
Probabilistic flood loss modelling for residential buildings

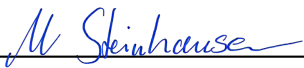
Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
im Fach Geographie
eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin
von
M.Sc. Max Jacob Steinhausen (13.06.1991, Berlin)

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

Declaration of Authorship

„Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 42/2018 am 11.07.2018 angegebenen Hilfsmittel angefertigt habe.“



Max Jacob Steinhausen

London, October 22, 2021

“To know what the future holds, in even the most general and probabilistic way, would serve as a new and marvelous guide for our actions, one that humanity has never before had. ”

— Prelude to Foundation, Isaac Asimov

Summary

Flooding poses great risks for residential buildings in Europe and is expected to increase in the future, driven by climatic and socio-economic change. Current flood risk models rely mostly on simple stage-damage curves for flood loss estimation. Especially continental applications rely on this approach due to low data requirements and established methodology. This approach oversimplifies flood damage processes, can be inaccurate and harbour large uncertainties that often are not quantified and transparently communicated.

This thesis presents research that integrates new data sources into probabilistic, multi-variable loss models to improve their transferability. Tree- and graph-based approaches are implemented to test novel data sources for flood loss estimation and included uncertainty quantification. These new data sources and approaches are used to estimate future fluvial flood risk change for residential buildings in Europe. Contributions of the three risk components, hazard, exposure, and vulnerability are analysed and compared independently and in combination.

OpenStreetMap (OSM) data are identified as a valuable source of information for flood loss modelling and enables model transfers while retaining high predictive performance. Integrating OSM derived building characteristics and flood experience information from flood event databases into the Bayesian Network Flood Loss Estimation Model for the private sector (BN-FLEMOps) enables the spatio-temporal and scale transformation of the model. The model is validated by comparison with reported losses in multiple case studies in Europe. Model updating with local data results in more accurate estimations with lower uncertainty for flood events with sufficient empirical data. A detailed comparison with a model ensemble of 20 internationally published flood loss models reveals that BN-FLEMOps has a higher average accuracy and covers the entire uncertainty range of the ensemble with its probabilistic results. Thus, providing valuable information to decision-makers for the evaluation of flood risk management measures. In a final study, the future flood risk changes for residential buildings in Europe are modelled. Comparisons with an historical baseline period reveal that the expected annual damage will increase up to 10-fold until the end of the 21st century. The high spatial resolution of the analysis shows risk hotspots in urban centres. Most of Central Europe and the British Isles have to expect strong risk increases. Parts of Scandinavia and the Mediterranean on the other hand will see stagnating or decreasing fluvial flood risk. A separate analysis of the risk drivers reveals that exposure increase has a stronger effect on flood risk compared to the impacts of climate change. Furthermore, the study shows that current flood protection levels reduce flood loss on average by 4.7 billion € per year in Europe. Improving private precaution could reduce flood risk by 15 % on average and up to 20 % in some European regions.

Zusammenfassung

Hochwasser stellt ein großes Risiko für Wohngebäude in Europa dar, und es wird erwartet, dass das Risiko in der Zukunft aufgrund klimatischer und sozioökonomischer Veränderungen zunehmen wird. Die derzeitigen Hochwasserrisikomodelle basieren meist auf einfachen Wasserstands-Schadenskurven. Vor allem bei großskaligen Anwendungen wird dieser Ansatz noch immer wegen seiner geringen Datenanforderungen und der etablierten Methodik genutzt. Diese Ansätze vereinfachen die Hochwasserschadensprozesse stark, können ungenau sein und bergen große Unsicherheiten, die oft nicht quantifiziert und transparent kommuniziert werden. Diese Doktorarbeit stellt die Integration neuer Daten in probabilistische, multivariable Schadensmodelle zur Verbesserung deren Übertragbarkeit vor. Entscheidungsbaum und Graphen Ansätze zur Schadensmodellierung werden implementiert, um neue Datenquellen und die in den Ansätzen integrierten Unsicherheitsschätzungen zu testen. Diese neuen Datenquellen und Modellierungsansätze werden verwendet, um zukünftige Veränderung des Hochwasserrisikos für Wohngebäude in Europa abzuschätzen. Die individuellen und kombinierten Effekte der drei Risikokomponenten, Gefährdung, Exposition und Vulnerabilität, werden analysiert.

OpenStreetMap (OSM) Daten liefern nützliche Informationen für die Modellierung von Hochwasserschäden und ermöglichen Modelltransfers ohne starke Leistungsabnahme. Die Integration von aus OSM abgeleiteten Gebäudeeigenschaften und Hochwassererfahrung aus Ereignisdatenbanken in das Bayes'sche Netzwerk basierte Hochwasserschadensmodelle für den privaten Sektor (BN-FLEMOps) ermöglichte die Implementierung des Modells auf der Mesoskala. Durch Vergleiche von Schadensschätzungen mit beobachteten Schäden in mehreren Fallstudien in Europa wurde das Modell validiert. Die Aktualisierung des Modells mit lokalen Daten führt zu genaueren Schätzungen mit geringerer Unsicherheit für Ereignisse mit ausreichender Datengrundlage. Ein detaillierter Vergleich mit einem Modellensemble aus 20 Hochwasserschadensmodelle zeigt, dass BN-FLEMOps eine höhere durchschnittliche Genauigkeit hat und mit seinen probabilistischen Ergebnissen den gesamten Unsicherheitsbereich des Ensembles abbildet. Entscheidungsträgern können somit wertvolle Informationen für die Bewertung möglicher Maßnahmen zum Risikomanagement zur Verfügung gestellt werden. In einer abschließenden Studie werden die zukünftigen Veränderungen des Hochwasserrisikos für Wohngebäude in Europa modelliert. Vergleiche mit einem historischen Zeitraum zeigen, dass die erwarteten jährlichen Schäden bis zum Ende des 21. Jahrhunderts um das 10-fache ansteigen werden. Die hohe räumliche Auflösung der Analyse offenbart, dass urbane Zentren einem besonders hohen Risiko ausgesetzt sein werden. Die Britischen Inseln und der größte Teil von Zentral-Europa müssen mit einer starken Risikozunahme rechnen. Teile Skandinaviens und des Mittelmeerraums werden dagegen ein stagnierendes oder abnehmendes Hochwasserrisiko verzeichnen. Eine separate Analyse der Risikotreiber zeigt, dass die Zunahme der Exposition, das Hochwasserrisiko stärker beeinflusst als die Auswirkungen des Klimawandels. Zudem zeigt die Studie, dass das derzeitige Hochwasserschutzniveau in Europa den jährlich erwarteten Schaden um 4,7 Milliarden € senkt. Eine Verbesserung der privaten Vorsorgemaßnahmen könnte das Hochwasserrisiko im Mittel um 15 % und in einigen europäischen Regionen um bis zu 20 % verringern.

Acknowledgements

During the years of research and writing of this thesis I was fortunate to receive a great deal of assistance and support from colleagues, friends, and family. I would like to express my gratitude to all of you, who were a part of this journey.

First and foremost, I would like to thank my supervisor Heidi Kreibich, who guided me throughout the entire PhD period. She always had an open ear for my ideas and questions and found a way to steer this work into the right direction. I would like to express my sincere gratitude to all the members of my PhD committee: Doris Dransch, Jeroen Aerts, Dagmar Haase and Dörthe Tetzlaff and Heidi Kreibich. I especially thank Doris Dransch for heading the committee and Jeroen Aerts, Dagmar Haase and Heidi Kreibich for reviewing the thesis.

I am deeply grateful to Kai Schröter for his role as my scientific mentor. His great organizational skills and critical mind are an inspiration and one of the reasons this thesis exists today. I am very thankful to Stefan Lüdtke, who has patiently taught me the benefits of open software and how to deal with a detached HEAD. A special thanks to Nivedita Sairam for her support during the writing and submission process. I would like to express my gratitude to my co-authors: Dominik Paprotny, Nivedita Sairam, Rui Figureido, Marco Cerri, Laura Weise, Lorenzo Alfieri, Lorenzo Mentaschi and Francesco Dottori. Without their invaluable advice and contribution, this work would have been impossible. I would like to thank my PhD fellows and young scientists for all the fruitful discussions and exchanges that always kept me motivated: Nivedita Sairam, Duha Metin, Fabio Brill, Viktor Rözer, Tobias Sieg and Sophie Ullrich. I further thank all of my colleagues at the GFZ Hydrology section for a wonderful time with great Christmas parties and works outings. I especially thank Bruno Merz for providing an open and supportive research environment at GFZs Hydrology section.

I would like to express my gratitude to the H2020 | Insurance project and all of its members. A special thank you to my colleagues Fred Hattermann and Michel Wortmann from PIK Potsdam and Martin Drews from DTU Copenhagen. I am thankful to my GFZ colleagues Danijel Schorlemmer and Thomas Beutin for the insights into OSM and the data they provided. I would also like to acknowledge the voluntary work of the reviews that helped improve the research in this thesis, and all the contributors to open data and software that build the technical foundation for so many research projects.

My warmest thanks go to my friends and family for their unwavering support and encouragement. Finally, I want to thank my beloved partner, Gökben Demir, for your support and patience during long days and nights of work. Thank you for all the joy you bring to my life.

Contents

Summary	v
Acknowledgements	ix
Contents	xiii
List of Figures	xvi
List of Tables	xvii
List of Abbreviations	xix
1 Research motivation and objectives	1
1.1 Flood loss and risk	1
1.1.1 Modelling approaches	3
1.1.2 New data sources	6
1.1.3 Uncertainty sources, quantification, and perspectives	8
1.2 Research objectives and outline	9
1.3 Author contributions	12
2 Are OSM building data useful for flood vulnerability modelling?	13
2.1 Introduction	15
2.2 Data	17
2.2.1 Computer aided telephone interview data	17
2.2.2 OpenStreetMap data	18
2.2.3 Data preparation	20
2.2.4 Numerical measures	20
2.3 Methods	20
2.3.1 Variable selection	23
2.3.2 Predictive model learning	24
2.3.3 Predictive model evaluation	24
2.3.4 Spatial-transfer evaluation	25
2.4 Results and Discussion	26
2.4.1 Variable selection and predictive model learning	26
2.4.2 Model predictive performance: model benchmarking	29
2.4.3 Spatial-transfer testing	29
2.5 Conclusions	33
3 Consistent Probabilistic Flood Loss modelling in Europe	35
3.1 Introduction	36
3.2 Consistent Approach for Flood Loss modelling in Europe	38
3.2.1 The Microscale BN-FLEMOps	38
3.2.2 Method and Data for Model Upscaling at the meso-scale	41

3.2.2.1	Input Data Available for the European Domain	41
3.2.2.2	meso-scale Loss Calculation	43
3.2.3	Results and Discussion of Model Upscaling	44
3.2.3.1	Model Upscaling—Suitability Test of Proxy Variables	44
3.2.3.2	Flood Loss Assessment at European Scale	48
3.3	Model Validation and Adaptation Test in Three Case Study Areas	49
3.3.1	Case Study Descriptions	49
3.3.1.1	Case Study Caldogno, Italy	50
3.3.1.2	Case Study Lech, Austria	50
3.3.1.3	Case Study Mulde, Germany	51
3.3.2	Updating Method Using Local Data	52
3.3.3	Result and Discussion of the Validation	53
3.3.4	Result and Discussion of Model Adaptation Test	55
3.4	Conclusion	56
4	Das probabilistische Hochwasserschadensmodell BN-FLEMOps	59
4.1	Einleitung	61
4.2	Methoden und Daten	63
4.2.1	BN-FLEMOps Entwicklung und Modellstruktur	63
4.2.2	Datengrundlage und -verarbeitung	65
4.2.3	Schadensmodelle des Ensembles	68
4.3	Ergebnisse und Diskussion	70
4.3.1	Schadenschätzung und Modellvergleich	70
4.3.2	Probabilistische Modelle in der Entscheidungsfindung	72
4.3.3	Anwendbarkeit	73
4.3.4	Implementierung im OASIS-LMF	74
4.4	Zusammenfassung. <i>Summary</i>	74
5	Drivers of flood risk change for residential buildings in Europe	77
5.1	Introduction	78
5.2	Data and Methods	80
5.2.1	Hazard	80
5.2.1.1	River flow projections	80
5.2.1.2	Inundation modeling	80
5.2.1.3	Structural flood defences	81
5.2.2	Exposure	81
5.2.3	Vulnerability	82
5.2.4	Risk	84
5.2.5	Scenarios	85
5.2.5.1	Baseline scenarios	85
5.2.5.2	Future scenarios	85
5.2.5.3	Risk reduction scenarios	86
5.3	Results	87
5.3.1	Baseline flood risk	87
5.3.2	Drivers of future fluvial flood risk	89
5.3.2.1	Climate change	89
5.3.2.2	Exposure change	91
5.3.2.3	Combined effects	92
5.3.3	Risk reduction	93
5.4	Discussion	95
5.4.1	Comparisons with previous studies	95

5.4.2	Comparison with reported losses	96
5.4.3	Uncertainties	97
5.5	Conclusions	98
6	Discussion, conclusions, and outlook	101
6.1	Summary of key findings	101
6.2	Discussion	104
6.2.1	New data sources and model transfer	104
6.2.2	Uncertainty in flood loss estimation	106
6.2.3	Future flood risk in Europe	108
6.3	Conclusion and outlook	109
A	Appendix to Chapter 2	111
A.1	Numerical spatial measures	111
A.2	Code and data availability	111
A.3	Competing interests	112
B	Appendix to Chapter 5	113
B.1	Exposure projections	113
B.1.1	Population growth	113
B.1.2	Economic growth	113
B.1.3	Wealth-to-income ratio	114
B.2	Flood loss model variables	115
B.3	Baseline period	115
B.4	Climate change	115
B.5	Exposure change	115
B.6	Combined effects	115
B.7	Private precaution effects	115
B.8	Comparison with reported loss	116
B.9	Private precaution scenarios under future conditions	117
B.10	Uncertainty of climate model ensemble	118
B.11	Uncertainty of loss estimates under climate scenarios	119
	Bibliography	121

List of Figures

1.1	Schema of the process chain from flood risk modelling to decision-making	2
1.2	Stage-damage curve	4
1.3	Schematic structure of multi-variable modelling approaches applied in this thesis	5
1.4	Overview of thesis chapters	11
2.1	Regional subdivision of the data set for spatial split-sample testing . .	21
2.2	Data pre-processing, model learning and model transfer workflow . .	23
2.3	Out-of-Bag error for variations of mtry and ntree RF parameters	27
2.4	Spearman's correlation of model variables	27
2.5	Predictive performance of models using an increasing number of variables in order of their importance	28
2.6	Performance metrics of OSM based models and benchmark models . .	30
2.7	Model performance metrics in regional transfer	31
2.8	Scatter plots of numerical spatial measures and relative loss in regional sub-samples (Danube, Dresden, Elbe)	33
3.1	Model structure of BN-FLEMOps with available variables.	39
3.2	Comparison between the variables and data source for the microscale and meso-scale application of BN-FLEMOps.	41
3.3	Flow chart of the input data preparation for BN-FLEMOps.	43
3.4	Comparison of building area distributions per NUTS-3 region	45
3.5	Comparison of the number of historic flood events in DFO and msdd.	46
3.6	European map of historic flood events from the DFO catalog.	47
3.7	European asset map displaying the unit area values of residential buildings in 2013 in ($\text{€}/\text{m}^2$) following reconstruction costs approach.	48
3.8	Maps of the cumulated flood loss of residential buildings in Europe.	49
3.9	Empirical cumulative density function of the absolute residential building loss for three case studies	54
3.10	Comparison of reported loss and estimated loss for the application of BN-FLEMOps with update.	55
4.1	Modellstruktur des BN-FLEMOps mit Variablennamen. <i>Model structure of the BN-FLEMOps with variable names.</i>	64
4.2	Übersichtskarte des Untersuchungsgebietes an der Mulde mit Gemeinden mit offiziellen Schadenswerten. <i>Overview map of the study area at the Mulde river with municipalities with official reported loss figures.</i>	66
4.3	Vergleich der Schadensschätzungen von BN-FLEMOps und Modellen-semble. <i>Comparrison of the loss estimation of BN-FLEMOps and modellen-semble.</i>	70
5.1	Directed acyclic graph structure of BN-FLEMOps	83

5.2	Overview of the scenarios for flood risk estimation	86
5.3	Expected Annual Damage per country for the baseline scenario	87
5.4	Expected Annual Damage per NUTS-3 region for the baseline scenario	88
5.5	Relative change in EAD per NUTS-3 region from the baseline to climate projections in three future periods and for two rcp scenarios	90
5.6	EAD in Europe for the baseline and scenarios of future flood risk change.	91
5.7	Relative change in EAD per NUTS-3 region from the baseline to exposure projections in three future periods	92
5.8	Combined effects of climate and exposure change on EAD per country	93
5.9	Relative change in EAD per NUTS-3 region from the baseline to climate projections combined with exposure projections in three future periods and for two rcp scenarios.	94
5.10	Exceedance probability curve of flood losses for Europe	95
A.1	Definition and examples for numerical spatial measures.	111
B.1	Comparison of baseline losses with the HANZE database for the period from 1981 to 2010 on the national level	116
B.2	Future EAD per country under rcp45 combining the effects of climate and exposure change with three levels of private precautionary measures	117
B.3	Future EAD per country under rcp85 combining the effects of climate and exposure change with three levels of private precautionary measures	117
B.4	Relative uncertainty of the model ensemble for discharge change projections under rcp45 and rcp85 scenarios	118
B.5	Ratio between IQR and median of EAD for climate projections in three future periods and under two rcp scenarios	119

List of Tables

2.1	Preselected variables from CATI surveys	19
2.2	Variables of the amended OSM data set for each building object	21
2.3	Computational experiments for transfer applications	26
2.4	Model performance metrics for models using an increasing number of variables	29
2.5	Model precision, accuracy, and reliability performance metrics for OSM based and benchmark models	30
3.1	Overview of the Variables of BN-FLEMOps	40
3.2	Overview of Case Study Characteristics and Available Variables for Model Updating in the Case Studies	52
3.3	Comparison of Average Loss Estimate From BN-FLEMOps and Official Loss Information	53
4.1	Übersicht der Modellvariablen von BN-FLEMOps. <i>Overview of the model variables of BN-FLEMOps.</i>	65
4.2	Übersicht des Ensembles der Hochwasserschadensmodelle. <i>Overview of the ensemble of flood loss models.</i>	69
4.3	Performance-Indikatoren für den Modellvergleich. <i>Performance indicators for model comparison</i>	71
B.1	Overview of BN-FLEMOps variables	115
B.2	Performance indicators for the comparison of baseline losses with the HANZE database EAD [million €] for the period from 1981 to 2010 on the national level	116

List of Abbreviations

API	Application Programming Interface
ATKIS	Amtliches Topographisch-Kartographisches Informationssystem
ba	building area
BDe	Bayesian Dirichlet equivalence
BMI	lower Benchmark Model
BMrm	randomized Benchmark Model
BMu	upper Benchmark Model
BN	Bayesian Network
BN-FLEMOps	Bayesian Network – Flood Loss Estimation Model for the private sector
bt	building type
CART	Classification And Regression Tree
CATI	computer-aided telephon interview
CLC	CORINE Land Cover
d	inundation duration
d2D	Model transferred from Dresden to Danube
D2E	Model transferred from Danube to Elbe
d2E	Model transferred from Dresden to Elbe
DAG	Directed Acyclic Graph
DEM	Digital Elevation Model
DFO	Dartmouth Flood Observatory
E2D	Model transferred from Elbe to Danube
EAD	Expected Annual Damage
EU	European Union
Fast-IAMB	Fast Incremental Association
fe	flood experience
GDP	Gross Domestic Product
GPS	Global Positioning System
HANZE	Historical Analysis of Natural Hazards in Europe
HR	Hit Rate
Inter-IAMB	Interleaved Incremental Association
IPCC	Intergovernmental Panel on Climate Change

IQR	Inter Quartile Range
IS	Interval Score
JRC	Joint Research Center
LiDAR	Light imaging, Detection And Ranging
MAE	Mean Absolute Error
MBE	Mean Bias Error
MCMC	Markov Chain Monte Carlo
msdd	microscale damage data
MSE	Mean Square Error
mtry	variable subset in a Random Forest
NPT	Node Probability Table
ntree	number of trees in a Random Forest
NUTS	Nomenclature of Territorial Units for Statistics
OASIS-lmf	OASIS loss modelling framework
ODbL	Open Data Commons Open Database Licence
OOB	Out-Of-Bag (error)
OSM	OpenStreetMap
PESETA VI	Projection of Economic impacts of climate change in Sectors of the EU based on bottom-up Analysis (No. 5)
Q ₂₀	20 % quantile
Q ₈₀	80 % quantile
qcv	quantile coefficient of variation
QR ₉₀	Quantile Range from 5 % to 95 %
rbloss	relative building loss
rcp	regional concentration pathway
RF	Random Forest
rloss	relative (building) loss
rp	return period
SAB	Sächsische AufbauBank
VGI	Volunteered Geographic Information
wd	water depth (same as wst)
wst	Wasserstand (gleichbedeutend mit wd)

Chapter 1

Research motivation and objectives

1.1 Flood loss and risk

In the sixth report, the Intergovernmental Panel on Climate Change (IPCC) concludes that global temperatures have already risen by 1 °C at present and are expected to increase to 2 °C as early as the middle of the 21st century in comparison to the pre-industrial climate. At 2 °C global warming, the magnitude, and frequency of heavy precipitation and flooding will increase globally and in most regions of Europe (IPCC, 2021). While this may seem like an abstract risk for the future, fluvial floods are already the prevalent natural disasters and have increased in number and reported economic losses in the past three decades (WMO, 2021). The number of people directly affected by flooding grew globally by 58-86 million in the period from 2000 to 2015 (Tellman et al., 2021). The most recent floods of 2021 in Germany, Belgium, Italy, Turkey, Russia, China and many other countries have again emphasized the relevance of flood risk research.

