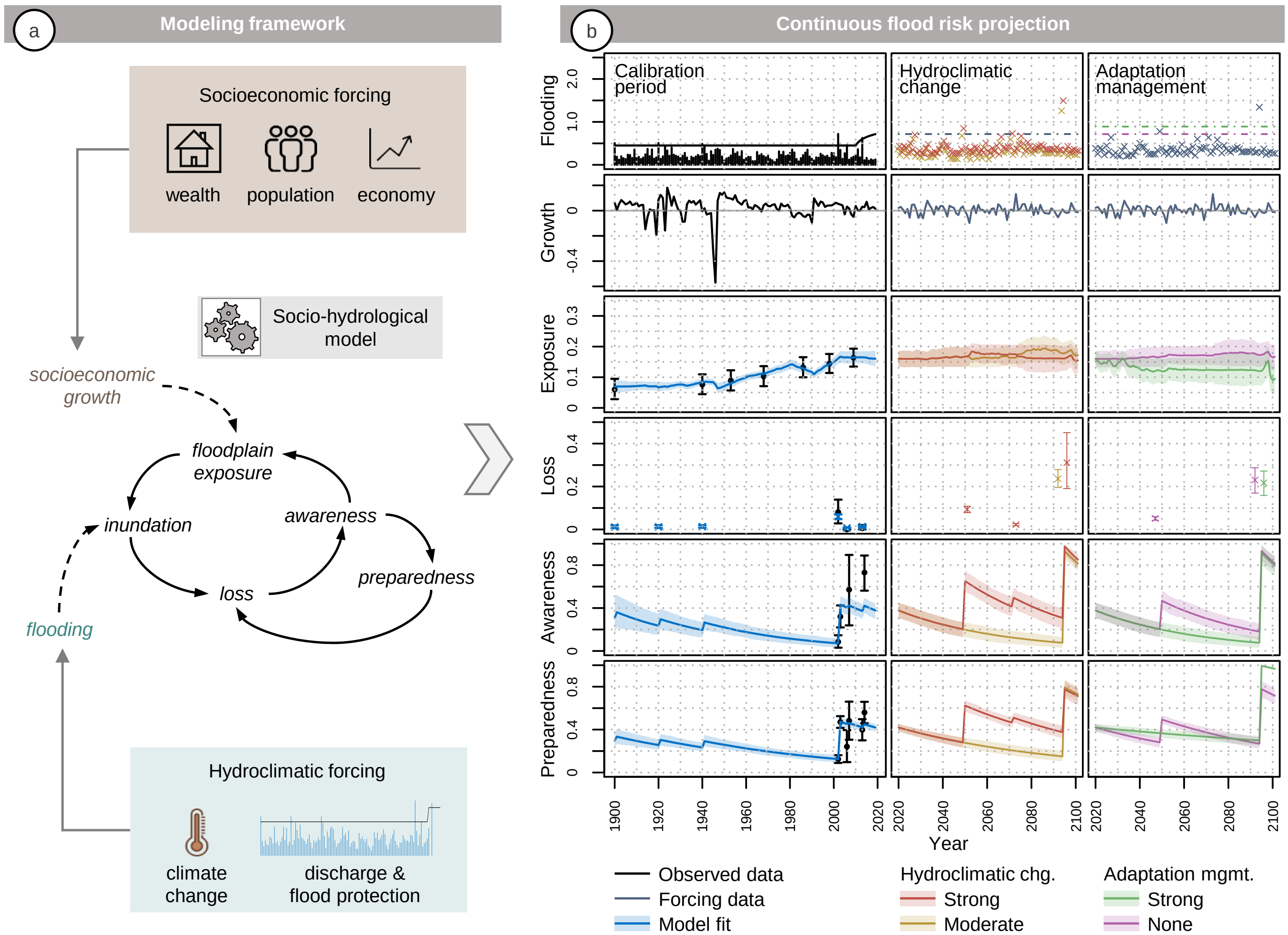

Figure 4.1: Our method for continuous flood risk projection. The socio-hydrological model, which we previously calibrated to observed data, uses socioeconomic growth and flood forcing data (a) to project continuous flood risk trajectories (b). In simulation experiments, we explore the influence of hydroclimatic change and adaptation management scenarios. We consider a set of adaptation measures (see Table 4.1) and combine them to identify effective risk reduction strategies; for example, dyke heightening in combination with increasing the longevity of precaution. Depending on the strength of the hydroclimatic change and the adaptation intervention (e.g., moderate or strong), the coupled flood risk system evolves differently.

Conditional on previous model calibration and validation on the observed time period, the historic hydroclimatological and socioeconomic forcing time series can be substituted by projection data to drive the model for the future period. This allows for the simulation of continuous trajectories for all system variables into the future, while maintaining the temporal interdependency of flood events and human choices. Figure 4.1(b) shows examples of such trajectories. Each trajectory represents one possible future of the flood risk system conditional on the forcing data and the local flood coping characteristic of the society. The model is capable of generating a large number of these trajectories facilitating the quantitative exploration of the possibility space; i.e., the set of future outcomes that could emerge from feedbacks between humans and flooding.

Our workflow uses Bayesian inference to capture uncertainties in the estimation of the socio-hydrological parameters. The modular setup also allows for a propagation of the uncertainty in the forcing data as hydroclimatological and socioeconomic projections can be passed to the model in probabilistic form. Consequently, the method can capture, combine, and communicate the different systemic and statistical sources of uncertainty of flood risk projections.

4.2.2 Application to the commercial sector in Dresden

We apply the socio-hydrological method for continuous flood risk projection to the commercial sector in the city of Dresden, Germany, which is situated at the river Elbe. After a comparably long, flood scarce period in the past century, Dresden faced a series of floods in the past 20 years. A major flood in 2002 caused severe losses and induced the society to adapt, which substantially reduced the losses in subsequent events (2006, 2013). Barendrecht et al. (2019) developed the socio-hydrological model described in section 4.2.1 to study the historical flood risk dynamics in Dresden for the residential sector. Schoppa et al. (2022) transferred the model to the commercial sector and advanced the loss estimation. For the projection study of this work, we use an updated version of the model by Schoppa et al. (2022) with an improved parameterization and adapted inundation estimation (see Text C.2 in supporting information). The model operates on an annual time step and was calibrated on a socio-hydrological dataset that covers the period 1900-2019. Moreover, we validated the model for the flood loss events with available loss reports (i.e., 2002, 2006, and 2013) using leave-one-out cross validation in the previous study.

For this projection exercise, we force the calibrated socio-hydrological model with annual time series of maximum flood return periods (hydroclimatic) from the Elbe and gross domestic product (GDP) growth rate (socioeconomic) in Dresden for the period 2020-2100. The hydroclimatic forcing data are generated on the basis of projected changes in flood frequency at the representative concentration pathways (RCP) 4.5 and 8.5. The data stem from the European Union's Joint Research Centre and were computed by an ensemble of coupled regional climate and hydrological models (Mentaschi et al. 2020). The socioeconomic forcing data is derived from an Eurostat projection of population growth (Eurostat 2021) and a Markov Chain Monte Carlo projection of GDP per capita for Dresden (Steinhausen et al. 2022). Further, we estimate the uncertainty in population growth from the probabilistic country level World Population Prospects 2019 of the UN (UNDESA 2019).

Beside the physical floodplain exposure, which is captured in the case study through the economic density variable (i.e., share of the floodplain area with commercial occupation), exposure dynamics are also influenced by variations in wealth (i.e., financial value). The socio-hydrological model expresses flood loss as a relative loss ratio; that is, the absolute flood loss divided by the replacement value of the commercial building assets. In this way, the loss estimates are independent of the wealth and can be compared directly between different time periods. For an evaluation of the influence of wealth changes on flood losses it is useful to project the replacement value (i.e., fixed assets) into the future. We assessed the future fixed assets in commercial buildings in the Dresden floodplain from the GDP projections, extrapolated wealth-to-income ratios (Paprotny et al. 2018a), and regional accounts data (Federal and State Statistical Offices of Germany 2021). In this way, we assumed the wealth dynamics to be exogenous to the socio-hydrological system but still included them in the analysis of the simulation results. In all calculations that involve monetary units, we used deflated, constant 2015 prices. The supporting information provide further explanation on the raw datasets and the data processing (Text C.3).

4.2.3 Simulation experiments

We projected 1000 trajectories of commercial flood risk in Dresden until 2100. To analyze the influence of the flood risk drivers and adaptation measures over time, we subdivided the projection period into three horizons corresponding to the time periods of the hydroclimatological forcing dataset: 2020-2040, 2041-2070, and 2071-2100. Moreover, we ran the simulations for RCP4.5 and 8.5 global warming levels using the respective 25%, 50%, and 75% percentiles of the ensemble prediction as hydroclimatic forcing. The ensemble percentiles reflect the uncertainty in the climate and hydrological models. For comparison, we also show simulations for a baseline scenario, which assumes constant hydroclimatic conditions as in the reference period (1981-2010) and wealth as in 2020. This projection study is subdivided into two parts:

In the first experiment, we assume that the flood coping characteristics of the companies in Dresden do not change, which can be considered as a 'business as usual' scenario (section 4.3.1). This means that we keep the calibrated socio-hydrological model parameters and, hence, the socio-hydrological dynamics fixed during the projection runs. This allows for an assessment of the influence of

Table 4.1: *Parameters in the socio-hydrological model and their interpretation. In the adaptation experiments, we incrementally increased the calibrated parameter values from Schoppa et al. (2022) to evaluate the potential of different adaptation measures to reduce flood risk.*

Model Variable	Model Parameter	Parameter interpretation	Calibrated value (median)	Adaptation measures	Parameter changes
flood discharge	protection level	return period	90 yr	levee heightening	100, 150, 200, 300, 500 yr
economic density	risk aversion	inclination to develop/abandon floodplain	2.89 [-]	relocation, building bans	+25%, +50%, +100%, +200%, +300%
awareness	anxiousness	increase in awareness per unit of loss	6.91 [-]	information campaigns, flood drills	
	flood memory	half time of awareness	32 yr		
preparedness	activeness	increase in implemented precautionary measures after a flood	1.17 [-]	building codes, subsidization of precaution	
	longevity of precaution	half time of preparedness	50 yr		

hydroclimatic and socioeconomic drivers on flood risk changes for the three future periods.

In the second experiment, we alter the socio-hydrological model parameters to quantify the sensitivity of the flood risk system to changes in the companies' coping characteristics through *adaptation* (section 4.3.2). The alteration of model parameters should be interpreted as adaptation measures by the government (e.g., increasing protection level, information campaigns) or the companies themselves (e.g., implementation of private precautionary measures, resettling) with the objective of flood risk reduction. For instance, measures that aim at sustaining or increasing the flood awareness of the companies in the floodplain can be considered in the model by increasing the respective flood memory parameter, which controls the decay rate of the flood awareness. Since all model variables have been calibrated to observed data, we can treat the adaptation measures quantitatively and measure their effectiveness. In this example, an increase of the flood memory parameter increases the share of companies that are aware that they are situated in a flood risk area. Empirical evidence on the awareness of companies can be obtained using survey campaigns or expert interviews. Based on the results of the sensitivity analysis, we finally compare the effectivity of a structural, integrated, and non-structural adaptation strategy in the context of the expected future flood risk.

Table 4.1 lists the socio-hydrological parameters and provides further details on the adaptation experiment. Figure 4.1(b) illustrates the two projection experiments and how hydroclimatic change and adaptation management can lead to different evolutions of the flood risk system.

Throughout the experiments, we evaluate the flood risk for the individual projection runs on the basis of risk curves, which are a standard method of quantitative risk assessment in science (Merz and Thieken 2009; Metin et al. 2018; Priestley et al. 2018) and the insurance industry (Khare et al. 2015; Prettenthaler et al. 2017). Risk curves summarize all projected annual maximum loss events in the simulation period and assign an occurrence exceedance probability to each event. Additionally, we derive three risk metrics from the risk curves: the expected annual damage (EAD), which indicates the average loss in any given year and, hence, distributes risk evenly over time; the value at risk (VAR) at a 99.5% confidence level corresponding to a loss event with a 200 year return period; and the tail value at risk (TVAR) at the same confidence level. VAR describes the maximum annual loss at the specified return period, while TVAR characterizes the upper tail of the risk curve by integrating losses beyond this return period (Sairam et al. 2021). The confidence level of 99.5% is the current industry standard for (re-)insurers as prescribed by the European Union Solvency II legislation (European Parliament and European Council 2009).

Further, we assess the statistical significance of the investigated effects using the continuous, Bayesian 'percentage in ROPE' index. It quantifies by how much the posterior distribution of a risk metric has shifted away from a Region of Practical Equivalence (ROPE) due to the effect of hydroclimatic change or adaptation. Here, we computed the ROPE on basis of the risk metric posteriors under a baseline or reference scenario. Depending on the percentage of the risk metric posteriors under the investigated scenario in the ROPE, we classify the scenario effect as: negligible/undecided significance ($\geq 2.5\%$ in ROPE), probably significant ($\geq 1\%$ & $< 2.5\%$ in ROPE), or significant ($< 1\%$ in ROPE) (see Text C.4 in supporting information for details).

4.3 Results and discussion

4.3.1 Projection of future flood risk

Our simulations show that global warming influences the co-evolution of the socio-hydrological flood risk system for companies in Dresden. The hydroclimatic model ensemble projects increasing flood hazard until the end of the century, which propagates through the coupled human-flood system (Figure 4.2). After an average decline in flood awareness and preparedness until the middle of the century (i.e., increasing vulnerability), awareness and preparedness rise towards the year 2100 (see dashed, black lines). The non-linear development of vulnerability can be explained by the relatively high awareness and preparedness levels at the start of the projection period (shortly after three loss events) and by the intensification of the flood hazard under global warming, which leads to a gradual accumulation of the companies' flood awareness and preparedness. The physical exposure of companies in the floodplain (i.e., economic density) remains nearly constant with only small variations in the median between the RCP4.5 simulations. The exposure is less sensitive to the hydroclimatic forcing than the vulnerability but is rather dominated by the socioeconomic forcing (i.e., GDP growth), which on average is projected to remain relatively stable throughout the century. In general, the projected trajectories across the different ensemble prediction percentiles (25%, 50%, 75%) reveal that stronger shifts in the flood regime cause more pronounced deviations from the baseline scenario. The development of vulnerability is relatively uncertain and strongly depends on the inherent stochasticity in the flood discharge series, especially the number and temporal succession of loss events. Still, the average tendency towards increased flood adaptation under more severe hydroclimatic forcing is evident. The projected differences in the system evolution between RCP4.5 and 8.5 are hardly distinguishable (see Figure C.4 in supporting information). This is due to the pronounced within-pathway variability of the hydroclimatological forcing data, which masks a possible between-pathway signal in flood change (Mentaschi et al. 2020).

A closer evaluation of the simulated losses via risk curves reveals that flood risk is expected to increase towards the end of the century (Figure 4.3). We computed risk curves accounting for hydroclimatic change under RCP4.5 and wealth growth (red lines). These risk curves clearly exceed the baseline scenario (grey line) that assumes constant climate and wealth conditions with a growing margin towards the far future (i.e., from left to right plot panel). The growing difference is also reflected by the risk metrics, which increase consistently over time and for all ensemble prediction percentiles. Accounting for changes in climate (RCP4.5

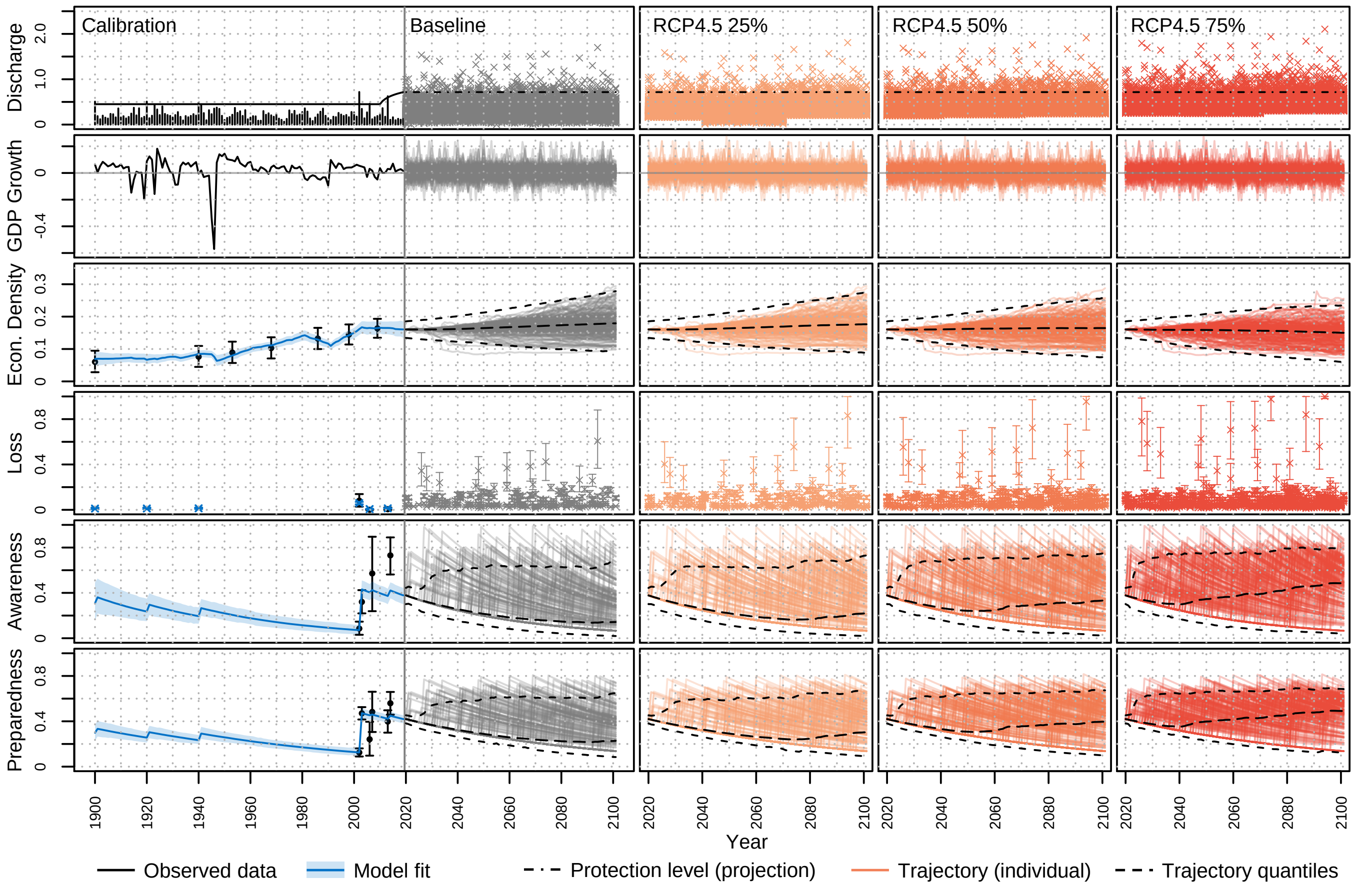

Figure 4.2: *Continuous evolution of the socio-hydrological system for the calibration (1900-2019) and projection (2020-2100) period. The plot visualizes the influence of different hydroclimatic forcing scenarios: baseline with present climate and ensemble percentiles under RCP4.5 climate. For the projections, the colored lines show 200 individual trajectories (median of model uncertainty) and dashed, black lines show the aggregate evolution across all 1000 simulated trajectories (median and 95% highest density interval of projection uncertainty). The supporting information (Figure C.4) contains a similar plot for RCP8.5.*

50%) and wealth, the median expected annual damage (EAD) is projected to double (€2.5M) in the near future (2020-2040) and quadruple (€7M) approximately until the end of the century (2071-2100) relative to the baseline (€1.1M and €1.6M respectively). The relative change in large loss events (VAR, TVAR) is smaller but still increases by approximately a factor of three in the far future (2071-2100). The increase in flood risk in a warmer climate can be explained by more frequent overtopping of the flood protection (risk curves shift towards the left) and higher flood magnitudes causing larger losses (risk curves shift towards the top).

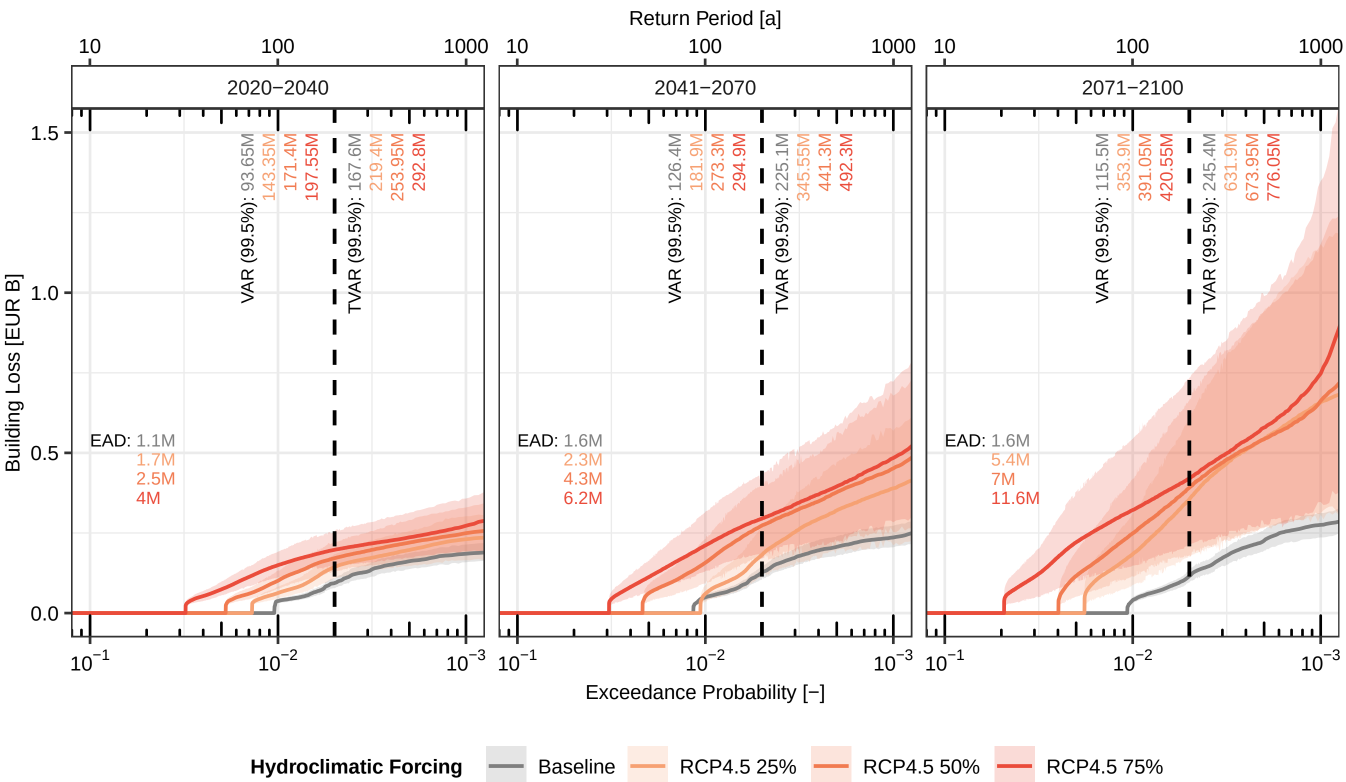

Figure 4.3: Risk curves (median and 90% highest density interval) and metrics for the three projection horizons considering RCP4.5 climate and wealth growth. For this plot, we multiplied the projected relative losses from Figure 4.2 with the fixed asset values (i.e. wealth) to receive absolute losses. The individual risk curves reflect the uncertainty in hydroclimatic forcing, while intervals summarize the uncertainty in the wealth projection, flood risk model, and stochastic flood series. The baseline scenario assumes constant hydroclimatic conditions (as in 1981-2010) and wealth (as in 2020). Risk metrics (median): expected annual damage (EAD), value at risk (VAR), tail value at risk (TVAR). The supporting information (Figure C.5) contains a similar plot for RCP8.5.

The risk curves show future flood risk considering the superimposed effects and uncertainties of hydroclimatic (flood forcing) and socioeconomic (GDP growth forcing and wealth) change, socio-hydrological model (parameter estimation), and inherent stochasticity (randomness in flood series). A decomposition of the different risk drivers for the individual projection horizons (Figure 4.4) reveals that the uncertainty in hydroclimatic forcing (across and within RCPs) dominates over wealth uncertainty in the near future (2020-2040; panels a, d, g). However, towards the end of the century (2071-2100; panels c, f, i) the growth of fixed assets in company building stock becomes increasingly influential. Predominantly, this applies to the metrics VAR and TVAR (panels d-i) which describe the flood risk for large loss events and therefore are more sensitive to differences in wealth. On the contrary, the hydroclimatic forcing remains comparably important for the average annual risk (EAD; panels a-c) until the end of the century since the frequency of dyke overtopping and, thus, loss events strongly depends on alterations of the flood regime. The projected changes in the risk metrics are statistically significant for the most part although large uncertainties in the wealth forcing and the tail of the risk curve (TVAR) mask robust signals until the far future.

Even though we kept the socio-hydrological parameters fixed in this experiment, the simulations capture the influence of changing physical exposure and vulnerability in form of the economic density, awareness, and preparedness trajectories (Figure 4.2). Under the baseline simulation with constant hydroclimatic and socioeconomic boundary conditions, all three risk metrics increase over time (Figure 4.4, grey intervals). This increase of flood risk solely traces back to the internal dynamics of the socio-hydrological system; namely, an average incline of the economic density and decline of awareness and preparedness across the individual trajectories. Yet, compared to the changes in the external hydroclimatic and wealth conditions (red and purple intervals), these socio-hydrological system dynamics only cause small differences in the resulting flood risk.

In summary, the simulations show that the flood risk of the commercial sector in Dresden is likely to increase until the end of the 21st century. This rise is mostly driven by intensifying flood patterns and growing wealth in the floodplain and is in line with projected large scale trends in flood risk (Jongman et al. 2012; Winsemius et al. 2016; Kinoshita et al. 2018). Under the assumption of constant risk coping characteristics of the companies, the influence of socio-hydrological dynamics on the resulting flood risk is almost negligible. This means that the adaptive behaviour of companies as in the past century will not suffice to counteract the expected increase in flood risk due to exacerbating hydroclimatic and socioeconomic pressure.

4.3.2 Effectiveness of flood adaptation

The projected positive trends in flood risk from section 4.3.1 underline the necessity of effective and optimized adaptation strategies that alter the risk coping behaviour (i.e., socio-hydrological parameters) of the commercial sector in Dresden.

Our sensitivity analysis shows that the risk mitigation potential of the different adaptation measures varies across the projection periods and risk metrics. Changes in the parameters protection level, risk aversion, activeness, and longevity of precaution have the largest reduction effect on commercial flood risk in Dresden

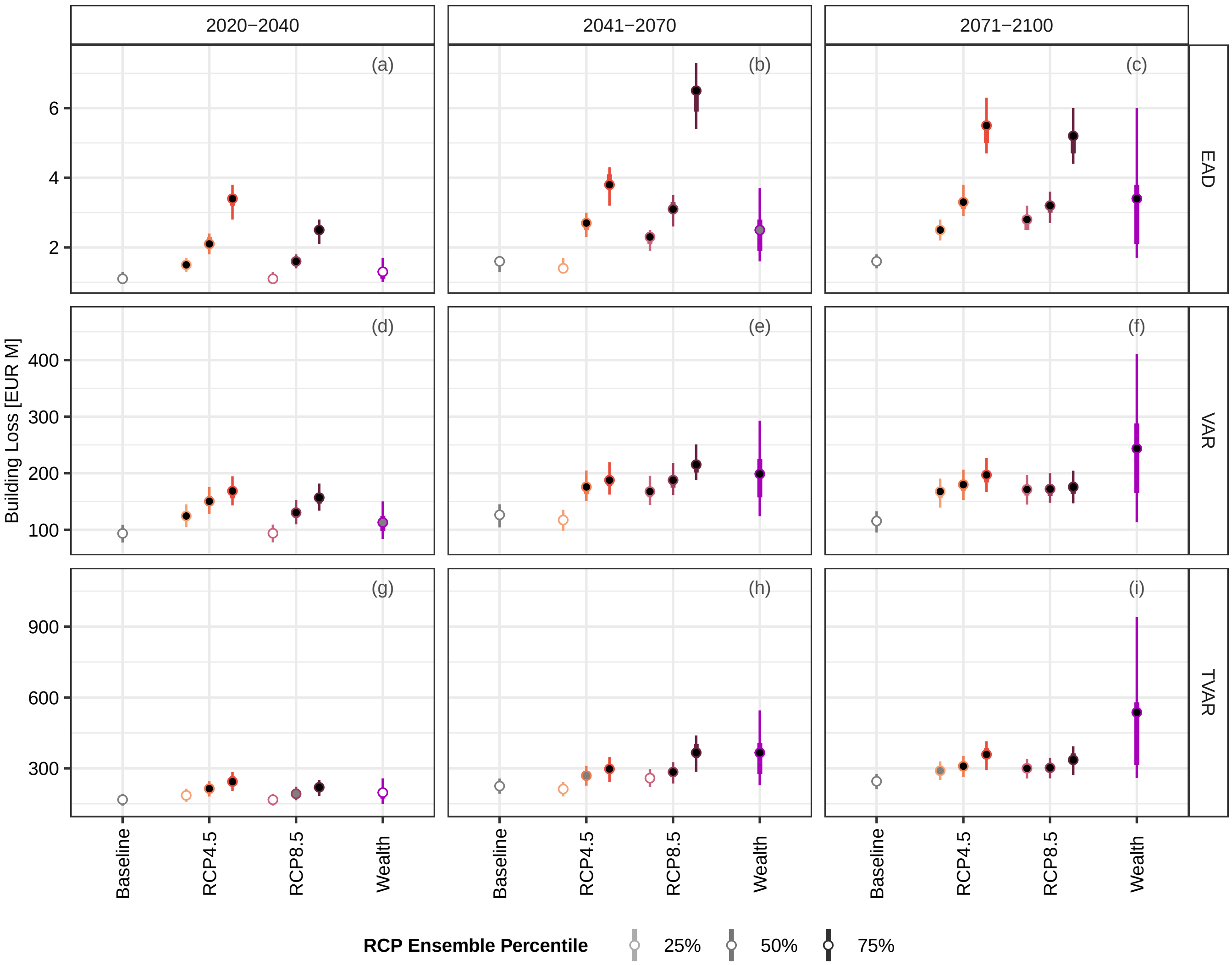

Figure 4.4: *Isolated effect of hydroclimatic (RCPs) and socioeconomic (Wealth) changes on the flood risk metrics relative to the baseline scenario. Each interval plot shows the median, and 50% and 90% highest density intervals across 1000 simulated trajectories. Statistical effect significance is indicated by the points' fill colors: negligible/undecided (white), probably significant (grey), significant (black). Risk metrics: expected annual damage (EAD), value at risk (VAR), tail value at risk (TVAR).*

(Figure 4.5). Increasing the protection level reduces the EAD most quickly and effectively over the entire projection period (panels a-c). Such structural flood protection can prevent single, severe loss events entirely until a clear return period threshold that depends on the protection level (see VAR; panels d-f), but the risk reduction effect shrinks when considering the entire tail of the risk curve (see discrepancy between VAR and TVAR; panels d-i). Non-structural measures that reduce the economic density (risk aversion) and maintain high levels of preparedness (activeness, longevity of precaution) take until the middle (2041-2070) or end (2071-2100) of the century to unfold their influence and steadily become more significant over time. On the long run, increasing the risk aversion, activeness, or longevity of precaution diminishes flood risk in the tail of the risk curve more effectively than an increase of the protection level (TVAR; panels h and i). The influence of the parameters that affect the companies' flood awareness (anxiousness, flood memory) on the risk metrics is not clearly significant for any projection horizon or metric, which potentially traces back to the indirect link between awareness and loss in the socio-hydrological model (Figure 4.1).