In the past two decades over 1100 lives and 100 billion € have been lost to floods in Europe (Munich-RE, 2019; Petrucci et al., 2020). The 2002 floods alone have caused over 20 billion € losses in Central Europe (European Environment Agency, 2010), but have also led to increased efforts to better understand and manage the risk posed by river floods. In the future, combined effects of climate and socio-economic changes are projected to considerably increase the number of people affected, and the losses generated by floods in Europe (Rojas et al., 2013; Dottori et al., 2018). Thus, it is vital to further improve our understanding of flood risk on the European continent by developing more accurate and reliable modelling approaches and transparently communicate uncertainties to better inform climate adaptation and flood risk management. Many fundamental insights into human-flood interactions and the risk they pose to the lives and properties of people in affected areas were already conceptualized in White (1936, 1945) and still influence flood hazard and risk research today (Macdonald et al., 2012). White's notion of human adjustment to floods is now more relevant than ever in our efforts to adapt not only to floods, but to all climate related risks.

Risk is commonly understood as the probability and magnitude of adverse effects associated with a hazardous event. Definitions of risk differ and heavily depend on the respective scientific field. Vlek (1996) gives a comprehensive overview of risk definitions used in research. The concept of risk as it applies to natural disasters can be expressed as the intersection of hazard, exposure, and vulnerability IPCC (2012):

$$\text{Hazard} \cap \text{Exposure} \cap \text{Vulnerability} = \text{Risk}$$

Hazard, in general, refers to a potentially damaging event such as a flood. All elements of a system that may be impacted by the hazard are subsumed under the term exposure, for example human health, buildings, vehicles, and infrastructure. Vulnerability describes the susceptibility of the exposed elements to suffer damaged by the hazard. Although the exact definitions may vary in literature (Jonkman, 2007), flood damage is commonly organized into four categories by distinguishing the origin of the damage and the ability to quantify it in monetary terms (Merz et al., 2010b):

- **Tangible:** When the flood loss can be easily assessed in monetary values. This includes for example damage to buildings, roads, vehicles, and agriculture.
- **Intangible:** Refers to damage to people, goods, and services that are not easily measurable in monetary terms because they are not traded on a market.
- **Direct:** All damages that occur in the flooded area as a direct consequence of the flood event.
- **Indirect:** All damages in time or space outside the flood area but attributed to the flood event.

Modelling flood damages, calculating losses and estimating the risk should serve to inform management decisions and ultimately contribute to foster a more resilient society. Figure 1.1 shows the schematic process of identifying relevant data and knowledge to develop and apply risk models that inform decision-making. Integrated risk management recognizes the beneficial and harmful interactions between all elements of the flood system, with the aim to retain positive and minimize negative effects of flooding (Hall et al., 2003). Flood risk can be reduced not only by physical flood defences (Willner et al., 2018), but also through improved flood resistance for exposed buildings by measures such as wet and dry-proofing (de Moel et al., 2013) and other means of private precaution (Kreibich et al., 2012), emergency response (Molinari et al., 2013) and behaviour (Aerts et al., 2018), higher (local) retention (Förster et al., 2005), more efficient drainage (Sohn et al., 2020) or relocation (Dottori et al., 2021b). Risk may also be efficiently redistributed through insurance (Tariq et al., 2014; Solín et al., 2018).

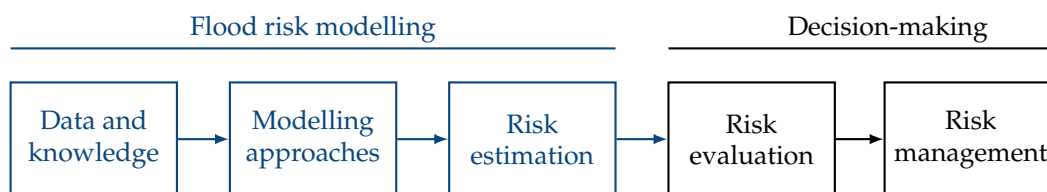


Figure 1.1: Schema of the process chain from flood risk modelling to decision-making (based on Aven (2016) and built using Walczak (2007))

In the context of this thesis, the three components of risk are more narrowly defined. Hazard is the likelihood and magnitude of fluvial flooding. Fluvial or river floods occur when the water level in a river rises until it overflows the river banks or breaches embankments and inundates surrounding areas. Exposure is the monetary value of residential building structures and the cost of repair or reconstruction. This includes all costs of material and labour required to restore the building structure after damages directly related to a flood. It does not represent a (speculative) market

value of real estate assets. The research presented here does explicitly not include losses originating from damage to household contents such as furniture, carpets, electronic devices and other mobile items. Vulnerability describes susceptibility of buildings and its residents to the adverse effects of the flood hazard. Vulnerability of a residential building to the damaging effects of flooding is influenced by building characteristics, the behaviour of residents and the level of adaptation (Few, 2003). This thesis touches on all three risk components, but focuses on the vulnerability aspect, as there is a clear need to better understand its influence and potential for flood risk reduction (Kreibich et al., 2017a). Damage to residential buildings or their complete destruction by a flood poses not only a financial and economic risk, but can also have intangibles negative effects on the physical (Hajat et al., 2005; Burton et al., 2016) and psychological (Stanke et al., 2012; Sedighi et al., 2021) health of residents. Risks for human health, though, are not actively discussed in the following.

This thesis focuses on flood risk as the probability and severity of flood damage at affected residential buildings, causing monetary loss. The need to advance the state of flood risk modelling becomes apparent when we consider the historical record of devastating flood losses and projections for a future climate in which extreme flooding becomes even more likely in a world with growing concentration of assets in dense urban environments. To improve flood risk estimations for residential buildings in Europe, the thesis identifies and integrates novel data sources into newly developed and existing multi-variable approaches. Thus, enabling the transfer of models from the local scale to a continent wide application with integrated uncertainty quantification. Hereby, this research directly addresses priority 1 of the Sendai Framework by the United Nations, "Understanding disaster risk" and informs disaster risk management and investments for risk reduction (priority 2 and 3) (UNISDR, 2015).

The following sections in [Chapter 1](#) introduce the current state of research in flood loss modelling approaches, data sources and their inherent uncertainties with a focus on vulnerability aspects. Subsequently, research objectives for the advancement of probabilistic flood loss modelling are developed based on identified research gaps and the opportunities of new data and modelling approaches. A concise overview of the thesis is presented in [Figure 1.4](#). Finally, the individual contributions of all co-authors to each chapter are listed.

1.1.1 Modelling approaches

Two main approaches for the development of flood loss modelling are established in literature and practice. Empirical models are based on the statistical analysis of damage data (e.g., Thielen et al., 2005; Zhai et al., 2005), while synthetic approaches consist of What-If scenarios combining expert knowledge and engineering information (e.g., Penning-Rowsell and Chatterton, 1977; Dottori et al., 2016). A distinction is also made between models that directly include exposure values and calculate absolute losses in monetary terms (e.g., NR&M, 2002; Penning-Rowsell et al., 2018) and models that provide relative degrees of damage that can be combined with external exposure values to derive monetary results (e.g., Scawthorn et al., 2006; Pistrika and Jonkman, 2010). Relative damage models are used in this study, because they are considered more suitable for model transfer (Merz et al., 2010b).

Flood loss models use mathematical methods that conceptualize the damaging interaction between flood hazard and exposed assets to quantify impacts in monetary

terms. The relationship of water depth to the severity of flood damage called “stage-damage curve” or “depth-damage function” is the most frequently used approach to model flood losses (Penning-Rowsell & Chatterton, 1977; Smith, 1994; Dutta et al., 2003; Meyer & Messner, 2005; Gerl et al., 2016). A schematic example of a stage-damage curve is depicted in Figure 1.2. Depending on the model application and asset at risk, a stage-damage curve can take many forms.

The complexity reduction of flood damage processes by utilizing only the inundation depth at a certain location to infer flood loss has the clear advantage of requiring few input data to train and apply a model. It is because of the low data requirements and simple structure, that stage-damage functions are still widely used (Meyer & Messner, 2005). The downside of this simplified approach are large inaccuracy and uncertainty in the loss estimates (Middelmann-Fernandes, 2010). In comparison with more complex multi-variable models, stage-damage curves underperform in most applications (Schröter et al., 2014; Rözer et al., 2019). Multi-variable models better capture the various processes influencing flood damage and, depending on the method, inherently provide uncertainty information (Kreibich et al., 2017b). Despite the limitations, the latest developments in continental flood risk modelling for current and future climate conditions in Europe still rely to a majority on simple stage-damage curves. These curves are often specifically developed for various asset types and countries (Huizinga, 2007; Alfieri et al., 2015b; Huizinga et al., 2017; Dottori et al., 2021a). Even for highly specialized applications that only cover one asset type such as rail (Bubeck et al., 2019), road (van Ginkel et al., 2021), or port infrastructure, (Izaguirre et al., 2021) stage-damage curves are the default method.

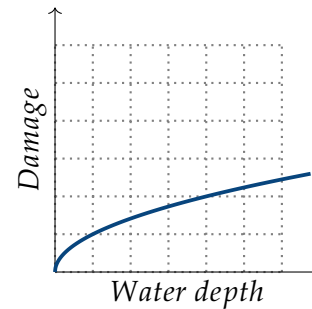


Figure 1.2: Stage-damage curve

Multi-variable models include a number of variables to represent flood damage processes. Water depth is also, in multi-variable approaches, usually considered the most important predictor of flood loss (Merz et al., 2013; Schoppa et al., 2020). Besides water depth, other variables such as the return period and duration of a flood, the flow velocity and water contamination are useful to characterize the flood hazard in a loss model. Variables describing socio-economic status, building characteristics, warning lead time and private precaution may represent the vulnerability of an exposed asset (Scawthorn et al., 2006; Elmer et al., 2010; Merz et al., 2013; Wagenaar et al., 2017). Recent studies by Paprotny et al. (2020a) and Schoppa et al. (2020) have shown the prospects of probabilistic, multi-variable approaches for flood risk modelling for the commercial sector. This thesis explores the potential of these approaches for flood loss estimation for residential buildings on multiple scales.

Data for natural events such as extreme floods are rare by definition (El Adlouni et al., 2008). Flood loss models are, therefore, often designed with data or knowledge of a specific region and event to later be applied in other areas where local data is not available or insufficient to create a dedicated model (Cammerer et al., 2013). Transferring models in space, time, and scale generally introduces new uncertainties and reduces estimation accuracy (Papathoma-Köhle et al., 2011; Meyer et al., 2013). Multi-variable models, however, have shown to deliver more robust estimations when transferred (Schröter et al., 2014; Wagenaar et al., 2018).

The use of model ensembles, that consist of multiple flood loss models, can deliver robust flood loss estimates and uncertainty quantification (Figueiredo et al., 2018). (Re-)Insurance companies often run multiple models from various vendors to get a picture of the uncertainty inherent in the estimations (Wüest, 2019). Implementing multiple models for a study is not always a viable approach, because of data and time constraints. Probabilistic multi-variable models can improve flood loss estimations in comparison to simple stage-damage curves and inherently provide uncertainty information without requiring the setup of multiple models (Merz et al., 2013; Schröter et al., 2014; Hasanzadeh Nafari et al., 2016c; Kreibich et al., 2017b). In this thesis, multi-variable models based on graphical networks and decision trees are developed and applied.

Tree-based models also referred to as classification and regression tree (CART) models use a number of split criteria to divide the feature space of predictor variables. At each split, a decision following a predefined rule is performed to successively narrow down the field of possible predictions in the leaf nodes of the tree (Figure 1.3a). Ji et al. (2013), Chinh et al. (2015), Hasanzadeh Nafari et al. (2016c), Kreibich et al. (2017b), and Sieg et al. (2017) provide examples for tree-based approaches in flood loss modelling. In this thesis, the Random Forest (RF) algorithm (Breiman, 2001) was used in its implementation in the R package “randomForest” by Liaw and Wiener (2002). RF is an ensemble of n decision trees, where each split decision is based on a subset of features ($mtry$). This approach combines bootstrap aggregation (bagging) with randomized subspace selection. RF can be used for classification and regression, as well as feature importance analysis. The results of all trees in a forest are aggregated by majority vote or averaging. Thereby, these large, automatically build ensembles generate quasi probabilistic results. RF is a non-parametric method that is robust against overfitting and can be used with correlated input data. The calculation of the so called Out-Of-Bag (OOB) error provides insight into the model performance (Biau & Scornet, 2016). RF is well established and applied for a diverse set of modelling tasks in research fields such as chemistry for predicting molecular interactions (Svetnik et al., 2003), remote sensing for land-use classification from satellite imagery (Waske & Braun, 2009; Steinhausen et al., 2018), and flood loss modelling for various economic sectors (Merz et al., 2013; Sieg et al., 2017). In Chapter 2 RF models are developed to test the importance of various input variables, and the transferability of flood loss models.

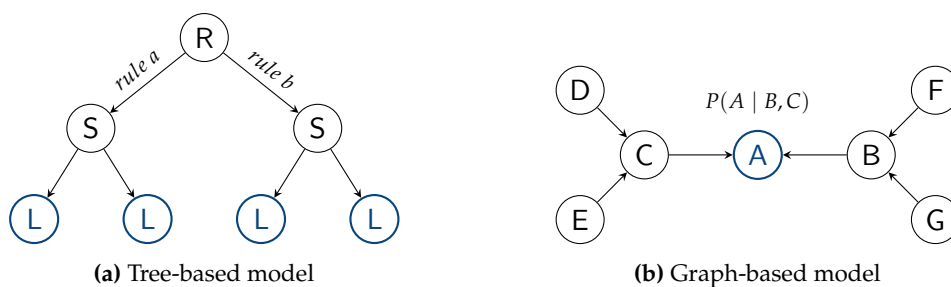


Figure 1.3: Schematic structure of multi-variable modelling approaches applied in this thesis. (a) Decision tree with root (R), split (S) and leaf (L) nodes. Two decisions from root to split nodes are indicated by split *rule a* and *b*. (b) Graph of a BN with seven variables (A–G) and the joined probability of A given B and C. Blue coloured leaves/nodes represent the model response values.

In graph-based models, variables are represented by nodes and connected via edges

(Figure 1.3b). Graph-based models can be constructed with expert knowledge (Pitchforth & Mengersen, 2013) and/or with the help of data analysis and machine learning techniques, e.g., (Heckerman et al., 1995; Yaramakala & Margaritis, 2005). Schröter et al. (2014), Vogel et al. (2014), Jäger et al. (2018), and Wagenaar et al. (2018) show implementations of graph-based models following a Bayesian approach for the estimation of flood loss. Bayesian (Belief) Networks (BN) are constructed as Directed Acyclic Graphs (DAG) in which the state of the interconnected nodes represent the joint probability of variables considering the conditional (in)dependency structure (Pearl, 1988). This approach offers great transparency for modellers and decision-makers, because each node of an established network contains the conditional probabilities in so-called node probability tables (NPT). Inference within a BN can be made in each direction of an edge and at all nodes. This approach therefore also allows for loss estimations with incomplete input data. The results of a BN are probabilistic and inherently provide uncertainty information (Hanea et al., 2006; Fenton & Neil, 2013; Hanea et al., 2015). BN based approaches are used to estimate flood loss in Chapter 3, 4, and 5.

1.1.2 New data sources

Empirical data from flood events are essential to the development of reliable empirical flood loss models. Details about the hazard characteristics as well as the vulnerability and exposure can be learned from data, acquired during, and after flood events. Data are typically obtained via surveys of the affected population and other stakeholders, insurance claims, expert assessment and remote sensing. Prominent examples for empirical event databases are the NatCatSERVICE from Munich Re and EM-DAT with data on multiple natural hazards globally (Choryński et al., 2012; EM-DAT, 2019; Munich-RE, 2019). The HANZE database contains Europe-wide data for various flood and asset types from 1870 to 2016 (Paprotny et al., 2018a). While these examples contain event-specific data that are useful for model validation, the HOWAS 21 database stores object-specific data on flood affected properties from 1978 to 2013 that can effectively be used for model development (Buck & Merkel, 1999; Kreibich et al., 2017c). The database includes flood damage data on private households as well as the commercial, industrial, and public sector assets (Kellermann et al., 2020).

Synthetic data for model development relies on engineering and expert knowledge to create hypothetical curves independently of empirical flood data for a specific area or event. Data is collected, for example, through surveys asking for expected damage under What-If flood scenarios. This approach does therefore not rely on the often resource demanding and difficult collection of in-situ damage data (Penning-Rowsell & Chatterton, 1977; Parker et al., 1987; Penning-Rowsell et al., 2018). However, synthetic data are not used in this thesis, because they are not consistently available for Europe, are considered subjective and may result in higher uncertainties in loss estimates (Merz et al., 2010b).

The development and training of the flood risk modelling approaches presented in Chapters 2 – 5 rely on empirical ex post flood survey data collected via computer aided telephone interviews (CATI) in Germany. CATI data is part of the HOWAS 21 database and was collected after flood events in 2002, 2005, 2006, 2010, 2011, and 2013 in the Rhine, Elbe and Danube catchments in Germany. In these survey campaigns, flood affected households were asked questions about the properties of the flood, their building and socio-economic status as well as the loss they had suffered. With an average response rate of 15 % a total of 3056 interviews were completed. Thielen

et al. (2005), Kienzler et al. (2015), and Thieken et al. (2017) report details on questions asked during the surveys and the variables collected and computed.

New sources of data have become more readily available in recent years in the form of Volunteered Geographic Information (VGI). VGI data is created by citizen contributors and is in most cases not intentionally designed for the use in any specific field of research. Sources for VGI include social media platforms, contributor driven mapping projects such as OpenStreetMap (OSM) (OSM contributors, 2020), articles and posts on news websites that include information on flooding in photographs or videos with geolocation (Goodchild, 2007).

OSM as a global database for geographic information has potential to enable the application and transfer of flood loss models. The OpenStreetMap project was created in the UK in 2004 and is, similar to Wikipedia, build on the contributions of volunteers with growing support of corporate editors. Users can contribute information to the map, edit existing data and verify other users additions. Over 1 million contributors have made at least one change, and over 700,000 have made more than ten changes to the map (Anderson et al., 2019). Geodata are typically created based on local knowledge, GPS surveys and from remote sensing imagery. Large datasets of official cadastral data in France provided by Direction Générale des Impôts (Mooney & Minghini, 2017), road networks for the Netherlands (Anderson et al., 2019) and building footprints for Individual cities (Hecht et al., 2013) have been integrated into OSM database. OSM as a user generated and not standardized open database relies on the community to review and verify changes. There have been many attempts at scrutinizing the quality and completeness of the OSM database, often by comparison with official reference data (Senaratne et al., 2017). Mooney and Minghini (2017) provide a general overview of the completeness of OSM data and Barrington-Leigh and Millard-Ball (2017, 2019) find that over 80 % of the world's road network are included in OSM. The positional accuracy, semantic information, and building footprint completeness of has also been analysed. Hecht et al. (2013) found in case studies in Saxony, Germany in 2012 large regional differences in completeness of building footprint data in OSM. Urban centres are generally more complete than rural regions. For the city of Munich, Fan et al. (2014) stated that OSM buildings cover almost all build up areas, but still lack many attributes such as name, type, or height. Brückner et al. (2021), estimate the completeness of shops in the OSM data for Saxony at 82 % and Baden-Württemberg at 88 %. Based on point and building geometry comparisons Brovelli and Zamboni (2018) found that the quality of OSM is comparable to the regional technical authoritative map for the Italian region of Lombardy. Map completeness and data quality are generally higher in more densely populated areas, with higher GDP and more contributors (Dorn et al., 2015; Zhou, 2018; Tian et al., 2019). Many of the recent advances in flood risk modelling make use of OSM data in their approaches (Figueiredo & Martina, 2016; Sieg et al., 2019; van Ginkel et al., 2021).

This thesis utilizes land use information included in OSM to extract all residential buildings for further processing. In [Chapter 2](#) the geometry of OSM building footprints is used to derive information for building flood loss estimation. For the loss modelling approaches in [Chapters 3 – 5](#), the building footprint area is calculated based on OSM data.

1.1.3 Uncertainty sources, quantification, and perspectives

Uncertainty can originate from many sources in approaches and data for modelling flood risk, as well as the natural processes of flooding itself. To conceptualize the elusive term uncertainty, a distinction into two general types is useful. Uncertainty may originate from incomplete knowledge (epistemic) or from the random nature of a process (aleatory). In other words, aleatory uncertainty describes the properties of a system and epistemic uncertainty results from the properties of its observer (Cullen et al., 1999). Epistemic uncertainty refers to the limited ability to measure and understand the processes of flooding and flood damage. Epistemic uncertainty is therefore associated with the quantities used to describe flood risk, such as, the water depth and the monetary loss. It further manifests in the structure of models that can not represent every detail and interaction in the flood system. Aleatory uncertainty describes the stochastic character of natural processes like the distribution of rainfall over a catchment (Merz & Thielen, 2005) or the hysteresis of waves and spatio-temporal course of inundation during floods (Apel et al., 2004).

The variability of flood damages even within one asset category is large and introduces uncertainty into flood risk modelling. Two buildings of the same size and type affected by the same water depth may show significant differences in loss. This may result from different levels of precaution, contamination, or flow velocities at the building location (Merz et al., 2010b). The impact of different uncertainty sources varies with scale. On the microscale (object level) epistemic uncertainty in the parametrization of hazard models is relevant. On the meso-scale (regional) many studies have found substantial uncertainties related to flood event probability and duration as well as the damage models used for flood loss estimation (de Moel et al., 2015). On the macro-scale (national) especially, the necessary simplifications in hazard and loss modelling and the limitations of coarse resolution in the underlying data contribute uncertainties (Alfieri et al., 2015b). Inadequate data on dykes, potential dyke breaches and general protection levels introduce considerable uncertainties on all scales considered (de Moel et al., 2015). Model transfers between the scales as well as in time and space are an additional source of uncertainty (Cammerer et al., 2013; Wagenaar et al., 2018).

Research aims to reduce epistemic and describe aleatory uncertainties, and has an obligation to acknowledge and report uncertainties where they can be identified (Willows et al., 2003). Uncertainty is an indicator of the quality of knowledge about flood risk. As described in [section 1.1](#) risk assessments serve to inform flood risk management and support decision-making. Information about uncertainties in modelling approaches and data is therefore essential to the understanding of reliability of modelling results and for decisions based on them.

This thesis expresses uncertainties in the language of probability and makes use of both, the classical frequentists, and the Bayesian interpretation of probability. The frequentist understanding builds on the relative frequency of an event observed in an experiment repeated multiple times under the same conditions. While the Bayesian probability, relies on a prior belief of likelihoods and the correction or updating of these beliefs in light of new evidence (Weise & Woger, 1993; Fenton & Neil, 2013). Bayes' theorem formulates the method as:

$$P(H | E) = \frac{P(E | H) \times P(H)}{P(E)} \quad (1.1)$$

The conditional probability $P(H | E)$ of a hypothesis H is evaluated under the assumption of evidence E being true. This term is referred to as the posterior (belief). It is derived from the likelihood of the evidence E given the hypothesis H expressed as $P(E | H)$ and the two marginal probabilities, the prior $P(H)$ and the data distribution $P(E)$ under the condition $P(E) \neq 0$. Bayes' theorem builds the foundation for the Bayesian Network approach used for integrated uncertainty quantification in Chapters 3 – 5. Uncertainty, when numerically quantifiable, is communicated in this thesis in terms of probability distributions or confidence intervals (quantiles) around the first moment of a probability distribution (mean, median) as recommended by van der Bles et al. (2019).

1.2 Research objectives and outline

Fluvial floods already pose a great risk to the communities living in flood plains. Climate change and socio-economic developments will further increase this risk. Simplified approaches to estimate flood risk based on limited available data are still widely used and contain large uncertainties in the results that often are not made transparent. However, probabilistic multi-variable models consider the most influential hazard and vulnerability variables that control flood damage and can therefore estimate flood losses more precisely while providing uncertainty information. Requirements for large and detailed input data limit the use of these models to small-scale applications. With the availability of new open data sources and greater computing power, the development and implementation of these approaches becomes viable for larger applications. Utilizing open data sources and implementing models in open source frameworks makes these advancements more reproducible and accessible. Uncertainty inherent in all data and models are directly quantifiable with probabilistic approaches and should be transparently communicated to stakeholders and decision-makers.

The main objective of this thesis is to advance flood risk modelling for residential buildings in Europe to improve our understanding of flood risk under current and future climatic and socio-economic conditions. This requires the identification and integration of new data sources into probabilistic multi-variable loss models. Thus, enabling the implementation and transfer of advanced loss models to the European continent and the transparent quantification of inherent uncertainties. To accomplish these objectives, the following four research questions are addressed in Chapters 2 – 6:

1. Which new data sources can advance multi-variable approaches for flood loss modelling? (Chapter 2 and 3)
2. What is required to improve the transferability of loss models with included uncertainty quantification? (Chapter 2 and 3)
3. How does a probabilistic, multi-variable model, built on novel data, perform in comparison with traditional approaches? (Chapter 4)
4. How will flood risk for residential buildings in Europe change throughout the 21st century? (Chapter 5)

This cumulative thesis is structured into six chapters. The first chapter introduces the principal concepts of flood loss data and modelling and formulates the main research objectives. The following four chapters present the conducted research and detailed

findings to answer the research questions. The final chapter discusses results, concludes the main findings, and gives an outlook of future research opportunities.

Chapters 2 and 3 describe the identification and integration of large new data sets into multi-variable loss models. Chapter 2 also explores novel variables to represent vulnerability in flood loss models. Both chapters present approaches to enable spatio-temporal and scale transfer of models. In Chapter 4 the performance of a multi-variable model is tested and the utility of probabilistic flood loss estimates for decision-making is examined. Chapter 5 combines the research presented in Chapters 2 – 4 into an application of a probabilistic multi-variable loss model to estimate present and future flood risk in Europe under climate and socio-economic change. Furthermore, the potential of private precaution to reduce flood risk for residential buildings in Europe is examined. The final Chapter 6 discusses the main findings presented in Chapters 2 – 5 and draws conclusions for the research questions posed in Chapter 1. Moreover, opportunities for further research are presented. All chapters except Chapter 4 are in written and published in English language. Chapter 4 is written and published in German language, but has translations into English for the title, abstract, all figure and table captions and the conclusion section. An overview of all thesis chapters and key findings is presented in Figure 1.4.

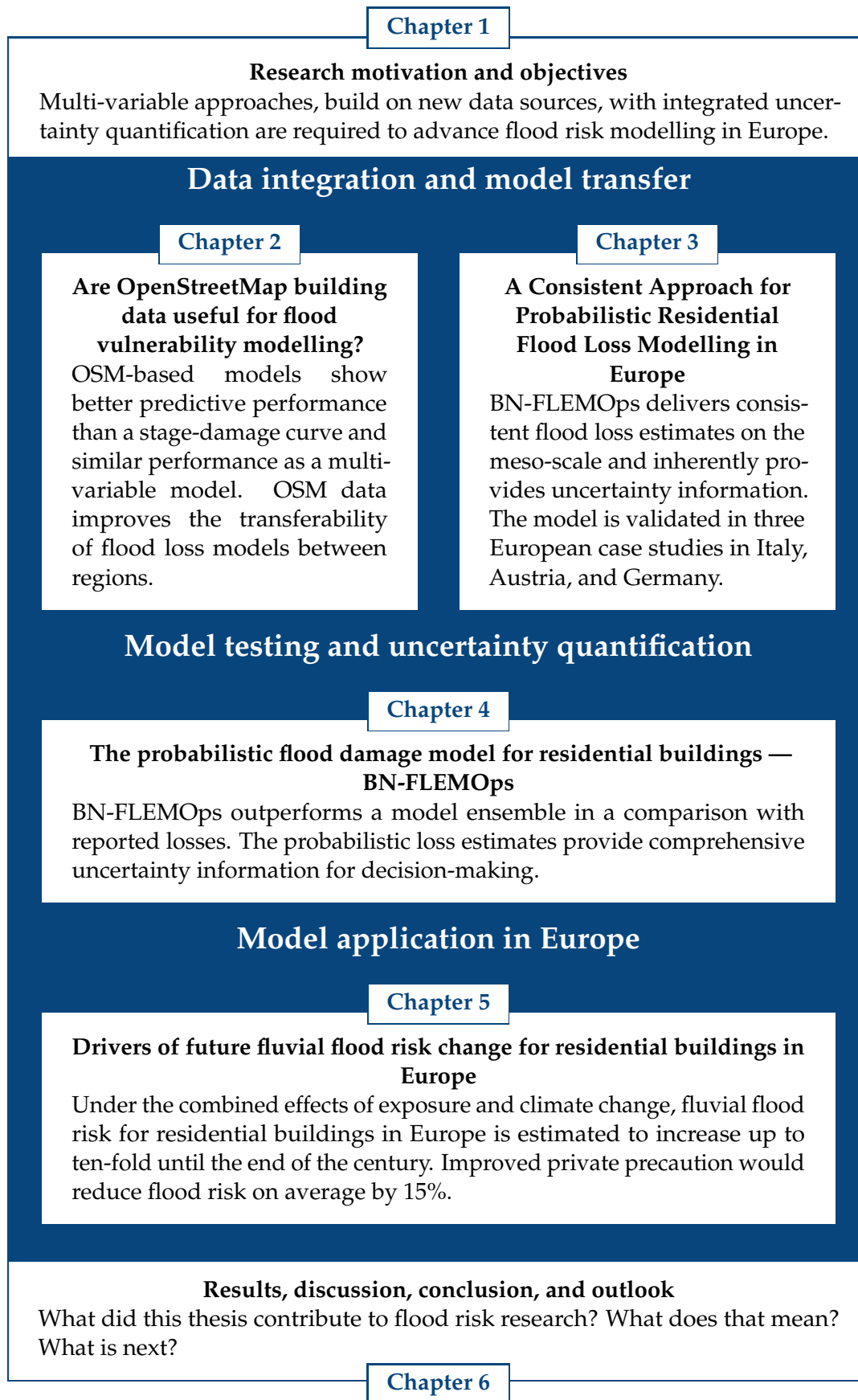


Figure 1.4: Overview of thesis chapters

1.3 Author contributions

This thesis comprises four manuscripts in chapters 2 – 5 which have been published in or submitted to peer-reviewed journals. These manuscripts have been created in a collaborative effort by the author of this thesis (M.S.) and several co-authors (initials). Individual contributions are outlined in detail for each chapter below. The introduction [Chapter 1](#) and the concluding [Chapter 6](#) are the sole work of the author of this thesis (M.S.).

Chapter 2: Are OpenStreetMap building data useful for flood vulnerability modelling?

M.C. and K.S. conceived and designed the study. M.C. prepared and analysed the data with major contributions from M.S. and K.S.. M.C., M.S. and K.S. wrote the first draft of the paper. H.K. helped guide the research through technical discussions. All authors reviewed the draft of the paper and contributed to the final version. M.S. provide and revised some of the figures.

Chapter 3: A Consistent Approach for Probabilistic Residential Flood Loss Modelling in Europe

S.L., K.S. and M.S. designed the research with support from H.K.. M.S. and S.L. analysed the data with support from L.W., developed the model and implement the case study applications. M.S. and S.L. prepared all figures, tables, and support material published as separate data publication in Lütke et al. (2019b). M.S., L.S., K.S., and H.K. interpreted the results and contributed together with R.F. to the original draft. The revision was performed by M.S., L.S., K.S. and H.K. M.S. contributed significantly to data analysis, model development and the presentation and interpretation of results. M.S. wrote large parts of the data and methods, results and discussion chapters.

Chapter 4: The probabilistic flood damage model for residential buildings — BN-FLEMOps

The study design was conceived by M.S., S.L., K.S. and H.K.. Data analysis and modelling was performed by M.S.. R.F. contributed data and advised in the writing of the draft. The original draft was mainly written by M.S.. K.S. and H.K. contributed with support from S.L.. All figures were created by M.S.. M.S. K.S. and H.K. revised the draft during review.

Chapter 5: Drivers of future fluvial flood risk change for residential buildings in Europe

The conceptualization of the study was led by M.S. in collaboration with D.P., F.D., N.S., L.A., H.K., and K.S.. M.S., S.L. and D.P. developed the methodology and software. Data analysis and curation was performed by M.S. with contributions of D.P.. D.P. created the exposure baseline and projections. The hazard maps and flood protection standards were provided by F.D.. All authors contributed to the original draft. M.S. wrote the majority of the text with contributions of the co-authors to the introduction, methods and discussion chapters. All but one figure was created by M.S. with supporting advice from the co-authors. Figure B.4 was provided by L.M..

Chapter 2

Are OpenStreetMap building data useful for flood vulnerability modelling?

Abstract

Flood risk modelling aims to quantify the probability of flooding and the resulting consequences for exposed elements. The assessment of flood damage is a core task that requires the description of complex flood damage processes, including the influences of flooding intensity and vulnerability characteristics. Multi-variable modelling approaches are better suited for this purpose than simple stage–damage functions. However, multi-variable flood vulnerability models require detailed input data and often have problems in predicting damage for regions other than those for which they have been developed. A transfer of vulnerability models usually results in a drop of model predictive performance. Here we investigate the questions whether data from the open-data source OpenStreetMap is suitable to model flood vulnerability of residential buildings and whether the underlying standardized data model is helpful for transferring models across regions. We develop a new data set by calculating numerical spatial measures for residential-building footprints and combining these variables with an empirical data set of observed flood damage. From this data set, random forest regression models are learned using regional subsets and are tested for predicting flood damage in other regions. This regional split-sample validation approach reveals that the predictive performance of models based on OpenStreetMap building geometry data is comparable to alternative multi-variable models, which use comprehensive and detailed information about preparedness, socio-economic status and other aspects of residential-building vulnerability. The transfer of these models for application in other regions should include a test of model performance using independent local flood data. Including numerical spatial measures based on OpenStreetMap building footprints reduces model prediction errors (MAE — mean absolute error — by 20 % and MSE — mean squared error — by 25 %) and increases the reliability of model predictions by a factor of 1.4 in terms of the hit rate when compared to a model that uses only water depth as a predictor. This applies also when the models are transferred to other regions which have not been used for model learning. Further, our results show that using numerical spatial measures derived from OpenStreetMap building footprints does not resolve all problems of model transfer. Still, we conclude that these variables are useful proxies for flood vulnerability modelling because these data are consistent (i.e., input variables and underlying data model have the same definition, format, units, etc.)

and openly accessible and thus make it easier and more cost-effective to transfer vulnerability models to other regions.

Published as:

Cerri, M., Steinhausen, M., Kreibich, H., & Schröter, K. (2021): Are OpenStreetMap building data useful for flood vulnerability modelling? *Natural Hazards and Earth System Sciences*, 21(2), 643–662. <https://doi.org/10.5194/nhess-21-643-2021>

2.1 Introduction

Floods have huge socio-economic impacts globally. Driven by increasing exposure, as well as increasing frequency and intensity of extreme weather events, consequences of flooding have sharply risen during recent decades (Lugeri et al., 2010; Hoeppe, 2016). Therefore, effective adaptation to growing flood risk is an urgent societal challenge (UNISDR, 2015; Jongman, 2018). With the transition to risk oriented approaches in flood management, flood risk models are important tools to conduct quantitative risk assessments as a support for decision-making from continental to local scales (Winsemius et al., 2013; de Moel et al., 2015; Alfieri et al., 2016b). While macro- or meso-scale risk assessment approaches target regional, national or continental studies, risk assessment on the microscale is needed to guide urban planning, optimize investment for protection and other mitigation measures considered in flood risk management plans (Meyer et al., 2013; de Moel et al., 2015; Rehan, 2018). Flood risk models include components to represent the key elements of flood risk: hazard, exposure and vulnerability (Kron, 2005). Flood hazard is usually modelled with high spatial resolutions in order to realistically capture variability in flood hazard intensity in consideration of local topographic characteristics (Apel et al., 2009; Teng et al., 2017). For consistent risk assessments, exposure, and vulnerability need to be analysed on similar scales and with appropriate spatial resolution. With an increasing availability of new exposure data sets including for instance information about the number, occupancy, and characteristics of exposed objects, (Figueiredo & Martina, 2016; Pittore et al., 2017; Paprotny et al., 2020b) microscale exposure and vulnerability modelling gains much traction (Schröter et al., 2018; Lüdtkke et al., 2019a; Sieg et al., 2019).

Both synthetic (e.g., Penning-Rowsell and Chatterton (1977), Blanco-Vogt and Schanze (2014), and Dottori et al. (2016)) and empirically based models (e.g., Thielen et al. (2005) and Zhai et al. (2005)) have been proposed for microscale vulnerability modelling. As flood damaging processes are complex, a large diversity of influencing factors needs to be taken into account to capture and appropriately represent flooding intensity and resistance characteristics of exposed elements in flood vulnerability models (Thielen et al., 2005). In this context, multi-variable modelling approaches are an important advance from simple stage-damage curves, which relate only water depth to flood loss. While multi-variable vulnerability models usually outperform traditional stage-damage functions (Merz et al., 2004; Schröter et al., 2014), the downside of these approaches is an increased need for detailed data on the level of individual objects, (Merz et al., 2010b; Merz et al., 2013) which are often not available in the target area of the analysis (Apel et al., 2009; Cammerer et al., 2013; Dottori et al., 2016). Missing standards for collecting comparable and consistent data are one reason for this problem (Changnon, 2003; Meyer et al., 2013). Hence, providing the input variables for multi-variable flood vulnerability models on the microscale is a key challenge for their practical applicability. Another challenge is the generalization of locally derived vulnerability models. A number of studies confirm a model performance mismatch between regions where models have been developed and the target areas for application (Jongman et al., 2012a; Cammerer et al., 2013; Schröter et al., 2016; Wagenaar et al., 2018). It is argued that the generalized application of vulnerability models to different geographic and socio-economic conditions needs to consider an adequate representation of local characteristics and damage processes (Felder et al., 2018; Figueiredo et al., 2018; Sairam et al., 2019b). Hence, consistency in input data is an important requirement for the spatial transfer of vulnerability

models (Lüdtke et al., 2019a; Molinari et al., 2020). The availability, accessibility, and consistency of data sources are important requirements for generalized vulnerability model applications, but also poses requirements on modelling approaches. With an increased number of input variables and an enlarged diversity of data sources used for vulnerability modelling, we usually deal with heterogeneous data in terms of different scaling, degrees of detail, resolution, and complex interdependencies (Schröter et al., 2016; Schröter et al., 2018).

Tree based algorithms are a suitable approach to handle heterogeneous data, represent non-linear and non-monotonic dependencies, and, as a non-parametric approach, do not require assumptions about independence of data (Merz et al., 2013; Schröter et al., 2014; Wagenaar et al., 2017; Carisi et al., 2018). The Random Forest (RF) algorithm (Breiman, 2001) is broadly used in many disciplines, due to its high predictive accuracy, simplicity in use and flexibility concerning input data. In the domain of flood risk modelling, Wang et al. (2015) have successfully applied RF for flood risk assessment and Bui et al. (2020) used RF for flood susceptibility mapping. Merz et al. (2013) demonstrated the suitability of tree-based algorithms for flood vulnerability modelling. Following this, Chinh et al. (2015), Hasanzadeh Nafari et al. (2016a), Sieg et al. (2017), Wagenaar et al. (2017), and Carisi et al. (2018) have used RF and other tree-based algorithms for flood loss estimation in flood prone regions in Vietnam, Australia, the Netherlands, and Italy. In these studies, vulnerability modelling using RF was based on site specific empirical data sets which had been collected ex-post major flood events. In contrast, the framework proposed by (Amirebrahimi et al., 2016) successfully used 3D building information for flood damage assessment of individual buildings. Gerl et al. (2016) and Schröter et al. (2018) investigated the suitability of alternative more general data sources for flood vulnerability modelling using urban structure type information derived from remote sensing images, virtual 3D city models and numerical spatial measures which describe the extent and shape complexity of residential buildings. It was shown that geometric information such as building area and height are useful variables to describe building characteristics relevant for estimating flood losses (Schröter et al., 2018). From these studies, it has been concluded that data about building geometry work as a proxy to describe resistance characteristics of buildings. However, further analyses are needed to understand whether building geometry data enable consistent flood vulnerability modelling with high resolution and are suitable to characterize differences in flood vulnerability across regions.