The plot also reveals that adaptation measures might even lead to unintended feedbacks and increases in risk. For instance, higher flood protection levels reduce the annual loss expectancy (EAD; panel c) but have the opposite effect on large loss events (VAR and TVAR; panels f and i). This 'levee effect' occurs when higher protection standards lead to reduced flood frequency and, in turn, to declining vulnerability and increasing exposure (Montz and Tobin 2008; Di Baldassarre et al. 2015). Haer et al. (2020) provided quantitative evidence for this effect, and our simulations indicate that this phenomenon also emerges in Dresden in the far future (2071-2100), though the results are only statistically significant in case of the VAR.

The sensitivity analysis shows that the individual adaptation measures have different advantages and drawbacks and, in some cases, only lead to significant risk reduction after strong intervention (i.e., parameter change). Therefore, combinations of adaptation measures could combine the strengths to optimize the risk reduction. We compared the potential of three adaptation strategies to reduce the projected increase in flood risk due to hydroclimatic and wealth changes (Figure 4.6). A structural strategy (i) that only focuses on a protection level increase, an integrated strategy (ii) that combines an increase in flood protection and the longevity of precautionary measures, and a non structural strategy (iii) that only relies on the reduction of physical floodplain exposure and increased preparedness through private precaution.

Across the three adaptation strategies, our simulations show a reduction potential in median EAD of up to 16-60% in the near, 44-63% in the middle, and 50-60% in the long term (panels a-c). The median reduction potential for the VAR ranges from 100% for the structural and integrated strategy to 63% for the non structural strategy (panel d-f). While the projected risk reduction for the EAD and VAR is statistically significant for strong interventions, the effects for TVAR are insignificant due to large uncertainty in both the wealth projections and the tail risk.

The results show that an integrated adaptation strategy is an alternative to a

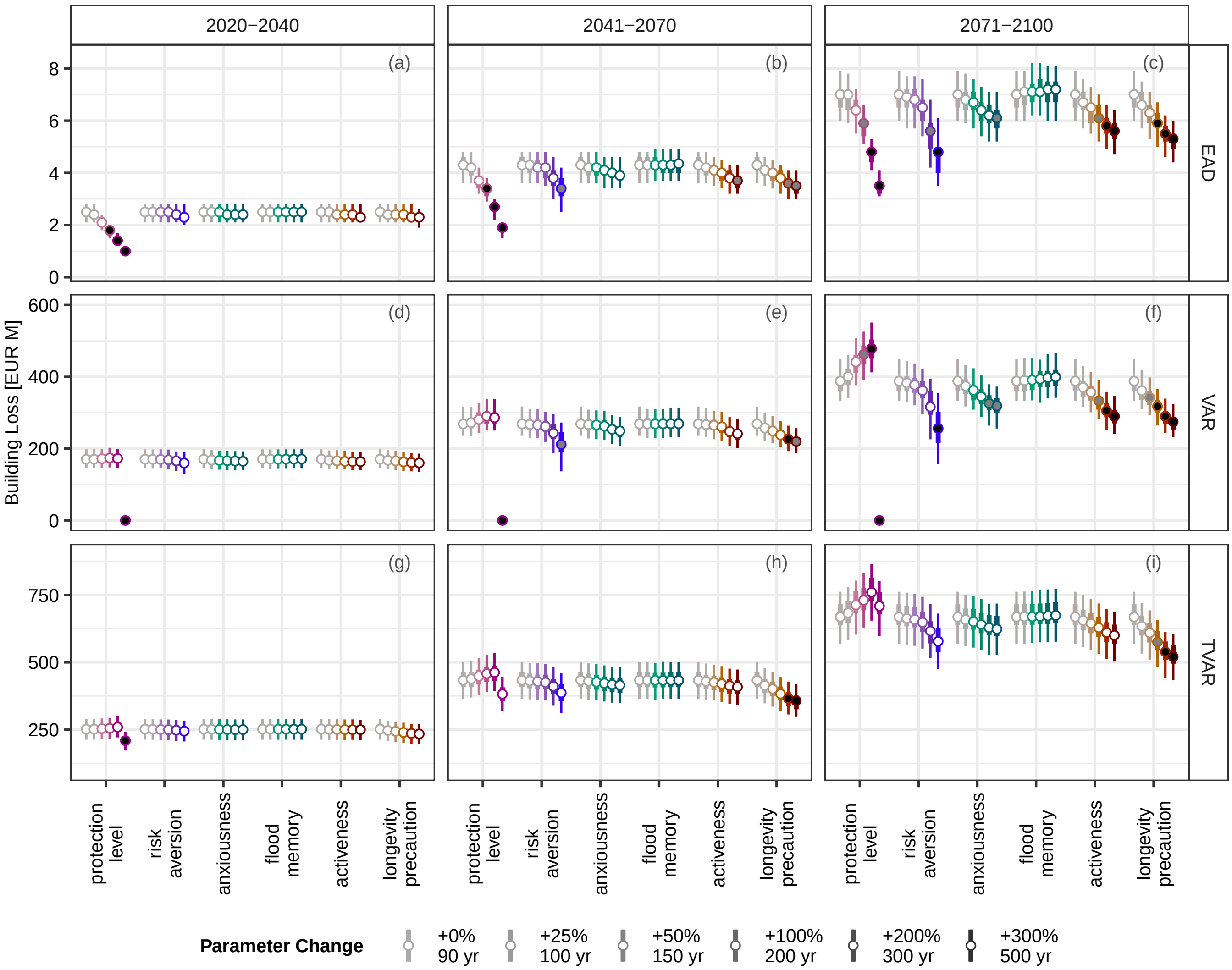

Figure 4.5: *Sensitivity of flood risk metrics towards adaptation measures (i.e., parameter changes). For each interval, we changed the respective parameter while keeping the other parameters fixed at their calibrated value. The interval colors allocate the socio-hydrological parameters to the model variables that they control (pink: protection level, blue: economic density, green: awareness, orange: preparedness). The simulations are based on RCP4.5 (50% ensemble percentile) and median wealth projections (i.e., deflated climate and wealth uncertainty). The strength of the intervention (i.e., parameter change) through a adaptation measure is visualized by the boxplot chroma. Statistical effect significance is indicated by the points' fill colors: negligible/undecided (white), probably significant (grey), significant (black). Risk metrics: expected annual damage (EAD), value at risk (VAR), tail value at risk (TVAR). The supporting information (Figure C.6) contains a similar plot for RCP8.5.*

purely structural strategy.

While the non structural strategy is less effective than the structural and integrated strategy in the near future, it reduces risk more efficiently by the end of the century. Over time, the structural strategy requires increasing intervention strength (i.e., parameter change) to counteract the adverse consequences of the levee effect. On the contrary, this unintended risk increase is attenuated or entirely avoided under the integrated and structural strategy, especially when considering the most severe loss events (VAR and TVAR). Consequently, adaptation strategies that (partially) aim at changing the behaviour of society promise more sustainable flood risk reduction and are less prone to unintended feedbacks and adverse consequences.

Eventually, the exact risk figures of this projection are of secondary importance compared to the overall response of the flood risk system over time, for example, the direction and relative magnitude of risk change. Our results support the finding of a study at European scale by Haer et al. (2019) that optimized adaptation on the governmental and private level carries the potential to outweigh the flood risk increase due to climate and exposure change (RCP4.5 50% + Wealth). However, while Haer et al. (2019) report that the risk reduction potential of structural measures grows over time relative to private adaptation (e.g., precautionary measures), we observe opposite trends. While this discrepancy might originate from differences in study scale, model configuration, or considered risk metric, it highlights that further research on the interplay between governmental and private adaptation is necessary.

4.3.3 Potential of socio-hydrological flood risk projection

The proposed socio-hydrological method addresses current challenges in flood risk assessment such as narrowly defined system boundaries (Merz et al. 2015) or the lack of holistic modeling solutions for small-scales (Jongman et al. 2015). As shown for the study site Dresden, the method can translate transient, large-scale hydroclimatic and socioeconomic boundary conditions into a response of a local, coupled flood risk system (e.g., continuous trajectories, changes in risk metrics). The approach also expands the system boundary by incorporating physical exposure (economic density) and vulnerability (awareness, preparedness) as intrinsic system components capturing potentially adverse non-linearity and feedbacks such as the levee effect. Additionally, it enhances the temporal scope (i.e., centuries) and considers temporal dependencies (continuous simulation) in the analysis, revealing the different time scales at which adaptation measures act. Ultimately, the parsimonious design of the underlying system dynamics model makes the approach an efficient solution for the exploration of the possibility space.

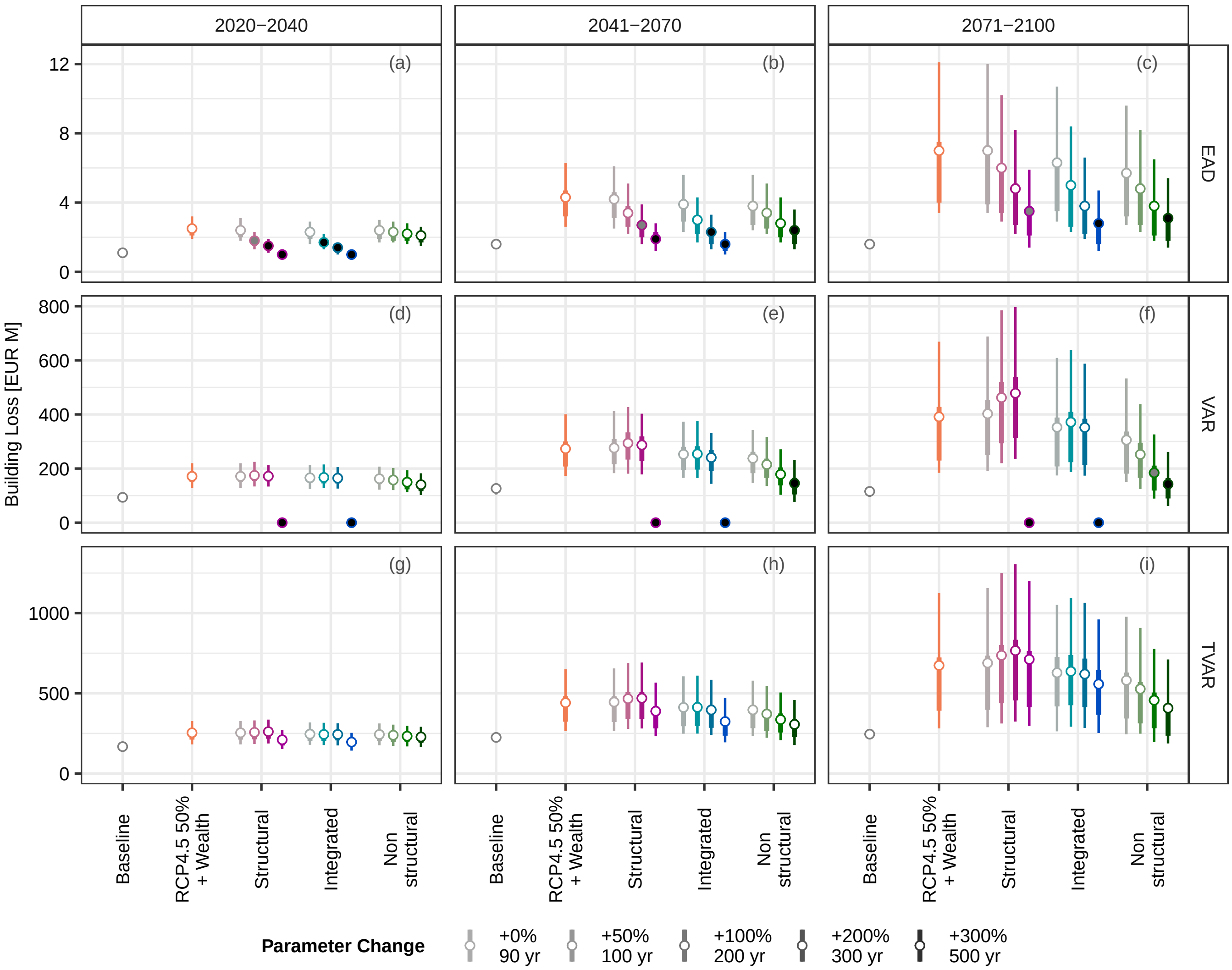

Figure 4.6: *Potential of competing adaptation strategies (structural, integrated, non-structural) to mitigate the expected increase in flood risk. The baseline scenario assumes fixed climate and wealth while the 'RCP4.5 50% + Wealth' scenario assumes hydroclimatic and wealth projections with uncertainty. The interval colors correspond to the color coding from Figure 4.4 and 4.5 for same simulation runs (i.e., 'RCP4.5 50% + Wealth' and 'Structural'). The strength of the intervention (i.e., parameter change) through a adaptation measure is visualized by the boxplot chroma. Statistical effect significance is indicated by the points' fill colors: negligible/undecided (white), probably significant (grey), significant (black). Risk metrics: expected annual damage (EAD), value at risk (VAR), tail value at risk (TVAR). The supporting information (Figure C.7) contains a similar plot for RCP8.5.*

A crucial aspect for the added value of the method is its usability in practice, for instance for adaptation planning. While the applicability of socio-hydrological methods is often limited due to high demands towards data or a lack of variable interpretability (Sivapalan et al. 2012; Troy et al. 2015a), the proposed method relies on a fully quantitative modeling framework, where variables and parameters are informed by empirical evidence. This enables the monitoring of changes in the flood risk system with benchmark data that is collected through remote sensing (e.g., physical floodplain exposure) or surveys (e.g., awareness). Such monitoring could reveal if flood risk actually develops as projected after the implementation of adaptation measures or if amendments to the risk reduction strategy are required. Further, in the Bayesian framework of the proposed method, newly available observations can be included in the inference facilitating ongoing updates of the calibrated model parameters and uncertainty estimates (Schoppa et al. 2022).

Nevertheless, the practical value of the adaptation experiment is limited as it does not consider the tangible and intangible costs and feasibility of the competing adaptation strategies. Differences in the implementation cost of protective or preventive measures (e.g., levee heightening, relocation) might make an adaptation strategy more favorable although its risk reduction potential is inferior, or measures might fail due to the resistance of the resident society. Thus, the combination of the presented method with economic instruments such as cost benefit analysis or expected utility theory (Haer et al. 2019; Dottori et al. 2020), could enhance the informative value of the simulations. Moreover, given the considerable uncertainties in the projections, it might be necessary to switch the adaptation strategy at certain points in the future. Risk-based decision making such as dynamic adaptive policy pathways (Kwakkel et al. 2015; Haasnoot et al. 2013) could be combined with socio-hydrological flood risk projection to deal with this deep uncertainty (Merz et al. 2021) and leverage the potential of adaptation even further. From a modeling perspective, the expansion of the systemic and temporal model scope comes at the cost of coarse spatial detail. While this simplification allows for the efficient exploration of the possibility space (Aerts et al. 2018), localized hydraulic effects or exposure hotspots can affect flood risk substantially and might be missed by the presented lumped method. An appropriate definition of the system boundary and detail is not trivial and usually a trade-off between model accuracy and operability (Sivapalan and Blöschl 2015; Troy et al. 2015b). Here, iterative model development and comparisons with different holistic modeling approaches such as agent based models could prove beneficial.

4.4 Conclusions

Converging global dynamics of flood hazard and exposure are intensifying flood risk at the local scale and necessitate holistic decision-support tools for flood adaptation planning. The presented method for socio-hydrological flood risk assessment represents a systemic and yet efficient solution for the long-term projection of flood risk dynamics. Our case study for the commercial sector in Dresden confirmed that greater efforts on flood adaptation are required to offset the expected four-fold increase in flood risk (i.e., expected annual damage) due to hydroclimatic change and accumulating wealth. We demonstrated that our continuous simulation method can identify effective adaptation strategies that are robust to unintended feedbacks such as the levee effect. According to our simulations, an integrated adaptation strategy, which combines levee heightening with an enhancement of private precaution, could reduce the average annual flood risk by up to 60% at the end of the century. By expanding the system boundary of conventional risk assessment by vulnerability dynamics, this approach can explore a wide range of probable outcomes instead of only the most plausible futures. This raises the chance of detecting risky system states before catastrophes occur.

Nevertheless, the enhanced perspective and computational efficiency comes along with process simplifications and spatial aggregation. Therefore, socio-hydrological flood risk assessment is particularly useful in combination with established risk assessment practices. We see clear advantages of a flood risk assessment workflow that combines coarse and holistic with detailed and focused modeling solutions. For instance, a socio-hydrological flood risk model could first identify effective and robust adaptation strategies from a large set of possible adaptation measures, considering potentially adverse consequences or the factor of surprise. Afterwards, spatially explicit risk assessment, including hydraulic or object specific loss modeling, could be used for further optimization or a selection process among the subset of efficient adaptation scenarios.

5 Discussion, Outlook, and Synthesis

5.1 Summary of findings

Key findings

- Multivariable flood loss models improve loss estimates at the object and systems-level and reduce predictive uncertainty by accounting for the complexity of damage processes (Chapters 2 and 3).
- Systemic flood risk models capture vulnerability and risk dynamics continuously and facilitate long-term projections (e.g., for adaptation management). The holistic system definition captures human-flood feedbacks and, hence, reduces the potential for surprise or maladaptation (e.g., levee effect) (Chapter 4).
- Probabilistic modeling is essential for flood loss and risk assessment since it quantifies the considerable uncertainty, propagates it through the risk system, and provides a basis for informed decision making and risk communication (Chapters 2-4).
- Applying Bayesian inference in socio-hydrological modeling allows for an integration of heterogeneous information (i.e., data, prior knowledge), reducing overall uncertainty in risk assessment (Chapters 3 and 4).
- Damage processes and vulnerability dynamics of companies differ across economic sectors and compared to private households, which calls for tailored modeling solutions and data collection (Chapters 2 and 3).

The representation of vulnerability in conventional flood risk assessment is often oversimplified, stationary, and deterministic. Therefore, this thesis explored possibilities to improve the process detail and treatment of uncertainty in vulnerability modeling. After the development and comparative evaluation of three multivariable, probabilistic flood loss models, I coupled the Bayesian regression loss model to a socio-hydrological flood risk model to investigate observed vulnerability dynamics. Ultimately, I applied the socio-hydrological model in a projection exercise to demonstrate the potential of systemic risk modeling for long-term adaptation management. The main contribution of this research is a systemic modeling solution for continuous, long-term flood risk assessment whit a rigorous loss estimation and often neglected simulation of vulnerability dynamics. By applying the models to the commercial sector, the thesis additionally provides novel insights on the influencing factors and temporal changes of company flood vulnerability as well as on the challenges of model development for the commercial sector.

In the following, I use the key findings to answer the three research questions of this thesis.

How can flood loss and risk models better account for the complexity of vulnerability processes?

The predictor importance of the model fits in Chapter 2 show that various predictors (e.g., precaution, sector, spatial situation) influence the damage grade and not only the inundation depth. Expanding the univariable loss model (i.e., inundation depth) by additional predictor variables improved the predictive skill of the object-level loss models and systemic socio-hydrological models. Similarly, accommodating the superposition of different data generating processes in the model proved to enhance the predictive capacity in both studies (i.e., inflated Beta regression).

Chapters 3 and 4 investigated vulnerability dynamics and revealed that factors that determine vulnerability such as awareness or preparedness fluctuate considerably at a decadal time scale. The augmentation of the socio-hydrological model with more process detail enhanced its capability to reproduce the variability in flood losses. Further, Chapters 3 and 4 demonstrated that holistic models are capable of capturing non-linear, long-term vulnerability and risk dynamics such as the adaptation or levee-effect effect.

My findings show that approaches such as **multivariable modeling, mixture models (e.g., inflated Beta distribution), and system dynamics** account for prevalent complexity factors in flood risk assessment such as heterogeneity, non-linearity, feedbacks, or non-stationarity. Consequently, these modeling solutions outperform simpler methods that only focus on dominant explanatory variables (e.g., inundation depth), falsely assume stationarity (e.g., in vulnerability), or parameterize crucial processes (e.g., inundation in socio-hydrological flood risk models).

What is required to enhance the scope of uncertainty analysis in flood risk assessment and to optimize the use of available information?

This thesis consistently employed probabilistic modeling for uncertainty quantification. In Chapter 3, I coupled the inflated Beta loss model and the socio-hydrological system dynamics model in a fully integrated Bayesian model , which propagates uncertainties (e.g., from the loss estimation) seamlessly and without loss of information through the entire risk system - including often neglected uncertainty in vulnerability and its change. The projection exercise in Chapter 4 demonstrated that this quantitative risk assessment method is interoperable with scenario-based approaches that are necessary when probabilities are difficult to assign - for instance, when considering the hardly quantifiable uncertainty in future adaptation decisions or hydroclimatic forcing.

Chapters 3 and 4 showed how the Bayesian framework efficiently integrates information that was collected from diverse sources, at different points in time and spatial scales, and with varying accuracy. The observations of the socio-hydrological variables jointly inform the probability distribution across all model parameters, enabling the flow of information - for example, historic land use maps reduce uncertainty in the unobserved level of preparedness in 1900, or object-level damage data inform loss estimates at the floodplain level. Additionally, the integration of results from a previous study for the residential sector into the model for the commercial sector via prior distributions diminished uncertainty further. Ultimately, Bayesian methods can include uncertain observations or forcing variables directly in the inference; for example, by treating socio-hydrological training data as random variables (Chapter 3) or driving a socio-hydrological projection with probabilistic input data (Chapter 4).

Altogether, this thesis highlights that a coherent probabilistic design of all model components fosters **integrated uncertainty estimation and propagation**. Where quantitative uncertainty analysis reaches its limits, **scenario-based experiments** with a socio-hydrological model enable the exhaustive exploration of possible system evolutions. Finally, Bayesian inference enables the **assimilation of diverse data sources** (via joint parameter estimation) and **accumulation of information across studies** (via the prior), which helps easing the problem of data scarcity in loss and socio-hydrological modeling.

Which new insights into company vulnerability does this model development provide?

Chapter 2 underlines that company vulnerability is highly heterogeneous. The predictor importance in the calibrated loss models points towards large variation in damage processes across economic sectors and considered assets. In addition, company flood losses are dispersed widely (e.g., bimodal distributions), with large loss cases being particularly difficult to predict. The study shows that the heterogeneity and uncertainty in damage processes is substantial and, hence, reinforces the necessity for multivariable, probabilistic loss estimation.

On the systems-scale, Chapter 3 revealed that companies at the study site Dresden cope differently with flood risk than private households. For instance, the estimated posterior parameters showed that companies forget less quickly about flood risk in comparison. On the contrary, differences in the flood coping and system evolution between companies from the manufacturing and the service sector were very small. Overall, companies preferred to take precautionary measures rather than retreat to safer locations outside the floodplain to reduce their risk. Furthermore, the results did not point towards an added value of the sector differentiation for the accuracy of the loss estimation.

The modeling experiments in this thesis showed that **vulnerability processes and dynamics differ between the commercial and the residential sector**. As a relevant entity within society, the commercial sector actively shapes the overall flood risk evolution. However, model development for companies is commonly more challenging as for the residential sector due to poorer availability of damage and vulnerability data and large intra-sectorial heterogeneity. These characteristics demand **tailored modeling solutions and data collection for flood loss and risk estimation for companies**.

5.2 Discussion and outlook

This chapter reflects on the methods and findings of this work and summarizes the lesson learned during conducting the studies. The following sections point out challenges, give recommendations on how to address them, and suggests potential directions for upcoming research.

5.2.1 Embracing complexity in vulnerability modeling

The model development and comparisons in Chapters 2 and 3 confirm that complexity affects both loss estimation and systems-level risk assessment and that comprehensive modeling solutions improve predictions and simulations.

Loss estimation

Chapter 2 demonstrated the advantages of multivariable probabilistic loss models that consider heterogeneity over univariable deterministic approaches (see e.g., Figure 2.3). The superior predictive performance of complex loss models has also been reported by other authors (Schröter et al. 2014; Rözer et al. 2019; Kreibich et al. 2017b; Wagenaar et al. 2017). However, the analysis of model errors also revealed the current limitations of the presented loss models. The variability in predictive errors is substantial as a result of enhancing the number of model variables and parameters (i.e., bias-variance trade-off), particularly for cases with large relative loss and the asset types equipment and goods and stock (Figure 2.5). Moreover, the different multivariable candidate models achieved relatively similar predictive performance scores (Figure 2.3) and returned wide, sometimes bimodal, predictive distributions (Figure 2.4). These results support previous findings (Sieg 2018; Spekkers et al. 2014; Merz et al. 2010b) that decisive predictors or processes are missing in the available damage datasets and models. The loss models developed in this work improved predictive capacity by accounting for heterogeneity in the flood vulnerability across economic sectors and asset types (Figure 2.2), but damage processes just as well vary across distinct events, flood types, or study regions (Mohor et al. 2020; Mohor et al. 2021; Vogel et al. 2018). Conclusively, additional efforts with respect to the representation of heterogeneity in models and the gathering of explanatory loss data are necessary.

However, a mere increase in sample size does not improve model skill as effectively as a deliberate setup of loss models that explicitly address differences in damage processes (e.g., multivariable and sector-specific) (Sieg et al. 2017). Instead, a simultaneous and iterative enhancement of the number of data and model complexity promises better predictive accuracy and could generate synergies between model development and data collection; for example, the predictor importance in complex loss models could be used to tailor survey campaigns to the distinct governing loss drivers of the sector or region under study (Vogel et al. 2018; Mohor et al. 2020). Yet in many practical problems flood loss data will remain sparse, imbalanced (i.e., few severe loss cases), or overly generalized (i.e., aggregated). An active research strand tries to make as much use of the available data records as possible and yielded effective modeling solutions that are mostly rooted in Bayesian inference (also see section 5.2.2). For example, graphical models such as Bayesian networks also take into account usually omitted incomplete data records in the process of model development and prediction (Vogel et al. 2018; Lüdtke et al. 2019). Multilevel models that account for the fact that data is usually structured hierarchically in groups (e.g., thematically in sectors or spatially in regions) allow for the flow of information across sub-samples and offer additional possibilities for incorporating heterogeneity (Mohor et al. 2021; Sairam et al. 2019a). Finally, Sairam et al. (2020) proposed to combine the usually competing synthetic and empirical loss modeling paradigms for problems where empirical data only becomes available gradually over time.

Systems-level dynamics

Chapters 3 and 4 moved the focus to vulnerability dynamics, which are closely tied to two additional factors of complexity: non-stationarity and feedbacks. The proposed socio-hydrological method for continuous flood risk assessment respects temporal dependencies and legacy effects in its system dynamics framework and allows for transient external forcing (Figure 4.1).

Nevertheless, in its current form the model makes assumptions that offer room for improvement. The lumped modeling approach omits spatial processes within the floodplain which might obscure relevant information. For instance, the location of company premises and inundation extents within the floodplain are only considered implicitly in the model functions. Further, the definition of the system boundary could be reconsidered. In contrast to other studies (Haer et al. 2019; Ridolfi et al. 2020; Di Baldassarre et al. 2013), the model design excludes the adaptive behavior of the government (i.e., structural flood protection such as levees) since the protection level is assumed to be an exogenous model variable. While these choices are justified against the background of unavailable empirical data for the calibration of more sophisticated sub-models, they contradict the objective of allowing for more heterogeneity in the simulation. Moreover, the simulated awareness trajectories (Figure 3.2) deviate noticeably from the observed values in 2006 and 2013. These results indicate that certain time-lags are at work in the real-world system which could be represented better in the system dynamics model; for instance, by introducing explicit delay functions to the coupled differential equations (Bala et al. 2017).

Flood risk assessment methods can be sorted along a continuous spectrum of complexity. Physically and spatially complex model chains (Falter et al. 2015; Sairam et al. 2021; Ward et al. 2013; Alfieri et al. 2017; Aznar-Siguan and Bresch 2019) link stand-alone models for hazard, exposure, and vulnerability assessment, but their linear design falls short of capturing risk and vulnerability feedbacks (Barendrecht et al. 2019). Agent-based models account for such feedbacks and are spatially flexible since computations are carried out with high granularity on the level of decision-making entities (e.g., households or companies). However, empirical data for the calibration of the underlying behavioral models are often not available so that modellers have to resort to assumptions about model parameters (e.g., on the basis of interviews) (Haer et al. 2017; Haer et al. 2019; Coates et al. 2019). The system dynamics approach proposed in this thesis exhibits the most parsimonious model structure, but the spatial aggregation facilitates extensive computations and the collection of empirical data for calibration. In general, there exists no universally valid level of abstraction in the analysis of flood risk systems. Instead, the appropriate complexity very much depends on the questions asked and the available resources (process-understanding, data availability, computational capacity).

Independent of the approach, systemic flood risk modeling faces the problem that the assumptions about the behavior of society or the temporal evolution of the human-flood system cannot be validated as rigorously as it is the case for other elements of the risk analysis (e.g., hydrological or hydraulic models). This is problematic as vulnerability changes and interactions between risk drivers are understood most poorly in the flood risk system. While recent studies (Barendrecht et al. 2019) and the work in Chapter 3 undertook first steps towards the calibration and validation of human-flood dynamics on the basis of empirical data, further research in this direction is necessary. In the absence of suitable validation time series, the mutual benchmarking of systemic modeling paradigms could enhance the confidence in the simulated system evolution or reveal biases when results disagree; for example, in a direct comparison of established agent-based and system dynamics models. Similarly, the application of socio-hydrological models at multiple study sites would contribute new insights about the transferability of holistic risk assessment tools. A promising development in the field is the shift towards longitudinal data collection (i.e., panel data) (Bubeck et al. 2020; Hudson et al. 2020). Repeated survey campaigns with flood affected households or companies could promote the tracing of temporal changes of flood vulnerability and improve the prospects of proper calibration and validation in systemic risk models.

Overall, given the large gaps in our understanding of coupled human-flood systems, data collection and model development also serve the purpose of gaining new insights into vulnerability and flood risk dynamics (Di Baldassarre et al. 2016). With respect to model development, the iterative fitting and evaluation of candidate models with varying complexity - as it is common practice in the Bayesian workflow (Gelman et al. 2020) and was carried out in Chapter 3 - is useful for identifying the required level of complexity in a system dynamics model (Chapter 3).

5.2.2 Uncertainty quantification - potential and limitations

Although the enhancement of complexity improved the skill of the flood loss and risk models in Chapters 2 and 3, the model outputs remain substantially uncertain; for example, the probabilistic loss estimates (Figure 2.4) or the projected evolution of the flood risk system (Figure 4.2). These results support the growing consensus that consequent uncertainty quantification and communication is essential for well-informed flood loss and risk assessment (e.g., Schröter et al. 2014; Rözer et al. 2019; Apel et al. 2008; Merz and Thieken 2009).

This thesis supplies many arguments in favor of probabilistic modeling and particularly the use of Bayesian methods. Chapter 2 shows that Bayesian predictive models are competitive to machine learning methods such as random forest in loss modeling. Further, the application of the socio-hydrological flood risk projection method (Chapter 4) showcases how Bayesian inference fosters the consistent combination, propagation, and communication of uncertainty in a fully probabilistic modeling framework. Among the already discussed strengths of the Bayesian framework (i.e., joint parameter estimation, probabilistic predictions, uncertainty propagation, incorporation of many levels of randomness) the capability to accumulate diverse information - such as prior knowledge from literature or comparable experiments, expert judgement, or data - presumably represents the largest benefit for flood risk assessment. The use of informative priors is sometimes criticized to be subjective. Yet when motivated, verified (e.g. via prior predictive checks), and communicated transparently, the inclusion of prior information represents a pragmatic solution for modellers that face the problem of data scarcity (Schoot et al. 2021; Gelman and Hennig 2017; Smid et al. 2020). Passing on information from previous to subsequent studies can also have positive effects with respect to uncertainty reduction as information accumulates over the course of repeated experiments. This 'Bayesian updating', as applied in Chapter 3 to make inferences about the vulnerability dynamics of companies, is not employed much in flood risk research at the moment (for exceptions see Sairam et al. 2020; Lüdtke et al. 2019) although it could leverage the little available evidence more rigorously. The flexibility of Bayesian inference might tempt developers to construct overly complex models. Nonetheless, the required definition and verification of generative prior models (i.e., model testing before data is taken into account) encourages developers to deliberate their modelling choices deeply (e.g., distributions, parameterization) and inappropriate specifications quickly lead to computational problems or erroneous outputs. Ultimately, the results of a Bayesian analysis communicate uncertainty intuitively in the form of probability statements (in contrast to frequentist statistics), which reduces the risk of misinterpretation and biased management decisions (McElreath 2018; Gelman et al. 2013).

In order to support risk management and decision making with reliable information, further efforts to reduce the associated uncertainty bounds are necessary. While the uncertainty in estimates of flood risk can be attributed to all the three determinants of risk (hazard, exposure, vulnerability), the necessity and potential to reduce it varies (de Moel et al. 2015; Sieg 2018; Steinhausen 2022) The physical understanding of flood hazard (i.e., atmospheric, hydrological, hydraulic drivers) is already far advanced and, hence, represented relatively accurately in process-based models (e.g., climate or river routing models). Recently proposed methods for the modeling of dyke failure or spatial dependencies of flood peaks promise further improvements (Vorogushyn et al. 2012; Metin et al. 2020). In the assessment of flood exposure, rapid developments in geoinformatics and remote sensing enhance the accessibility of comprehensive and open data (e.g., satellite imagery or crowd-sourced data such as OpenStreetMap), which facilitates the identification and specification of affected assets with high accuracy (e.g., monetary value, building type) (Pittore et al. 2017; Cerri et al. 2021) as well as the upscaling of loss estimates (Sieg et al. 2019b; Lüdtke et al. 2019). In contrast, vulnerability assessment represents a large challenge from the object to the (supra-)national level, especially when temporal dynamics are also taken into account (de Moel and Aerts 2011; Winter et al. 2018; Merz et al. 2014a). Loss estimates will underlie large aleatory uncertainty for the foreseeable future (Sieg 2018), and the simulation of vulnerability change depends heavily on insufficiently verified hypotheses about system dynamics (e.g., differential equations in Chapter 3) or behavioral models that are rarely calibrated empirically (e.g., in agent-based models) (Haer et al. 2017).

Nevertheless, systemic risk models such as the continuous, socio-hydrological flood risk assessment method (Chapter 4) advance our understanding of the human-flood system which, in turn, reduces the epistemic uncertainty about vulnerability dynamics (Di Baldassarre et al. 2016). In addition to quantifiable uncertainty in form of probability distributions (known probabilities, e.g. occurrence probability of a loss event), the presented approach also takes into account non-quantifiable uncertainties in form of scenarios (unknown probabilities, e.g. future adaptation decisions). Moreover, the expansion of the system boundary to vulnerability dynamics improves the predictability of "known unknowns"; that is, well-known phenomena that are difficult to capture in risk analyses such as the adaptation (Figure 3.2) or levee effect (Figure 4.5).

However, even the most accurate systemic risk model is incapable of capturing so-called "unknown unknowns" or "black swans": unexpected, high-impact events that are unpredictable because they have not been observed before or originate from surprising and fundamental shifts in the functioning of the system. Such deep uncertainty is characteristic for flood risk assessment that usually considers long planning horizons (e.g., flood adaptation under climate change) and impedes effective risk management. Under such circumstances, approaches that go beyond probabilistic or scenario analysis are necessary (Merz et al. 2015; Di Baldassarre et al. 2016; Hall and Solomatine 2008; Cox 2012). Methods such as info-gap analysis (Ben-Haim 2019; Hall and Solomatine 2008), dynamic adaptive policy pathways (Haasnoot et al. 2013; Kwakkel et al. 2015), or resilience-focused strategies (Disse et al. 2020) aim at identifying management strategies that maintain favorable conditions (e.g., low flood losses) under a wide range of possible futures. For instance, Haasnoot et al. (2012) used a simplistic meta-model in a hypothetical floodplain to simulate storylines of the future as a basis for identifying effective dynamic adaptive policy pathways. It could be worthwhile to explore if the socio-hydrological risk assessment method from Chapter 4 would also be applicable in such an overarching framework for robust decision making. A potential study design could look as follows: the system dynamics model could generate reliable flood risk trajectories (similar to the simplistic meta model) that are grounded in empirical evidence and process-oriented loss estimation and provide information on the risk reduction potential of adaptation measures (as in Chapter 4). Afterwards, the resulting trajectories could be evaluated using dynamic adaptive policy pathways to map out viable adaptation strategies. In this way, the advancements of this thesis in loss and dynamic vulnerability modeling could feed into practice and support operational flood risk management.

5.3 Synthesis

This thesis advances the representation of vulnerability in flood risk assessment, especially with respect to temporal dynamics. Using the example of companies, this work presents fully probabilistic modeling approaches for loss estimation and socio-hydrological risk assessment, which embrace the complexity in damage processes and the human-flood system.

Flood loss estimation is often carried out with univariable and deterministic stage-damage functions that fall short of describing the multicausality and variability of damage processes. Multivariable models consider various explanatory variables, which improves the predictive performance as heterogeneity can be reflected more accurately. However, even in state-of-the-art loss models, influential controls of flood damage remain unresolved, which causes considerable uncertainty. Probabilistic loss models quantify this predictive uncertainty and offer great structural flexibility to accommodate superimposing data generating processes (e.g., using mixture models). Probabilistic loss estimation not only communicates the reliability of loss estimates more transparently but also facilities the integration of loss models into overarching flood risk assessment frameworks.

Flood risk assessment has adhered to hazard-centric and static approaches in the past. It is increasingly recognized that a shift towards systems-thinking is necessary to take into account socioeconomic risk drivers (exposure, vulnerability) and interactions between flooding and humans more strongly. Yet data constraints and limited understanding of human-flood systems prevent that systemic approaches such as socio-hydrology are applied more widely. The presented augmentation procedure incorporates process-detail and previously unused data sources in socio-hydrological flood risk models to enhance the accuracy and reliability of simulations. In contrast to inconsistent, linear, and computationally expensive model chains, systemic risk models yield continuous trajectories of all risk drivers including often neglected vulnerability dynamics. The model implementation in the Bayesian framework enables the joint evaluation of all involved uncertainties and the formalized integration of prior knowledge.

Ultimately, this thesis demonstrated possible applications of socio-hydrological flood risk assessment. Simulation-based experiments with the augmented socio-hydrological model allow for an efficient and plausible exploration of a wide set of possible futures and the identification of effective adaptation measures. Additionally, the method facilitates more reliable risk assessment as unintended long-term feedbacks that might lead to maladaptation can be detected (e.g., levee effect). The seamless integration of all governing risk drivers in one system dynamics model enables the comprehensive treatment and propagation of uncertainty. Eventually, socio-hydrological flood risk projections could provide an information-rich basis for decision making under deep uncertainty.

As previous research and model development in the field of flood vulnerability predominantly addressed private households, I applied the newly developed models to companies. The model fits provide new insights into company flood vulnerability. Damage processes of companies vary strongly across economic sectors and assets types, and companies cope differently with flood risk than households, which gives rise to distinct vulnerability and risk dynamics. These findings highlight the need for tailored data collection and modeling solutions for companies, which take into account the pronounced heterogeneity.

In conclusion, vulnerability remains the most challenging element of flood risk assessment as it originates from the interplay of natural and human processes, where causal mechanisms are still poorly understood. Therefore, the value of comprehensive risk assessment approaches such as the socio-hydrological method of this thesis is not primarily in prediction, but rather in advancing our understanding of human-flood systems.

A Appendix to Chapter 2

A.1 Introduction

This Supplementary Information (SI) contains further details on the implementation of the flood loss models, that is, random forest (A.2), Bayesian networks (A.3), Bayesian regression (A.4), and the probabilistic stage-damage function (A.5). Tables A.1 and A.2 provide the prior specification of the Bayesian regression and the probabilistic stage-damage function.

A.2 Random forest

Following Sieg et al. (2019b) and Sieg et al. (2017) we incorporated recent advancements in tree-based learning in our random forest models. Meinshausen (2006) response variable but are capable of estimating its entire conditional distribution. This method is referred to as quantile regression forests and tracks the prediction of each individual tree in the random forest instead of only the mean prediction. Hothorn et al. (2006) developed an alternative recursive partitioning algorithm, called conditional inference trees, which avoids a variable selection bias in the conventional classification and regression tree algorithm towards continuous predictors (Breiman et al. 1984). We implemented the random forest models in the programming-language R using the 'partykit'-package (Hothorn and Zeileis 2015). The package supports both quantile regression forests and the conditional inference tree algorithm.

A.3 Bayesian network

We learned the directed acyclic graph structure of the Bayesian networks from the company loss data with the score-based 'Tabu Search' algorithm (Bouckaert 1995). This optimization routine searches the space of candidate Bayesian network structures for a directed acyclic graph that maximizes a predefined goodness-of-fit score. We opted for the Bayesian Dirichlet equivalent score (Heckerman et al. 2013). After obtaining the Bayesian network structure, we learn the conditional probability tables from the survey dataset through Bayesian parameter estimation. In order to make predictions with the fitted Bayesian network, we employ exact Bayesian inference by querying the conditional probability of the target variable, relative loss, conditioned on the observed predictors. For further details on Bayesian network theory we refer to the literature (e.g., Jensen and Nielsen 2007; Koller and Friedman 2009; Nagarajan et al. 2013; Scutari and Denis 2014). We implemented the proposed Bayesian networks in R using the packages 'bnlearn' (Scutari 2010) and 'gRain' (Højsgaard 2012).

A.4 Bayesian regression

In Bayesian parameter estimation, we make probability statements about a parameter θ conditional on observations $\mathbf{y}$ by applying Bayes theorem (Gelman et al. 2013):

$$p\left(\theta|\mathbf{y}\right) = \frac{p(\theta)p\left(\mathbf{y}|\theta\right)}{p\left(\mathbf{y}\right)} \propto p(\theta)p\left(\mathbf{y}|\theta\right), \tag{A.1}$$

where $p\left(\theta|\mathbf{y}\right)$ is the posterior distribution, $p(\theta)$ is the prior distribution, $p\left(\mathbf{y}|\theta\right)$ is the likelihood and $p\left(\mathbf{y}\right) = \int p(\theta)p\left(\mathbf{y}|\theta\right) d\theta$ is the marginal likelihood. The marginal likelihood is a normalizing constant which guarantees that the posterior distribution integrates to one (McElreath 2018).

Priors describe the initial plausibility of each possible value for a parameter, where each parameter in a Bayesian model requires its own prior. Commonly, priors are classified as non-informative, weakly informative, or informative depending on how strongly the prior restricts the domain of plausible parameter values. In general, priors should place the majority of their mass on parameter values which lead to reasonable model response (Gelman and Hennig 2017; Simpson et al. 2017). Further, priors are useful for ruling out unrealistic parameter estimates such as negative values for a standard deviation (McElreath 2018).

Following recommendations from the literature (Gelman and Hennig 2017), we assign weakly informative priors for the regression parameters based on our understanding of damage processes. For instance, for the mean of the beta distribution, μ, we assigned the prior $Normal(1, 1.5)$ to the regression coefficient of water depth, since it places most probability mass on positive values. This reflects our expectation to observe a positive correlation between water depth and flood loss without ruling out the possibility of obtaining a negative parameter value. Following Gelman and Hennig (2017), we list and motivate the prior choice

Table A.1: *Prior distributions for the regression coefficients of the zero-and-one-inflated beta model classified according to distribution parameters and predictors. Each prior contains a note motivating the prior choice.*

Distribution Parameters	Predictors	Prior	Notes
μ – beta mean γ – conditional one-inflation	water depth, return period, inundation duration	$Normal(1, 1.5)$	More severe floods cause larger flood loss.
	precaution, flood experience	$Normal(-1, 1.5)$	Precaution and flood experience decrease vulnerability and, in turn, attenuate flood loss.
	spatial situation, sector, size	$Normal(0, 1.5)$	The expected effect of these predictors is less clear. Hence, we center the prior at zero.
λ – zero-and-one-inflation	water depth	$Normal(-1, 1.5)$	Increasing water depth reduces the chances to have no loss.
	precaution	$Normal(1, 1.5)$	Increasing precaution increases the chances to have no loss.
	spatial situation, sector	$Normal(0, 1.5)$	The expected effect of these predictors is less clear. Hence, we center the prior at zero.
ϕ – beta precision	water depth, precaution	$Normal(0, 1.5)$	The expected effect of predictors on the precision of the beta distribution is hard to assess. Hence, we center the prior at zero.

for the regression coefficients in Table A.1. Apart from the presented priors, we adopted the weakly informative default priors for regression intercepts from the brms-package (see Bürkner 2017). We also tested other, more informative and less informative, prior settings and found the data to dominate over the priors, which means that the results of the parameter estimation were affected little by the prior

Table A.2: *Prior distributions for probabilistic square root stage-damage function classified according to model parameters and predictors. Each prior contains a note motivating the prior choice.*

Model Parameters	Predictors	Prior	Notes
α – Intercept	-	$Student - t(3, 0, 10)$	Weakly informative standard prior for intercepts in brms-package.
β – regression parameter	$\sqrt{waterdepth}$	$Normal(1, 1.5)$	Higher water depth causes larger flood loss.
λ – zero-and-one-inflation	-	$Beta(1, 1)$	The parameter represents a probability and, hence, is constrained to the interval [0,1].
γ – conditional one-inflation	-	$Beta(1, 1)$	The parameter represents a probability and, hence, is constrained to the interval [0,1].
ϕ – beta precision	-	$Gamma(0.01, 0.01)$	The precision of the beta distribution has to be positive.

choice.

In the computation of the posterior, Bayesian inference usually relies on numerical techniques, since analytical solutions to Equation A.1 only rarely exist in real applications. We employ Markov chain Monte Carlo (MCMC) simulation, which is the most common way of approximating a posterior by means of sampling. For an introduction to MCMC in the context of Bayesian modeling we refer to Gelman et al. (2013) and McElreath (2018). We implemented the proposed zero-and-one-inflated beta regression model in the statistical programming language Stan (Carpenter et al. 2017) using the R-package 'brms' (Bürkner 2017; Bürkner 2018), which allows for Bayesian modeling from R. For model fitting and prediction, we used the 'No-U-Turn' sampler (Hoffman and Gelman 2014), which is included in Stan, to generate 4000 samples from the posterior distribution using two chains. In each chain, the first 1500 samples were omitted (burn-in phase of the sampler) resulting in 1000 samples from the posterior predictive distribution of relative loss, which we used for model evaluation and comparison. The MCMC diagnostics and experiments with more samples showed that 1000 samples estimate the posterior distribution of the model reliably.

A.5 Comparison to stage-damage functions

The theory and computational implementation of the probabilistic stage-damage function followed the methodology of the Bayesian regression model as described in the previous paragraph. The MCMC specification remained unchanged, yielding 1000 posterior samples of relative loss for the subsequent analyses. Table A.2 contains the priors for the parameters of the probabilistic stage-damage function.

B Appendix to Chapter 3

B.1 Introduction

This Supplementary Information (SI) contains further details on the implementation of the socio-hydrological flood risk models. Text B.2 introduces the concept of Bayesian inference. Texts B.3 to B.5 provide details on all model data and the treatment of observational uncertainty. Text B.6 discusses the reparameterization of the sector differentiating models, while Text B.7 gives insight on the computational implementation of the candidate models in the probabilistic programming language Stan. Text B.8 introduces the continuous ranked probability score. Tables B.1 to B.3 provide an overview of the prior specifications for the socio-hydrological parameters and the parameters of the process-oriented loss estimation (i.e., inundation and loss regression). Finally, Figures B.3 visualizes the prior and posterior densities of the inundation and loss regression parameters for the candidate models that contain the process-oriented loss estimation.

B.2 Bayesian parameter estimation

We estimated the parameter values of the four socio-hydrological model versions by means of Bayesian inference (Gelman et al. 2013; McElreath 2018; Schoot et al. 2021). The objective of Bayesian inference is to learn the posterior distribution $p\left(\theta|Y\right)$, which is the distribution of the model parameters θ conditional on the data Y. The posterior is computed by Bayes' theorem

$$p\left(\theta|Y\right) = \frac{p\left(Y|\theta\right)p(\theta)}{p(Y)}, \tag{B.1}$$

where $p\left(Y|\theta\right)$ is the likelihood, $p(\theta)$ is the joint prior distribution over the parameters, and $p(Y)$ is the marginal likelihood. The likelihood encompasses the probabilistic model and describes our understanding of the data generating process, while the prior encodes the initial plausibility of the parameter values. The marginal likelihood serves as a normalizing constant that guarantees that the posterior is a proper probability distribution. Bayesian parameter estimation produces an entire posterior parameter distribution and, hence, fully discloses the uncertainty in the socio-hydrological parameters and the flood loss predictions of the candidate models.

As described in the main text, we adopted the posteriors from the residential model of Barendrecht et al. (2019) as priors for the new company model. We used prior predictive checks (Gelman et al. 2020) to verify that the priors, which we transferred from the residential model, are also reasonable for the company model, which resulted in some adjustments to the acquired posteriors. Namely, the posteriors from the residential model for the anxiousness (α_A) and the effectiveness of the preparedness (α_R) proved to have a different magnitude than the respective company parameter so that they contradicted with the information that the data provided. See Table B.1 for a full list of the socio-hydrological priors.

Furthermore, we also fitted the candidate models with the weakly informative priors that Barendrecht et al. (2019) used to ensure that the prior choice does not bias the Bayesian inference. Figures B.1 and B.2 compare the prior and posterior distributions of the four candidate models for the two prior configurations (weakly informative as in Barendrecht et al. (2019) and informative as in this study). The median posterior parameter estimates are very similar for the two prior configurations, which indicates that the choice of the prior has small influence on the inference. The initial awareness and preparedness, A_0 and P_0, are influenced most by the prior, i.e., their dispersion is similar to the spread of the prior. Yet, as there is no mismatch between the median parameter estimates for A_0 and P_0 between the prior configurations, we conclude that the priors do not contradict with the information in the company data. The overall observed reduction in uncertainty for the models with the informative priors is the product of combining the a-posteriori information from the residential model (through informative priors) with the new company data (through Bayes theorem).

Apart from that, the informative priors are advantageous from a computational perspective. Weakly informative priors like the ones in the study from Barendrecht et al. (2019) work well for comparably simple model structures like in the parsimonious model. However, as models become more complex (e.g., through the presented augmentations), narrower/informative priors are required to obtain robust inferences and sufficient MCMC sampling efficiency (Gelman et al. 2020). We also observed this phenomenon and were able to improve the model fitting in Stan significantly by means of the informative priors (faster sampling and fewer problems with divergences).

In summary, we intentionally used informative priors to obtain more robust models and inferences. The 'sanity checks' described above, show that the prior choice does not bias our findings but rather makes efficient use of the Bayesians workflow and, hence, facilitates more precise statements about the vulnerability dynamics of companies.

B.3 Historical socio-hydrological data

In accordance to Barendrecht et al. (2019), we defined the spatial bound of the model as the area that a flood with a return period of 500 years would inundate. The corresponding discharge of such a flood at the Elbe river in Dresden is 6255 m^3/s, which is, in turn, the value for the model parameter $W_{\max}$. We derived the series of annual discharge maxima from a daily discharge record at the gauge Dresden covering the period 1853-2017 of the Global Runoff Data Centre (BfG 2021). Based on this annual maxima series, we computed the return periods V for each annual maximum discharge through the L-moments method (Hosking 1990). For the flood protection level H of the Elbe at Dresden, we employed the reconstruction of Barendrecht et al. (2019), which was derived from historical reports of flooding (Weikinn 2000; Weikinn 2002; Pohl 2004), previous studies (Kreibich and Thieken 2009), and authority reports (Federal Dam Operation Authority of Saxony 2013). The discharge and the protection level both have the unit $(m^3/s)/(m^3/s)$ and were scaled from zero to one by dividing through the maximum discharge.

The flood awareness and preparedness of the companies in Dresden are fuzzy variables and their state cannot be measured directly. Therefore, we used survey data (GFZ 2021) from computer-aided telephone interviews to quantify the state of the two variables in time. The surveys have been conducted with companies that were affected by major flood events in Germany in the period 2002-2013 and contain 1346 completed interviews (Kreibich et al. 2007; Thieken et al. 2016). Out of these interviews, 111 were conducted with companies that were located in the 500-year floodplain in Dresden during the Elbe floods in 2002 (n=82), 2006 (n=5), and 2013 (n=24). In the interviews, the companies (i.e., the managing director or a responsible employee) were asked (i) whether they were affected by flooding before, (ii) whether they knew that the company premises were situated in the flood risk area, and (iii) how likely they deemed that the company will incur loss due to flooding again. The share of companies that answered ‘yes’ to the first and second question is the awareness A before the observed flood events (i.e., in the years of their occurrence: 2002, 2006, 2013). The third question collected data on a categorical scale ranging from one (‘not very likely’) to six (‘very likely’). We used this question to estimate the awareness after the events (i.e., in 2003, 2007, 2014) and assumed that all companies that answered with values from 4 to 6 were aware of the flood risk. Barendrecht et al. (2019) observed that the first and second question overestimated the flood awareness before the 2006 and 2013 flood events, and, therefore, only used these questions for the approximation of the awareness in 2002. We found the same overestimation pattern in the company survey data and, hence, only used the first and second question for the estimation of the 2002 flood awareness as well.

The preparedness P of the companies is approximated similarly based on a list of eight private precautionary measures that a company could implement such as the installation of water barriers or the relocation of vulnerable assets to higher stories. The interviewers asked the company representative, whether the company had already implemented the respective measure before the flood, during/after the flood, or whether it intends to install the measure in the upcoming six months. We then computed the preparedness before the floods (i.e., in 2002, 2006, 2013) from the number of previously implemented measures, divided by the number of possible or relevant measures. In accordance, the preparedness after the floods (i.e., in 2003, 2007, 2014) was computed from the sum of the measures that were implemented before, during, and after the flood, divided the total number possible or relevant measures.

The flood loss L is the relative loss to company building structures in the floodplain. The relative loss is calculated from the flood loss to building values divided by the replacement value of the company buildings and, hence, is expressed in €/€. For the flood loss in 2002 and 2013, we used loss data from the Saxonian Relief Bank, which is responsible for the financial compensation of companies that incurred damages due to flooding. This dataset contains the reported company flood losses on the municipality level and, in case of the 2002 flood, indicates how these losses are distributed across economic sectors. The 2006 flood caused minor losses and no corresponding company flood loss dataset is available. Therefore, we estimated the company flood loss from the reported compensation of the Saxonian Relief Bank for residential buildings (Saxonian Relief Bank 2007), assuming that the compensation requirements and the proportion of residential to company flood loss was the same as in 2002. We derived the replacement value of the company building structures for the two economic sector in the floodplain from data on tangible fixed assets, which is contained in the HANZE exposure dataset (Paprotny et al. 2018a). The fixed asset data are available as grids and cover the evolution of asset values in five year steps until 2020. We intersected these fixed asset grids with the areas within the 500-year floodplain that were occupied by companies at the respective time and computed the area weighted sum of the corresponding grid cell values. This resulted in aggregated company fixed asset values within the system boundary with a temporal resolution of five years, which we interpolated to yearly values. Since the loss and asset values not only comprise the building values of the companies but also other assets such as equipment or goods and stock, we assessed the relative loss specifically to buildings on basis of the company survey data. The survey data include information on how the flood loss and asset values are distributed across buildings, equipment, and goods and stock. By transferring these ratios from the surveys to the aggregated loss and fixed asset values, we received relative loss values specifically for building values of companies in the Dresden floodplain. During these calculations we adjusted for price changes due to inflation, using the deflator time series from the HANZE dataset so that all monetary values of this study refer to 2011 prices.

The economic density D is the floodplain area that is covered by company premises and has a unit of m^2/m^2. We obtained historical spatial information on city areas that were covered by manufacturing and service companies for seven points in time (1900, 1940, 1953, 1968, 1986, 1998, 2009) from reconstructed land use maps of Dresden (Gruner 2012). The 2009 land use map did not resolve between different economic sectors so that we used the dataset of Rosina et al. (2020) to classify manufacturing and service areas in the floodplain. Then, we computed the sector specific economic densities by dividing the sum of the company areas that belong to the respective sector by the total area of the 500-year floodplain. For the sector aggregated model, the economic density is the sum of the densities of the two sectors. The economic density in the floodplain is driven by the economic growth rate U. Here, we used the historic growth rate of the gross domestic product in Dresden from the HANZE dataset to force the model. For the sector differentiating candidate models, we calculated individual growth rates for each sector. In order to obtain annual economic densities and growth rates, we linearly interpolated between the points in time for which observations were available.

B.4 Inundation and survey loss data

We estimated the regression coefficients of the inundation and loss regressions from intersections of historical land use maps and inundation maps for the city of Dresden. The seven available land use maps refer to a certain year and provide information on the economic density in the floodplain at that time. The inundation maps (Saxonian Environmental Agency 2012) correspond to flood return periods (20, 50, 100, 300, 500 years) and account for public flood protection structures such as levees. In the following, we refer to the intersections of the land use and inundation maps as inundation scenarios. For each inundation scenario, we computed the economic area that is flooded and the mean inundation depth in the flooded economic areas. The inundation model presumes that the flooded area and the mean inundation depend on the flood return period and the current economic density. The flooded area F_{obs} is the share of the company area in the 500-year flood extent that is inundated given the return period and the economic density. The mean inundation depth I_{obs} required more processing because the inundation maps provide the inundation depths in form of binned water depths (e.g., 0-0.5 m) instead of continuous values. This binning of inundation depths censors the underlying inundation data and complicates the calculation of a mean inundation depth. We estimated the mean inundation on the basis of a distributional assumption. Namely, we presumed that the inundation depth is gamma distributed and estimated the mean of the distribution by means of maximum likelihood estimation using the approach by Delignette-Muller and Dutang (2015). This resulted in 28 inundation scenarios for the sector aggregated models (56 for the sector specific models) which we used to estimate the regression parameters in equation 3.2 of the main text.