With new data sources emerging from crowdsourcing projects and open data initiatives, detailed building data are increasingly available and accessible (Irwin, 2018). Open and/or standardized building data are a promising data source to coherently describe exposure and characterize vulnerability of residential buildings, and to improve the spatial transfer of vulnerability models given a consistent underlying data model and clear specification of input variables across regions. Data science methods are predestined to make use of these data in flood vulnerability modelling. Against this backdrop, we investigate the suitability of the open data source OpenStreetMap (OSM) (OSM contributors, 2020) for flood vulnerability modelling of residential buildings. OSM is a geographic database with a worldwide coverage which is nowadays considered as reliable (Barrington-Leigh & Millard-Ball, 2017). The information about building footprints is freely available and straightforward to obtain from public online servers. The OSM contributors' community is constantly growing and assures regular updates in terms of accuracy and completeness of the data (Hecht et al., 2013).

We test the hypothesis that numerical spatial measures derived from OSM building footprints provide useful information for the estimation of flood losses to residential buildings. From the underlying consistent OSM data model and standardized calculation of spatial measures, we expect an improvement of the spatial transfer of flood vulnerability models across regions. Accordingly, the research objectives are (i) to understand which building-geometry-related variables are useful to describe building vulnerability, (ii) to learn predictive flood vulnerability models, and (iii) to test and evaluate model transfer across regions. In [section 2.2](#) the data sources, the derived variables and the preparation of data sets are described. [Section 2.3](#) introduces the methods to identify predictor variables and to derive predictive models. Further, it describes the set-up for testing and evaluating model performance in spatial transfers. The results from these analyses are reported and discussed in [section 2.4](#). Conclusions are drawn in [section 2.5](#).

2.2 Data

We use an empirical data set of relative loss to residential buildings and influencing factors, which has been collected via computer-aided telephone interview (CATI) data during survey campaigns after major floods in Germany since 2002. Another data source is OSM (OSM contributors, 2020), providing information about building locations, geometries, occupancy and other characteristics. OSM data are complemented with numerical spatial measures calculated from geometries of OSM building footprints. Details on data and software used in this study are listed in [Appendix A.2](#).

2.2.1 Computer aided telephone interview data

CATI surveys were conducted with affected private households ex post major floods in Germany. The regional focal points of flood impacts were the Elbe catchment in Eastern Germany and the Danube catchment in southern Germany. Particularly noteworthy are the floods of 2002 and 2013, which caused economic losses of 11.6 billion € (reference year 2005) and 8 billion € respectively in Germany (Thielen et al., 2006b; Thielen et al., 2016). With 1 billion € in economic damage, the city of Dresden at the Elbe River in Saxony had been a hotspot of flood impacts during the August 2002 flood (Kreibich & Thielen, 2009). In August 2002, flash floods triggered by record-breaking precipitation and numerous dyke failures caused widespread flooding along the Elbe River and its tributaries in Saxony and Saxony-Anhalt, as well as along the Regen River and other southern tributaries to the Danube River in Bavaria (Schröter et al., 2015). The magnitude of flood peak discharges along these rivers well exceeded a statistical return period of 100 years (Ulbrich et al., 2003). In May 2013 a pronounced precipitation anomaly with subsequent extreme precipitation at the end of May and beginning of June caused severe flooding in June 2013, especially along the Elbe and Danube rivers, with new water level records and major dike breaches both at the Elbe and Danube rivers (Conradt et al., 2012; Merz et al., 2014b; Schröter et al., 2015). The magnitude of flood peak discharges exceeded statistical return periods of 100 years along the Elbe, Mulde and Saale tributaries and along the Danube and Inn River in Bavaria (Blöschl et al., 2013; Schröter et al., 2015). With 180 questions, the CATI surveys cover a broad range of flood-impact-related factors including building characteristics, effects of warnings, precaution and the socio-economic background of households. The survey campaigns for different floods

are consistent in terms of acquisition methodology, type, and scope of questions. The interviewees were randomly selected from lists of potentially affected households along inundated streets which have been identified from satellite data, flood reports and press releases. With an average response rate of 15 %, in total 3056 interviews have been completed. For further details about the surveys and data processing, refer to (Thieken et al., 2005; Kienzler et al., 2015; Thieken et al., 2017).

Building on the findings of previous work (Merz et al., 2013; Schröter et al., 2014), for this study 23 variables have been preselected with a focus on building characteristics, flood intensity at the building and socio-economic status as well as warning, precaution and previous flood experience (Table 2.1). In addition, relative loss to the building has been determined as the ratio of reported actual losses and the building value (replacement cost) at the time of the flood event (Elmer et al., 2010). Hence, it describes the degree of building damage on a scale from 0 (no damage) to 1 (total damage). Building values are based on the standard actuarial valuation method of the insurance industry in Germany (Dietz et al., 2014), which estimates replacement costs using information about the floor space, basement area, number of storeys, roof type, etc. that are available from CATI data. Relative loss to the building and water depth (“wst”) at the building are the key variables from the CATI data set used in this study. The variable *rloss* is used to learn predictive models and to evaluate their performance. Consequently, the records in the CATI data set without values for *rloss* are removed. This reduces the number of available records from 3056 to 2203. The variable *wst* is the most commonly used predictor in flood vulnerability modelling (Gerl et al., 2016) because it is a highly relevant characteristic of flood intensity, and it is usually available from hydrodynamic numerical simulations; *wst* from CATI is a continuous variable with a length unit in centimetres. Negative values represent a water level below the ground surface, which affects only the basement of a building.

2.2.2 OpenStreetMap data

OSM is a free web-based map service built on the activity of registered users who contribute to the database by adding, editing or deleting features based on their local knowledge. The contributors use GPS devices and satellite, as well as aerial imagery, to verify the accuracy of the map. OSM is an open-data project, and the cartographic information can be downloaded, altered and redistributed under the Open Data Commons Open Database Licence (ODbL) (OSM contributors, 2020). Among the so-called volunteered geographic information (VGI) projects (Goodchild, 2007), OSM is the most widely known. OSM data provide information about building locations, footprint geometries, occupancy and other characteristics. The positional accuracy of OSM data, as well as the completeness of the database in respect to the number of mapped objects present in the real world, is nowadays considered satisfactory for most developed countries and urban areas (Hecht et al., 2013; Barrington-Leigh & Millard-Ball, 2017). On the contrary, information on object attributes such as road names or building types is often scarce and inconsistent. The tag “building” is used to identify the outline of a building object in OSM. The majority of buildings (82 %) have no further description, and only 12 % are specified as primarily “residential” or a single-family “house” (<https://taginfo.openstreetmap.org/keys/building#values>, last access: 28 February 2020). Therefore, the filtering for residential buildings from the OSM database uses the underlying “residential” land use information of OSM.

Table 2.1: Preselected variables from CATI surveys; C: continuous, O: ordinal, N: nominal scaled variables

Variable		Type and range
Warning, precaution and previous experience		
1	wt	Early warning lead time
2	wq	Quality of warning
3	ws	Indicator of flood warning source
4	wi	Indicator of flood warning information
5	wte	Lead time period not used for emergency
6	em	Emergency measures indicator
7	epre	Perception of efficiency of private precaution
8	pre	Precautionary measures indicator
9	fe	Flood experience indicator
10	kh	Knowledge of flood hazard
Hydraulic characteristics of the inundation		
11	wst	Water depth
Building characteristics		
12	bt	Building type
13	nfb	Number of flats in building
14	fsb	Floor space of building
15	bq	Building quality
16	bv	Building value
Socio-economic status of the residents		
17	age	Age of the interviewed person
18	hs	Household size, i.e., number of persons
19	chi	Number of children (<14 years) in household
20	eld	Number of elderly persons (>65 years) in household
21	own	Ownership structure
22	inc	Monthly net income in classes
23	socP	Socio-economic status according to Plapp (2003)
Experienced damage		
-	rloss	Relative loss of the residential building

By joining the land use information to the building polygons, those of residential occupation can be identified and selected.

2.2.3 Data preparation

The OSM and CATI data sets have been conflated in order to link the empirically observed variables *rloss* and *wst* with OSM data for individual residential buildings. This operation uses the geolocation information of both data sources. The CATI data are provided with address details including community, postal code, street name and the house number ranges in blocks of five numbers. Geocoding algorithms including open web API (application programming interface) services like Google (<https://developers.google.com/maps/documentation/geolocation/overview>, last access: 3 February 2021), Photon (<https://photon.komoot.io/>, last access: 3 February 2021) and Nominatim (<https://nominatim.org/>, last access: 3 February 2021) were applied to obtain geocoordinates for the address information from the interview data.

OSM is a spatial data set including georeferenced building outlines. The geolocated interviews are spatially matched with OSM building polygons using an overlay operation which merges interview points with OSM building polygons. In view of limited address details regarding the building house number ranges and inherent inaccuracies of geocoding databases and algorithms (Teske, 2014), a buffer radius of 5 m has been used to correct for offsets between geocoding points and building polygons. CATI records which still could not be matched with OSM geometries and with obviously erroneous geolocations, e.g., position is far away from flood-affected areas or urban settlements, have been removed from the data set. After these steps, 1649 records remain from the original set of CATI surveys. The spatial distribution of these data points highly concentrates on the Elbe catchment (1234 records) including Dresden (310 records) and on the Danube catchment (105 records) (Figure 2.1).

2.2.4 Numerical measures

Information about the building geometry is useful to support the estimation of flood losses to residential buildings (Schröter et al., 2018). Building on this knowledge, numerical spatial measures are calculated for OSM building footprints with the aim to add potential explanatory variables to the estimation of relative loss to residential buildings. For this purpose, image analysis algorithms typically used in landscape ecology are adopted. These algorithms calculate numerical spatial measures like area, perimeter, elongation, and complexity based on the analysis of geometries identified in aerial or remote sensing images (Lang & Tiede, 2003; Jung, 2016; Rusnack, 2021). The numerical spatial measures are calculated for each OSM building polygon and are compiled in Table 2.2 along with the other CATI variables that are used to derive flood vulnerability models. The meaning of these spatial measures, the equations, and range of values and examples are listed in the Appendix A1 (Figure A.1).

2.3 Methods

We analyse the created data set with two main objectives. First, we strive to identify those variables from Table 2.2 which are most useful for explaining relative loss to residential buildings. Second, we aim to derive flood vulnerability models for residential buildings and to test these models for spatial transfers across regions. The



Figure 2.1: Regional subdivision of the data set for spatial split-sample testing (Dresden municipality, the Elbe catchment and the Danube catchment).

Table 2.2: Variables of the amended OSM data set for each building object

Empirical variables from the CATI interviews		Range	
—	Relative loss of the residential building (rloss)	Relative loss	0 = no damage to 1 = total damage
—	Water depth (wst)	Water level in respect to ground level	-248 cm to 670 cm
Numerical spatial measures calculated for OSM building geometries		Range	
1	Area (Area)	Area of the building	0 m ² to ∞
2	Perimeter (Perimeter)	Perimeter of the building	0 m to ∞
3	Degree of compactness (DegrComp)	Compactness of the building shape, relative vicinity of the internal points, normalized to a circle	0 to 1
4	Perimeter-area ratio (PARatio)	Shape complexity, biased by building size	0 to ∞
5	Shape index (ShapeIndex)	Shape complexity, adjusted to building size normalized to a square	1 to ∞
6	Fractal dimension index (FracDimInd)	Shape complexity, adjusted to building size scaled between	1 to 2
7	Radius of gyration (RadGyras)	Building extent and compactness	0 m to ∞
8	Linear segment indicator (LinSegInd)	Elongation of the polygon, normalized to a square	1 to ∞
9	Ratio of bounding rectangle area (BoundRatio)	Shape complexity, normalized to the hypothetical simplest polygon	1 to ∞

workflow for data analysis including data pre-processing, model learning, model selection and model transfer is illustrated in [Figure 2.2](#). The data pre-processing steps

with data preparation and numerical spatial measures have been described in the previous section. For model learning and model transfer, we use the random forest (RF) machine learning algorithm introduced by (Breiman, 2001).

RF is an extension of the classification and regression tree (CART) algorithm, (Breiman, 1998) which aims to identify a regression structure among the variables in the data set. Regression trees recursively subdivide the space of predictor variables to approximate a non-linear regression structure. This subdivision is driven by optimizing the accuracy of local regression in these regions, which, by repeated partitioning, leads to a tree structure. Predictions are made by following the division criteria along the nodes and branches from the root node to the leaves, which finally contain the predicted value for a given set of input variables. RFs make predictions based on numerous decision trees, i.e., a forest, which is learned by randomly selecting the variables considered for splitting the features space of the data. RFs incorporate bootstrap aggregation (bagging) as a simple and powerful ensemble method to reduce the variance of the CART algorithm. In comparison to single trees, RFs are more suitable to identify complex patterns and structures in the data (Basu et al., 2018). As an ensemble approach, RFs learn a regression tree for a number of bootstrap replicas of the learning data. This results in a number of trees (*ntree*) forming a forest of regression trees. To reduce correlation between trees, the RF algorithm randomly selects a subset of variables (*mtry*) which are evaluated for dividing the space of predictor variables. This efficiently reduces overfitting and makes RF less sensitive to changes in the underlying data. Each bootstrap replica is created by randomly sampling with replacement about two-thirds of observations from the original data set. The remaining data are indicated as out of-bag (OOB) observations and are used for evaluating the predictive accuracy of the tree, in terms of the OOB error. For regression trees, the OOB error is the mean squared sum of residuals. For loss estimation, the predictions of all trees are combined by aggregating the individual predictions as the mean prediction from the forest. The predictions of the individual trees, i.e., from the ensemble of models, provide an estimate of predictive uncertainty.

For variable selection and predictive model learning, RFs provide a concept to quantify the importance of candidate explanatory variables, which allow for selecting the subset of most relevant variables. RFs are also an efficient algorithm to learn predictive models from heterogeneous data sets with complex interactions and with different scales like continuous or categorical information (Huang & Boutros, 2016).

RF predictive model performance is sensitive to specifications of the algorithm parameters *mtry* and *ntree* (Huang & Boutros, 2016). Therefore, the optimum values for both parameters are identified as those which yield minimum OOB errors on an independent data set. For parameter tuning, we pursue the variation approach implemented by (Schröter et al., 2018) by selecting parameters from a broad and comprehensive range of values $ntree \in [100, 500, 1000, 2000, 3000, \dots, 15000]$ and $mtry \in [p/6, p/3, 2p/3]$ with p as the number of candidate predictors, and derive RF models for each combination. For each pair of chosen values, the algorithm is repeated 100 times to account for inherent data variability. The optimum parameters will minimize the prediction error on the OOB sample data. Using the optimum RF parameter settings, we derive predictive models for *rloss*.

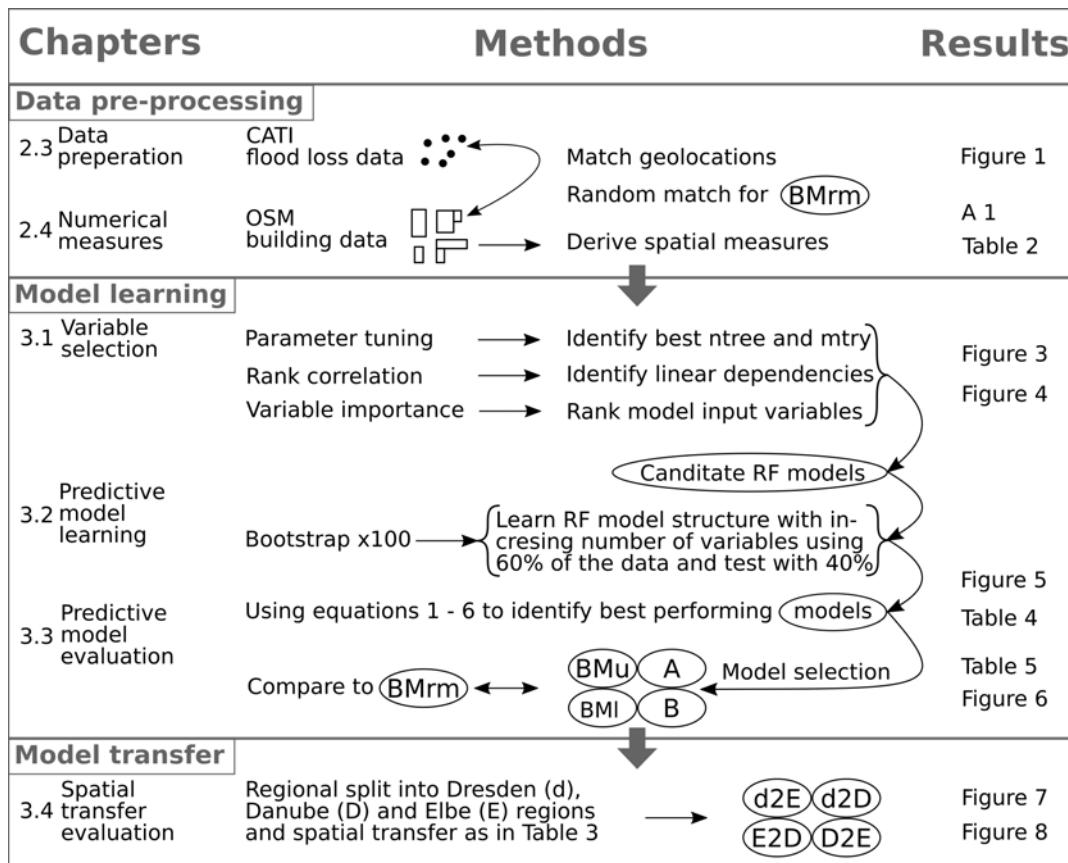


Figure 2.2: Data pre-processing, model learning and model transfer workflow, with **BMu** (upper benchmark model), **BMI** (lower benchmark model), **BMrm** (Benchmark model with random match of interview locations with OSM building data), **A** (Random Forest model using eight predictors), **B** (Random Forest model using eight predictors), and model transfers **d2E** (learning with Dresden and predictions for Elbe), **d2D** (learning with Dresden and predictions for Danube), **E2D** (learning with Elbe and predictions for Danube), **D2E** (learning with Danube and predictions for Elbe)

2.3.1 Variable selection

The first step in model learning is the selection of variables to be used as predictors in the model. The analysis of the Spearman's rank correlation between the variables gives a first insight into the linear dependency structure of the data set. Furthermore, RF supports the evaluation and ranking of potential predictors by quantification of variable importance, which also accounts for variable interaction effects. The importance of a selected variable is evaluated by calculating the changes of the squared error of the predictions when the values of that variable are randomly permuted in the OOB sample. The increase of the average error will be larger for more important variables and smaller for less important variables. On this basis, it is possible to decide which variables to include in a predictive model. The outcomes of variable importance evaluations are sensitive to the RF algorithm parameters *mtry* and *ntree* (Genuer et al., 2010). Therefore, to achieve stable results for these analyses, we implement a robust approach which averages the outcomes of multiple runs with variations in RF parameters (Schröter et al., 2018): $ntree \in [500, 1000, 1500, 2000, \dots, 5000]$ whereby each tree is repeatedly built for $mtry \in [p/6, p/3, 2p/3]$, with p as number of candidate predictors, which correspond to the lower limit, the default

value and the upper limit, suggested by (Breiman, 2001). Following this procedure, the potential explanatory variables of our data set (Table 2.2) are evaluated and ranked according to their relative importance to predict $rloss$.

2.3.2 Predictive model learning

Variable selection needs to be considered as an essential part of the model evaluation process. Therefore, candidate RF models using different numbers of variables are assessed in terms of predictive performance for independent data.

The OSM-based numerical spatial measures differentiate building form and shape complexity. To gain further insights into the suitability of these variables for flood vulnerability modelling, we incrementally add explanatory variables to the learning data set. Based on the outcomes of variable importance ranking, the learning set is expanded variable by variable, and models of increasing complexity are learned (cf. Table 2.2). From the comparison of model predictive performance between these candidate models, the best balance between model performance and number of input variables is assessed. This is implemented by bootstrapping the splitting of the data into sub-sets for learning (60 %) and testing (40 %) with 100 iterations.

Further, for an independent assessment of OSM-based vulnerability model performance, we consider two benchmark models. We argue that the set of CATI variables (Table 2.1) represents the most detailed data set available for flood loss estimation of residential buildings (Merz et al., 2013; Schröter et al., 2014; Thielen et al., 2016). Therefore, a RF model is learned using all 23 CATI predictors as an upper benchmark (BMu). In contrast, a RF model using only wst as a predictor is learned as a lower benchmark. The reasoning is that using extra variables in addition to wst will improve the predictive performance of the models (Schröter et al., 2016; Schröter et al., 2018). As described in subsection 2.2.3, the detail of geolocation information from CATI data is limited to ranges of house numbers. Therefore, we face uncertainty in whether CATI data and OSM building footprints have been matched correctly. To assess the potential implications of this source of uncertainty, we derive a model (BMrm) which is based on a data set with $rloss$ and wst observations randomly assigned to OSM building footprints. We keep the RF modelling approach for the benchmark models consistent to ensure that any observed difference in model performance stems from differences in the underlying input variables.

2.3.3 Predictive model evaluation

Model predictive performance is evaluated by comparing predicted (P) and observed (O) $rloss$ values from the validation sample using the following metrics. In these metrics, RF predictions are evaluated for the median prediction (P_{50}) derived from the ensemble of individual tree predictions.

Mean Absolute Error (MAE) quantifies the precision of model predictions, with smaller values indicating higher precision:

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_{50_i} - O_i| \quad (2.1)$$

Mean bias error (**MBE**) is a measure of accuracy, i.e., systematic deviation from the observed value. Unbiased predictions yield a value of 0; underestimation results in negative; and overestimation in positive values:

$$MBE = \frac{1}{n} \sum_{i=1}^n (P_{50_i} - O_i) \quad (2.2)$$

Mean Squared Error (**MSE**) combines the variance of the model predictions and their bias. Again, smaller values indicate better model performance:

$$MSE = \frac{1}{n} \sum_{i=1}^n (P_{50_i} - O_i)^2 \quad (2.3)$$

The ensemble of model predictions from the RF models offers insight into prediction uncertainty. This property is analysed by evaluating the 90 % quantile range, i.e., the difference between the 5 % quantile and 95 % quantile in relation to the median, as a measure of ensemble spread:

$$QR_{90} = \frac{1}{n} \sum_{i=1}^n (P_{95_i} - P_{5_i}) / P_{50_i} \quad (2.4)$$

with 95 % quantile, 5 % quantile and the 50 % quantile, i.e., the median of the predictions. QR_{90} (Quantile Range) is a measure of sharpness, with smaller values indicating a smaller prediction uncertainty.

Reliability of model predictions is quantified in terms of the hit rate (**HR**) (Gneiting & Raftery, 2007):

$$HR = \frac{1}{n} \sum_{i=1}^n h_i ; h_i = \begin{cases} 1, & \text{if } O_i \in [P_{95_i}, P_{5_i}] \\ 0, & \text{otherwise} \end{cases} \quad (2.5)$$

HR calculates the ratio of observations within the 95 % — 5 % quantile range of model predictions. For a reliable prediction HR should correspond to the expected nominal coverage of 0.9.

HR and QR_{90} are combined to the interval score (**IS**) which accounts for the trade-off between HR values and QR_{90} ranges (Gneiting & Raftery, 2007):

$$IS = QR_{90} + \frac{1}{n} \sum_{i=1}^n \frac{2}{\beta} (P_{05_i} - O_i) | \{O_i < P_{05_i}\} + \frac{2}{\beta} (O_i - P_{95_i}) | \{O_i > P_{95_i}\} \quad (2.6)$$

2.3.4 Spatial-transfer evaluation

We investigate whether the consistent data basis of OSM-derived numerical spatial measures supports the transfer of flood vulnerability models across regions by splitting the available data set into subsets for different regions affected by major floods. The CATI data are mainly located in the Elbe and Danube catchments in Germany, which are the regions mostly affected by inundations and flood impacts. This suggests a regional subdivision of the empirical data set according to these

river basins for the investigation of spatial model transfer. In detail, we partition the data set between the metropolitan area of Dresden (Saxony), the Elbe catchment (Saxony, Saxony-Anhalt, and Thuringia) and the Danube catchment (Bavaria and Baden-Württemberg); see Figure 2.1. This split is applied irrespective of the CATI survey campaign year, and thus the regional subsets contain records from different flood events. The idea is to investigate examples with a small set of learning data for a small specific region (Dresden), a large learning data set from an extended region (Elbe catchment) and a small set of learning data from an extended region (Danube catchment). The details for the learning and transfer applications are listed in Table 2.3. For these three regions, we learn RF models using the selected variables and assess their predictive performance when transferred to the other regions. As we use a completely independent data set for model transfer testing, no additional bootstrap on top of RF internal bootstrapping is required.

Table 2.3: Computational experiments for transfer applications

Transfer experiment	Implementation	Learned on/applied to # buildings
d2E	Learned from Dresden and applied to Elbe	310/1234
d2D	Learned from Dresden and applied to Danube	310/105
E2D	Learned from Elbe and applied to Danube	1234/105
D2E	Learned from Danube and applied to Elbe	105/1234

2.4 Results and Discussion

Random forest OOB errors are sensitive to the choice of RF parameters $mtry$ and $ntree$. From the variation of RF parameters, we observe that OOB errors decrease with smaller values for $mtry$ and larger numbers of trees in a forest $ntree$; see Figure 2.3.

The coloured bands represent the 90 % quantile range of OOB values from the 100 bootstrap repetitions for each RF algorithm configuration and illustrate the inherent variability of input variables in the learning data set. The colour code distinguishes the number of variables used to determine splits at each node ($mtry$). For $mtry = 2$ the smallest OOB errors are achieved throughout the variations in the number of trees ($ntree$). This value represents the lower bound of recommended values for $mtry$ in RF regression models (Breiman, 2001). For smaller values of $mtry$ fewer variables are considered for splitting the space of predictor variables, which reduces the correlation between individual trees of the forest. Further, increasing values of $ntree$ asymptotically approximate smaller OOB values. It appears that for the given data set, OOB values are virtually stable above $ntree = 7000$. As the computational effort increases with larger forests, it has to be balanced with improvements regarding predictive performance. Building on these results, we use RF parameters $mtry = 2$ and $ntree = 7000$, which are comparable to those used by (Schröter et al., 2018).

2.4.1 Variable selection and predictive model learning

The numerical spatial measures (Table 2.2, and Appendix A1) evaluate properties of the building footprints including area, perimeter, and elongation of main building axes. Accordingly, some of these variables are strongly correlated (Figure 2.4). The Spearman’s rank correlation matrix of the variables confirms a high degree of correlation in the data set, as for instance between Area, Perimeter and RadGyras. In contrast, the spatial measures are only slightly correlated with wst and $rloss$. The presence

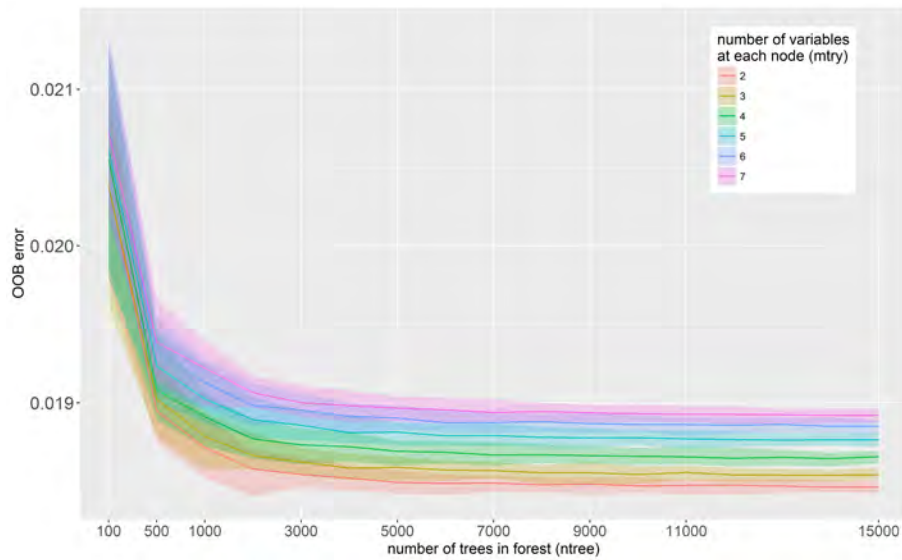


Figure 2.3: Out-of-Bag error for variations of mtry and ntree RF parameters. Colour bands represent the variation range of OOB errors obtained from 100 bootstrap repetitions

of multicollinearity may influence the analysis of variable importance (Gregorutti et al., 2017). The robust importance analysis uses different RF parameter settings and reports an average importance rank, which alleviates this problem.

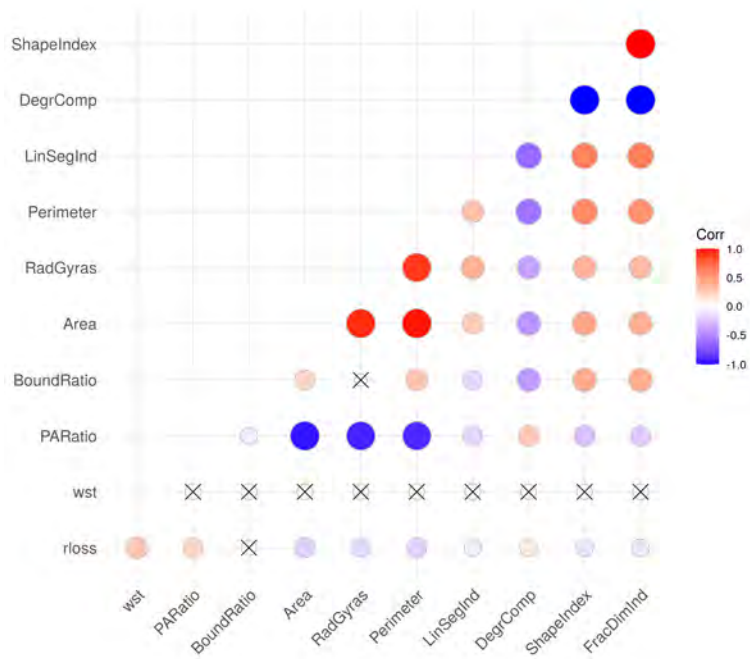


Figure 2.4: Spearman’s correlation of model variables (significance level 1 %), non-significant correlations are crossed out.

The variable *wst* ranks first in the importance analysis (results not shown) which confirms common knowledge in flood loss modelling (Smith, 1994; Gerl et al., 2016).

In comparison to *wst*, the numerical spatial measures of OSM building footprints have clearly smaller importance values, with relatively small differences between them. In terms of building characteristics, both spatial measures which express the size and extension of the building (e.g., Area and Perimeter) and spatial measures which describe building compactness and shape complexity (e.g., PARatio, RadGyras, LinSegInd, and BoundRatio) seem to add information to better estimate relative building loss. The following order of importance was determined for the variables: *wst*, *PARatio*, *RadGyras*, *Area*, *LinSegInd*, *BoundRatio*, *Perimeter*, *DegrComp*, *FracDimInd*, *ShapeIndex*. Predictive performance tests for models with two to ten variables (Figure 2.5, Table 2.4) build on this order of importance.

However, the outcome of the variable importance analysis does not suggest a clear selection of features to be included in a predictive flood vulnerability model. The model predictive-performance-based assessment of variables uses an increasing number of variables following their ranking order of variable importance in the RF modelling. The predictive performance is quantified in terms of MAE, MBE and MSE (Equation 2.1, 2.2 and 2.3) for 100 bootstrap repetitions. While the MAE is decreasing when additional variables are used with an overall minimum for a model using six variables, including more than six variables tends to increase MAE again (Figure 2.5). However, regarding MBE these changes go in an opposite direction. We observe the smallest MBE when only two variables are included. MBE then grows continuously for using up to seven variables and then slightly reduces when more variables are used. The increase in precision expressed by the smaller MAE is accompanied by a reduction in accuracy reflected by an increasing MBE. This yields an almost-balanced performance in terms of MSE for all models tested.

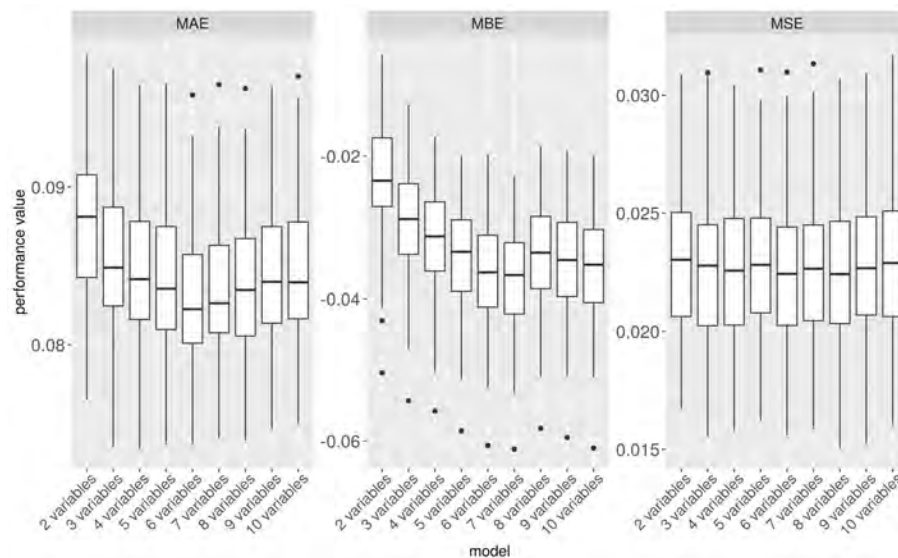


Figure 2.5: Predictive performance of models using an increasing number of variables in order of their importance. Smaller MAE and MSE values and MBE values close to 0 indicate better performance; cf. Equation 2.1, 2.2 and 2.3.

Looking into the sharpness of model predictions, the quantile range (QR_{90}) is getting larger with an increasing number of model variables, which reflects larger uncertainty (Table 2.4). In terms of model reliability (HR), an increasing number of model variables achieves better performance statistics, up to using eight variables. The combination of both, QR and HR , in the interval score (IS) shows a similar pattern.

Table 2.4: Model performance metrics for models using an increasing number of variables including arranged in order: *wst*, *PARatio*, *RadGyras*, *Area*, *LinSegInd*, *BoundRatio*, *Perimeter*, *DegrComp*, *FracDimInd*, *ShapeIndex*. Best performance values and selected models in bold.

Model	MAE	MBE	MSE	QR	HR	IS
2 variables	0.0878	-0.0234	0.0230	0.2765	0.5864	7.9402
3 variables	0.0853	-0.0293	0.0226	0.2992	0.6301	7.1154
4 variables	0.0843	-0.0316	0.0224	0.3070	0.6433	6.8440
5 variables	0.0840	-0.0348	0.0227	0.3182	0.6533	6.7166
6 variables	0.0826	-0.0364	0.0222	0.3270	0.6622	6.5728
7 variables	0.0830	-0.0373	0.0225	0.3302	0.6614	6.5715
8 variables	0.0839	-0.0337	0.0224	0.3314	0.6640	6.3757
9 variables	0.0841	-0.0349	0.0226	0.3346	0.6639	6.3766
10 variables	0.0844	-0.0357	0.0228	0.3365	0.6631	6.4000

On the basis of these assessments, two model alternatives are selected for further analysis: model A using eight variables, as it provides the most reliable model predictions, and model B using six variables, which provide the highest precision and balance between accuracy and precision. In detail, Model B uses the variables *wst*, *PARatio*, *RadGyras*, *Area*, *LinSegInd*, and *BoundRatio*. Model A, in addition, uses *Perimeter* and *DegrComp* as predictors.

2.4.2 Model predictive performance: model benchmarking

The OSM models A and B are benchmarked with a model that uses all information available from the CATI surveys as an upper benchmark (**BMu**) and a model that uses only water depth as predictor as a lower benchmark (**BMI**). The performance statistics achieved by models A and B for the complete data set (all events and regions) are slightly inferior to BMu but clearly better than the outcomes of BMI (Figure 2.6). Both models A and B give very similar performance statistics with slightly higher precision (smaller MAE) but larger bias (MBE) for model B. In contrast, model A provides more reliable predictions indicated by larger HR and smaller IS (Table 2.5). The randomized benchmark model (**BMrm**) achieves a better performance than BMI but is inferior to the models A and B (Figure 2.6, Table 2.5). Hence, we are confident that the remaining uncertainty associated with the mapping of geolocations to building geometries does not affect the outcomes of our analyses. Overall, we note that including numerical spatial measures based on OSM building footprints add useful information to predict loss to residential buildings. The numerical spatial measures included in the models are all directly calculated using building footprints. Therefore, a larger number of variables used for loss estimation does not imply increased efforts to collect data. From this perspective, the cost of using model A or B is equal. The RF algorithm strives to reduce overfitting when large numbers of predictors are included, and thus the parsimonious modelling principle can be relaxed. A possible negative effect of overfitting when using more predictors should manifest in spatial-transfer applications.

2.4.3 Spatial-transfer testing

The predictive performance of RF models is tested in regional-transfer applications. For this purpose, the RF models A and B as well as the benchmark models BMu and BMI, as specified in the previous section, are learned using regional subsets of the

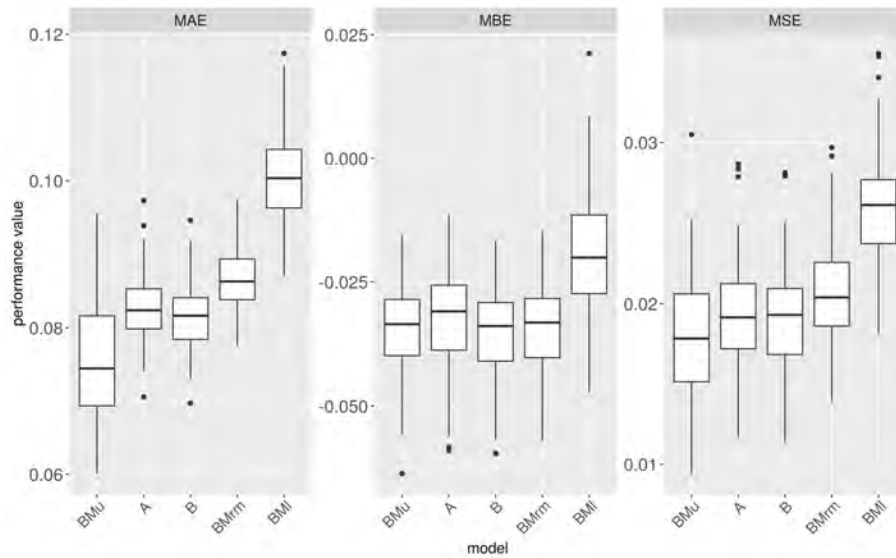


Figure 2.6: Performance metrics of OSM based models and benchmark models

Table 2.5: Model precision, accuracy, and reliability performance metrics for OSM based and benchmark models

Model	MAE	MBE	MSE	QR	HR	IS
BMu (Upper benchmark, all 23 predictors from CATI interviews)	0.075	-0.034	0.018	0.336	0.733	3.573
A (7 numerical spatial measures derived from OSM plus water depth)	0.083	-0.032	0.019	0.322	0.699	6.022
B (5 most important numerical spatial measures plus water depth)	0.081	-0.035	0.019	0.319	0.698	6.238
BMrm (random match of CATI geolocation with OSM building polygons)	0.087	-0.034	0.021	0.319	0.688	6.535
BMI (Lower benchmark, only water depth as predictor)	0.100	-0.019	0.026	0.177	0.490	10.107

data and applied to predict flood losses in a different region; see [subsection 2.3.4](#) and [Table 2.3](#) for details about the regional subdivision of data and spatial-transfer experiments. Learning models with a regional subset of data and applying the models to other regions results in a drop of predictive performance in comparison to the case when the entire data set is used for model learning, except for the case of d2E ([Figure 2.7](#)). In most of the learning or transfer cases, BMu scores best in terms of precision and reliability, represented by the performance metrics MAE, MSE, HR and IS. Using only wst as a predictor (BMI) produces less precise and less reliable predictions, as indicated by larger MAE and MSE, as well as smaller HR and larger IS. While the performance of models A and B is very similar, model A, using eight predictors, more reliably predicts residential loss (larger HR and smaller IS), and model B, using six predictors, provides more accurate (MBE closer to 0) and more precise predictions (smaller MAE and MSE). Hence, overfitting does not seem to be an issue when more input variables are used. In contrast to the model benchmark comparison ([subsection 2.3.4](#)) BMu and BMI do not entirely frame the RF model performance values. Instead, models A and B in some cases achieve better and in other cases worse performance statistics.