The flood loss regression was informed by the same survey data, from which we obtained the awareness and preparedness data. We used all interviews that contain information on the relative building loss, the inundation depth at the company premises, the preparedness before the flood, and the company sector to inform the parameter estimation. These complete interviews amount to 597 cases. In accordance to Schoppa et al. (2020), we centered and scaled the predictor variables of the loss regression model before running the model as similar variable scales improve the efficiency of the MCMC sampling.

B.5 Uncertainty in socio-hydrological data

The definition of a Bayesian model involves the selection of distributional forms for all model variables, which constitute the likelihood in equation B.1. We already introduced the gamma and beta likelihoods that we used for the inundation and loss regressions in section 3.2.3 in the main text. For the socio-hydrological variables in equations 3.1-3.5 of the main text, we chose beta distributions since the variables are scaled from zero to one. In the sector differentiating candidate models, we substituted the beta distribution by a Dirichlet distribution (also see Text B.6).

We employed a parameterization of the beta and Dirichlet distribution, where the distributions of the data Y are defined by their mean values μ and a precision parameter ϕ, which quantifies the dispersion (Ferrari and Cribari-Neto 2004; Maier 2014; Sennhenn-Reulen 2018). The precision parameter relates to the variance of the distribution as follows

$$\phi = \frac{\mu(1-\mu)}{var(Y)} - 1. \tag{B.2}$$

This allowed us to express the uncertainty in the socio-hydrological data in form of their variance. We either based these variances on domain knowledge (economic density, loss) or derived it from the data (awareness, preparedness). Experiments with larger and smaller variance revealed that the assumptions with respect to the data uncertainty do not affect the findings of this study.

We assumed that recent land use maps are more reliable than those that correspond to earlier points in time. Starting in the year 1900, we fixed the variance to 0.005 times the observed mean value and decreased the multiplication factor linearly to 0.001 in 2020. As a result, the uncertainty in the economic density data is the largest in the year 1900 and lowest in the year 2009. Reported flood losses commonly vary, depending on the source of information, so that we considered the loss data to be more uncertain in comparison to the land use data. For 2002 and 2013, loss data are available and, therefore, we set the variance to 0.01 times the mean for these events. As the loss value for 2006 is not based on data but on a report and assumptions, we chose a larger variance for this event (0.1 times the mean). The observations of awareness and preparedness come from the survey data and allow for a more informed assessment of uncertainties. Under the assumption that the binary survey answers on the awareness and preparedness of the single companies are Bernoulli trials, the overall awareness and preparedness can be interpreted as the success probability of the trials, which follows a beta distribution. Since we know the number of interviewed companies in each year, we were able to determine the variance from the survey answers analytically. In accordance to the number of surveyed companies in the respective years, the associated uncertainty is largest in 2006, followed by the events in 2013, and 2002. In general, the confidence in the awareness is lower than in the preparedness because the preparedness was computed from several survey questions on different precautionary measures, while the awareness was estimated from one single question.

We did not fix the uncertainty in the inundation and survey loss data as we did for the socio-hydrological data. Instead, as more data were available for the inundation and loss regression, we also estimated the respective precision parameters (ϕ_F, ϕ_I, ϕ_L) in the Bayesian inference.

B.6 Parameterization of sector differentiating models

Two candidate models (int_sd, aug) differentiate between companies that operate in the manufacturing ('man') and the service ('ser') sector. We reparameterized these socio-hydrological models so that the economic density is modelled by a Dirichlet distribution instead of a beta distribution, which we used for the sector aggregated models. The Dirichlet distribution generalizes the beta distribution for more than one random variable. Maier (2014) formalized regression models based on a reparameterized Dirichlet distribution, which Sennhenn-Reulen (2018) transferred into the framework of Bayesian inference. Based on these two papers, we adapted the model and replaced equation 3.5 in the main text by the following expression

$$D_{obs} \sim Dirichlet\left(softmax\left(\boldsymbol{\eta}\right) \cdot \phi_D\right) \tag{B.3}$$

$$\frac{d\eta_{\text{sec}}}{dt} = U_{\text{sec}}\left(1 - \alpha_{D,\text{sec}}A\right). \tag{B.4}$$

The observed economic density $\mathbf{D_{obs}} = \left(D_{man},\ D_{ser}, \left(1 - D_{man} - D_{ser}\right)\right)^T$ is a vector, with one component for each considered economic sector (manufacturing, service) plus one component for all areas that are covered by other land uses. The individual components of the vector have to sum to one, corresponding to the total floodplain area in the non-dimensional model. The economic density is modelled by a Dirichlet distribution, which can be defined by a vector of expected values $\mathbf{J}$ and a global precision parameter ϕ_D (i.e., dispersion). The individual expected values are predicted by the terms η_{sec}. The softmax function guarantees that the resulting expected values sum to one. During the parameter estimation, we fixed the predictor term that corresponds to other land uses, so that the parameters can be identified properly. As the softmax transformation maps the predictor terms to the unit interval [0,1], the saturation term of equation 3.5 in the main text, i.e. $D(1 - D/D_{max})$, becomes redundant. Hence, we dropped it from the equation. Tests with different parameterizations confirmed that this does not affect the behavior of the model.

In the inundation and loss regressions, we added the economic sector as a discrete predictor variable. This predictor is denoted as S and equals one when a company belongs to the service sector and zero when it operates in the manufacturing sector. The adapted model formulation of the inundation model substitutes equations 3.6 and 3.7 in the main text and reads

$$
\begin{aligned}
F_{obs} &\sim Beta\,(F,\ \phi_F) && \text{(B.5)}\\
logit(F) &= \alpha_F + \beta_{F,D}D + \beta_{F,V}V + \beta_{F,S}S, && \text{(B.6)}
\end{aligned}
$$

$$
\begin{aligned}
I_{obs} &\sim Gamma\,(I,\ \phi_I) && \text{(B.7)}\\
\log(I) &= \alpha_I + \beta_{I,D}D + \beta_{I,V}V + \beta_{I,S}S, && \text{(B.8)}
\end{aligned}
$$

For the process-oriented loss estimation, we adopted the building loss model from Schoppa et al. (2020) and only included the sector as a predictor for the mean of the zero-and-one-inflated beta distribution. This leaves us with the following regression model

$$
\begin{aligned}
L_{obs} &\sim BEINF\,(\mu_L, \phi_L, \lambda, \gamma) && \text{(B.9)}\\
logit\,(\mu_L) &= \alpha_{\mu_L} + \beta_{\mu_L,I}I + \beta_{\mu_L,P}P + \beta_{\mu_L,S}S && \text{(B.10)}\\
logit(\lambda) &= \alpha_\lambda + \beta_{\lambda,I}I + \beta_{\lambda,P}P && \text{(B.11)}\\
logit(\gamma) &= \alpha_\gamma + \beta_{\gamma,I}I + \beta_{\gamma,P}P. && \text{(B.12)}
\end{aligned}
$$

The computation of the sector specific loss estimates is carried out according to equation 4 in the main text. Finally, the sector specific loss estimates are weighted, based on the contribution of the sector to the overall economic area, and summed up to determine the total loss of an event

$$
L = \sum_{n_{\text{sec}}} L_{\text{sec}} \left(\frac{D_{\text{sec}}}{\sum_{n_{\text{sec}}} D_{\text{sec}}} \right). \qquad [€/€] \quad \text{(B.13)}
$$

B.7 Computational implementation in Stan

We solved the differential equations of the socio-hydrological model numerically using the forward Euler approach (for details see Barendrecht et al. (2019)). In the candidate models with the standard loss estimation (pars, int_sd), two parameters control the relative loss; that is, the effectiveness of preparedness α_R and the discharge to loss relationship β_R. In order to avoid non-identifiability of the model, we fixed the parameter β_R to the value one.

In practical problems, the posterior distribution in equation B.1 cannot be derived analytically so that the parameters are usually obtained through Markov-Chain Monte Carlo methods (MCMC). These techniques approximate the posterior distribution by simulating large numbers of randomized samples. Here, we used the Bayesian inference software 'Stan' (Carpenter et al. 2017), which is compatible with standard statistical software such as R. For the candidate models that feature the process-oriented loss estimation, we combined the socio-hydrological and the loss model component into one Stan model. In this way, the information from the different data sources is shared, leading to one coupled inference process. Stan uses an efficient 'No-U-Turn' sampler (Hoffman and Gelman 2014) to generate draws from the posterior distribution. We configured the sampler with four MCMC-chains with 2000 iterations and a burn-in phase of 1000 iterations. Thus, each model run yielded 4000 posterior samples. Standard diagnostics measures (i.e., mixture of chains, effective sample size) that are distributed with Stan, confirmed the model convergence and that the number of samples was sufficiently high.

B.8 Continuous ranked probability score

The continuous ranked probability score (CRPS) (Matheson and Winkler 1976; Gneiting and Katzfuss 2014) is a proper scoring rule and judges the sharpness and calibration of a predictive distribution with respect to a point observation. It generalizes the absolute error and can be compared directly to the MAE. That is, the CRPS has the same unit as the observation and a value of zero indicates a perfect fit (i.e., no error). The CRPS for one observation y_i is

$$CRPS_i\left(F_i, y_i\right) = \int_{-\infty}^{\infty} \left(F_i(x) - \mathbf{1}\left\{y_i \leq x\right\}\right)^2 \, dx \tag{B.14}$$

where $F_i(x)$ is the cumulative distribution function of the predictive distribution $f_i(x)$, and $\mathbf{1}\{\cdot\}$ is the indicator function. In this study, y_i are the reported losses of the three observed floods (i.e., 2002, 2006, 2013) and $f_i(x)$ are the corresponding probabilistic loss estimates. We compute the CRPS with an empirical cumulative distribution function estimated from the MCMC-samples of $f_i(x)$. Details on the numerical implementation of the CRPS for simulated forecasts are explained in Jordan et al. (2019) and Krüger et al. (2016).

Table B.1: *Prior distributions of the socio-hydrological model parameters. The majority of the priors are the adopted posteriors from the residential model of Barendrecht et al. (2019). Each prior contains a note with further details on the prior choice.*

Model Parameters	Description	Prior	Notes
α_A	Anxiousness	$Lognormal(2,\ 0.5)$	Cannot be adopted directly from residential models as the parameter value depends on the magnitude of economic density (i.e., lower than for households). Compromise between sector differentiating and aggregating models.
μ_A	Forgetfulness	$Lognormal(-3.393,\ 0.497)$	Posterior residential model.
α_P	Activeness	$Normal(0.871,\ 0.244)$	Posterior residential model.
μ_P	Deterioration rate of precautionary measures	$Normal(0.016,\ 0.005)$	Posterior residential model.
α_D	Risk-taking attitude	$Lognormal(0.9077,\ 0.378)$	Posterior residential model.
α_R	Effectiveness of preparedness	$Gamma(1.5,\ 2.969)$	Residential posterior was not in line with the data and produced too low losses. Hence, we chose a wider prior with the same mode.
A_0	Initial awareness	$Beta(6.01,\ 12.771)$	Posterior residential model in the year 1900.
P_0	Initial preparedness	$Beta(41.7,\ 97.3)$	Posterior residential model in the year 1900.
D_0	Initial economic density	Sector aggregated: $Beta(10.916,\ 125.534)$ Sector differentiated: $Dirichlet(6.823, 4.094, 125.534)$	We assumed that in the year 1900 companies covered less floodplain area than residential buildings. Further we expected the economic density of the manufacturing sector to be larger than for the service sector.

Table B.2: *Prior distributions for the regression coefficients of the inundation model component. Priors are classified according to the variable (flooded area and mean inundation depth), which they predict. Each prior contains a note with further details on the prior choice.*

Distributional parameter	Predictors	Prior	Notes
F – flooded area beta mean	intercept	$Student - t(3,\ 0,\ 2.5)$	Adopted from brms-package (Bürkner 2017).
	economic density	$Normal(1,\ 1.5)$	In a more populated floodplain, companies might have to move to more exposed areas
	flood return period	$Normal(1,\ 1.5)$	More intense floods are associated with larger flood extent
	economic sector	$Normal(0,\ 1.5)$	The expected effect of the sector is less clear. Hence, we center the prior at zero.
ϕ_F – flooded area beta precision	not predicted	$Gamma(0.01,\ 0.01)$	Rules out implausible negative values and covers a wide range of possible parameter magnitudes.
I – inundation depth gamma mean	intercept	$Student - t(3,\ 0,\ 2.5)$	Adopted from brms-package (Bürkner 2017).
	economic density	$Normal(0,\ 1.5)$	For inundation depth, the effect of the density is less clear. Hence, we center the prior at zero.
	flood return period	$Normal(1,\ 1.5)$	More intense floods are associated with larger inundation depth.
	economic sector	$Normal(0,\ 1.5)$	The expected effect of the sector is less clear. Hence, we center the prior at zero.
ϕ_I – inundation depth gamma precision	not predicted	$Gamma(0.01,\ 0.01)$	Rules out implausible negative values and covers a wide range of possible parameter magnitudes.

Table B.3: *Prior distributions for the regression coefficients of the process-oriented loss model component. Priors are classified according to the distributional parameter of the zero-and-one-inflated beta distribution which they predict. Each prior contains a note with further details on the prior choice.*

Distribution Parameters	Predictors	Prior	Notes
μ_L – loss beta mean	intercept	$Student - t(3,\ 0,\ 2.5)$	Adopted from brms-package (Bürkner 2017).
	inundation depth	$Normal(1,\ 1.5)$	Larger inundation depth causes more severe damages.
	preparedness	$Normal(-1,\ 1.5)$	Preparedness decreases vulnerability and attenuates flood loss.
	economic sector	$Normal(0,\ 1.5)$	The expected effect of the sector is less clear. Hence, we center the prior at zero.
γ – conditional one-inflation	intercept	$Logistic(0,\ 1)$	Adopted from brms-package (Bürkner 2017).
	inundation depth	$Normal(1,\ 1.5)$	Larger inundation depth increases probability of total loss.
	preparedness	$Normal(-1,\ 1.5)$	Larger preparedness decreases probability of total loss.
λ – zero-and-one-inflation	intercept	$Logistic(0,\ 1)$	Adopted from brms-package (Bürkner 2017).
	inundation depth	$Normal(0,\ 1.5)$	Larger inundation depth reduces probability of zero loss but increases probability of total loss.
	preparedness	$Normal(0,\ 1.5)$	Larger preparedness increases probability of zero loss but decreases probability of total loss.
ϕ_L – loss beta precision	not predicted	$Gamma(0.01,\ 0.01)$	Rules out implausible negative values and covers a wide range of possible parameter magnitudes.

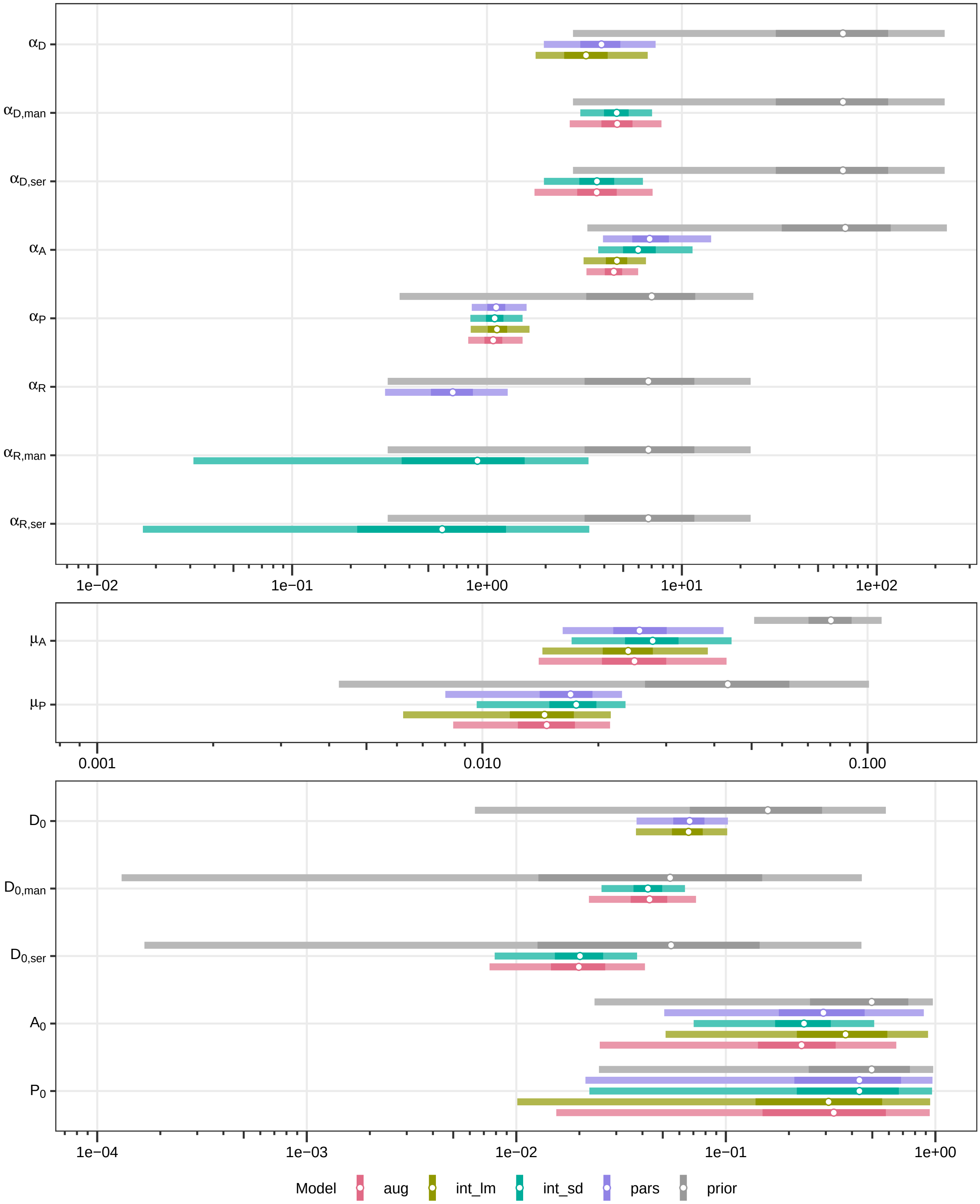

Figure B.1: *Marginal posterior and prior distributions (log-scale) of the socio-hydrological parameters in the four candidate models (color coded). Here, the priors are weakly informative and correspond to the priors that Barendrecht et al. (2019) used in their residential model.*

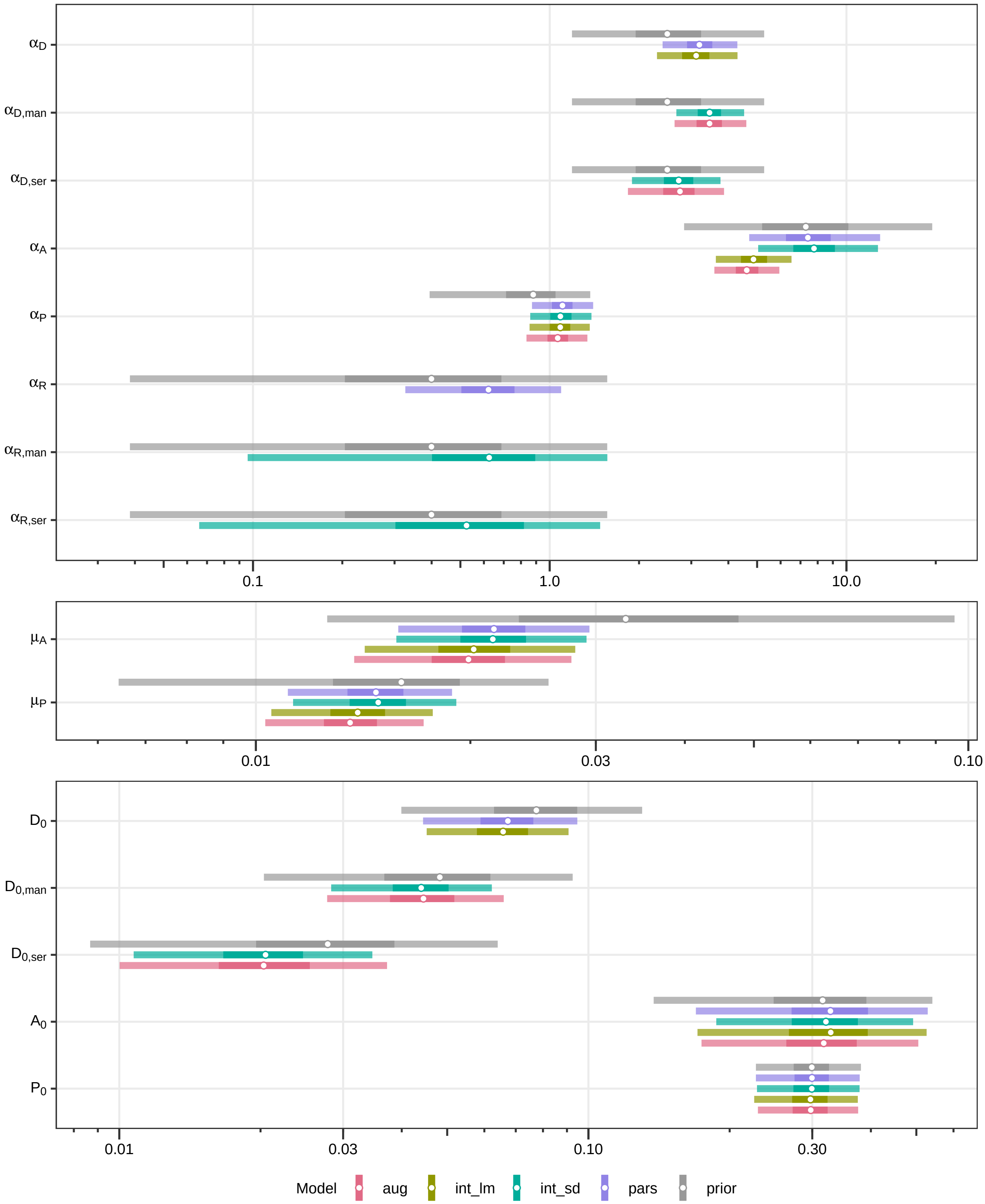

Figure B.2: *Marginal posterior and prior distributions (log-scale) of the socio-hydrological parameters in the four candidate models (color coded). Here, the priors are informative and correspond to the posteriors from Barendrecht et al. (2019). Same as Figure 3.3 in the manuscript.*

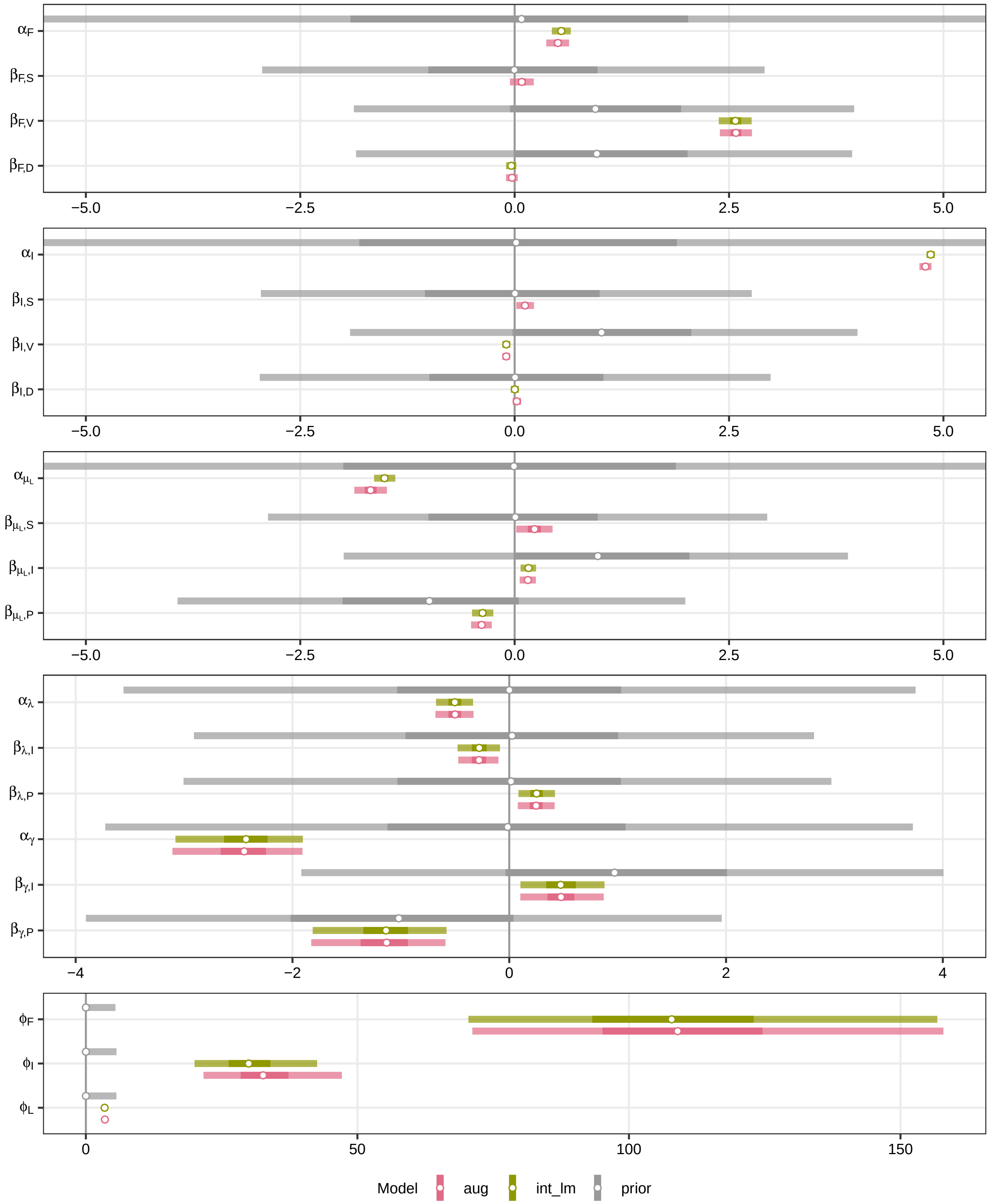

Figure B.3: *Marginal prior (grey) and posterior (color coded) distributions of the inundation and loss regression parameters in the candidate models with the process-oriented loss estimation (int_lm, aug). Dots represent medians, while bars show 50%, and 95% confidence intervals.*

C | Appendix to Chapter 4

C.1 Introduction

This supporting information presents details on the socio-hydrological flood risk model and explains and motivates the alterations, which we applied to the original model by Schoppa et al. 2022 for this projection study. Further, we elaborate on the construction and processing of the forcing data, which we used to drive the socio-hydrological model for the future time period. Finally, we demonstrate how we conducted the Bayesian significance testing and present additional simulation results for the radiative concentration pathway 8.5.

C.2 Socio-hydrological flood risk model

The socio-hydrological flood risk model is based on the works by Di Baldassarre et al. (2013) and Viglione et al. (2014) and was developed by Barendrecht et al. (2019) in the Bayesian probabilistic programming language Stan (Carpenter et al. 2017). In Schoppa et al. (2022) we transferred the model from the residential to the commercial (i.e., manufacturing and service) sector and augmented the loss estimation in the model. In this study, we adopted the model by Schoppa et al. (2022) (i.e., the intermediate candidate model with the process-oriented loss estimation 'int_lm') and introduced additional improvements. For a detailed explanation of the original model, we refer to Schoppa et al. (2022). In the following, we explain and motivate the alterations to the original model, which concern the calibration data and the model equations.

With respect to the calibration data, we first substituted the historic GDP growth rate forcing data for the NUTS3 region of Dresden from the original study by a new, annual time series from the HANZE v2.0 database by Paprotny and Mengel (2022) (available from Paprotny (2022)). As the data before year 2000 were only available in 5- or 10-yearly time steps, we used annual time series of growth rates for all of Germany (1900-2000) from the Maddison Project Database (Inklaar et al. 2018) and modified the growth rates proportionally so that the 5- or 10-yearly growth rates match Dresden's, as defined in HANZE. The resulting annual GDP growth series offers more detail in the simulation and enhances the consistency between the historic and projected forcing data. Secondly, we replaced the flood hazard maps that are used to infer the relationship between flood discharge and average inundation depth at the commercially used areas. That is, we exchange the official flood hazard maps of the Environmental Agency of Saxony by flood hazard maps of the Environmental Agency of the city of Dresden (Nuremberg Institute of Technology 2019). These new inundation maps cover a larger range of flood return periods (i.e., 2-690 years instead of 20-300 years), facilitating a more robust estimation of flood inundation at high flood return periods. Thirdly, we updated the historic time series of the flood protection level, which was composed by Barendrecht et al. (2019). Here, we used geospatial data on the protection level of flood protection structures in Dresden (City of Dresden 2018) to assess the current average protection level in the city (i.e., a 90 year flood). Finally, we reduced the calibration data point for the commercial flood loss of the 2002 event based on flood footprints of the event since a considerable share of the city was inundated by Elbe tributaries rather than the Elbe itself. The remaining raw data for model calibration data comprise time series of discharge and water level (BfG 2021; WSV 2021), land use maps (Gruner 2012), flood loss reports (Saxonian Relief Bank 2007), and company survey data (Kreibich et al. 2007; Thieken et al. 2016; GFZ 2021).