Generally speaking, the predictive performance differs more strongly between the regional-transfer settings than between the models ([Figure 2.7](#)). This is more pronounced for precision and accuracy metrics (MAE, MBE and MSE) than for sharpness

and reliability indicators (QR, HR and IS). Learning from the Dresden subset and transferring the model to the Elbe region (d2E) works best, as is shown by the smallest MAE and MSE as well as an MBE closest to 0. Learning the models with the Danube subset and transferring them to the Elbe region (D2E) yields comparably small MAE and MSE values, but this is also the only case with a tendency to overestimate loss resulting in a positive MBE. The models are struggling most to predict loss when they are learned with the Dresden subset and transferred to the Danube region (d2D), showing the lowest precision and accuracy. In turn, extending the learning subset to the Elbe region improves the transfer to the Danube (E2D). Concerning predictive uncertainty and reliability, learning with the Danube subset yields large QRs, which however only partly cover the observed loss values reflected in comparably low HR and high IS (D2E). Learning from Dresden or Elbe and transferring to Elbe or Danube (d2E, d2D and E2D) produces sharper predictions, but still the models differ in reliability, i.e., covering the observed values within their predictive uncertainty ranges (HR). In this respect, the upper benchmark model (BMu) performs best. The differences between models A and B are small, and both are better than the lower benchmark model (BMi) and almost similar to BMu for the transfer cases between the regions Elbe and Danube (E2D and D2E).

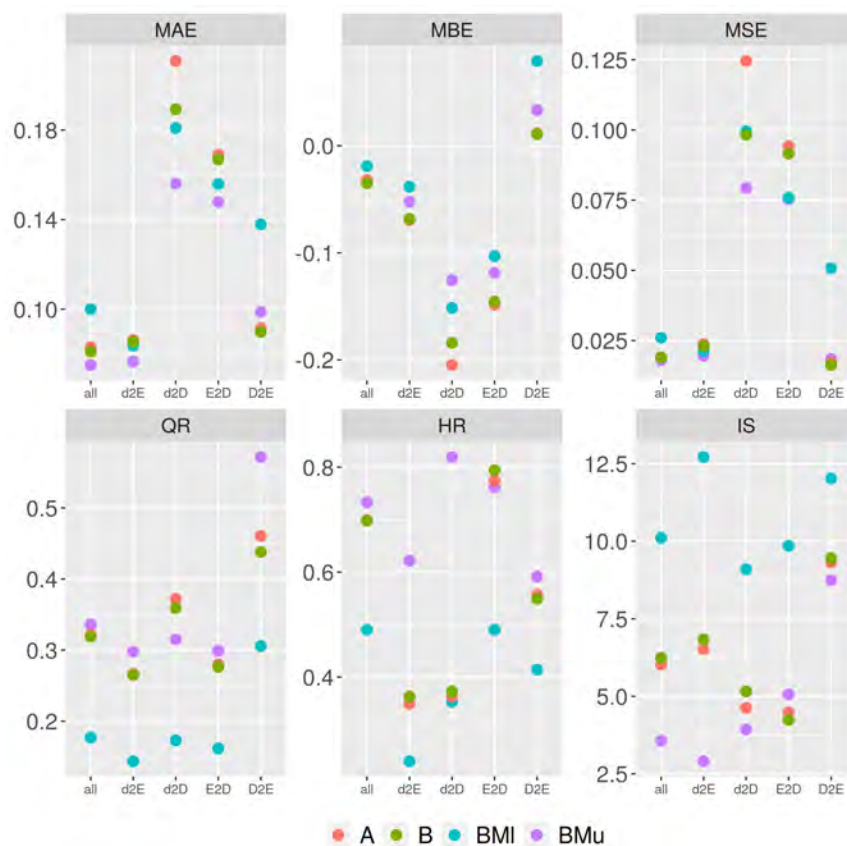


Figure 2.7: Model performance metrics in regional transfer. Models A and B based on spatial numerical measures calculated for OSM building footprints, benchmark models BMi and BMu based on CATI survey data. Transfer experiments d2E, d2D, E2D, and D2E as described in Table 2.3. 'all' refers to using all records from all regions, cf. Table 2.5.

With 105 records, the Danube data set is the smallest subsample. It has a smaller variability and range of values for most numerical spatial measures in comparison to

the Dresden and Elbe regional sets (Figure 2.8).

The geometric properties of the flood-affected residential buildings in the Danube region seem to differ from the affected residential buildings in the Elbe region. In the Danube subset, the area and perimeter of buildings tends to be smaller than in the Elbe region. Also, the values for spatial measures representing building shape complexity, for instance RadGyras, DegrComp and BoundRatio, indicate more compact building footprints in the Danube region than in the Elbe region. These differences can be attributed to different socio-economic characteristics as well as building practices in former East and West Germany and regional differences in building types (Thieken et al., 2007). With only 310 records, the Dresden sub-sample covers comparable ranges of observed variables as the Elbe sub-set (1234 records). Both sub-sets show largely similar relations between individual variables and $rloss$. Still, the Danube sub-set includes relatively many records with high $rloss$ values, which are distributed along the whole spectrum of above ground-level water depths (Figure 2.8). In comparison, the Dresden subset consists of very few cases with high relative loss, which is partly related to differing inundation processes. In the Elbe and Danube catchments, large areas have been flooded as a consequence of dyke failures. Hence, the relationship of model variables to high $rloss$ values cannot be learned from this sub-set, and thus is not represented well by the model. Therefore, this difference in the learning data may explain the positive bias introduced by learning the model in the Danube and transferring it to the Elbe and, vice versa, the pronounced negative bias introduced by learning the model in Dresden and transferring it to the Danube region. Viewed from a model performance perspective, the transfer applications show that a good agreement between learning and transfer data sets (e.g., d2E) produces more precise and reliable predictions than the transfer to regions with pronounced differences (e.g., d2D and D2E). Still, from the Danube region with limited ranges of variable values, it is possible to obtain relatively precise and accurate predictions of relative building loss. This suggests that a broad variability of observed $rloss$ values in the learning data set is an important control for the predictive capability of the model in other regions. In contrast, small samples with limited variability and only few records with high $rloss$ values struggle with predicting $rloss$ in other regions. This confirms insights that a model based on more heterogeneous data performs better when transferred in space (Wagenaar et al., 2018).

Our findings also reveal that using numerical spatial measures derived from OSM building geometries does not resolve all problems of model transfer. As not many variables of building characteristics are available from OSM data, the spatial measures calculated from building footprints serve as proxy variables for these unavailable details. These proxies achieve comparable predictive performance as specific property level data sets, as for instance collected via computer-aided telephone interview surveys represented by the BMu model. This model uses a broad range of variables to characterize vulnerability of residential buildings, including details of building characteristics; socio-economic status of the household; and flood warning, precaution and previous flood experience (cf. Table 2.1). Still, this more comprehensive information does not result in a clearly better model predictive performance in transfer applications. Additional improvements can be expected from including local expert knowledge about inundation duration, flood experience and return period of the event into the modelling process (Sairam et al., 2019b). Flood-event-related variables including flood type appear to be important information for estimating the degree of building loss because they describe differences in the damaging processes (Vogel et al., 2018). Other data sources have been used to enrich empirical datasets

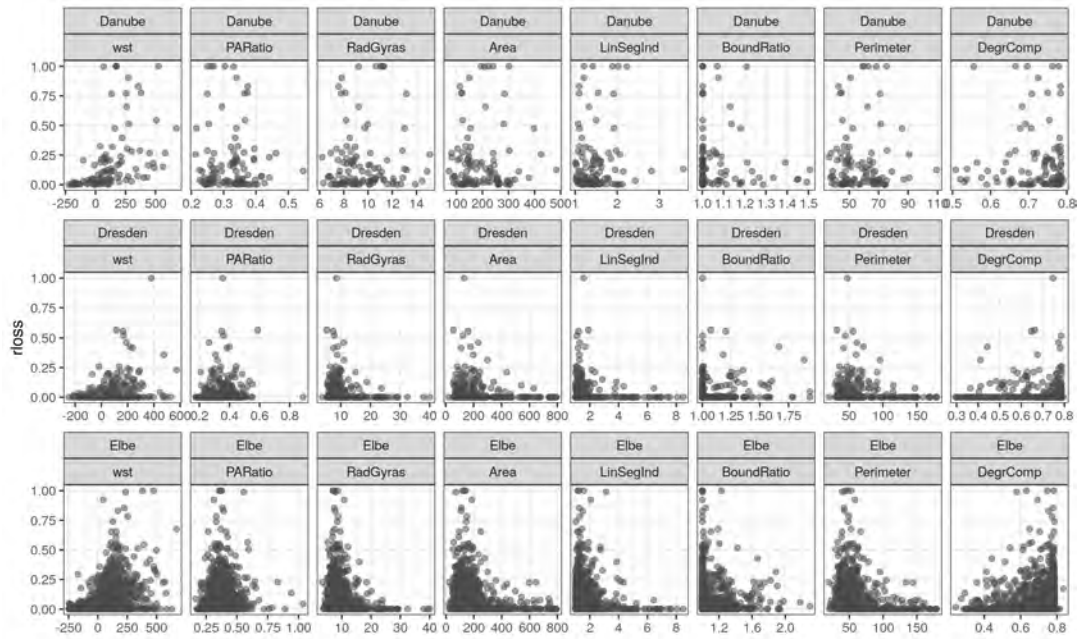


Figure 2.8: Scatter plots of numerical spatial measures and relative loss in regional sub-samples (Danube, Dresden, Elbe)

for learning flood loss models. This includes for instance information about building age and floor area for living from cadastral data (Wagenaar et al., 2017), number of storeys, building type, building structure, finishing level and conservation status from census data (Amadio et al., 2019). However, using these data did not result in a clear improvement in spatial model transfer. Using variables derived from OSM data increases the flexibility of the models to be applied in other regions because the accessibility and availability of OSM data reduces the effort of data collection, simplifies the preparation of input variables and ensures consistency of input data. The latter point is an important advantage because achieving consistency of input data has been stressed to cause large efforts in model transfers (Jongman et al., 2012a; Molinari et al., 2020). The suggested RF models are based on an ensemble approach and thus provide a view to the predictive uncertainty of the model outputs. We have shown this to be a valuable detail in assessing the reliability of model predictions in spatial transfers. In cases where model performance cannot be tested with local empirical evidence, using model ensembles has been shown to provide more skilful loss estimates (Figueiredo et al., 2018).

2.5 Conclusions

The transfer of flood vulnerability models to regions other than those for which they have been developed for often comes with reduced predictive performance. In this study we investigated the suitability of numerical spatial measures calculated for residential building footprints, which are accessible from OpenStreetMap, to predict flood damage. Further, we tested potential benefits from using this widely available and consistent input data source for the transfer of vulnerability models across regions. We develop a new data set based on OpenStreetMap data, which comprises variables representing building footprint dimensions and shape complexity, and we devise novel flood vulnerability models for residential buildings.

The geometric characteristics of building footprints serve as proxy variables for building resistance to flood impacts and prove useful for flood loss estimation. These model input variables are easily extracted by an automated process applicable to every type of building polygon. Hence, the models can be applied to areas where information about the footprint geometry of residential buildings is available. Also, other data sources, e.g., cadastral data or data derived from remote sensing, can be used besides the OpenStreetMap data source. While the variables derived from building footprints ensure consistency and support transferability of models, the models remain context specific and should only be transferred to regions with comparable building geometric features as the learning data set.

The vulnerability models have been validated using empirical data of relative loss to residential buildings. Further, a benchmark comparison of the models has been conducted in spatial-transfer applications. The models give comparable performance to alternative multi-variable models, which use comprehensive and detailed information about preparedness, socio-economic status and other aspects of building vulnerability. In comparison to a model which uses only water depth as a predictor, they reduce model prediction errors (MAE by 20 % and MSE by 25 %) and increase the reliability of model predictions by a factor of 1.4.

OpenStreetMap is a highly popular and evolving data source with constantly increasing completeness and up-to-date data. In the future, the attributes of residential buildings are expected to provide additional details which are of interest for the characterization of building resistance to flooding. This includes for instance information about the building type, roof type, number of floors and building material and opens up further possibilities to refine the variables used for vulnerability modelling. These data could be further amended with other open-data sources, including socio-economic statistical data. In view of a large variability of flood loss on the individual-building level, vulnerability modelling for individual buildings remains challenging and is subject to large uncertainty. Advances in the understanding of damage processes and the improvement of flood vulnerability modelling hence require an improved and extended monitoring of flood losses.

Chapter 3

A Consistent Approach for Probabilistic Residential Flood Loss Modelling in Europe

Abstract

In view of globally increasing flood losses, a significantly improved and more efficient flood risk management and adaptation policy are needed. One prerequisite is reliable risk assessments on the continental scale. Flood loss modelling and risk assessments for Europe are until now based on regional approaches using deterministic depth-damage functions. Uncertainties associated with the risk estimation are hardly known. To reduce these shortcomings, we present a novel, consistent approach for probabilistic flood loss modelling for Europe, based on the upscaling of the Bayesian Network Flood Loss Estimation MOdel for the private sector, BN-FLEMOps. The model is applied on the meso-scale in the whole of Europe and can be adapted to regional situations. BN-FLEMOps is validated in three case studies in Italy, Austria, and Germany. The officially reported loss figures of the past flood events are within the 95 % quantile range of the probabilistic loss estimation, for all three case studies. In the Italian, Austrian, and German case studies, the median loss estimate shows an overestimation by 28 % (2.1 million €) and 305 % (5.8 million €) and an underestimation by 43 % (104 million €), respectively. In two of the three case studies, the performance of the model improved, when updated with empirical damage data from the area of interest. This approach represents a step forward in European wide flood risk modelling, since it delivers consistent flood loss estimates and inherently provides uncertainty information. Further validation and tests with respect to adapting the model to different European regions are recommended.

Published as:

Lüdtke, S., Schröter, K., Steinhausen, M., Weise, L., Figueiredo, R., & Kreibich, H. (2019): A Consistent Approach for Probabilistic Residential Flood Loss modelling in Europe. *Water Resources Research*, 55(12), 10616–10635. <https://doi.org/10.1029/2019WR026213>

3.1 Introduction

Floods are a global hazard with high socio-economic impacts. In Europe, floods caused about 1,000 fatalities and 52 billion € overall losses between 1998 and 2009 (European Environment Agency, 2010). The flood in central Europe in 2002 caused the largest economic losses, with over 20 billion € (European Environment Agency, 2010). Continuous effort is necessary to further reduce flood risks. The basis of efficient flood risk management is a reliable risk assessment on various spatial scales (Merz et al., 2010a; de Moel et al., 2015). Continental flood risk analyses are important for the (re)insurance industry to assess accumulation risk and manage their risk portfolios, the financial sector to rate creditworthiness for investments (Kron, 2005), and for multinational companies to identify possible risks in their supply chains. Furthermore, European wide flood risk assessments are essential for governments to support climate change adaptation policies (van Renssen, 2013) and to manage the European Union solidarity fund (Hochrainer et al., 2010). The European Flood Directive (European Union, 2007) requests the European Union member states to provide risk management plans for areas with potentially significant flood risk. Decisions between alternative risk management options should be taken based on of flood risk analyses which are integrated with decision-support frameworks like cost-benefit analysis, multicriteria analysis, or robust decision-making (Kunreuther et al., 2013). Risk analyses combine flood hazard modelling with loss modelling and provide quantitative estimates of expected flood losses in monetary terms. Current flood loss assessments at European scale have shortcomings because they are based on heterogeneous, deterministic depth-damage functions for the individual countries (Jongman et al., 2014; Alfieri et al., 2015a; Alfieri et al., 2016b; Dottori et al., 2017). Deterministic, depth-damage functions, which use only the water depth to estimate loss, are not able to adequately describe complex damage processes (Meyer et al., 2013; Gerl et al., 2016), so that associated uncertainties might be high but are unknown due to a lack of validation and missing uncertainty quantification. Differences in vulnerability between countries, for example, due to differences in building stock, are considered via different depth-damage functions, which leads to an inconsistent, fragmented loss assessment approach.

Depth-damage functions estimate the loss from the type or use of the element at risk (e.g., residential building) and the inundation depth (Smith, 1994; Wind et al., 1999). They are associated with high uncertainties, since flood damage processes depend on many more factors besides water depth (Merz et al., 2004; Merz et al., 2013). Several studies identified additional loss determining variables like duration of inundation, sediment concentration, contamination of floodwater, flood experience, availability, and content of flood warning, precautionary measures, and the quality of external response in a flood situation (e.g., Smith, 1994; Penning-Rowsell and Green, 2000; Thieken et al., 2005; Kreibich and Thieken, 2009; Elmer et al., 2010; Hudson et al., 2014; Vogel et al., 2018). Particularly, resistance factors, such as the level of precautionary measures, are rarely taken into account by current loss models but are considered a precondition for the evaluation and development of effective risk mitigation strategies (Kreibich et al., 2015). Some multi-variable models have been developed, for example for Japan by Zhai et al. (2005), for the UK by Penning-Rowsell et al. (2005), and for Germany (Elmer et al., 2010; Kreibich et al., 2010). Studies have shown that the application of multi-variable models that take several loss influencing factors into account are better able to describe complex damage processes and improve the reliability of flood loss modelling (Apel et al., 2009; Merz et al., 2013; Dottori

et al., 2016). Loss modelling is subject to considerable uncertainty (de Moel & Aerts, 2011), which results from various sources, including the incomplete knowledge and representation as well as the stochastic nature of the damage processes. It is, therefore, crucial to quantify uncertainties in flood loss estimates and hereby support reliable risk assessment as well as informed and robust decision-making (Pappenberger & Beven, 2006; de Brito & Evers, 2016).

Recent studies developed ensemble approaches to provide uncertainty information for flood loss modelling (Hasanzadeh Nafari et al., 2016a; Wagenaar et al., 2017; Figueiredo et al., 2018). Merz et al. (2013), Kreibich et al. (2017b) and Schröter et al. (2018) demonstrated that tree-based model ensembles, such as Bagging Decision Trees or Random Forests, are suitable for flood loss modelling on the microscale (e.g., individual buildings) and the meso-scale (e.g., land use units), as they are able to capture non-linear and non-monotonous dependencies between predictor and response variables, and they take interactions between the predictors into account. Bayesian networks were used to develop probabilistic, multi-variable flood loss models (Vogel et al., 2013; Schröter et al., 2014; Wagenaar et al., 2018; Sairam et al., 2019a) for estimating flood loss of residential buildings on the microscale. Ensemble and probabilistic approaches inherently provide quantitative information on uncertainty associated with the variability of input data and model structure. Capturing and providing quantifications of the uncertainty in flood loss estimations is crucial for reliable risk assessment as well as informed and robust decision-making (Pappenberger & Beven, 2006). Two main distinct approaches exist to develop flood loss models (Merz et al., 2010b): empirical approaches which use loss data from flood events (e.g., HOWAS21 the German flood damage database; Kreibich et al., 2017c) and synthetic approaches which use loss estimates or functions collected by building experts (e.g., the Multi Coloured Manual for the UK; Penning-Rowsell et al., 2005; Penning-Rowsell et al., 2018). Loss models can be based on empirical loss data surveyed on the one hand by science (e.g., Schröter et al., 2014; Sairam et al., 2019a) and on the other hand by governmental agencies (e.g., Merz et al., 2004; Hasanzadeh Nafari et al., 2016b) and insurance companies (e.g., Spekkers et al., 2015; Cortès et al., 2018) in the frameworks of loss compensation.

Spatial scales are an important aspect of loss modelling (Apel et al., 2009; de Moel et al., 2015). Microscale models calculate the loss for single objects, for example, residential buildings, while meso-scale models estimate loss for aggregated land use units. On continental scale, loss is usually modelled on the basis of land use units. Commonly, a bottom-up approach is used for model development, which starts with a detailed analysis and modelling of losses to individual buildings (microscale) and develops an upscaling procedure for application based on land use units (Kreibich et al., 2010; Kreibich et al., 2016). That is, the structure of the microscale model is preserved, but for the input variables suitable proxy data, which are available area-wide, have to be acquired. For multi-variable models, this is particularly challenging, since the use of proxies for the diverse input variables may introduce additional uncertainty (Kreibich et al., 2017b).

The lack of flood loss data and models in many regions requires the transfer of models to contexts for which they had not been developed. This is often done with insufficient justification and without reviewing the model suitability so that models usually show decreased predictive capability under these new circumstances (Cammerer et al., 2013; Schröter et al., 2014). This results in a patchwork of approaches with low comparability and consistency, also in respect to model validation (Jongman et al.,

2012b). Previous studies on continental and global flood loss estimation utilized deterministic, country-specific, stage damage curves and a more heterogeneous database (Huizinga, 2007; Alfieri et al., 2015a; Huizinga et al., 2017). We propose an advanced approach for a consistent flood risk assessment on the continental scale and to overcome the current challenges (uncertainty quantification, spatial transferability, scale transfer), we suggest a probabilistic, multi-variable model based on coherent data sources across Europe, which is adaptable to different regional situations, via model updating with local empirical data.