As for the socio-hydrological model, we reparameterized the system dynamics model and introduced a non-linear inundation sub-model to enhance the model robustness and allow for a more realistic loss estimation for large flood events. The socio-hydrological model is dimensionless and describes the continuous interactions between society and flooding through four coupled differential equations:

$$L = F\left(\lambda\gamma + (1-\lambda\mu_L)\right) \qquad [€/€] \quad \text{(C.1)}$$

$$\frac{dA}{dt} = \tanh\left(\alpha_A L\right)\left(1-\frac{A}{A_{max}}\right) - \mu_A A \qquad [n_c/n_c] \quad \text{(C.2)}$$

$$\frac{dP}{dt} = \begin{cases} \tanh\left(\alpha_P \frac{A}{A_{max}}\right)\left(1-\frac{P}{P_{max}}\right) - \mu_P P, & L>0 \\ -\mu_P P, & L=0 \end{cases} \qquad [n_m/n_m] \quad \text{(C.3)}$$

$$\frac{dD}{dt} = U\left(1-\alpha_D A\right) D\left(1-\frac{D}{D_{max}}\right). \qquad [m^2/m^2] \quad \text{(C.4)}$$

Table C.1 provides an explanation of the socio-hydrological model variables and parameters. The parameters μ_L, λ, and γ are distributional parameters in the loss regression sub-model, and where not affected by the changes in this study. For meaningful projection results, we introduced a saturation term using the tangens hyperbolicus in the differential equations for the awareness and preparedness (Equations C.2 and C.3), which replaced the original linear term. In this way, it is guaranteed that the simulated awareness and preparedness values stay within their defined range [0,1] during the projection simulation. This reparameterization also required an adaptation of the marginal prior distributions for the affected parameters. We list the priors of the socio-hydrological parameters in Table C.1. The remaining priors stayed the same as in Schoppa et al. (2022).

The inundation sub-model estimates the average inundation depth I at company premises in the flooded areas F of the model domain from the flood return period V and the current economic density D. In the original configuration, the sub-model consisted of a linear Gamma and Beta regression model, respectively. We improved the estimation of the flooded area, by accounting for inflation in the probabilistic beta model (i.e., the presence of the values 0 and 1 in the data). This accounts for the unlikely but possible case that the floodplain is completely abandoned by companies and, hence, no premises can be flooded any more. The regression for the flooded area F is defined as:

$$F_{obs} \sim BEINF(\mu_F, \phi_F, \lambda_F, \gamma_F) \tag{C.5}$$

$$logit(\mu_F) = \alpha_{\mu_F} + \beta_{\mu_F,D} D + \beta_{\mu_F,V} V \tag{C.6}$$

$$logit(\lambda_F) = \alpha_{\lambda_F} + \beta_{\lambda_F,D} D + \beta_{\lambda_F,V} V \tag{C.7}$$

$$logit(\gamma_F) = \alpha_{\gamma_F} + \beta_{\gamma_F,D} D + \beta_{\gamma_F,V} V, \tag{C.8}$$

where μ_F, ϕ_F, λ_F, and γ_F are the distribution parameters (i.e., mean, precision, zero-one inflation, conditional one-inflation) of the zero-one-inflated Beta distribution $BEINF$. The α and β parameters are the regression intercepts and coefficients, respectively.

Additionally, the new inundation data revealed that the relationship between flood return period and average inundation depth at exposed commercial areas is highly non-linear. Therefore, we replaced the original linear inundation regression sub-model by a more flexible spline regression. A spline model is composed of a family of basis splines which, in turn, are defined by the order of the basis spline and a sequence of knots (i.e., the points at which a set of polynomials are joined together to form a basis splines). We used cubic basis splines with knots at the data points of the two predictors plus outer knots for extrapolation (i.e., k_D and k_V). The spline-based inundation sub-model reads as follows:

Table C.1: *Variables and parameters of the socio-hydrological system dynamics model. For the parameters, we also list the corresponding marginal prior distributions.*

Variable	Parameter	Description	Unit	Prior
V		flood return period	[a]	
U		economic growth rate	[1/a]	
A		awareness	$[n_c/n_c]$	
P		preparedness	$[n_m/n_m]$	
D		economic density	$[m^2/m^2]$	
L		building loss	[€/€]	
F		flooded area	$[m^2/m^2]$	
	α_A	anxiousness	$[1/(€/€)]$	$LogNormal(2, 0.2)$
	α_P	activeness	$[(n_m/n_m)/(n_c/n_c)]$	$Normal(1.25, 0.5)$
	α_D	risk aversion	$[1/(n_c/n_c)]$	$LogNormal(1.1, 0.25)$
	μ_A	flood memory	$[1/a]$	$LogNormal(-3.3927, 0.3)$
	μ_P	longevity precaution	$[1/a]$	$Normal(0.016, 0.005)$

$$I_{obs} \sim Normal(I, \sigma_I) \tag{C.9}$$
$$I = a_{0,D}D + \sum_{n=i} a_{i,D}B_{D_{i,k_D}} + a_{0,V}V + \sum_{n=j} a_{j,V}B_{V_{j,k_V}} \tag{C.10}$$
$$a_{0,D} \sim Normal(0, 1) \tag{C.11}$$
$$a_{i,D} \sim Normal(a_{i-1,D}, \tau) \tag{C.12}$$
$$a_{0,V} \sim Normal(0, 1) \tag{C.13}$$
$$a_{j,V} \sim Normal(a_{i-1,V}, \tau) \tag{C.14}$$
$$\tau \sim Normal(0, 0.25) \tag{C.15}$$
$$\sigma_I \sim Half - Normal(0, 1), \tag{C.16}$$

where σ_I is the standard deviation of the Normal distribution, B are the families of cubic basis splines (one per predictor), a_0 are the spline intercepts, and $a_{i/j}$ are the spline coefficients. We assigned a random walk smoothing prior τ on the spline coefficients, which penalizes wiggliness and, hence, prevents overfitting and reduces the influence of the knot number and location. In the computational implementation of the model in Stan, we followed Kharratzadeh (2017) and Simpson (2020). Figure C.1 illustrates the improved fit of the new inundation sub-model as compared to the original version. In particular, the new model provides more realistic estimates of inundation for very large flood return periods (i.e., beyond 500 years). We implemented the socio-hydrological flood

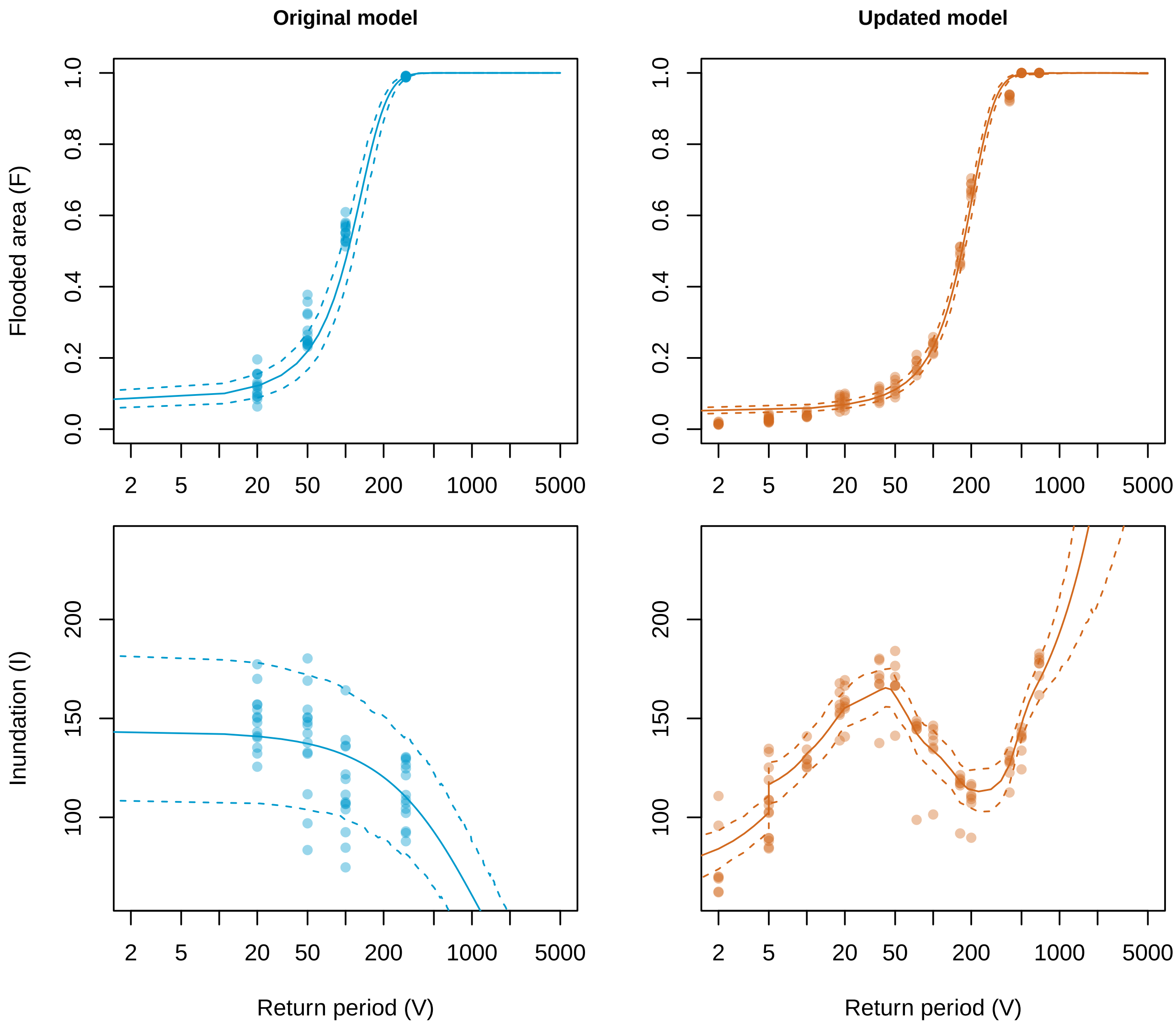

Figure C.1: *Comparison of the original, linear inundation sub-model (blue) as in Schoppa et al. (2022) and the updated inundation sub-model (orange) that uses a zero-one-inflated (flooded area) and a spline regression (inundation). The updated model accounts for the non-linearity in the new calibration data and extrapolates more realistically to large flood return periods (note the logarithmic x-axis). Calibration data is shown as points while lines show median and 95% credible intervals of the model fits. For this plot, the second predictor variable, economic density, was fixed to the value 0.15.*

risk model in the Bayesian probabilistic programming language Stan (Carpenter et al. 2017), which approximates the posterior distribution via an adaptive variant of Hamiltonian Monte Carlo sampling ('no-U-turn' sampler). For this study, we ran the sampler with four chains and 2000 samples (burn-in phase: 1500 samples). For computational reasons, we thinned the resulting samples by a factor of two so that 1000 samples of the calibrated model remained for further analysis. We extracted the parameter samples from the Stan model and continued with the simulations for the projection period in the programming language R.

C.3 Hydroclimatic and socioeconomic projection data

For the projection of future flood risk we forced the model with annual maximum flood return periods (hydroclimatic forcing) and gross domestic product (GDP) growth rates (socioeconomic forcing).

The hydroclimatic projections were generated by Mentaschi et al. (2020) using the hydrological model LISFLOOD (Van Der Knijff et al. 2010) under forcing of 11 EURO-CORDEX regional climate models (Jacob et al. 2014). The resulting gridded, pan-European simulations of future flood discharge were computed for the radiative concentration pathways (RCPs) 4.5 and 8.5, and feature changes in flood frequency in form of shifts in return period for three time hoizons - 2011-2040, 2041-2070, and 2071-2100 - relative to the baseline 1980-2010. We extracted the simulated shifts in flood return period (available for 10, 20, 50, 100, 200, 500 years) for the river Elbe at the gauge Dresden. To generate stochastic series of annual flood maxima under future climate, we first sampled flood events according to their occurrence probability (i.e., inverse of the return period) under the baseline climate and then shifted the return period on basis of the projected ensemble change under future climate scenarios. In this way, the number and succession of flood events in a simulated time series is the same across climate scenarios, isolating the effect of hydroclimatic change on flood risk. Since changes in flood frequency were only available for selected flood return periods, we approximated the flood frequency shifts for flood events of other return periods by linear inter- and extrapolation. The uncertainty in the coupled climate and hydrological models is reflected by the two RCP scenarios and the 25%, 50%, 75% percentile of the respective ensemble projection. The resulting series of flood events under different hydroclimatic forcing is then fed into the inundation regression as a predictor in form of the event return period.

For the GDP growth rates, we combined population projections with simulations of future GDP per capita for the NUTS3 region Dresden ('DED21') as described in Steinhausen et al. (2022). The basis for the population projection is the regional EUROPOP2019 projection for 2019-2100 (Eurostat 2021), which we augmented with probabilistic country level population projections from the

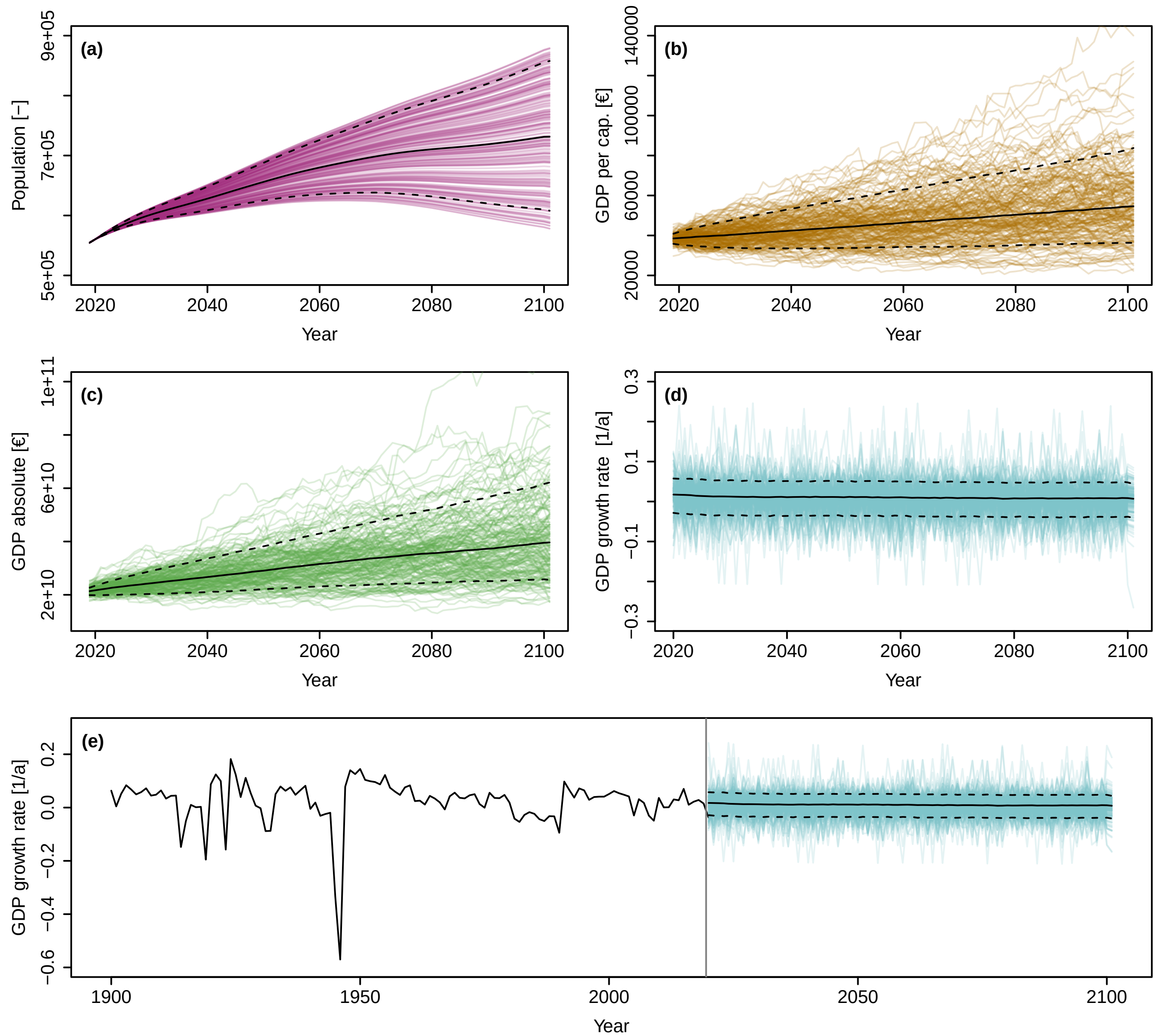

Figure C.2: *Construction of the socioeconomic forcing data for the socio-hydrological model. We multiplied population (a) with GDP per capita (b) projections to obtain probabilistic projections of absolute GDP (c) in Dresden. Then, we derived GDP growth rate trajectories (d) which drove the socio-hydrological system dynamics model. Panel (e) compares the historic GDP growth rate series with the projected forcing data. The plot displays 200 out of the 1000 trajectories, which we used for the experiments. Colored lines shows individual trajectories, while solid and dashed black lines show the median and 90% confidence intervals.*

United Nations (UNDESA 2019) to obtain estimates about the projection uncertainty. The probabilistic projection of GDP per capita stems from Steinhausen et al. (2022) and was generated by a Markov Chain Monte Carlo (MCMC) simulation on basis of observed GDP per capita data. As shown in Figure C.2, we multiplied the population and GDP per capita trajectories to obtain estimates of future absolute GDP. Afterwards, we derived GDP growth rate series, which acted as socioeconomic forcing for the socio-hydrological system dynamics model. For a more detailed methodological explanation of the population and GDP per capita projections, we refer to Steinhausen et al. (2022).

The socio-hydrological flood loss projections are expressed relative to the current building asset value (as a ratio on the interval [0,1]). To account for changes in wealth, we also projected the fixed assets in commercial buildings by using the previously calculated absolute GDP projections and wealth-to-income ratios. Namely, we extrapolated the historic trend in wealth-to-income ratio of the commercial sector in Germany (period: 1870-2020) from Paprotny et al. (2018a) and multiplied it by the absolute GDP trajectories to assess the future wealth of companies. Subsequently, we estimated the share of company wealth that is fixed in building structures from national account data of the Federal state of Saxony for the period 1995-2018 (Federal and State Statistical Offices of Germany 2021). Since the fixed asset share of buildings remained stable in the analyzed period, we assumed a constant ratio as in the year 2018. Furthermore, we downscaled the computed fixed assets in buildings as the preceding computations were carried out on the NUTS3-level, whereas the socio-hydrological model (and our analysis) is spatially limited to the 500-yr floodplain of the Elbe river. To this end, we analyzed the share of commercial building wealth in the floodplain relative to the entire NUTS3 region from observed spatial data in the period 1990-2020 (Paprotny et al. 2018a). Since this share also remained constant in the recent past (2005-2020), we assumed a constant concentration of assets in the floodplain for the projection study.

In reality, the fixed assets in commercial buildings in the floodplain is influenced by the physical floodplain exposure. That is, the presence/absence of companies in the floodplain, which is described by the economic density variable in the socio-hydrological model. We tried to couple these variables for this study but did not observe a detectable statistical influence of the economic density on the share of fixed assets in the floodplain (i.e., close to zero parameter effects). Therefore, we decided to consider these two dynamics as detached for simplification. Future efforts could improve this aspect by including dedicated model components and data that resolve the spatial distribution of wealth in the floodplain. Similarly, we assumed independence between the hydroclimatic and socioeconomic forcing data for simplification. The coupling of the hydroclimatic and GDP growth projections could lead to more accurate simulations of the flood risk evolution.

C.4 Bayesian significance testing

We evaluated the statistical significance of changes in flood risk (i.e. the metrics EAD, VAR, TVAR) due to climate, wealth, and adaptation by means of Bayesian significance testing. Bayesian significance testing offers different indices for describing the existence and significance of effects, all of which are based on the posterior distribution of a parameter or, as in this case, the simulated risk metric values (Makowski et al. 2019b). We used the 'percentage in ROPE' index that quantifies by how much the posterior distribution of a risk metric shifted away from a Region of Practical Equivalence (ROPE) due to an effect of interest (i.e., climate, wealth, adaptation). The ROPE defines the range of risk metric values in which we considered the risk reduction effect of the adaptation measure to be negligible and is the Bayesian equivalent to the frequentist point null hypothesis (Kruschke 2015; Makowski et al. 2019b). Figure C.3 shows the concept of Bayesian significance testing for an example. Based on recommendations from the literature (Kruschke 2018; Makowski et al. 2019a), we defined the ROPE as ±0.1 times the standard deviation of the risk metric value under the baseline simulation around its median (i.e., neglecting the effect of

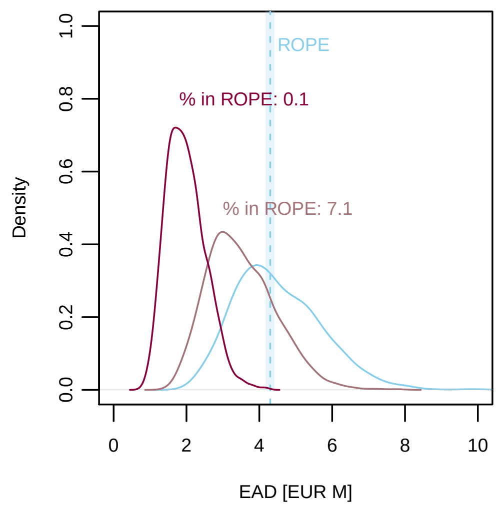

Figure C.3: *Example of Bayesian significance testing for the risk metric expected annual damage (EAD). Significance is evaluated based on the percentage of the posterior that falls within the region of practical equivalence (ROPE). For our application, we define the ROPE as ±0.1 times the standard deviation of the baseline EAD around its median (blue). The red distributions show two exemplary simulations with minor and large adaptation intervention. The risk reduction of the light red simulation is considered insignificant under our decision criterion (7.1% of posterior in ROPE), whereas the dark red simulation is evaluated as significant (0.1% of posterior in ROPE).*

interest). The amount by which the risk metric distribution under the simulation with the effect considered shifts away from the baseline simulation is quantified by the percentage of the posterior samples that fall inside the ROPE. This results in a continuous index of significance, where a smaller percentage of posterior samples in the ROPE means that the change in risk is more significant. For the interpretation of the results, we classify the computed continuous significance indices according to (Makowski et al. 2019b):

- $\geq 2.5\%$ in ROPE: undecided or negligible significance
- $\geq 1\%$ & $< 2.5\%$ in ROPE: probably significant
- $< 1\%$ in ROPE: significant

For the computational implementation of the Bayesian significance testing we used the R-package 'bayestestR' (Makowski et al. 2019a).

C.5 Additional figures simulation experiments

On the following pages, we show additional risk curves and system trajectory plots for simulations under RCP8.5.

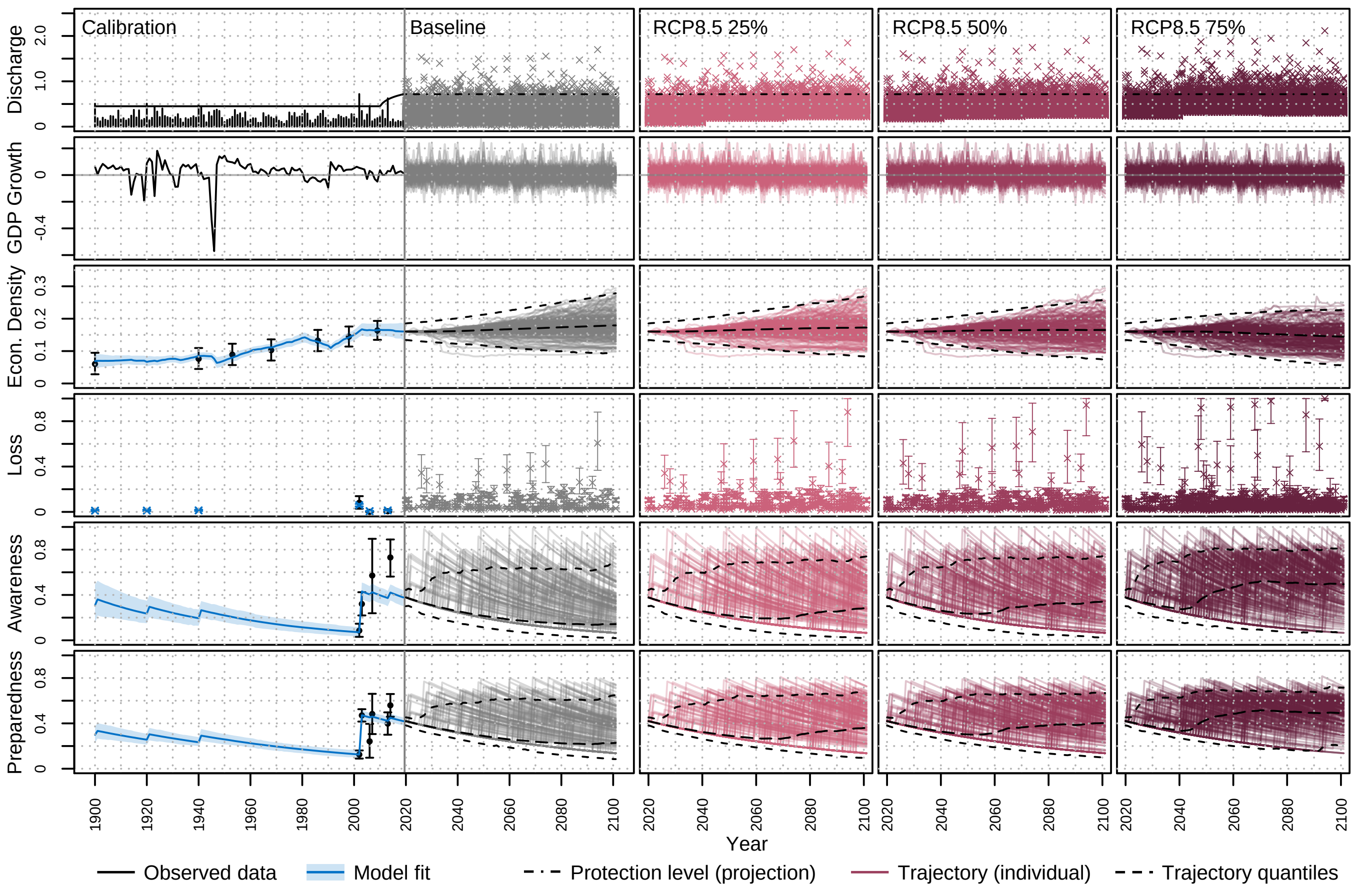

Figure C.4: *Continuous evolution of the socio-hydrological system for the calibration (1900-2019) and projection (2020-2100) period. The plot visualizes the influence of different hydroclimatic forcing scenarios: baseline with present climate and ensemble percentiles under RCP8.5 climate. For the projections, the colored lines show 200 individual trajectories (median of model uncertainty) and dashed, black lines show the aggregate evolution across all 1000 simulated trajectories (median and 95% highest density interval of projection uncertainty). The main text (Figure 4.2) contains a similar plot for RCP4.5.*

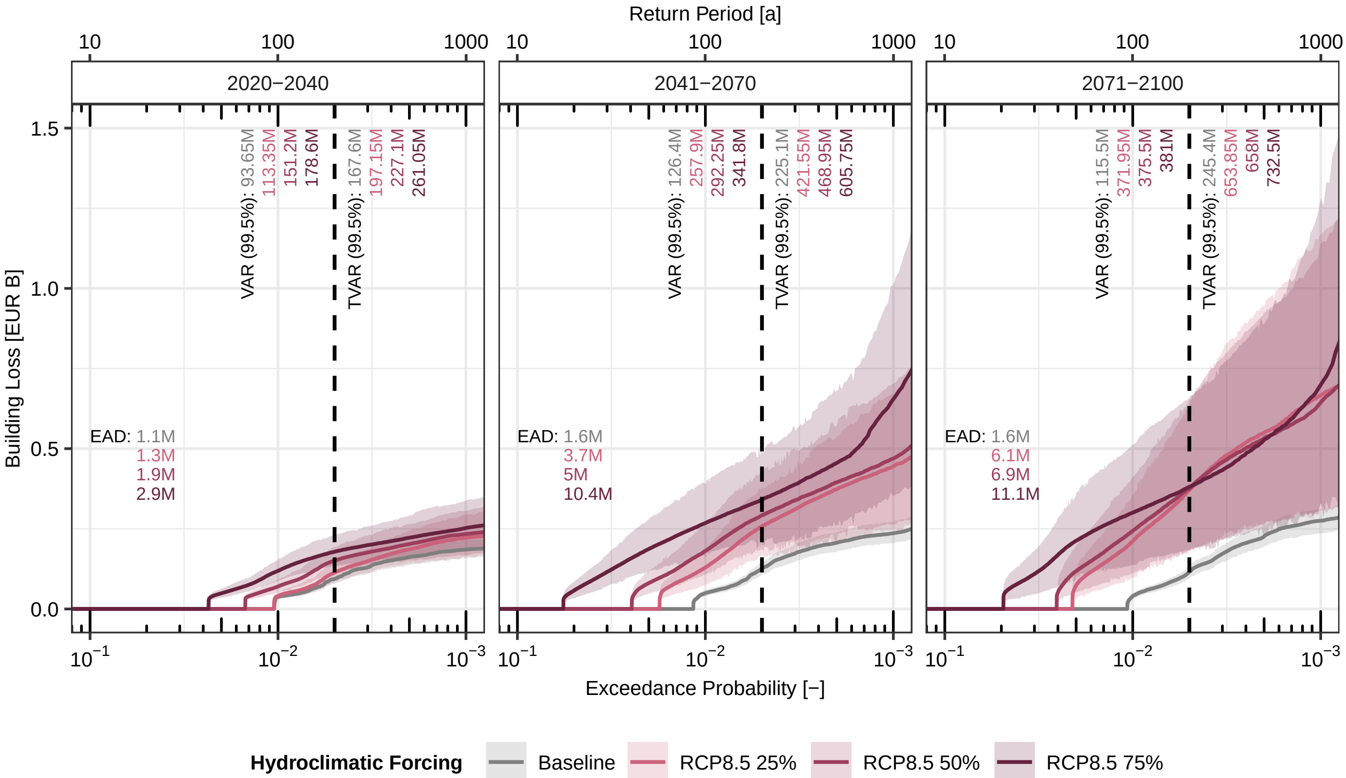

Figure C.5: *Risk curves (median and 90% highest density interval) and metrics for the three projection horizons considering RCP8.5 climate and wealth growth. For this plot, we multiplied the projected relative losses from Figure C.4 with the fixed asset values (i.e. wealth) to receive absolute losses. The individual risk curves reflect the uncertainty in hydroclimatic forcing, while intervals summarize the uncertainty in the wealth projection, flood risk model, and stochastic flood series. The baseline scenario assumes constant hydroclimatic conditions (as in 1981-2010) and wealth (as in 2020). Risk metrics (median): expected annual damage (EAD), value at risk (VAR), tail value at risk (TVAR). The main text (Figure 4.3) contains a similar plot for RCP4.5.*