The objective of this study is to develop a consistent approach for probabilistic flood loss modelling for residential buildings in Europe, which is based on the upscaling of the microscale multi-variable flood loss model **Bayesian Network - Flood Loss Estimation MOdel for the private sector (BN-FLEMOps)** presented in Wagenaar et al. (2018). That is, the structure of the microscale **BN-FLEMOps** model is preserved, and European-wide proxy data are acquired and tested for the meso-scale application of the model. The approach is applied in the whole of Europe and validated using official loss figures of past flood events in three case study areas of varying spatial scale in Germany, Italy, and Austria. As a second objective, the possibility of adapting the loss model to different regions in Europe via updating the model with local empirical data is tested in the three case studies. The paper is structured in two parts. [Section 3.2](#) contains the upscaling of the model **BN-FLEMOps** and the application of the approach in the whole of Europe. [Section 3.3](#) contains the validation and adaptation test in the three case study areas. Both sections start with the descriptions of methods and data, followed by the presentation and discussion of the results, respectively.

3.2 Consistent Approach for Flood Loss modelling in Europe

We start this section with an introduction to the **BN-FLEMOps** which presents the basis and starting point of the European loss modelling approach ([subsection 3.2.1](#)). Next, the method and data for the upscaling of the model from microscale to meso-scale and the calculation of residential building loss on the meso-scale in Europe is explained ([subsection 3.2.2](#)). [Section 3.2.3](#) contains the results and discussion of the model upscaling and the application of the consistent approach for probabilistic flood loss modelling for residential buildings in Europe.

3.2.1 The Microscale **BN-FLEMOps**

The **BN-FLEMOps** (see [Figure 3.1](#)) has been developed for flood loss estimation on the level of individual residential buildings, that is, microscale applications, and was first presented by Wagenaar et al. (2018). It estimates relative building loss of residential buildings, that is, the relation between the absolute building loss and the replacement value of the building. Building loss includes all costs (e.g., costs of wages and material) that are associated with repairing the damage to the building structure caused by flooding. In this study, this model is further developed to be applicable on the meso-scale across Europe, preserving the structure of the Bayesian network ([Figure 3.1](#)) and using consistent proxy data sets. A Bayesian network is represented by a Directed Acyclic Graph (**DAG**) which consists of nodes and arcs. The single variables in the network are represented by nodes, and the direct dependencies between variables are represented by arcs between these nodes. The

node at the tail of an arrow is called the parent node, and the node at the head of an arrow is called the child node (Lauritzen, 1996; Fenton & Neil, 2013). As an example from the structure of BN-FLEMOps (Figure 3.1): The variable inundation duration (*d*) is the child node with the variable return period (*rp*) as the parent node. The relation between these two variables reads as duration depends on return period. Wagenaar et al. (2018) described the derivation of the Bayesian network structure using a combined data and expert-driven approach. The resulting DAG with a given direction of arcs does not necessarily present a causal relationship Vogel et al. (2018). For the calculation of relative building loss (*rbloss*), it is irrelevant in which direction the arcs are pointing.

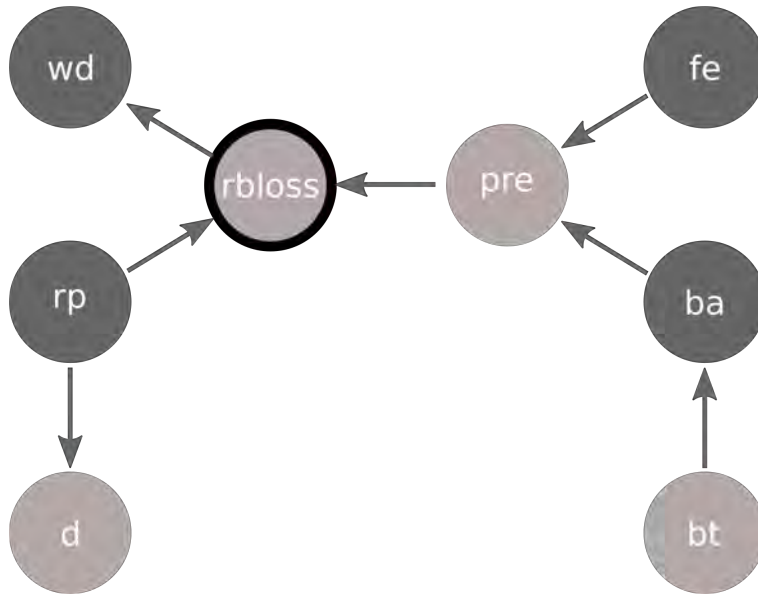


Figure 3.1: Model structure with the variables of BN-FLEMOps (adapted from Wagenaar et al. (2018)) *rbloss* = relative building loss; *wd* = water depth; *rp* = return period; *d* = inundation duration; *pre* = precautionary measures; *fe* = flood experience; *ba* = footprint area of the building; *bt* = building type). The highlighted node “*rbloss*” indicates the target variable relative building loss. The dark grey nodes represent the variables available for meso-scale applications in Europe; see subsection 3.2.2

Inference from the Bayesian network can be made in any direction and for any variable included (Fenton & Neil, 2013). As the DAG represents the conditional (in)dependency structure of the variables, this translates into a conditional probability for duration given the return period, formalized as $P(d|rp)$, where P denotes the probability. Indirect dependencies between variables are shown by the sequence of arcs between nodes. The DAG represents the joint probability of all variables in the Bayesian network respecting these dependencies and independencies. The joint probability for the variables of the BN-FLEMOps nodes is given according to the DAG as Equation 3.1.

$$P(\text{rbloss} \mid \text{bt}, \text{ba}, \text{fe}, \text{d}, \text{rp}, \text{wd}, \text{pre}) = P(\text{fe}) \times P(\text{bt}) \times P(\text{rp}) \times P(\text{ba} \mid \text{bt}) \times P(\text{d} \mid \text{rp}) \times P(\text{rbloss} \mid \text{rp}, \text{pre}) \times P(\text{wd} \mid \text{rbloss}) \times P(\text{pre} \mid \text{fe}, \text{b}) \quad (3.1)$$

BN-FLEMOps uses discretized variables and the joint probability distributions are represented as conditional probability tables or Node Probability Table (NPT). These

tables contain the conditional probabilities for each parent node and their associated child node(s). The underlying empirical microscale damage database contained information of damaged residential buildings, which were collected in surveys via computer-aided telephone interviews after the floods in 2002, 2005, 2006, 2010, 2011, and 2013 in Germany. A total of 1522 data sets with complete observations of all model variables were available. The conditional probability tables of the Bayesian network were derived from these 1522 data sets using maximum likelihood estimation. The variables were discretized on an equal frequency basis with water depth (*wd*) and relative building loss (*rbloss*) in 10 classes, return period and inundation duration in five classes and footprint area of the building in three classes (Table 3.1). The number of classes was set based on expert judgement and data availability, aiming at a compromise between a stable network, a detailed representation of variables and their presumed importance. The other variables, that is, building type *bt*, flood experience, and precautionary measures, are discrete by definition (Wagenaar et al., 2018).

Table 3.1: Overview of the Variables of BN-FLEMOps

Variable	Description	Unite	Discrete classes	Variable type
<i>wd</i>	Water depth relative to ground level	Meters	10	Continuous
<i>d</i>	Inundation duration at the affected building	Hours	5	Continuous
<i>rp</i>	Return period — calculated for peak flood discharges in sub-catchments with extreme value statistics on the basis of annual maximum series of discharge for gauges in the flood-affected areas (see Elmer et al., 2010)	Years	5	Continuous
<i>fe</i>	Flood experience — number of floods experienced before the respective damaging flood event	Score	6	Ordinal
<i>bt</i>	Building type: 1 = single-family houses, 2 = (semi)detached houses, 3 = multifamily houses	Index	3	Nominal
<i>ba</i>	Footprint area of residential buildings	Square meter	3	Continuous
<i>pre</i>	Precautionary measures — indicator (no, good, very good precaution) considering the number and type of private precautionary measures undertaken	Score	3	Ordinal
<i>rbloss</i>	Relative building loss of residential buildings — relation between the absolute building loss and the replacement value of the building as at event year	Relative	10	Continuous

Figure 3.1 and Equation 3.1 for the joint probability show that the occurrence of the variable of interest, relative building loss, is directly conditioned on the return period, and precautionary measures and that water depth is conditioned on relative building loss. The latter does not indicate a causal relationship since it is rather the other way around, with higher water depth causing higher relative building losses. However, since the directions are not of interest for the calculation of relative building loss (Fenton & Neil, 2013), we kept the structure as reported in Wagenaar et al. (2018) derived from the data-driven approach. The variables water depth, return period, and precaution constitute the so-called Markov blanket of relative building loss. This is of special interest because the availability of observations for the variables within the Markov blanket of relative building loss makes it conditionally independent of all other variables in the DAG (Pearl, 1988; Murphy, 2012; Fenton & Neil, 2013). For example, computing the relative building loss for a given combination of water depth, return period, inundation duration and precautionary measures, the value of inundation duration is not of interest because having observations for the variable return period, relative building loss is conditionally independent of duration. Hence, if all variables of the Markov blanket of relative building loss are available, other variables are no longer considered in the estimation of *rbloss*. On the other

hand, if data are missing for a node of the Markov blanket, the model can still be applied, but the uncertainty of the model results will increase. Information about precautionary measures are missing for the meso-scale application of the model and is treated as an extra unknown variable (together with relative building loss); thus, the Bayesian network calculates a conditional probability for the missing variable precaution according to the NPTs conditioned on flood experience (fe) and building area (ba).

3.2.2 Method and Data for Model Upscaling and Application at the meso-scale in Europe

3.2.2.1 Input Data Available for the European Domain

In its original application at the microscale, all variables used to estimate the flood loss with BN-FLEMOps were derived from the empirical microscale damage database (Wagenaar et al., 2018). For the meso-scale application in Europe, such detailed information are not available. Instead, suitable proxy data sources have to be used (Figure 3.2). The compilation and preparation of proxy data focuses on the Markov-blanket variables of relative building loss and besides has to consider the availability of data sources. Hence, only those variables marked in dark grey in Figure 3.1 (wd , rp , fe and ba) are included in the meso-scale model. Accordingly, data sources are needed to quantify the flood intensity and the resistance characteristics of the buildings.

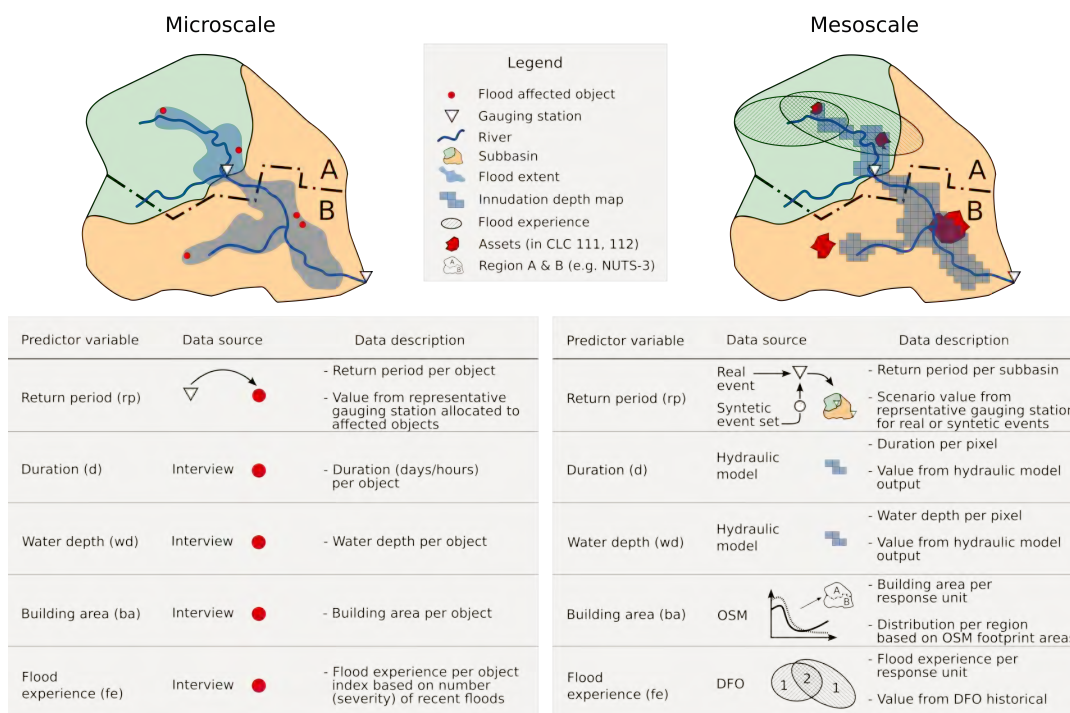


Figure 3.2: Comparison between the variables and data source for the microscale and meso-scale application of BN-FLEMOps.

Flood impact intensity is described by water depth, return period, and inundation duration. These variables are commonly estimated by hydrodynamic models, which are either used to calculate flood inundation scenarios for various return periods or to model inundation areas of real flood events. For the application of our loss modelling approach on the European domain, we used the inundation scenario for a continent-wide flood with 100 years return period provided by the Joint Research

Center (Alfieri et al., 2014). The Joint Research Center pan-European flood maps were computed using a model chain that includes hydrological modelling, derivation of flood hydrographs, and hydrodynamic modelling. The maps have a 100 m resolution of water-depth and cover catchments with areas larger than 500 km². Flood protection infrastructure such as dikes, embankments or flood walls, etc. are not considered in this hazard scenario.

The presence of precautionary measures is very heterogeneous within and across regions, and to our knowledge, no database or register of these measures exists for Europe. For this reason, the parent nodes flood experience and building area are used to infer the state (or class) of precautionary measures within the Bayesian network model.

The variables building area and flood experience describe the resistance characteristics of the residential buildings. OpenStreetMap (OSM, 2018) was identified as a suitable source to provide information on the building footprint area. OSM is a geographic database with worldwide coverage. It relies on a community of contributors to constantly add information and assure regular updates to enhance accuracy and completeness. The OSM project provides freely available open data and is nowadays considered a reliable source for most civil and common use cases (Barrington-Leigh & Millard-Ball, 2017). To obtain the building footprint area of residential buildings, the database was filtered to exclude objects with non-residential usage. In previous studies (Thieken et al., 2005; Merz et al., 2013; Schröter et al., 2014), flood experience is used as an indicator which is based on different factors such as the number of experienced floods, the associated losses with the latest flood experienced and the time period since the last flood event. These details are available from the empirical microscale damage database and are not available on the same level of detail in Europe. Thus, the proxy data for the variable flood experience had to be simplified for the meso-scale application. Based on the assumption that the more floods individuals have experienced during the last years, the greater is their flood experience, the meso-scale variable for fe is represented by the number of floods that occurred during the last 25 years in the particular region. To derive the number of floods people were exposed to at their home location, we use the database of the Dartmouth Flood Observatory (DFO) (Brakenridge, 2018). The DFO catalogue is an archive of historic flood events starting in the year 1985. This archive comprises maps of flood-affected areas with a set of additional characteristics like severity, start and end date, and the cause for flooding. One drawback of this data source is that the spatial extent of flood-affected areas is very coarse, as it often consists of an outline of the affected area instead of a detailed flood footprint. Nevertheless, to our knowledge, this data set is the only one that provides a comprehensive record of flood events in space and time across Europe. With these European wide data sets for the variables water depth, return period, building area and flood experience, relative building loss can be estimated for residential buildings using the BN-FLEMOps model. To further obtain absolute building losses, replacement cost values of residential buildings are needed.

The monetary values of the exposed residential buildings are taken from a European asset map developed for this study. This asset map was created by adapting the approach by Huizinga et al. (2017) who found a relationship between construction cost (material and labour cost) and Gross Domestic Product (GDP) per capita by comparing national socio-economic parameters with construction cost surveys from multinational construction companies. The replacement values of buildings for the

European asset map are calculated using the relation of construction cost and GDP per capita on the NUTS-3 level to account for regional differences. To develop the European asset map for this study, we used the GDP per capita information for all NUTS-3 regions available from Eurostat (2018) for the year 2013. The GDP per capita values can be adjusted to different flood event years by the GDP per capita growth rate of the respective region affected, such as a subbasin or municipality. The resulting asset values for residential buildings reflect tangible monetary assets and are based on the concept of reconstruction cost. The reconstruction costs of residential buildings are translated to unit area values in ($\text{€}/\text{m}^2$) for NUTS-3 regions by disaggregation on the CORINE Land Cover (CLC) 2012 (European Environment Agency, 2016) classes: continuous urban fabric (111) and discontinuous urban fabric (112) using asymmetric mapping following Huizinga et al. (2017). The European asset map created by this approach contains distinct unit area values ($\text{€}/\text{m}^2$) for residential buildings in NUTS-3 regions and is available for all European countries covered by CLC data.

3.2.2.2 meso-scale Loss Calculation

Loss estimation includes three steps: (i) data preprocessing, (ii) computation of relative loss estimates, and (iii) the conversion to monetary values. *Data Preprocessing and Spatial Intersection*. Data preprocessing is a two-step procedure consisting of the spatial intersection of data sets and joining features to resulting spatial units. Figure 3.3 illustrates the individual steps in a flow chart of input data preparation. First, the watershed subbasins holding the return period value (rp), the DFO data set providing the value for flood experience (fe), the administrative boundaries (required for the second step), the CLC classes holding the unit area values (assets) and the polygonized water depth raster providing the water depth values (wd) are spatially intersected, see second row of Figure 3.3. In the intersection process, the data sets are handled as polygons to ensure that no spatial details are lost—that means that even cells of the water depth raster might be subdivided by one or more other input data sets (see Figure 3.3).

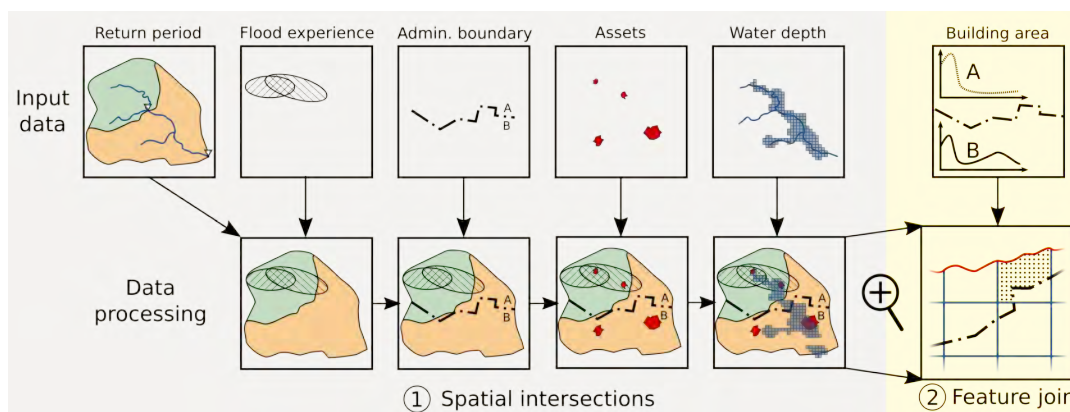


Figure 3.3: Flow chart of the input data preparation for the meso-scale loss calculation in BN-FLEMOps. 1 = spatial intersections: Shows the stepwise process of the spatial intersection of model input data. In subsequent processing steps, the data sets are overlaid and intersected to combine their information in one data layer. The response units are the areas created by the intersection and within these areas; all variable values are constant. 2 = feature join: Displays the magnified response units and the join of building area values based on administrative boundaries.

The result of the intersection is a set of response units which are spatial polygons holding single values for water depth, return period, flood experience as input variables for the BN-FLEMOps model and in addition the unit area value of residential buildings. The code of the administrative regions is not directly used as a model input, but is required for joining features to the response unit within the second step of the data preprocessing. Based on the OSM data set, a distribution of building footprint area is available for each administrative region (NUTS-3). For each response unit, we sample one value of building area from the distribution of the corresponding NUTS-3 region. A response unit is exemplary, shown on the bottom right of [Figure 3.3](#) as the dotted polygon area. In this example, the cells of the water mask (blue lines) are only separated by the administrative boundary (black dashed line) and the asset map (red line) and hold the same values for the other variables (*rp*, *fe*). As a result, each response unit holds the values for the input variables of the BN-FLEMOps model to estimate relative building loss per response unit.

Computation of Relative Loss Estimates. Applying BN-FLEMOps means inferring from the Bayesian network following the joint probability distribution ([Equation 3.1](#)). The joint probability distribution is depicted by the NPTs which were generated following the dependency structure of the Bayesian network and the empirical loss data. The conditional probability distribution of our target variable *rbloss* can therefore be queried from the NPTs given the available input variables. The application of the Bayesian network results in a probability distribution of relative building loss for each response unit. Summary statistics are reported as model results. The book “Bayesian networks—With examples in R” by Scutari and Denis (2014) provides a comprehensive introduction to the application of discrete Bayesian networks.