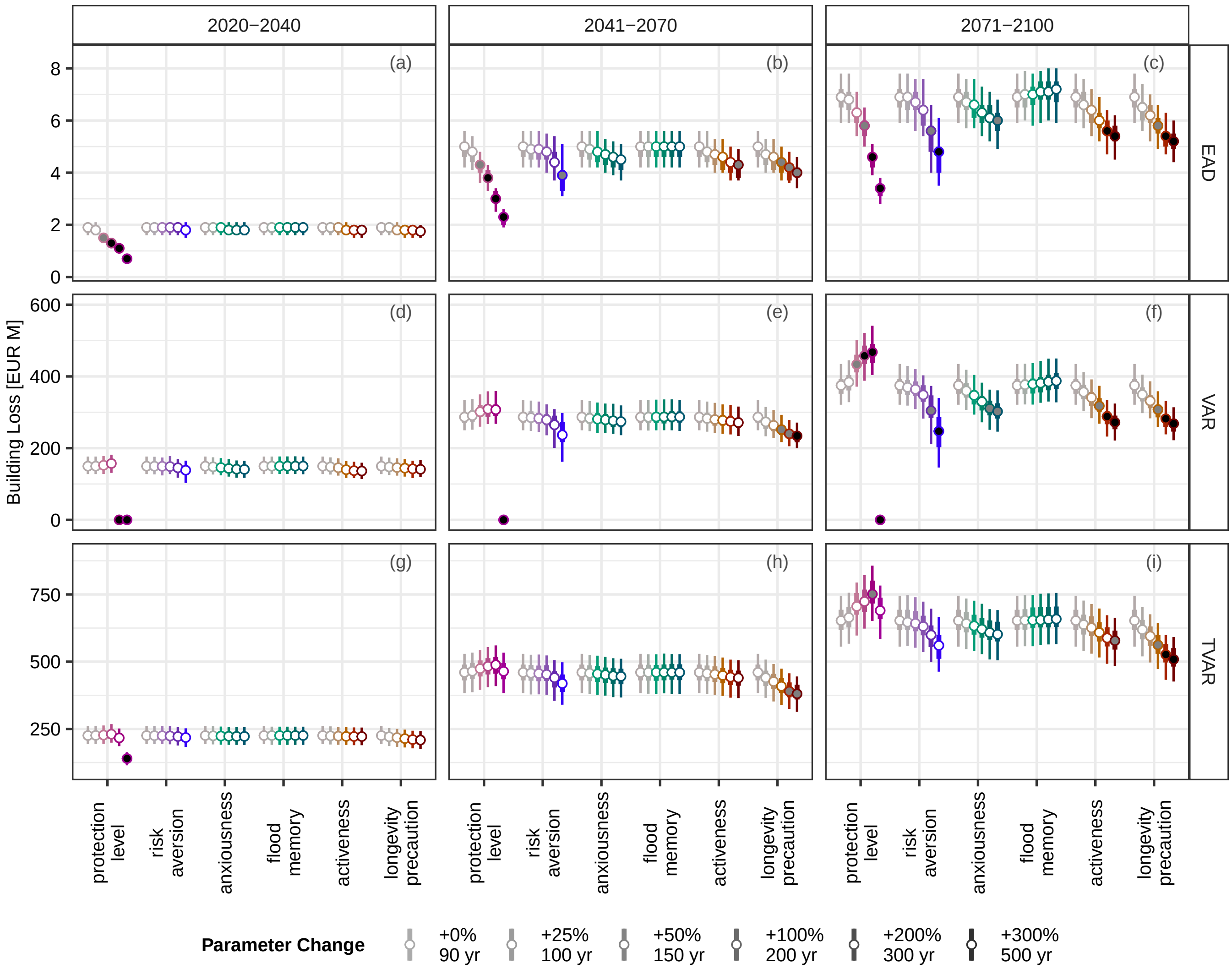

Figure C.6: *Sensitivity of flood risk metrics towards adaptation measures (i.e., parameter changes). For each interval, we changed the respective parameter while keeping the other parameters fixed at their calibrated value. The interval colors allocate the socio-hydrological parameters to the model variables that they control (pink: protection level, blue: economic density, green: awareness, orange: preparedness). The simulations are based on RCP8.5 (50% ensemble percentile) and median wealth projections (i.e., deflated climate and wealth uncertainty). The strength of the intervention (i.e., parameter change) through a adaptation measure is visualized by the boxplot chroma. Statistical effect significance is indicated by the points' fill colors: negligible/undecided (white), probably significant (grey), significant (black). Risk metrics: expected annual damage (EAD), value at risk (VAR), tail value at risk (TVAR). The main text (Figure 4.5) contains a similar plot for RCP4.5.*

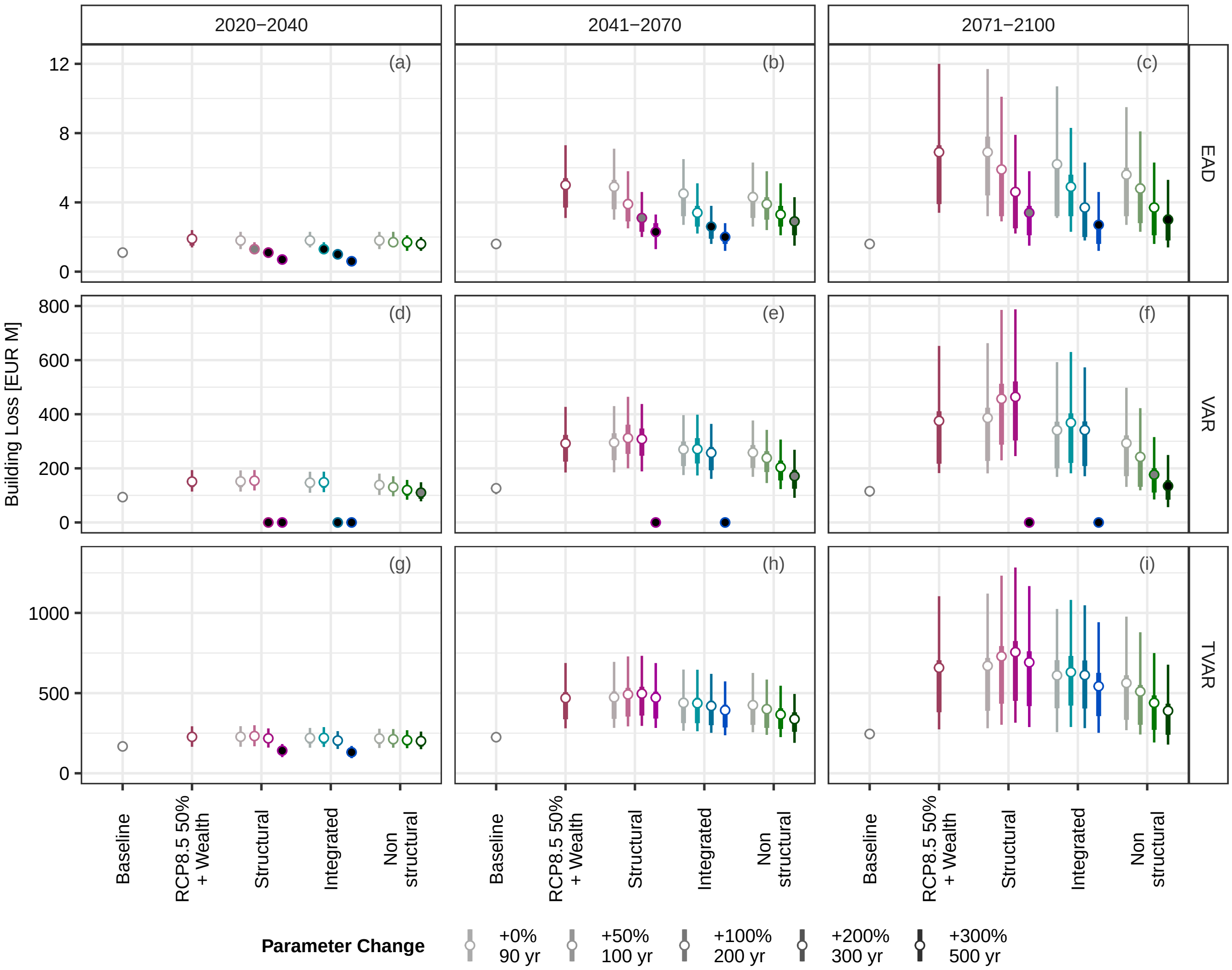

Figure C.7: *Potential of competing adaptation strategies (structural, integrated, non-structural) to mitigate the expected increase in flood risk. The baseline scenario assumes fixed climate and wealth while the 'RCP8.5 50% + Wealth' scenario assumes hydroclimatic and wealth projections with uncertainty. The interval colors correspond to the color coding from Figure 4 and C.6 for same simulation runs (i.e., 'RCP8.5 50% + Wealth' and 'Structural'). The strength of the intervention (i.e., parameter change) through a adaptation measure is visualized by the boxplot chroma. Statistical effect significance is indicated by the points' fill colors: negligible/undecided (white), probably significant (grey), significant (black). Risk metrics: expected annual damage (EAD), value at risk (VAR), tail value at risk (TVAR). The main text (Figure 4.6) contains a similar plot for RCP4.5.*

Bibliography

Abebe, Yared Abayneh et al. (Jan. 2019). "A coupled flood-agent-institution modelling (CLAIM) framework for urban flood risk management". In: *Environmental Modelling & Software* 111.February, pp. 483–492. ISSN: 13648152. DOI: 10.1016/j.envsoft.2018.10.015.

Aerts, J. C. J. H. et al. (Mar. 2018). "Integrating human behaviour dynamics into flood disaster risk assessment". In: *Nature Climate Change* 8.3, pp. 193–199. ISSN: 1758-678X. DOI: 10.1038/s41558-018-0085-1. URL: http://www.nature.com/articles/s41558-018-0085-1.

Alfieri, Lorenzo et al. (June 2016). "Modelling the socio-economic impact of river floods in Europe". In: *Natural Hazards and Earth System Sciences* 16.6, pp. 1401–1411. ISSN: 1684-9981. DOI: 10.5194/nhess-16-1401-2016. URL: https://www.nat-hazards-earth-syst-sci.net/16/1401/2016/.

Alfieri, Lorenzo et al. (Feb. 2017). "Global projections of river flood risk in a warmer world". In: *Earth's Future* 5.2, pp. 171–182. ISSN: 23284277. DOI: 10.1002/2016EF000485. URL: http://doi.wiley.com/10.1002/2016EF000485.

Alfieri, Lorenzo et al. (Jan. 2018). "Multi-Model Projections of River Flood Risk in Europe under Global Warming". In: *Climate* 6.1, p. 6. ISSN: 2225-1154. DOI: 10.3390/cli6010006. URL: http://www.mdpi.com/2225-1154/6/1/6.

Alonso Vicario, S. et al. (Oct. 2020). "Unravelling the influence of human behaviour on reducing casualties during flood evacuation". In: *Hydrological Sciences Journal* 65.14, pp. 2359–2375. ISSN: 0262-6667. DOI: 10.1080/02626667.2020.1810254. URL: https://doi.org/10.1080/02626667.2020.1810254.

Apel, Heiko, Bruno Merz, and Annegret H. Thieken (June 2008). "Quantification of uncertainties in flood risk assessments". In: *International Journal of River Basin Management* 6.2, pp. 149–162. ISSN: 1571-5124. DOI: 10.1080/15715124.2008.9635344. URL: http://www.tandfonline.com/doi/abs/10.1080/15715124.2008.9635344.

Apel, Heiko et al. (Apr. 2004). "Flood risk assessment and associated uncertainty". In: *Natural Hazards and Earth System Sciences* 4.2, pp. 295–308. ISSN: 1684-9981.

DOI: 10.5194/nhess-4-295-2004. URL: https://www.nat-hazards-earth-syst-sci.net/4/295/2004/nhess-4-295-2004.pdf.

Ara Begum, R. et al. (2022). "Point of Departure and Key Concepts". In: *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by H.-O. Pörtner et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Chap. 4, pp. 551–712. DOI: 10.1017/9781009325844.006. URL: https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_FullReport.pdf.

Arias, P. A. et al. (2021). "Technical Summary". In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson-Delmotte et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Chap. 1. URL: https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_TS.pdf.

Aznar-Siguan, Gabriela and David N. Bresch (July 2019). "CLIMADA v1: a global weather and climate risk assessment platform". In: *Geoscientific Model Development* 12.7, pp. 3085–3097. ISSN: 1991-9603. DOI: 10.5194/gmd-12-3085-2019. URL: https://gmd.copernicus.org/articles/12/3085/2019/.

Bala, Bilash Kanti, Fatimah Mohamed Arshad, and Kusairi Mohd Noh (2017). *System Dynamics*. Springer Texts in Business and Economics. Singapore: Springer Singapore, pp. 25–85. ISBN: 978-981-10-2043-8. DOI: 10.1007/978-981-10-2045-2. URL: http://link.springer.com/10.1007/978-981-10-2045-2.

Barendrecht, M. H. et al. (Feb. 2019). "The Value of Empirical Data for Estimating the Parameters of a Sociohydrological Flood Risk Model". In: *Water Resources Research* 55.2, pp. 1312–1336. ISSN: 0043-1397. DOI: 10.1029/2018WR024128. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024128.

Barendrecht, M.H. et al. (Dec. 2020). "Needed: A systems approach to improve flood risk mitigation through private precautionary measures". In: *Water Security* 11.February 2020, p. 100080. ISSN: 24683124. DOI: 10.1016/j.wasec.2020.100080. URL: https://doi.org/10.1016/j.wasec.2020.100080.

Barendrecht, Marlies H., Alberto Viglione, and Günter Blöschl (July 2017). "A dynamic framework for flood risk". In: *Water Security* 1, pp. 3–11. ISSN: 24683124. DOI: 10.1016/j.wasec.2017.02.001. URL: http://dx.doi.org/10.1016/j.wasec.2017.02.001.

Barthel, Fabian and Eric Neumayer (July 2012). "A trend analysis of normalized insured damage from natural disasters". In: *Climatic Change* 113.2, pp. 215–237. ISSN: 0165-0009. DOI: 10.1007/s10584-011-0331-2. URL: http://link.springer.com/10.1007/s10584-011-0331-2.

Ben-Haim, Yakov (2019). *Decision Making under Deep Uncertainty*. Ed. by Vincent A. W. J. Marchau et al. Cham: Springer International Publishing, pp. 93–115. ISBN: 978-3-030-05251-5. DOI: 10.1007/978-3-030-05252-2. URL: http://link.springer.com/10.1007/978-3-030-05252-2.

Bevere, Lucia and Federica Remondi (2022). *Natural catastrophes in 2021: the floodgates are open*. Tech. rep. 1. Zurich, Switzerland. URL: https://www.swissre.com/institute/research/sigma-research/sigma-2022-01.html.

BfG (2021). *Global Runoff Data Centre*. Dataset. URL: https://www.bafg.de/GRDC/.

Blair, P. and W. Buytaert (Jan. 2016). "Socio-hydrological modelling: a review asking "why, what and how?"". In: *Hydrology and Earth System Sciences* 20.1, pp. 443–478. ISSN: 1607-7938. DOI: 10.5194/hess-20-443-2016. URL: https://hess.copernicus.org/articles/20/443/2016/.

Bouckaert, R.R. (1995). "Bayesian Belief Networks: From Construction to Inference". PhD thesis. Utrecht University. ISBN: 90-393-0848-9. URL: https://dspace.library.uu.nl/handle/1874/845.

Bouwer, Laurens M. (Jan. 2011). "Have Disaster Losses Increased Due to Anthropogenic Climate Change?" In: *Bulletin of the American Meteorological Society* 92.1, pp. 39–46. ISSN: 0003-0007. DOI: 10.1175/2010BAMS3092.1. URL: http://journals.ametsoc.org/doi/10.1175/2010BAMS3092.1.

Brand, Amy et al. (Apr. 2015). "Beyond authorship: attribution, contribution, collaboration, and credit". In: *Learned Publishing* 28.2, pp. 151–155. ISSN: 09531513. DOI: 10.1087/20150211. URL: http://doi.wiley.com/10.1087/20150211.

Breiman, Leo (2001). "Random Forests". In: *Machine Learning* 45.1, pp. 5–32. DOI: https://doi.org/10.1023/A:1010933404324.

Breiman, Leo et al. (Oct. 19, 1984). *Classification And Regression Trees*. Vol. 1. 1. Monterey, CA: Routledge. ISBN: 9781315139470. DOI: 10.1201/9781315139470.

Brito, Mariana Madruga de and Mariele Evers (Apr. 2016). "Multi-criteria decision-making for flood risk management: a survey of the current state of the art". In: *Natural Hazards and Earth System Sciences* 16.4, pp. 1019–1033. ISSN: 1684-9981. DOI: 10.5194/nhess-16-1019-2016. URL: https://nhess.copernicus.org/articles/16/1019/2016/.

Bubeck, Philip et al. (2018). "Insights into Flood-Coping Appraisals of Protection Motivation Theory: Empirical Evidence from Germany and France". In: *Risk Analysis* 38.6, pp. 1239–1257. ISSN: 15396924. DOI: 10.1111/risa.12938.

Bubeck, Philip et al. (2020). "Using Panel Data to Understand the Dynamics of Human Behavior in Response to Flooding". In: *Risk Analysis* 40.11, pp. 2340–2359. ISSN: 15396924. DOI: 10.1111/risa.13548.

Bürkner, Paul-Christian (2017). "brms: An R Package for Bayesian Multilevel Models Using Stan". In: *Journal of Statistical Software* 80.1. ISSN: 1548-7660. DOI: 10.18637/jss.v080.i01.

Bürkner, Paul-Christian (2018). "Advanced Bayesian Multilevel Modeling with the R Package brms". In: *The R Journal* 10.1, p. 395. ISSN: 2073-4859. DOI: 10.32614/RJ-2018-017. arXiv: 1705.11123. URL: https://journal.r-project.org/archive/2018/RJ-2018-017/index.html.

Caretta, M.A. et al. (2022). "Water". In: *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by H.-O. Pörtner et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Chap. 4, pp. 551–712. DOI: 10.1017/9781009325844.006. URL: https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_FullReport.pdf.

Carisi, Francesca et al. (July 2018). "Development and assessment of uni- and multivariable flood loss models for Emilia-Romagna (Italy)". In: *Natural Hazards and Earth System Sciences* 18.7, pp. 2057–2079. ISSN: 1684-9981. DOI: 10.5194/nhess-18-2057-2018. URL: https://www.nat-hazards-earth-syst-sci.net/18/2057/2018/.

Carpenter, Bob et al. (2017). "Stan : A Probabilistic Programming Language". In: *Journal of Statistical Software* 76.1. ISSN: 1548-7660. DOI: 10.18637/jss.v076.i01. URL: http://www.jstatsoft.org/v76/i01/.

Cerri, Marco et al. (Feb. 2021). "Are OpenStreetMap building data useful for flood vulnerability modelling?" In: *Natural Hazards and Earth System Sciences* 21.2, pp. 643–662. ISSN: 1684-9981. DOI: 10.5194/nhess-21-643-2021. URL: https://nhess.copernicus.org/articles/21/643/2021/.

City of Dresden (2018). *Siedlungsbereiche mit einem verbleibenden Schutzgrad kleiner HQ100 an der Elbe (2010)*. Dataset. URL: https://kommisdd.dresden.de/net4/public/ogcapi/collections/L890.

Ciullo, Alessio et al. (Apr. 2017). "Socio-hydrological modelling of flood-risk dynamics: comparing the resilience of green and technological systems". In: *Hydrological Sciences Journal* 62.6, pp. 880–891. ISSN: 0262-6667. DOI: 10.1080/02626667.2016.1273527. URL: https://www.tandfonline.com/doi/full/10.1080/02626667.2016.1273527.

Coates, G. et al. (June 2014). "Agent-based modelling and inundation prediction to enable the identification of businesses affected by flooding". In: *WIT Transactions on Ecology and the Environment*. Vol. 184, pp. 13–22. ISBN: 9781845647841. DOI: 10.2495/FRIAR140021. URL: http://library.witpress.com/viewpaper.asp?pcode=FRIAR14-002-1.

Coates, Graham et al. (Feb. 2019). "Agent-based modeling and simulation to assess flood preparedness and recovery of manufacturing small and medium-sized enterprises". In: *Engineering Applications of Artificial Intelligence* 78.November

2018, pp. 195–217. ISSN: 09521976. DOI: 10.1016/j.engappai.2018.11.010. URL: https://doi.org/10.1016/j.engappai.2018.11.010.

Cox, Louis Anthony Tony (Oct. 2012). “Confronting Deep Uncertainties in Risk Analysis”. In: *Risk Analysis* 32.10, pp. 1607–1629. ISSN: 02724332. DOI: 10.1111/j.1539-6924.2012.01792.x. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.2012.01792.x.

de Moel, H. and J. C. J. H. Aerts (July 2011). “Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates”. In: *Natural Hazards* 58.1, pp. 407–425. ISSN: 0921-030X. DOI: 10.1007/s11069-010-9675-6. URL: http://link.springer.com/10.1007/s11069-010-9675-6.

de Moel, H. et al. (Aug. 2015). “Flood risk assessments at different spatial scales”. In: *Mitigation and Adaptation Strategies for Global Change* 20.6, pp. 865–890. ISSN: 1381-2386. DOI: 10.1007/s11027-015-9654-z. URL: http://link.springer.com/10.1007/s11027-015-9654-z.

de Moel, Hans, Mathijs van Vliet, and Jeroen C. J. H. Aerts (Feb. 2013). “Evaluating the effect of flood damage-reducing measures: a case study of the unembanked area of Rotterdam, the Netherlands”. In: *Regional Environmental Change* 14.3, pp. 895–908. ISSN: 1436-3798. DOI: 10.1007/s10113-013-0420-z. URL: http://link.springer.com/10.1007/s10113-013-0420-z.

Delignette-Muller, Marie Laure and Christophe Dutang (2015). “fitdistrplus : An R Package for Fitting Distributions”. In: *Journal of Statistical Software* 64.4, pp. 1–34. ISSN: 1548-7660. DOI: 10.18637/jss.v064.i04.

Devkota, Rohini et al. (2020). “Assessing the past and adapting to future floods: a hydro-social analysis”. In: *Climatic Change* 163.2, pp. 1065–1082. ISSN: 15731480. DOI: 10.1007/s10584-020-02909-w.

Di Baldassarre, Giuliano, Luigia Brandimarte, and Keith Beven (2016). “The seventh facet of uncertainty: Wrong assumptions, unknowns and surprises in the dynamics of human–water systems”. In: *Hydrological Sciences Journal* 61.9, pp. 1748–1758. ISSN: 21503435. DOI: 10.1080/02626667.2015.1091460. URL: http://dx.doi.org/10.1080/02626667.2015.1091460.

Di Baldassarre, Giuliano et al. (Aug. 2013). “Socio-hydrology: conceptualising human-flood interactions”. In: *Hydrology and Earth System Sciences* 17.8, pp. 3295–3303. ISSN: 1607-7938. DOI: 10.5194/hess-17-3295-2013. URL: https://hess.copernicus.org/articles/17/3295/2013/.

Di Baldassarre, Giuliano et al. (June 2015). “Debates-Perspectives on socio-hydrology: Capturing feedbacks between physical and social processes”. In: *Water Resources Research* 51.6, pp. 4770–4781. ISSN: 00431397. DOI: 10.1002/2014WR016416. URL: http://doi.wiley.com/10.1002/2014WR016416.

Di Baldassarre, Giuliano et al. (Mar. 2017). “Drought and flood in the Anthropocene: feedback mechanisms in reservoir operation”. In: *Earth System Dy-*

namics 8.1, pp. 225–233. ISSN: 2190-4987. DOI: 10.5194/esd-8-225-2017. URL: https://esd.copernicus.org/articles/8/225/2017/.

Disse, M. et al. (2020). "Exploring the relation between flood risk management and flood resilience". In: *Water Security* 9.January, p. 100059. ISSN: 24683124. DOI: 10.1016/j.wasec.2020.100059. URL: https://doi.org/10.1016/j.wasec.2020.100059.

Dottori, F et al. (2020). *Adapting to rising river flood risk in the EU under climate change : JRC PESETA IV project : Task 5*. Publications Office of the European Union. DOI: 10.2760/14505.

Dottori, Francesco et al. (Sept. 2018). "Increased human and economic losses from river flooding with anthropogenic warming". In: *Nature Climate Change* 8.9, pp. 781–786. ISSN: 1758-678X. DOI: 10.1038/s41558-018-0257-z. URL: http://dx.doi.org/10.1038/s41558-018-0257-z%20http://www.nature.com/articles/s41558-018-0257-z.

Elmer, F. et al. (Oct. 11, 2010). "Influence of flood frequency on residential building losses". In: *Natural Hazards and Earth System Science* 10.10, pp. 2145–2159. ISSN: 1684-9981. DOI: 10.5194/nhess-10-2145-2010.

European Parliament and European Council (Feb. 2009). *Directive 2009/138/EC on the Taking up and Pursuit of the Business of Insurance and Reinsurance (Solvency II)*. URL: http://data.europa.eu/eli/dir/2009/138/oj.

Eurostat (2008). *NACE Rev. 2 – Statistical classification of economic activities in the European Community*. Luxembourg: Office for Official Publications of the European Communities. ISBN: 978-92-79-04741-1.

— (2021). *EUROPOP2019 - Population projections at regional level*. Dataset. URL: https://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?file=data/proj_19rp3.tsv.gz.

Falter, Daniela et al. (May 2015). "Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain". In: *Journal of Hydrology* 524, pp. 182–193. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2015.02.021. URL: http://dx.doi.org/10.1016/j.jhydrol.2015.02.021.

Fanta, Václav, Miroslav Šálek, and Petr Sklenicka (Dec. 2019). "How long do floods throughout the millennium remain in the collective memory?" In: *Nature Communications* 10.1, p. 1105. ISSN: 2041-1723. DOI: 10.1038/s41467-019-09102-3. URL: http://www.nature.com/articles/s41467-019-09102-3.

Federal and State Statistical Offices of Germany (2021). *Regional accounts of the German states*. Dataset. URL: https://www.statistikportal.de/sites/default/files/2021-06/vgrdl_r1b4_bs2020_1.xlsx.

Federal Dam Operation Authority of Saxony (2013). *Neue Hochwasserschutzanlagen retten Dresden*. URL: https://www.medienservice.sachsen.de/medien/news/184938.

Ferrari, Silvia and Francisco Cribari-Neto (Aug. 2004). "Beta Regression for Modelling Rates and Proportions". In: *Journal of Applied Statistics* 31.7, pp. 799–815. ISSN: 0266-4763. DOI: 10.1080/0266476042000214501.

Figueiredo, Rui et al. (May 2018). "Multi-model ensembles for assessment of flood losses and associated uncertainty". In: *Natural Hazards and Earth System Sciences* 18.5, pp. 1297–1314. ISSN: 1684-9981. DOI: 10.5194/nhess-18-1297-2018. URL: https://nhess.copernicus.org/articles/18/1297/2018/.

Gabry, Jonah et al. (Feb. 15, 2019). "Visualization in Bayesian workflow". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 182.2, pp. 389–402. ISSN: 0964-1998. DOI: 10.1111/rssa.12378.

Gelman, Andrew and Christian Hennig (Oct. 2017). "Beyond subjective and objective in statistics". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 180.4, pp. 967–1033. ISSN: 0964-1998. DOI: 10.1111/rssa.12276. arXiv: 1508.05453. URL: https://onlinelibrary.wiley.com/doi/10.1111/rssa.12276.

Gelman, Andrew et al. (Nov. 27, 2013). *Bayesian Data Analysis*. 3rd ed. Chapman and Hall/CRC. ISBN: 9781439898208. DOI: 10.1201/b16018.

Gelman, Andrew et al. (Nov. 3, 2020). "Bayesian Workflow". In: DOI: 10.48550/arXiv.2011.01808. URL: https://arxiv.org/abs/2011.01808.

Gerl, Tina et al. (July 2016). "A Review of Flood Loss Models as Basis for Harmonization and Benchmarking". In: *PLOS ONE* 11.7. Ed. by Guy J-P. Schumann, e0159791. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0159791. URL: https://dx.plos.org/10.1371/journal.pone.0159791.

German Federal Institute of Hydrology (BfG) (2021). *Global Runoff Data Centre*. URL: https://www.bafg.de/GRDC/.

German Federal Ministry of the Interior (2013). *Bericht zur Flutkatastrophe 2013: Katastrophenhilfe, Entschädigung, Wiederaufbau*. Tech. rep. URL: https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/bevoelkerungsschutz/kabinettsbericht-fluthilfe.pdf?__blob=publicationFile&v=3.

GFZ (2021). *German flood damage database HOWAS21*. Dataset. DOI: 10.1594/GFZ.SDDB.HOWAS21.

Gissing, Andrew and Russell Blong (July 2004). "Accounting for variability in commercial flood damage estimation". In: *Australian Geographer* 35.2, pp. 209–222. ISSN: 0004-9182. DOI: 10.1080/0004918042000249511. URL: https://www.tandfonline.com/doi/full/10.1080/0004918042000249511.

Gneiting, Tilmann and Matthias Katzfuss (Jan. 3, 2014). "Probabilistic Forecasting". In: *Annual Review of Statistics and Its Application* 1.1, pp. 125–151. ISSN: 2326-8298. DOI: 10.1146/annurev-statistics-062713-085831.

Gober, Patricia and Howard S. Wheater (June 2015). "Debates-Perspectives on socio-hydrology: Modeling flood risk as a public policy problem". In: *Water Resources Research* 51.6, pp. 4782–4788. ISSN: 00431397. DOI: 10.1002/2015WR016945. URL: http://doi.wiley.com/10.1002/2015WR016945.

Grames, Johanna et al. (Sept. 2016). "Modeling the interaction between flooding events and economic growth". In: *Ecological Economics* 129, pp. 193–209. ISSN: 09218009. DOI: 10.1016/j.ecolecon.2016.06.014. URL: https://linkinghub.elsevier.com/retrieve/pii/S0921800915303980.

Grigg, Neil S. and Otto J. Helweg (Apr. 1975). "STATE-OF-THE-ART OF ESTIMATING FLOOD DAMAGE IN URBAN AREAS". In: *Journal of the American Water Resources Association* 11.2, pp. 379–390. ISSN: 1093-474X. DOI: 10.1111/j.1752-1688.1975.tb00689.x.

Grothmann, Torsten and Fritz Reusswig (2006). "People at risk of flooding: Why some residents take precautionary action while others do not". In: *Natural Hazards* 38.1-2, pp. 101–120. ISSN: 0921030X. DOI: 10.1007/s11069-005-8604-6.

Gruner, T. (2012). *WebGIS-basierte Visualisierung der Flächennutzungsentwicklung der Stadtregion Dresden auf Grundlage von MapServer*. Dataset. URL: http://maps.ioer.de/FNDD2/.

Haasnoot, Marjolijn et al. (Dec. 2012). "Exploring pathways for sustainable water management in river deltas in a changing environment". In: *Climatic Change* 115.3-4, pp. 795–819. ISSN: 0165-0009. DOI: 10.1007/s10584-012-0444-2. URL: http://link.springer.com/10.1007/s10584-012-0444-2.

Haasnoot, Marjolijn et al. (Apr. 2013). "Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world". In: *Global Environmental Change* 23.2, pp. 485–498. ISSN: 09593780. DOI: 10.1016/j.gloenvcha.2012.12.006. URL: http://dx.doi.org/10.1016/j.gloenvcha.2012.12.006%20https://linkinghub.elsevier.com/retrieve/pii/S095937801200146X.

Haeffner, Melissa and Dana Hellman (2020). "The social geometry of collaborative flood risk management: a hydrosocial case study of Tillamook County, Oregon". In: *Natural Hazards* 103.3, pp. 3303–3325. ISSN: 15730840. DOI: 10.1007/s11069-020-04131-4.

Haer, Toon, W J Wouter Botzen, and Jeroen C J H Aerts (Apr. 2019). "Advancing disaster policies by integrating dynamic adaptive behaviour in risk assessments using an agent-based modelling approach". In: *Environmental Research Letters* 14.4, p. 044022. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab0770. URL: https://iopscience.iop.org/article/10.1088/1748-9326/ab0770.

— (June 2016). "The effectiveness of flood risk communication strategies and the influence of social networks—Insights from an agent-based model". In: *Environmental Science & Policy* 60, pp. 44–52. ISSN: 14629011. DOI: 10.1016/j.

envsci.2016.03.006. URL: http://dx.doi.org/10.1016/j.envsci.2016.03.006.

Haer, Toon et al. (Oct. 2017). "Integrating Household Risk Mitigation Behavior in Flood Risk Analysis: An Agent-Based Model Approach". In: *Risk Analysis* 37.10, pp. 1977–1992. ISSN: 02724332. DOI: 10.1111/risa.12740. URL: https://onlinelibrary.wiley.com/doi/10.1111/risa.12740.

Haer, Toon et al. (Jan. 2020). "The safe development paradox: An agent-based model for flood risk under climate change in the European Union". In: *Global Environmental Change* 60.December 2018, p. 102009. ISSN: 09593780. DOI: 10.1016/j.gloenvcha.2019.102009. URL: https://doi.org/10.1016/j.gloenvcha.2019.102009%20https://linkinghub.elsevier.com/retrieve/pii/S0959378018314079.

Hall, Jim and Dimitri Solomatine (June 2008). "A framework for uncertainty analysis in flood risk management decisions". In: *International Journal of River Basin Management* 6.2, pp. 85–98. ISSN: 1571-5124. DOI: 10.1080/15715124.2008.9635339. URL: http://www.tandfonline.com/doi/abs/10.1080/15715124.2008.9635339.

Hasanzadeh Nafari, Roozbeh, T. Ngo, and W. Lehman (Aug. 2016a). "Development and evaluation of FLFAcs – A new Flood Loss Function for Australian commercial structures". In: *International Journal of Disaster Risk Reduction* 17, pp. 13–23. ISSN: 22124209. DOI: 10.1016/j.ijdrr.2016.03.007.

Hasanzadeh Nafari, Roozbeh, Tuan Ngo, and Priyan Mendis (July 9, 2016b). "An Assessment of the Effectiveness of Tree-Based Models for Multi-Variate Flood Damage Assessment in Australia". In: *Water* 8.7, p. 282. ISSN: 2073-4441. DOI: 10.3390/w8070282.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (Nov. 7, 2009). *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY: Springer New York. ISBN: 978-0-387-84857-0. DOI: 10.1007/978-0-387-84858-7.

Heckerman, David, Dan Geiger, and David Maxwell Chickering (Feb. 2013). "Learning Bayesian Networks: The Combination of Knowledge and Statistical Data". In: *Machine Learning* 20.3, pp. 197–243. ISSN: 15730565. DOI: 10.1023/A:1022623210503. eprint: 1302.6815. URL: https://link.springer.com/article/10.1023/A:1022623210503.

Hintze, Jerry L and Ray D Nelson (May 1998). "Violin Plots: A Box Plot-Density Trace Synergism". In: *The American Statistician* 52.2, pp. 181–184. ISSN: 0003-1305. DOI: 10.1080/00031305.1998.10480559. URL: http://www.tandfonline.com/doi/abs/10.1080/00031305.1998.10480559.

Hipsey, Matthew R. et al. (Sept. 2015). "Predicting the resilience and recovery of aquatic systems: A framework for model evolution within environmental observatories". In: *Water Resources Research* 51.9, pp. 7023–7043. ISSN: 00431397. DOI: 10.1002/2015WR017175.

Hirabayashi, Yukiko et al. (Sept. 2013). "Global flood risk under climate change". In: *Nature Climate Change* 3.9, pp. 816–821. ISSN: 1758-678X. DOI: 10.1038/nclimate1911. URL: http://www.nature.com/articles/nclimate1911.

Hoffman, Matthew D. and Andrew Gelman (2014). "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo". In: *Journal of Machine Learning Research* 15, pp. 1593–1623. ISSN: 15337928. URL: http://jmlr.org/papers/v15/hoffman14a.html.

Højsgaard, Søren (2012). "Graphical Independence Networks with the gRain Package for R". In: *Journal of Statistical Software* 46.10, pp. 37–44. ISSN: 1548-7660. DOI: 10.18637/jss.v046.i10. URL: http://www.jstatsoft.org/v46/i10/.

Holschneider, Matthias, Gert Zöller, and Sebastian Hainzl (2011). "Estimation of the maximum possible magnitude in the framework of a doubly truncated Gutenberg-Richter model". In: *Bulletin of the Seismological Society of America* 101.4, pp. 1649–1659. ISSN: 00371106. DOI: 10.1785/0120100289.

Hosking, J. R. M. (Sept. 1990). "L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 52.1, pp. 105–124. ISSN: 00359246. DOI: 10.1111/j.2517-6161.1990.tb01775.x.

Hothorn, Torsten, Kurt Hornik, and Achim Zeileis (Sept. 2006). "Unbiased Recursive Partitioning: A Conditional Inference Framework". In: *Journal of Computational and Graphical Statistics* 15.3, pp. 651–674. ISSN: 1061-8600. DOI: 10.1198/106186006X133933.

Hothorn, Torsten and Achim Zeileis (2015). "partykit: A Modular Toolkit for Recursive Partytioning in R". In: *Journal of Machine Learning Research* 16.118, pp. 3905–3909. URL: http://jmlr.org/papers/v16/hothorn15a.html.

Hudson, P. et al. (July 2014). "Evaluating the effectiveness of flood damage mitigation measures by the application of propensity score matching". In: *Natural Hazards and Earth System Sciences* 14.7, pp. 1731–1747. ISSN: 1684-9981. DOI: 10.5194/nhess-14-1731-2014. URL: https://nhess.copernicus.org/articles/14/1731/2014/.

Hudson, Paul, Annegret H. Thieken, and Philip Bubeck (May 2020). "The challenges of longitudinal surveys in the flood risk domain". In: *Journal of Risk Research* 23.5, pp. 642–663. ISSN: 1366-9877. DOI: 10.1080/13669877.2019.1617339. URL: https://doi.org/10.1080/13669877.2019.1617339.

Huizinga, Jan, Hans de Moel, and Wojciech Szewczyk (2017). *Global flood depth-damage functions: Methodology and the Database with Guidelines*. Tech. rep. Joint Research Centre (JRC), pp. 1–108. DOI: 10.2760/16510. arXiv: ISSN1743-3541.

URL: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC105688/global_flood_depth-damage_functions__10042017.pdf.

Inklaar, Robert et al. (Jan. 2018). *Rebasing 'Maddison': new income comparisons and the shape of long-run economic development*. English. WorkingPaper. Groningen Growth and Development Center. URL: https://research.rug.nl/en/publications/fecbca82-9752-40a9-a737-d524e629b6e7.

International Commission for the Protection of the Rhine (2016). *Tool and Assessment Method for Determining Flood Risk Evolution or Reduction - Technical Report*. Tech. rep. 237. URL: https://www.iksr.org/en/public-relations/documents/archive/technical-reports/reports-and-brochures-individual-presentation/237-technical-report-tool-and-assessment-method-to-determine-flood-risk-evolution-reduction.

Ishtiaque, Asif, Nikhil Sangwan, and David J. Yu (2017). "Robust-yet-fragile nature of partly engineered social-ecological systems: A case study of coastal Bangladesh". In: *Ecology and Society* 22.3. ISSN: 17083087. DOI: 10.5751/ES-09186-220305.

Jacob, Daniela et al. (Apr. 2014). "EURO-CORDEX: new high-resolution climate change projections for European impact research". In: *Regional Environmental Change* 14.2, pp. 563–578. ISSN: 1436-3798. DOI: 10.1007/s10113-013-0499-2. URL: http://link.springer.com/10.1007/s10113-013-0499-2.

James, Gareth et al. (2013). *An Introduction to Statistical Learning*. Vol. 103. Springer Texts in Statistics 9-12. New York, NY: Springer New York. ISBN: 978-1-4614-7137-0. DOI: 10.1007/978-1-4614-7138-7.

Jehmlich, Caroline, Paul Hudson, and Annegret H. Thieken (2020). "Short contribution on adaptive behaviour of flood-prone companies: A pilot study of Dresden-Laubegast, Germany". In: *Journal of Flood Risk Management* 13.4, pp. 1–7. ISSN: 1753318X. DOI: 10.1111/jfr3.12653.

Jensen, V. Finn and Tohmas D. Nielsen (2007). *Bayesian Networks and Decision Graphs*. 2nd ed. Springer-Verlag New York. ISBN: 9780387682815. DOI: 10.1007/978-0-387-68282-2.

Jongman, Brenden (Dec. 2018). "Effective adaptation to rising flood risk". In: *Nature Communications* 9.1, p. 1986. ISSN: 2041-1723. DOI: 10.1038/s41467-018-04396-1. URL: http://www.nature.com/articles/s41467-018-04396-1.

Jongman, Brenden, Philip J. Ward, and Jeroen C.J.H. Aerts (2012). "Global exposure to river and coastal flooding: Long term trends and changes". In: *Global Environmental Change* 22.4, pp. 823–835. ISSN: 09593780. DOI: 10.1016/j.gloenvcha.2012.07.004. URL: http://dx.doi.org/10.1016/j.gloenvcha.2012.07.004.

Jongman, Brenden et al. (May 2015). "Declining vulnerability to river floods and the global benefits of adaptation". In: *Proceedings of the National Academy of*

Sciences 112.18, E2271–E2280. ISSN: 0027-8424. DOI: 10.1073/pnas.1414439112. URL: https://pnas.org/doi/full/10.1073/pnas.1414439112.

Jordan, Alexander, Fabian Krüger, and Sebastian Lerch (2019). "Evaluating Probabilistic Forecasts with scoringRules". In: *Journal of Statistical Software* 90.12. ISSN: 1548-7660. DOI: 10.18637/jss.v090.i12.

Kelman, Ilan and Robin Spence (June 2004). "An overview of flood actions on buildings". In: *Engineering Geology* 73.3-4, pp. 297–309. ISSN: 00137952. DOI: 10.1016/j.enggeo.2004.01.010. URL: https://linkinghub.elsevier.com/retrieve/pii/S0013795204000262.

Khare, S. et al. (June 2015). "Modelling clustering of natural hazard phenomena and the effect on re/insurance loss perspectives". In: *Natural Hazards and Earth System Sciences* 15.6, pp. 1357–1370. ISSN: 1684-9981. DOI: 10.5194/nhess-15-1357-2015. URL: https://nhess.copernicus.org/articles/15/1357/2015/.

Kharratzadeh, Milad (2017). *Splines in Stan*. Computer Software. URL: https://mc-stan.org/users/documentation/case-studies/splines_in_stan.html.

Kienzler, S. et al. (Mar. 10, 2015). "After the extreme flood in 2002: changes in preparedness, response and recovery of flood-affected residents in Germany between 2005 and 2011". In: *Natural Hazards and Earth System Sciences* 15.3, pp. 505–526. ISSN: 1684-9981. DOI: 10.5194/nhess-15-505-2015.

Kinoshita, Youhei et al. (2018). "Quantifying the effect of autonomous adaptation to global river flood projections: Application to future flood risk assessments". In: *Environmental Research Letters* 13.1. ISSN: 17489326. DOI: 10.1088/1748-9326/aa9401.

Koller, Daphne and Nir Friedman (2009). *Probabilistic Graphical Models: Principles and Techniques*. MIT Press. ISBN: 9780262013192. URL: https://mitpress.mit.edu/9780262013192/probabilistic-graphical-models/.

Kreibich, H. and Annegret H. Thieken (2009). "Coping with floods in the city of Dresden, Germany". In: *Natural Hazards* 51.3, pp. 423–436. ISSN: 0921030X. DOI: 10.1007/s11069-007-9200-8.

Kreibich, H. et al. (Jan. 25, 2005). "Flood loss reduction of private households due to building precautionary measures – lessons learned from the Elbe flood in August 2002". In: *Natural Hazards and Earth System Sciences* 5.1, pp. 117–126. ISSN: 1684-9981. DOI: 10.5194/nhess-5-117-2005.

Kreibich, Heidi et al. (Mar. 2007). "Flood precaution of companies and their ability to cope with the flood in August 2002 in Saxony, Germany". In: *Water Resources Research* 43.3, pp. 1–15. ISSN: 00431397. DOI: 10.1029/2005WR004691. URL: http://doi.wiley.com/10.1029/2005WR004691.

Kreibich, Heidi et al. (Nov. 2010). "Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector". In: *Hydrological Sciences Journal* 55.8, pp. 1302–1314. ISSN: 0262-6667. DOI: 10.1080/02626667.

2010.529815. URL: https://www.tandfonline.com/doi/full/10.1080/02626667.2010.529815.

Kreibich, Heidi et al. (May 2014). "Costing natural hazards". In: *Nature Climate Change* 4.5, pp. 303–306. ISSN: 1758-678X. DOI: 10.1038/nclimate2182. URL: http://www.nature.com/articles/nclimate2182.

Kreibich, Heidi et al. (Oct. 2017a). "Adaptation to flood risk: Results of international paired flood event studies". In: *Earth's Future* 5.10, pp. 953–965. ISSN: 23284277. DOI: 10.1002/2017EF000606. URL: http://doi.wiley.com/10.1002/2017EF000606.

Kreibich, Heidi et al. (Apr. 2017b). "Probabilistic, Multivariable Flood Loss Modeling on the Mesoscale with BT-FLEMO". In: *Risk Analysis* 37.4, pp. 774–787. ISSN: 02724332. DOI: 10.1111/risa.12650.

Kreibich, Heidi et al. (Aug. 2022). "The challenge of unprecedented floods and droughts in risk management". In: *Nature* 608.7921, pp. 80–86. ISSN: 0028-0836. DOI: 10.1038/s41586-022-04917-5. URL: https://www.nature.com/articles/s41586-022-04917-5.

Kron, Wolfgang (Mar. 2005). "Flood Risk = Hazard • Values • Vulnerability". In: *Water International* 30.1, pp. 58–68. ISSN: 0250-8060. DOI: 10.1080/02508060508691837. URL: http://www.tandfonline.com/doi/abs/10.1080/02508060508691837.

Krüger, Fabian et al. (Aug. 24, 2016). "Predictive Inference Based on Markov Chain Monte Carlo Output". In: pp. 23–24. DOI: 10.48550/arXiv.1608.06802. URL: https://arxiv.org/abs/1608.06802.

Kruschke, John K. (2015). *Doing Bayesian Data Analysis - A Tutorial with R, JAGS, and Stan*. 2nd ed. Elsevier. ISBN: 978-0-12-405888-0. URL: https://www.sciencedirect.com/book/9780124058880/doing-bayesian-data-analysis.

— (June 2018). "Rejecting or Accepting Parameter Values in Bayesian Estimation". In: *Advances in Methods and Practices in Psychological Science* 1.2, pp. 270–280. ISSN: 2515-2459. DOI: 10.1177/2515245918771304. URL: http://journals.sagepub.com/doi/10.1177/2515245918771304.

Kundzewicz, Zbigniew W. et al. (Jan. 2014). "Flood risk and climate change: global and regional perspectives". In: *Hydrological Sciences Journal* 59.1, pp. 1–28. ISSN: 0262-6667. DOI: 10.1080/02626667.2013.857411. URL: http://dx.doi.org/10.1080/02626667.2013.857411.

Kunreuther, Howard et al. (May 2013). "Risk management and climate change". In: *Nature Climate Change* 3.5, pp. 447–450. ISSN: 1758-678X. DOI: 10.1038/nclimate1740. URL: http://www.nature.com/articles/nclimate1740.

Kwakkel, Jan H., Marjolijn Haasnoot, and Warren E. Walker (Oct. 2015). "Developing dynamic adaptive policy pathways: a computer-assisted approach for developing adaptive strategies for a deeply uncertain world". In: *Climatic*

Change 132.3, pp. 373–386. ISSN: 0165-0009. DOI: 10.1007/s10584-014-1210-4. URL: http://link.springer.com/10.1007/s10584-014-1210-4.

Li, C. and G. Coates (Apr. 1, 2016). “Design and development of an agent-based model for business operations faced with flood disruption”. In: *International Journal of Design & Nature and Ecodynamics* 11.2, pp. 97–106. ISSN: 1755-7437. DOI: 10.2495/DNE-V11-N2-97-106.

Liaw, Andy and Matthew Wiener (2002). “Classification and Regression by randomForest”. In: *R News* 2.3, pp. 18–22. ISSN: 1609-3631. URL: https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf.

Liu, Jianguo et al. (2007). “Complexity of coupled human and natural systems”. In: *Science* 317.5844, pp. 1513–1516. ISSN: 00368075. DOI: 10.1126/science.1144004.

Lüdtke, Stefan et al. (Dec. 2019). “A Consistent Approach for Probabilistic Residential Flood Loss Modeling in Europe”. In: *Water Resources Research*, 2019WR026213. ISSN: 0043-1397. DOI: 10.1029/2019WR026213. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2019WR026213.

Maier, Marco J (2014). *DirichletReg: Dirichlet Regression for Compositional Data in R*. Tech. rep. 125. URL: https://research.wu.ac.at/en/publications/ad3142d3-2fcd-4c37-aec6-8e0bd7d077e1.

Makowski, Dominique, Mattan Ben-Shachar, and Daniel Lüdecke (Aug. 2019a). “bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework”. In: *Journal of Open Source Software* 4.40, p. 1541. ISSN: 2475-9066. DOI: 10.21105/joss.01541. URL: https://joss.theoj.org/papers/10.21105/joss.01541.

Makowski, Dominique et al. (Dec. 2019b). “Indices of Effect Existence and Significance in the Bayesian Framework”. In: *Frontiers in Psychology* 10.December, pp. 1–14. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2019.02767. URL: https://www.frontiersin.org/article/10.3389/fpsyg.2019.02767/full.

Marks, Danny (2019). “Assembling the 2011 Thailand floods: Protecting farmers and inundating high-value industrial estates in a fragmented hydro-social territory”. In: *Political Geography* 68.August 2017, pp. 66–76. ISSN: 09626298. DOI: 10.1016/j.polgeo.2018.10.002.

Matheson, James E. and Robert L. Winkler (June 1976). “Scoring Rules for Continuous Probability Distributions”. In: *Management Science* 22.10, pp. 1087–1096. ISSN: 0025-1909. DOI: 10.1287/mnsc.22.10.1087.

McElreath, Richard (Jan. 3, 2018). *Statistical Rethinking*. Chapman and Hall/CRC. ISBN: 9781315372495. DOI: 10.1201/9781315372495.

Mechler, Reinhard and Laurens M. Bouwer (Nov. 2015). “Understanding trends and projections of disaster losses and climate change: is vulnerability the missing link?” In: *Climatic Change* 133.1, pp. 23–35. ISSN: 0165-0009. DOI: 10.

1007/s10584-014-1141-0. URL: http://link.springer.com/10.1007/s10584-014-1141-0.

Mechler, Reinhard and Jürgen Weichselgartner (2003). *Disaster Loss Financing in Germany – The Case of the Elbe River Floods 2002*. Tech. rep. IR-03-021. URL: https://pure.iiasa.ac.at/id/eprint/7060/1/IR-03-021.pdf.

Meinshausen, Nicolai (Sept. 2006). "Quantile Regression Forests". In: *Journal of Machine Learning Research* 7, pp. 983–999. URL: http://jmlr.org/papers/v7/meinshausen06a.html.

Mentaschi, Lorenzo et al. (Jan. 2020). "Independence of Future Changes of River Runoff in Europe from the Pathway to Global Warming". In: *Climate* 8.2, p. 22. ISSN: 2225-1154. DOI: 10.3390/cli8020022. URL: https://www.mdpi.com/2225-1154/8/2/22.

Merz, B., H. Kreibich, and U. Lall (Jan. 2013). "Multi-variate flood damage assessment: a tree-based data-mining approach". In: *Natural Hazards and Earth System Sciences* 13.1, pp. 53–64. ISSN: 1684-9981. DOI: 10.5194/nhess-13-53-2013. URL: https://www.nat-hazards-earth-syst-sci.net/13/53/2013/.

Merz, B. and Annegret H. Thieken (2009). "Flood risk curves and uncertainty bounds". In: *Natural Hazards* 51.3, pp. 437–458. ISSN: 0921030X. DOI: 10.1007/s11069-009-9452-6.

Merz, B. et al. (Mar. 9, 2004). "Estimation uncertainty of direct monetary flood damage to buildings". In: *Natural Hazards and Earth System Science* 4.1, pp. 153–163. ISSN: 1684-9981. DOI: 10.5194/nhess-4-153-2004.

Merz, B. et al. (Mar. 2010a). "Fluvial flood risk management in a changing world". In: *Natural Hazards and Earth System Sciences* 10.3, pp. 509–527. ISSN: 1684-9981. DOI: 10.5194/nhess-10-509-2010. URL: https://nhess.copernicus.org/articles/10/509/2010/.

Merz, B. et al. (Aug. 2010b). "Review article "Assessment of economic flood damage"". In: *Natural Hazards and Earth System Science* 10.8, pp. 1697–1724. ISSN: 1684-9981. DOI: 10.5194/nhess-10-1697-2010. URL: http://www.nat-hazards-earth-syst-sci.net/10/1697/2010/.

Merz, B. et al. (July 2014a). "Floods and climate: emerging perspectives for flood risk assessment and management". In: *Natural Hazards and Earth System Sciences* 14.7, pp. 1921–1942. ISSN: 1684-9981. DOI: 10.5194/nhess-14-1921-2014. URL: https://nhess.copernicus.org/articles/14/1921/2014/.

Merz, B. et al. (Feb. 24, 2014b). "The extreme flood in June 2013 in Germany". In: *La Houille Blanche* 1.1, pp. 5–10. ISSN: 0018-6368. DOI: 10.1051/lhb/2014001.

Merz, B. et al. (Aug. 2015). "Charting unknown waters-On the role of surprise in flood risk assessment and management". In: *Water Resources Research* 51.8, pp. 6399–6416. ISSN: 00431397. DOI: 10.1002/2015WR017464. URL: http://doi.wiley.com/10.1002/2015WR017464.

Merz, Bruno et al. (2021). “Causes, impacts and patterns of disastrous river floods”. In: *Nature Reviews Earth & Environment* 0123456789. DOI: 10.1038/s43017-021-00195-3.

Metin, Ayse Duha et al. (2018). “How do changes along the risk chain affect flood risk?” In: *Natural Hazards and Earth System Sciences* 18.11, pp. 3089–3108. ISSN: 16849981. DOI: 10.5194/nhess-18-3089-2018.

Metin, Ayse Duha et al. (Apr. 2020). “The role of spatial dependence for large-scale flood risk estimation”. In: *Natural Hazards and Earth System Sciences* 20.4, pp. 967–979. ISSN: 1684-9981. DOI: 10.5194/nhess-20-967-2020. URL: https://nhess.copernicus.org/articles/20/967/2020/.

Meyer, V. et al. (May 2013). “Review article: Assessing the costs of natural hazards – state of the art and knowledge gaps”. In: *Natural Hazards and Earth System Sciences* 13.5, pp. 1351–1373. ISSN: 1684-9981. DOI: 10.5194/nhess-13-1351-2013. URL: https://nhess.copernicus.org/articles/13/1351/2013/.

Middelmann-Fernandes, M.H. (Mar. 2010). “Flood damage estimation beyond stage-damage functions: an Australian example”. In: *Journal of Flood Risk Management* 3.1, pp. 88–96. ISSN: 1753318X. DOI: 10.1111/j.1753-318X.2009.01058.x. URL: http://doi.wiley.com/10.1111/j.1753-318X.2009.01058.x.

Mohor, G. S., P. Hudson, and A. H. Thieken (Apr. 2020). “A Comparison of Factors Driving Flood Losses in Households Affected by Different Flood Types”. In: *Water Resources Research* 56.4. ISSN: 0043-1397. DOI: 10.1029/2019WR025943. URL: https://onlinelibrary.wiley.com/doi/10.1029/2019WR025943.

Mohor, Guilherme S., Annegret H. Thieken, and Oliver Korup (May 2021). “Residential flood loss estimated from Bayesian multilevel models”. In: *Natural Hazards and Earth System Sciences* 21.5, pp. 1599–1614. ISSN: 1684-9981. DOI: 10.5194/nhess-21-1599-2021. URL: https://nhess.copernicus.org/articles/21/1599/2021/.

Molinari, D. et al. (2014). “Ex post damage assessment: An Italian experience”. In: *Natural Hazards and Earth System Sciences* 14.4, pp. 901–916. DOI: 10.5194/nhess-14-901-2014.

Molinari, Daniela et al. (Nov. 2020). “Are flood damage models converging to “reality”? Lessons learnt from a blind test”. In: *Natural Hazards and Earth System Sciences* 20.11, pp. 2997–3017. ISSN: 1684-9981. DOI: 10.5194/nhess-20-2997-2020. URL: https://nhess.copernicus.org/articles/20/2997/2020/.

Montz, Burrell E. and Graham A. Tobin (Aug. 2008). “Livin’ Large with Levees: Lessons Learned and Lost”. In: *Natural Hazards Review* 9.3, pp. 150–157. ISSN: 1527-6988. DOI: 10.1061/(ASCE)1527-6988(2008)9:3(150).

Munich Re (2018). *A Stormy Year: TOPICS Geo Natural Catastrophes 2017*. Tech. rep. Minich, Germany: Münchener Rückversicherungs-Gesellschaft. (Visited on 09/28/2022).

— (2022a). *Factsheet natural catastrophes in 2021*. URL: https://www.munichre.com/content/dam/munichre/mrwebsiteslaunches/natcat-2022/2021_Figures-of-the-year.pdf/_jcr_content/renditions/original./2021_Figures-of-the-year.pdf (visited on 09/27/2022).

— (2022b). *Flood risks on the rise - Underestimated natural hazard, devastating damage*. URL: https://www.munichre.com/en/risks/natural-disasters-losses-are-trending-upwards/floods-and-flash-floods-underestimated-natural-hazards.html (visited on 09/27/2022).

— (2022c). *Hurricanes, cold waves, tornadoes: Weather disasters in USA dominate natural disaster losses in 2021*. Media Release. URL: https://www.munichre.com/content/dam/munichre/mrwebsiteslaunches/natcat-2022/20220110-nat-cat-2021-EN.pdf/_jcr_content/renditions/original./20220110-nat-cat-2021-EN.pdf (visited on 09/27/2022).

Nagarajan, Radhakrishnan, Marco Scutari, and Sophie Lèbre (2013). *Bayesian Networks in R*. New York, NY: Springer New York. ISBN: 978-1-4614-6445-7. DOI: 10.1007/978-1-4614-6446-4.

Neumayer, Eric and Fabian Barthel (Feb. 2011). "Normalizing economic loss from natural disasters: A global analysis". In: *Global Environmental Change* 21.1, pp. 13–24. ISSN: 09593780. DOI: 10.1016/j.gloenvcha.2010.10.004. URL: http://dx.doi.org/10.1016/j.gloenvcha.2010.10.004.

Nuremberg Institute of Technology (2019). *Ermittlung potentieller Überschwemmungsgebiete der Elbe im Stadtgebiet von Dresden mittels aktualisiertem 2D-Modell Elbe (Strom-km 30,0 bis 85,0), Forschungsbericht 2017-08*. Tech. rep. Nuremberg, Germany. URL: https://www.dresden.de/media/pdf/umwelt/ua_4_38_text.pdf.

O'Connell, P. E. and G. O'Donnell (Jan. 14, 2014). "Towards modelling flood protection investment as a coupled human and natural system". In: *Hydrology and Earth System Sciences* 18.1, pp. 155–171. ISSN: 1607-7938. DOI: 10.5194/hess-18-155-2014.

Olsen, Anders et al. (Jan. 12, 2015). "Comparing Methods of Calculating Expected Annual Damage in Urban Pluvial Flood Risk Assessments". In: *Water* 7.12, pp. 255–270. ISSN: 2073-4441. DOI: 10.3390/w7010255.

Ospina, Raydonal and Silvia L. P. Ferrari (Jan. 11, 2010). "Inflated beta distributions". In: *Statistical Papers* 51.1, pp. 111–126. ISSN: 0932-5026. DOI: 10.1007/s00362-008-0125-4.

Pappenberger, F. and K. J. Beven (May 2006). "Ignorance is bliss: Or seven reasons not to use uncertainty analysis". In: *Water Resources Research* 42.5, pp. 1–8. ISSN: 00431397. DOI: 10.1029/2005WR004820. URL: http://doi.wiley.com/10.1029/2005WR004820.

Paprotny, D. (2022). *HANZE v2.0 exposure model input data*. Dataset. DOI: 10.5281/zenodo.6783023.

Paprotny, D. and M. Mengel (2022). " Reconstruction of past exposure to natural hazards driven by historical statistics: HANZE v2.0s". In: *Geoscientific Model Development*. in review. DOI: https://doi.org/10.5194/gmd-2022-194.

Paprotny, Dominik, Oswaldo Morales-Nápoles, and Sebastiaan N Jonkman (Mar. 2018a). "HANZE: a pan-European database of exposure to natural hazards and damaging historical floods since 1870". In: *Earth System Science Data* 10.1, pp. 565–581. ISSN: 1866-3516. DOI: 10.5194/essd-10-565-2018. URL: https://essd.copernicus.org/articles/10/565/2018/.

Paprotny, Dominik et al. (Dec. 2018b). "Trends in flood losses in Europe over the past 150 years". In: *Nature Communications* 9.1, p. 1985. ISSN: 2041-1723. DOI: 10.1038/s41467-018-04253-1. URL: http://dx.doi.org/10.1038/s41467-018-04253-1.

Paprotny, Dominik et al. (Oct. 2020). "Exposure and vulnerability estimation for modelling flood losses to commercial assets in Europe". In: *Science of The Total Environment* 737, p. 140011. ISSN: 00489697. DOI: 10.1016/j.scitotenv.2020.140011. URL: https://linkinghub.elsevier.com/retrieve/pii/S0048969720335312.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge University Press. ISBN: 9780521895606.

Penning-Rowsell, Edmund et al. (2005). *The Benefits of Flood and Coastal Risk Management: A Handbook of Assessment Techniques*. Middlesex University Press. ISBN: 1 904750 51 6. URL: https://repository.tudelft.nl/islandora/object/uuid:33f2d216-c9bf-419c-b3b1-415a6f6fd881?collection=research.

Pisarenko, V. and M. Rodkin (2010). *Heavy-Tailed Distributions in Disaster Analysis*. Vol. 30. Advances in Natural and Technological Hazards Research 9. Dordrecht: Springer Netherlands. ISBN: 978-90-481-9170-3. DOI: 10.1007/978-90-481-9171-0.

Pittore, Massimiliano, Marc Wieland, and Kevin Fleming (Mar. 2017). "Perspectives on global dynamic exposure modelling for geo-risk assessment". In: *Natural Hazards* 86.S1, pp. 7–30. ISSN: 0921-030X. DOI: 10.1007/s11069-016-2437-3. URL: http://link.springer.com/10.1007/s11069-016-2437-3.

Pohl, Reinhard (2004). "Historische Hochwasser aus dem Erzgebirge: von der Gottleuba bis zur Mulde." PhD thesis. Dresden. ISBN: 3860054287.

Poussin, Jennifer K., W.J. Wouter Botzen, and Jeroen C.J.H. Aerts (Mar. 2015). "Effectiveness of flood damage mitigation measures: Empirical evidence from French flood disasters". In: *Global Environmental Change* 31, pp. 74–84. ISSN: 09593780. DOI: 10.1016/j.gloenvcha.2014.12.007. URL: http://dx.doi.org/10.1016/j.gloenvcha.2014.12.007.

Prettenthaler, Franz et al. (Aug. 2017). “On flood risk pooling in Europe”. In: *Natural Hazards* 88.1, pp. 1–20. ISSN: 0921-030X. DOI: 10.1007/s11069-016-2616-2. URL: http://link.springer.com/10.1007/s11069-016-2616-2.

Priestley, Matthew D. K. et al. (Nov. 2018). “The role of serial European windstorm clustering for extreme seasonal losses as determined from multi-centennial simulations of high-resolution global climate model data”. In: *Natural Hazards and Earth System Sciences* 18.11, pp. 2991–3006. ISSN: 1684-9981. DOI: 10.5194/nhess-18-2991-2018. URL: https://nhess.copernicus.org/articles/18/2991/2018/.

Ridolfi, E., F. Albrecht, and Giuliano Di Baldassarre (Jan. 2020). “Exploring the role of risk perception in influencing flood losses over time”. In: *Hydrological Sciences Journal* 65.1, pp. 12–20. ISSN: 0262-6667. DOI: 10.1080/02626667.2019.1677907. URL: https://www.tandfonline.com/doi/full/10.1080/02626667.2019.1677907.

Ridolfi, Elena, Elena Mondino, and Giuliano Di Baldassarre (Feb. 2021). “Hydrological risk: modeling flood memory and human proximity to rivers”. In: *Hydrology Research* 52.1, pp. 241–252. ISSN: 0029-1277. DOI: 10.2166/nh.2020.195. URL: https://iwaponline.com/hr/article/52/1/241/77667/Hydrological-risk-modeling-flood-memory-and-human.

Rosina, Konštantín et al. (2020). “Increasing the detail of European land use/cover data by combining heterogeneous data sets”. In: *International Journal of Digital Earth* 13.5, pp. 602–626. ISSN: 17538955. DOI: 10.1080/17538947.2018.1550119.

Ross, Alexander and Heejun Chang (July 2020). “Socio-hydrology with hydrosocial theory: two sides of the same coin?” In: *Hydrological Sciences Journal* 65.9, pp. 1443–1457. ISSN: 0262-6667. DOI: 10.1080/02626667.2020.1761023. URL: https://doi.org/10.1080/02626667.2020.1761023.

Rözer, Viktor et al. (Apr. 2019). “Probabilistic Models Significantly Reduce Uncertainty in Hurricane Harvey Pluvial Flood Loss Estimates”. In: *Earth's Future* 7.4, pp. 384–394. ISSN: 2328-4277. DOI: 10.1029/2018EF001074. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2018EF001074.

Sairam, Nivedita et al. (Oct. 2019a). “Hierarchical Bayesian Approach for Modeling Spatiotemporal Variability in Flood Damage Processes”. In: *Water Resources Research* 55.10, pp. 8223–8237. ISSN: 0043-1397. DOI: 10.1029/2019WR025068. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2019WR025068.

Sairam, Nivedita et al. (Mar. 2019b). “Quantifying Flood Vulnerability Reduction via Private Precaution”. In: *Earth's Future* 7.3, pp. 235–249. ISSN: 2328-4277. DOI: 10.1029/2018EF000994. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2018EF000994.

Sairam, Nivedita et al. (Oct. 2020). “Bayesian Data-Driven approach enhances synthetic flood loss models”. In: *Environmental Modelling & Software* 132.July,

p. 104798. ISSN: 13648152. DOI: 10.1016/j.envsoft.2020.104798. URL: https://linkinghub.elsevier.com/retrieve/pii/S136481522030178X.

Sairam, Nivedita et al. (Oct. 2021). "Process-Based Flood Risk Assessment for Germany". In: *Earth's Future* 9.10, pp. 1–12. ISSN: 2328-4277. DOI: 10.1029/2021EF002259. URL: https://onlinelibrary.wiley.com/doi/10.1029/2021EF002259.

Saxonian Environmental Agency (2012). *Flood Hazard Maps for Saxony [Dataset]*. URL: https://www.wasser.sachsen.de/hochwassergefahrenkarte-11915.html#a-12452.

Saxonian Relief Bank (2007). *SAB-Förderbericht 2006 Wirtschaft, Technologie, Arbeit*. Tech. rep. Dresden.

Sayers, Paul B., Gerald E. Galloway, and Jim W. Hall (Jan. 2012). "Robust decision-making under uncertainty – towards adaptive and resilient flood risk management infrastructure". In: *Flood Risk*. ICE Publishing, pp. 281–302. ISBN: 9780727741561. DOI: 10.1680/fr.41561.281. URL: https://www.icevirtuallibrary.com/doi/10.1680/fr.41561.281.

Scawthorn, Charles et al. (May 2006). "HAZUS-MH Flood Loss Estimation Methodology. II. Damage and Loss Assessment". In: *Natural Hazards Review* 7.2, pp. 72–81. ISSN: 1527-6988. DOI: 10.1061/(ASCE)1527-6988(2006)7:2(72).

Schoot, Rens van de et al. (Dec. 2021). "Bayesian statistics and modelling". In: *Nature Reviews Methods Primers* 1.1, p. 1. ISSN: 2662-8449. DOI: 10.1038/s43586-020-00001-2. URL: http://www.nature.com/articles/s43586-020-00001-2.

Schoppa, Lukas et al. (Sept. 11, 2020). "Probabilistic Flood Loss Models for Companies". In: *Water Resources Research* 56.9. ISSN: 0043-1397. DOI: 10.1029/2020WR027649.

Schoppa, Lukas et al. (Aug. 2022). "Augmenting a socio-hydrological flood risk model for companies with process-oriented loss estimation". In: *Hydrological Sciences Journal* 00.00, pp. 1–17. ISSN: 0262-6667. DOI: 10.1080/02626667.2022.2095207. URL: https://www.tandfonline.com/doi/full/10.1080/02626667.2022.2095207.

Schröter, Kai et al. (Apr. 2014). "How useful are complex flood damage models?" In: *Water Resources Research* 50.4, pp. 3378–3395. ISSN: 00431397. DOI: 10.1002/2013WR014396. URL: http://doi.wiley.com/10.1002/2013WR014396.

Schröter, Kai et al. (Dec. 2021). "Large-scale flood risk assessment and management: Prospects of a systems approach". In: *Water Security* 14, p. 100109. ISSN: 24683124. DOI: 10.1016/j.wasec.2021.100109. URL: https://linkinghub.elsevier.com/retrieve/pii/S2468312421000250.

Scutari, Marco (2010). "Learning Bayesian Networks with the bnlearn R Package". In: *Journal of Statistical Software* 35.3. ISSN: 1548-7660. DOI: 10.18637/jss.v035.i03.

Scutari, Marco and Jean-Baptiste Denis (June 20, 2014). *Bayesian Networks With Examples in R*. Boca Raton: Chapman and Hall/CRC. ISBN: 9781482225594. DOI: 10.1201/b17065.

Seifert, Isabel et al. (Nov. 29, 2010). "Application and validation of FLEMOcs – a flood-loss estimation model for the commercial sector". In: *Hydrological Sciences Journal* 55.8, pp. 1315–1324. ISSN: 0262-6667. DOI: 10.1080/02626667.2010.536440.

Sennhenn-Reulen, Holger (Aug. 20, 2018). "Bayesian Regression for a Dirichlet Distributed Response using Stan". In: *arXiv* 1, pp. 1–13.

Sieg, Tobias (2018). "Reliability of flood damage estimations across spatial scales". doctoralthesis. Universität Potsdam, pp. XIII, 115. DOI: 10.25932/publishup-42616.

Sieg, Tobias et al. (July 2017). "Tree-based flood damage modeling of companies: Damage processes and model performance". In: *Water Resources Research* 53.7, pp. 6050–6068. ISSN: 00431397. DOI: 10.1002/2017WR020784. URL: http://doi.wiley.com/10.1002/2017WR020784.

Sieg, Tobias et al. (Apr. 2019a). "Integrated assessment of short-term direct and indirect economic flood impacts including uncertainty quantification". In: *PLOS ONE* 14.4. Ed. by Guy J-P. Schumann, e0212932. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0212932. URL: https://dx.plos.org/10.1371/journal.pone.0212932.

Sieg, Tobias et al. (May 2019b). "Seamless Estimation of Hydrometeorological Risk Across Spatial Scales". In: *Earth's Future* 7.5, pp. 574–581. ISSN: 2328-4277. DOI: 10.1029/2018EF001122. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2018EF001122.

Silverman, B.W. (Feb. 2018). *Density Estimation for Statistics and Data Analysis*. 1st ed. New York: Routledge. ISBN: 9781315140919. DOI: 10.1201/9781315140919. URL: https://www.taylorfrancis.com/books/9781351456173.

Simpson, Daniel et al. (Feb. 2017). "Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors". In: *Statistical Science* 32.1, pp. 1–28. ISSN: 0883-4237. DOI: 10.1214/16-STS576. arXiv: 1403.4630. URL: https://projecteuclid.org/journals/statistical-science/volume-32/issue-1/Penalising-Model-Component-Complexity--A-Principled-Practical-Approach-to/10.1214/16-STS576.full.

Simpson, Gavin (2020). *Extrapolating with B splines and GAMs*. Computer Software. URL: https://fromthebottomoftheheap.net/2020/06/03/extrapolating-with-gams/.

Sivapalan, M. et al. (Apr. 2014). "Socio-hydrology: Use-inspired water sustainability science for the Anthropocene". In: *Earth's Future* 2.4, pp. 225–230. ISSN:

2328-4277. DOI: 10.1002/2013EF000164. URL: https://onlinelibrary.wiley.com/doi/10.1002/2013EF000164.

Sivapalan, Murugesu and Günter Blöschl (Sept. 2015). "Time scale interactions and the coevolution of humans and water". In: *Water Resources Research* 51.9, pp. 6988–7022. ISSN: 00431397. DOI: 10.1002/2015WR017896. URL: http://doi.wiley.com/10.1002/2015WR017896.

Sivapalan, Murugesu, Hubert H. G. Savenije, and Günter Blöschl (Apr. 2012). "Socio-hydrology: A new science of people and water". In: *Hydrological Processes* 26.8, pp. 1270–1276. ISSN: 08856087. DOI: 10.1002/hyp.8426. URL: https://onlinelibrary.wiley.com/doi/10.1002/hyp.8426.

Smid, Sanne C. et al. (Jan. 2020). "Bayesian Versus Frequentist Estimation for Structural Equation Models in Small Sample Contexts: A Systematic Review". In: *Structural Equation Modeling: A Multidisciplinary Journal* 27.1, pp. 131–161. ISSN: 1070-5511. DOI: 10.1080/10705511.2019.1577140. URL: https://www.tandfonline.com/doi/full/10.1080/10705511.2019.1577140.

Song, Shuang et al. (Mar. 2021). "Improving representation of collective memory in socio-hydrological models and new insights into flood risk management". In: *Journal of Flood Risk Management* 14.1, pp. 1–12. ISSN: 1753-318X. DOI: 10.1111/jfr3.12679. URL: https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12679.

Spekkers, M. H. et al. (Sept. 2014). "Decision-tree analysis of factors influencing rainfall-related building structure and content damage". In: *Natural Hazards and Earth System Sciences* 14.9, pp. 2531–2547. ISSN: 1684-9981. DOI: 10.5194/nhess-14-2531-2014. URL: https://nhess.copernicus.org/articles/14/2531/2014/.

Srinivasan, V. et al. (2017). "Prediction in a socio-hydrological world". In: *Hydrological Sciences Journal* 62.3, pp. 338–345. ISSN: 21503435. DOI: 10.1080/02626667.2016.1253844.

Steinhausen, Max (2022). "Probabilistic flood loss estimation for residential buildings in Europe". Doctoral Thesis. Humboldt-Universität zu Berlin. DOI: 10.18452/24315.

Steinhausen, Max et al. (2022). "Drivers of future fluvial flood risk change for residential buildings in Europe". In: *Global Environmental Change* 76, p. 102559. ISSN: 0959-3780. DOI: https://doi.org/10.1016/j.gloenvcha.2022.102559. URL: https://www.sciencedirect.com/science/article/pii/S0959378022000978.

Sultana, Zakia et al. (Aug. 7, 2018). "Assessment of Business Interruption of Flood-Affected Companies Using Random Forests". In: *Water* 10.8, p. 1049. ISSN: 2073-4441. DOI: 10.3390/w10081049.

Tanoue, Masahiro, Yukiko Hirabayashi, and Hiroaki Ikeuchi (Dec. 2016). "Global-scale river flood vulnerability in the last 50 years". In: *Scientific Reports* 6.1, p. 36021. ISSN: 2045-2322. DOI: 10.1038/srep36021. URL: http://www.nature.com/articles/srep36021.

Tanoue, Masahiro et al. (2021). "Residual flood damage under intensive adaptation". In: *Nature Climate Change* 11.10, pp. 823–826. ISSN: 17586798. DOI: 10.1038/s41558-021-01158-8. URL: http://dx.doi.org/10.1038/s41558-021-01158-8.

Thieken, Annegret H. et al. (Dec. 2005). "Flood damage and influencing factors: New insights from the August 2002 flood in Germany". In: *Water Resources Research* 41.12, pp. 1–16. ISSN: 00431397. DOI: 10.1029/2005WR004177. URL: http://doi.wiley.com/10.1029/2005WR004177.

Thieken, Annegret H. et al. (2007). "Coping with floods: Preparedness, response and recovery of flood-affected residents in Germany in 2002". In: *Hydrological Sciences Journal* 52.5, pp. 1016–1037. ISSN: 02626667. DOI: 10.1623/hysj.52.5.1016.

Thieken, Annegret H. et al. (June 2008). "Development and evaluation of FLEMOps – a new F lood L oss E stimation MO del for the p rivate s ector". In: *Flood Recovery, Innovation and Response I*. Vol. I. WIT Transactions on Ecology and the Environment. Southampton, UK: WIT Press, pp. 315–324. ISBN: 9781845641320. DOI: 10.2495/FRIAR080301. URL: http://library.witpress.com/viewpaper.asp?pcode=FRIAR08-030-1.

Thieken, Annegret H. et al. (July 2016). "The flood of June 2013 in Germany: how much do we know about its impacts?" In: *Natural Hazards and Earth System Sciences* 16.6, pp. 1519–1540. ISSN: 1684-9981. DOI: 10.5194/nhess-16-1519-2016. URL: https://www.nat-hazards-earth-syst-sci.net/16/1519/2016/.

Thompson, S. E. et al. (Dec. 2013). "Developing predictive insight into changing water systems: use-inspired hydrologic science for the Anthropocene". In: *Hydrology and Earth System Sciences* 17.12, pp. 5013–5039. ISSN: 1607-7938. DOI: 10.5194/hess-17-5013-2013. URL: https://hess.copernicus.org/articles/17/5013/2013/.

Troy, T. J. et al. (Aug. 2015a). "Moving sociohydrology forward: a synthesis across studies". In: *Hydrology and Earth System Sciences* 19.8, pp. 3667–3679. ISSN: 1607-7938. DOI: 10.5194/hess-19-3667-2015. URL: https://hess.copernicus.org/articles/19/3667/2015/.

Troy, Tara J., Mitchell Pavao-Zuckerman, and Tom P. Evans (June 2015b). "Debates-Perspectives on socio-hydrology: Socio-hydrologic modeling: Tradeoffs, hypothesis testing, and validation". In: *Water Resources Research* 51.6, pp. 4806–4814. ISSN: 00431397. DOI: 10.1002/2015WR017046. URL: http://doi.wiley.com/10.1002/2015WR017046.

Ulbrich, Uwe et al. (Oct. 2003). “The central European floods of August 2002: Part 1 – Rainfall periods and flood development”. In: *Weather* 58.10, pp. 371–377. ISSN: 00431656. DOI: 10.1256/wea.61.03A.

UNDESA (2019). *World Population Prospects 2019*. Dataset. URL: https://population.un.org/wpp/Download/Archive/Standard/.

UNDRR (2022). *Global Assessment Report on Disaster Risk Reduction 2022*. Tech. rep. Geneva. URL: www.undrr.org/GAR2022.

Van Der Knijff, J. M., J. Younis, and A. P. J. De Roo (Feb. 2010). “LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation”. In: *International Journal of Geographical Information Science* 24.2, pp. 189–212. ISSN: 1365-8816. DOI: 10.1080/13658810802549154. URL: http://www.tandfonline.com/doi/abs/10.1080/13658810802549154.

Van Ootegem, Luc et al. (Sept. 2015). “Multivariate pluvial flood damage models”. In: *Environmental Impact Assessment Review* 54, pp. 91–100. ISSN: 01959255. DOI: 10.1016/j.eiar.2015.05.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S0195925515000505.

Viglione, Alberto et al. (2014). “Insights from socio-hydrology modelling on dealing with flood risk - Roles of collective memory, risk-taking attitude and trust”. In: *Journal of Hydrology* 518.PA, pp. 71–82. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2014.01.018. URL: http://dx.doi.org/10.1016/j.jhydrol.2014.01.018.

Visser, Hans, Arthur C. Petersen, and Willem Ligtvoet (Aug. 2014). “On the relation between weather-related disaster impacts, vulnerability and climate change”. In: *Climatic Change* 125.3-4, pp. 461–477. ISSN: 0165-0009. DOI: 10.1007/s10584-014-1179-z. URL: http://link.springer.com/10.1007/s10584-014-1179-z.

Vogel, Kristin et al. (2012). “Flood damage and influencing factors: A Bayesian network perspective”. In: *Proceedings of the 6th European Workshop on Probabilistic Graphical Models, PGM 2012*, pp. 347–354. URL: https://leo.ugr.es/pgm2012/proceedings/eproceedings/vogel_flood.pdf.

Vogel, Kristin et al. (Jan. 2014). “Challenges for Bayesian network learning in a flood damage assessment application”. In: *Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infrastructures*. CRC Press, pp. 3123–3130. ISBN: 9781138000865. DOI: 10.1201/b16387-452. URL: http://www.crcpress.com/product/isbn/9781138000865.

Vogel, Kristin et al. (2018). “Identifying Driving Factors in Flood-Damaging Processes Using Graphical Models”. In: *Water Resources Research* 54.11, pp. 8864–8889. ISSN: 19447973. DOI: 10.1029/2018WR022858.

Vorogushyn, Sergiy et al. (May 2012). “Analysis of a detention basin impact on dike failure probabilities and flood risk for a channel-dike-floodplain system along the river Elbe, Germany”. In: *Journal of Hydrology* 436-437, pp. 120–

131. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2012.03.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022169412001928.

Wagenaar, D. J. et al. (Jan. 2016). "Uncertainty in flood damage estimates and its potential effect on investment decisions". In: *Natural Hazards and Earth System Sciences* 16.1, pp. 1–14. ISSN: 1684-9981. DOI: 10.5194/nhess-16-1-2016. URL: https://nhess.copernicus.org/articles/16/1/2016/.

Wagenaar, Dennis, Jurjen de Jong, and Laurens M. Bouwer (Sept. 2017). "Multi-variable flood damage modelling with limited data using supervised learning approaches". In: *Natural Hazards and Earth System Sciences* 17.9, pp. 1683–1696. ISSN: 1684-9981. DOI: 10.5194/nhess-17-1683-2017. URL: https://nhess.copernicus.org/articles/17/1683/2017/.

Wagenaar, Dennis et al. (May 2018). "Regional and Temporal Transferability of Multivariable Flood Damage Models". In: *Water Resources Research* 54.5, pp. 3688–3703. ISSN: 0043-1397. DOI: 10.1029/2017WR022233. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2017WR022233.

Ward, Philip J. et al. (Dec. 2013). "Assessing flood risk at the global scale: model setup, results, and sensitivity". In: *Environmental Research Letters* 8.4, p. 044019. ISSN: 1748-9326. DOI: 10.1088/1748-9326/8/4/044019. URL: https://iopscience.iop.org/article/10.1088/1748-9326/8/4/044019.

Ward, Philip J. et al. (Sept. 2017). "A global framework for future costs and benefits of river-flood protection in urban areas". In: *Nature Climate Change* 7.9, pp. 642–646. ISSN: 1758-678X. DOI: 10.1038/nclimate3350. URL: http://www.nature.com/articles/nclimate3350.

Wedawatta, G., B. Ingirige, and D. Proverbs (2014). "Small businesses and flood impacts: The case of the 2009 flood event in Cockermouth". In: *Journal of Flood Risk Management* 7.1, pp. 42–53. ISSN: 1753318X. DOI: 10.1111/jfr3.12031.

Wedawatta, G. and Bingunath Ingirige (2012). "Resilience and adaptation of small and medium-sized enterprises to flood risk". In: *Disaster Prevention and Management: An International Journal* 21.4, pp. 474–488. ISSN: 09653562. DOI: 10.1108/09653561211256170.

Weikinn, Curt (2000). *Quellentexte zur Witterungsgeschichte Europas von der Zeitwende bis zum Jahr 1850 - Hydrographie Band 1, Teil 5 (1751-1800)*. Ed. by Hrsg.: Michael Börngen; Gerd Tetzlaff. Stuttgart, Germany: Schweizerbart Science Publishers. ISBN: 9783443010447.

— (2002). *Quellentexte zur Witterungsgeschichte Europas von der Zeitwende bis zum Jahr 1850 - Hydrographie Band 1, Teil 6 (1801-1850)*. Ed. by Hrsg.: Michael Börngen; Gerd Tetzlaff. Stuttgart, Germany: Schweizerbart Science Publishers. ISBN: 9783443010478.

Wheeler, Russell L. (2009). "Methods of Mmax Estimation East of the Rocky Mountains, Open File Report 2009-1018, USGS". In: *Open-File Report*. ISSN: 2331-1258. DOI: 10.3133/OFR20091018.

White, Allan P and Wei Zhong Liu (June 1994). "Bias in information-based measures in decision tree induction". In: *Machine Learning* 15.3, pp. 321–329. ISSN: 0885-6125. DOI: 10.1007/BF00993349. URL: http://link.springer.com/10.1007/BF00993349.

White, G. F. (1945). "Human Adjustment to Floods". PhD thesis.

Willner, Sven N. et al. (Jan. 2018). "Adaptation required to preserve future high-end river flood risk at present levels". In: *Science Advances* 4.1, pp. 1–9. ISSN: 2375-2548. DOI: 10.1126/sciadv.aao1914. URL: https://www.science.org/doi/10.1126/sciadv.aao1914.

Wing, Oliver E.J. et al. (2020). "New insights into US flood vulnerability revealed from flood insurance big data". In: *Nature Communications* 11.1, pp. 1–10. ISSN: 20411723. DOI: 10.1038/s41467-020-15264-2.

Winsemius, Hessel C. et al. (Apr. 2016). "Global drivers of future river flood risk". In: *Nature Climate Change* 6.4, pp. 381–385. ISSN: 1758-678X. DOI: 10.1038/nclimate2893. URL: http://www.nature.com/articles/nclimate2893.

Winter, B et al. (Mar. 2018). "Sources of uncertainty in a probabilistic flood risk model". In: *Natural Hazards* 91.2, pp. 431–446. ISSN: 0921-030X. DOI: 10.1007/s11069-017-3135-5. URL: http://link.springer.com/10.1007/s11069-017-3135-5.

World Meteorological Organization (2022). *WMO Atlas of Mortality and Economic Losses from Weather, Climate and Water Extremes (1970–2019)*. Tech. rep. WMO-No. 1267. Geneva, p. 90. URL: https://library.wmo.int/index.php?lvl=notice_display&id=21930 (visited on 09/30/2022).

WSV (2021). *Water level data at gauge Dresden*. Dataset.

Yeo, I.-K. (Dec. 1, 2000). "A new family of power transformations to improve normality or symmetry". In: *Biometrika* 87.4, pp. 954–959. ISSN: 0006-3444. DOI: 10.1093/biomet/87.4.954.

Yu, David J. et al. (Feb. 2017). "Incorporating institutions and collective action into a sociohydrological model of flood resilience". In: *Water Resources Research* 53.2, pp. 1336–1353. ISSN: 0043-1397. DOI: 10.1002/2016WR019746. URL: https://onlinelibrary.wiley.com/doi/10.1002/2016WR019746.

Zhai, Guofang, Teruki Fukuzono, and Saburo Ikeda (Feb. 2005). "MODELING FLOOD DAMAGE: CASE OF TOKAI FLOOD 2000". In: *Journal of the American Water Resources Association* 41.1, pp. 77–92. ISSN: 1093-474X. DOI: 10.1111/j.1752-1688.2005.tb03719.x.

Zöller, Gert (2013). "Convergence of the frequency-magnitude distribution of global earthquakes: Maybe in 200 years". In: *Geophysical Research Letters* 40.15, pp. 3873–3877. ISSN: 00948276. DOI: 10.1002/grl.50779.