Conversion to Monetary Values. The final processing step is the calculation of the absolute monetary loss by multiplying the distribution of loss ratios with the unit area value for each response unit and the surface area of the response unit. The result is a distribution of absolute loss in euro per response unit. Absolute monetary loss values for response units can be spatially aggregated, for instance within their respective NUTS-3 regions or river basins.

3.2.3 Results and Discussion of Model Upscaling and Application at European Scale

3.2.3.1 Model Upscaling—Suitability Test of Proxy Variables

To assess the suitability of this proxy data source, we compare the distribution of building footprint areas from OSM and the empirical microscale damage data (*msdd*) in 25 German NUTS-3 regions where 30 or more observations are available to have a reasonable number of data points for comparison. Buildings larger than 500 m² were excluded from the building data set, to avoid not correctly confined buildings (block of houses not separated into the individual buildings) and wrongly assigned objects (like large scale industry buildings). The threshold to exclude buildings larger than 500 m² from the OSM data set was derived by comparison with the empirical *msdd* database.

The median building footprint area in the empirical microscale damage database aggregated for the 25 NUTS-3 regions is 154 m² in comparison to 136 m² in the OSM geometries. This divergence of building areas between the empirical microscale damage data and the proxy data from OSM is also present at a disaggregated NUTS-3 level. [Figure 3.4](#) compares the distribution of building area values from both data

sources in 25 NUTS-3 regions in Germany. The OSM data set often displays slightly lower mean values and narrower interquartile ranges. The building area distributions of the empirical microscale damage data differ considerably more between the individual NUTS-3 regions in comparison to the more homogeneous building area distributions of the OSM data. Even though these differences between empirical microscale damage data and meso-scale proxy data from OSM exist, OSM seems to provide suitable information about the variable building area for the application of BN-FLEMOps on the meso-scale. Besides, it is currently the best pan-European data set available to describe building footprint area, and it is continuously improved by mapping efforts of OSM contributors and the integration of existing data sets.

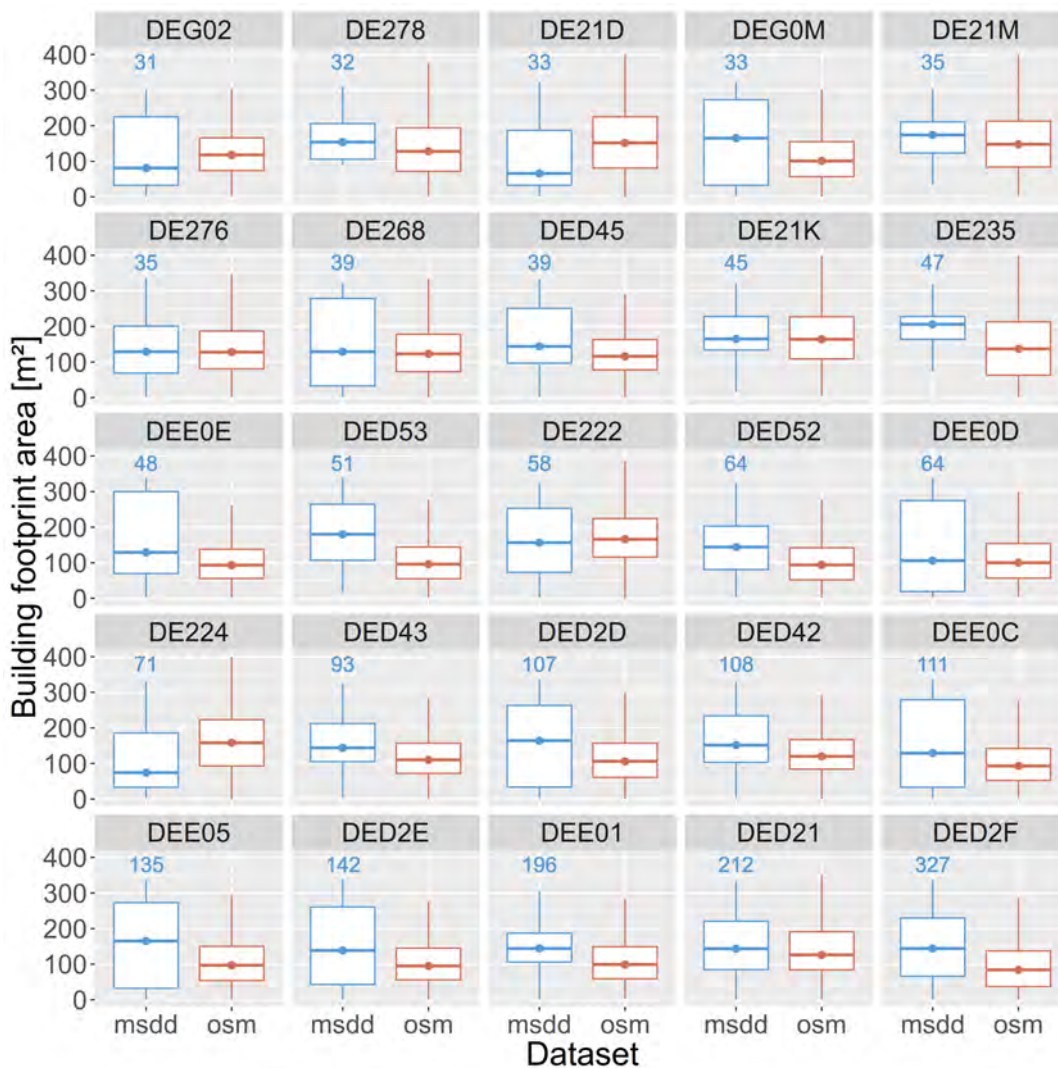


Figure 3.4: Comparison of box plots of building area distributions (y-axis) forming empirical microscale damage data (msdd) and meso-scale proxy data derived from OSM data (x-axis) in 25 NUTS-3 regions in Germany (panels). The number of empirical data points in the respective NUTS-3 region is given in blue.

The distribution of the footprint area of residential buildings in the NUTS-3 region in Europe is sampled to populate the building area variable in each response unit. Higher median building areas in peripheral regions like the west coast of Norway may partly be attributed to low mapping coverage. Prominent landmark buildings are more likely to be mapped first and are therefore dominant in regions with low

mapping completeness. This may result in a higher median building area compared to other European regions. France and the Netherlands, on the contrary, have a very complete building map since official data sets such as cadastral data provided by Direction Générale des Impôts and data from AND (<https://geojunction.com>) were integrated into the OSM database (Mooney & Minghini, 2017; OSM contributors, 2018). This more consistent representation of individual buildings may have contributed to the relatively low mean building areas in France and the Netherlands. The building area variable in the meso-scale application of BN-FLEMOps is sampled from the distribution of building footprint areas per NUTS-3 region, while the median value is solely used to present the European data set on the map.

European wide proxy data of the variable flood experience for the meso-scale application of BN-FLEMOps is derived from the DFO catalogue. To test the suitability of this proxy data to capture the state of flood experience and its variation across regions, we compare the number of floods experienced during the last 25 years as reported in the empirical microscale damage data to the flood occurrences according to the DFO catalogue per NUTS-3 region. Assuming a minimum threshold of 30 or more empirical data points per NUTS-3 region, we compared the aggregated data in 37 NUTS-3 regions in Germany (Figure 3.5).

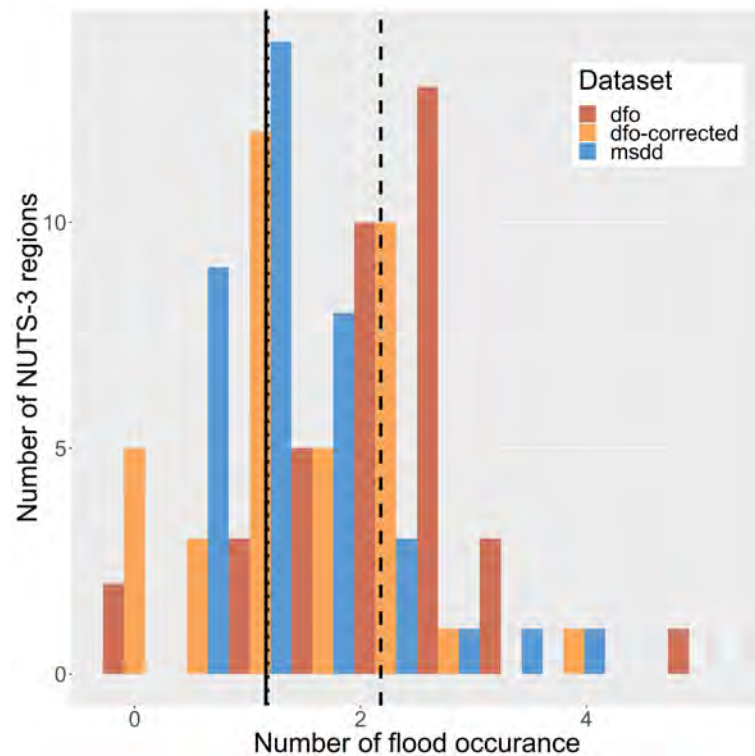


Figure 3.5: Histogram of the mean values of the number of floods that have occurred according to the empirical microscale damage data (msdd), meso-scale proxy data derived from dfo in 37 German NUTS-3 regions, and dfo data with bias correction factor minus 1. The black lines indicate the overall mean of the data sets (dfo = dashed; dfo-corrected = dotted; msdd = solid).

The threshold of 30 empirical data points is a compromise between the availability of a sufficiently large number of observations in a NUTS-3 region, and a sufficiently large number of NUTS-3 regions so that comparisons are meaningful. Most people have reported that they had been affected once or twice by floods before the event they were asked about in the telephone surveys (Figure 3.5). The DFO archive counts

two to three floods in the same areas on average. The reason for this mismatch may be the coarse spatial resolution of the DFO data, which contains a significantly larger area than what was impacted by the flood. To avoid an overestimation of flood experience in the meso-scale application of BN-FLEMOps this offset between both data sets is corrected by introducing a bias correction factor of minus one which improves the agreement of the histograms of both data sets (Figure 3.5). Figure 3.6 shows the spatial footprints of all large historic floods registered in the DFO catalogue that occurred in Europe in the time span from 1985 to 2015. In the model, the number of historic floods is counted for each response unit and is used on the meso-scale to describe flood experience in the model.

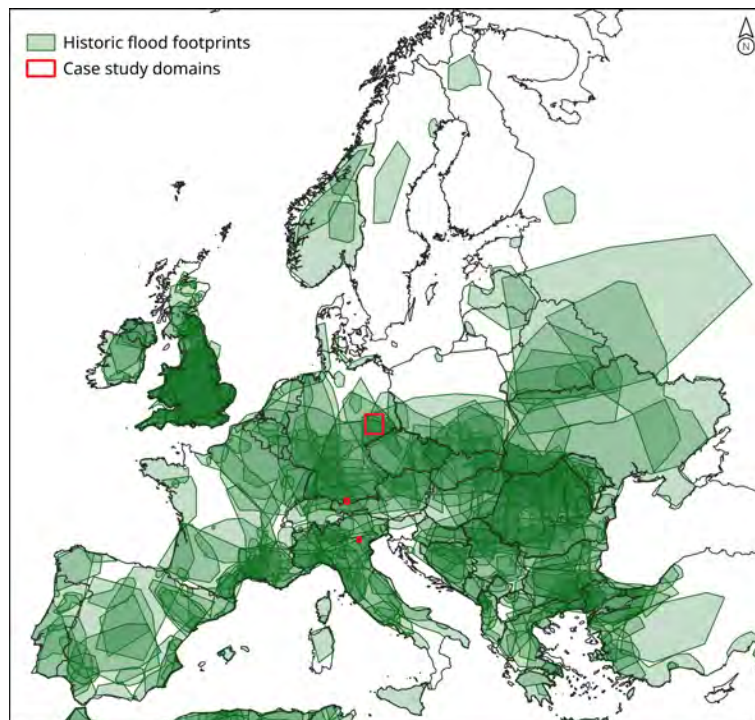


Figure 3.6: European map of historic flood events from the DFO catalogue for the period of 1985 to 2015. The red squares indicate the case study areas in Germany, Austria, and Italy.

The European asset map (Figure 3.7) displays regional differences in unit area values of residential buildings, giving the reconstruction costs. Metropolitan areas such as London, Paris, Madrid, Rome, and Berlin are associated with higher unit values than their surrounding areas. Scandinavian countries, Switzerland and the Benelux-countries show overall high unit values of residential buildings, while eastern and south-eastern European countries display lower unit area values of residential buildings. To use the asset map for the calculation of absolute losses on the meso-scale, the building values need to be translated to the respective year of the analyses. This is done via the relation between construction cost and GDP per capita, developed by Huizinga et al. (2017) on construction cost data from 2010 and 2013. Uncertainties associated with this translation of values might be higher for earlier or later years than from 2010 to 2013, as then the GDP per capita might not have the same correlation with the construction cost as was found for this time period. The spatial resolution of the European asset map is determined by the accuracy and minimum mapping distance of the CLC data set. Individual buildings in predominantly rural areas are not always represented, and thus are not accounted for in the loss estimates.

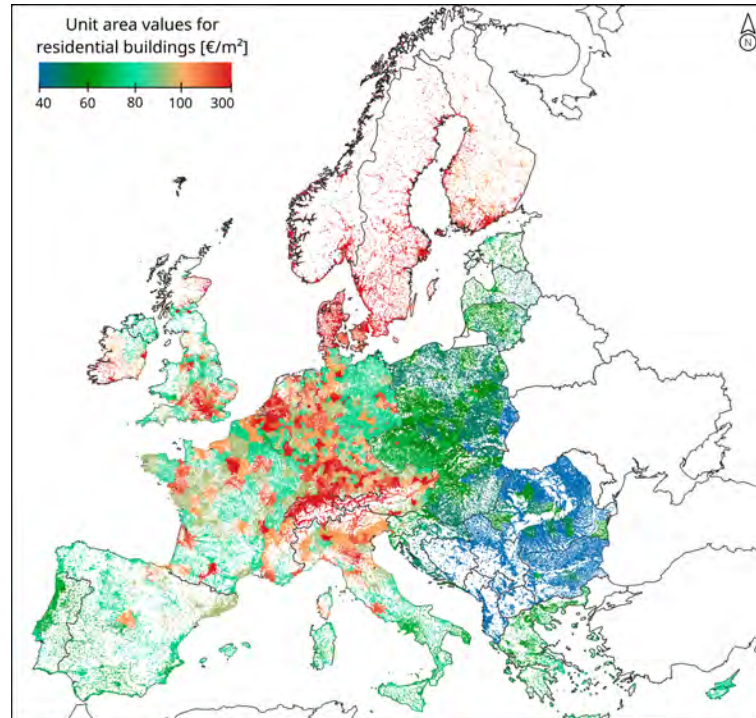


Figure 3.7: European asset map displaying the unit area values of residential buildings in 2013 in (€/m²) following reconstruction costs approach.

3.2.3.2 Flood Loss Assessment at European Scale

The results of the loss estimation for the 100-year European flood map are shown in [Figure 3.8](#) in terms of the median, the 20 % quantile (Q_{20}) and the 80 % quantile (Q_{80}) loss based on the distribution of absolute monetary losses per NUTS-3 region across Europe. For this flood scenario, the highest flood losses are expected in the flood plains of major European rivers, such as the Rhine and Meuse, Danube, Seine, Loire, and Po. Note that the flood scenario does not account for any flood protection infrastructure in place. NUTS-3 regions in the Netherlands, France Austria, Hungary, the Czech Republic, and Belgium would be most affected by a 100 years flood scenario. Groot-Rijnmond (NL339) in the Netherlands would suffer the overall highest losses with 2.6 billion € ($Q_{20} = 1.1$; $Q_{80} = 7.7$) loss to all residential buildings in the region. [Figure 3.8](#) depicts the accumulated flood loss for residential buildings in NUTS-3 regions in Europe. Therefore, large NUTS regions such as in Sweden tend to show high accumulated losses, whereas smaller regions in Germany have lower accumulated losses. On a national level, Germany and France are estimated to have the highest losses with 14.5 billion € ($Q_{20} = 5.8$; $Q_{80} = 38.9$) and 13.2 billion € ($Q_{20} = 5.4$; $Q_{80} = 36.1$) respectively. A data set of the flood loss distribution per NUTS-3 region in 10 % quantile steps and a detailed description is published in (Lüdtke et al., 2019b). The total accumulated loss for residential buildings in Europe is estimated to 79.0 billion € ($Q_{20} = 32.3$; $Q_{80} = 213.8$) for the continent-wide 100 years flood hazard map. A comparison with the aggregated direct flood losses estimated for 45 land use classes in European countries by Alfieri et al. (2015a) showed that the losses estimated by BN-FLEMOps were lower for most countries. This is a plausible result, since BN-FLEMOps only considers residential buildings in the CORINE land use classes for urban fabric.

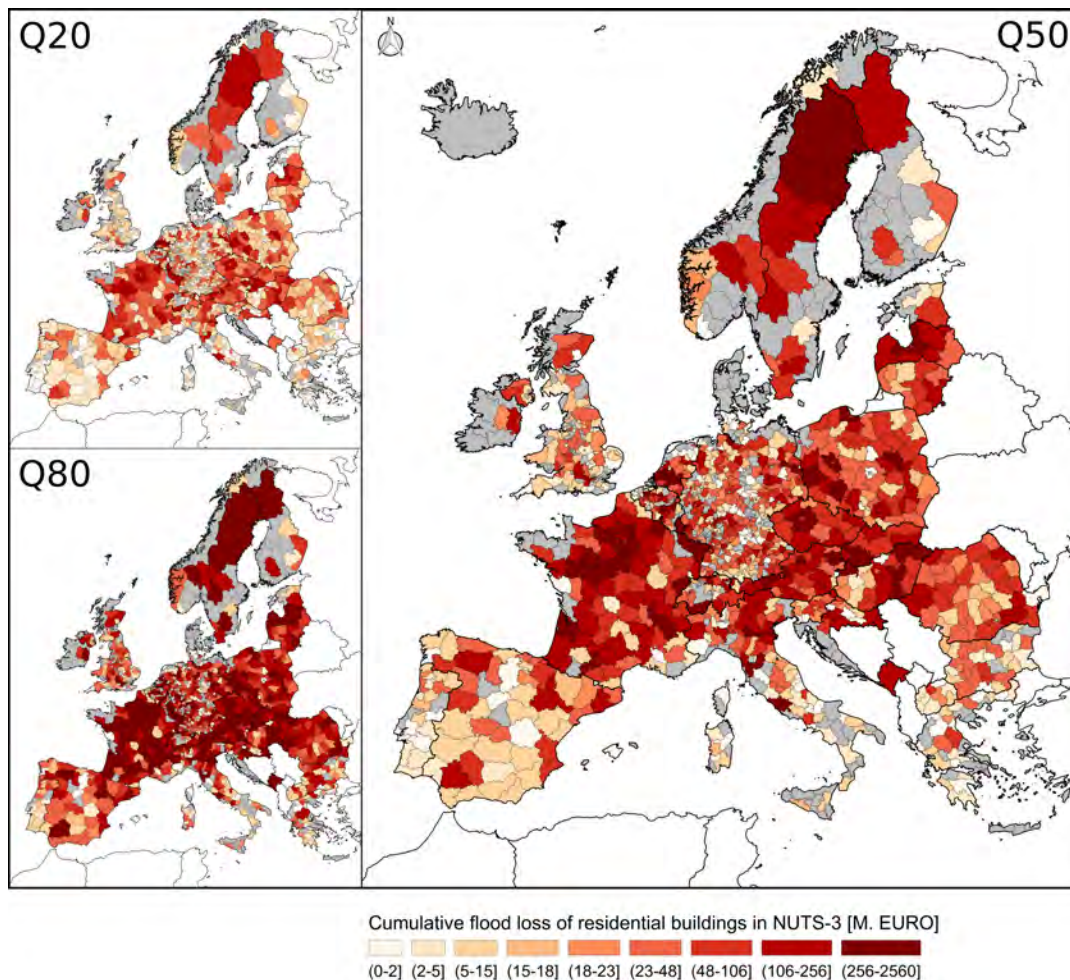


Figure 3.8: Maps of the cumulated flood loss of residential buildings for NUTS-3 regions in million €. Gray NUTS-3 regions were not affected by the 100-year flood scenario or were not calculated due to lack of input data. Map of 20 % quantile (left top), map of 80 % quantile (left bottom), and map of 50 % quantile (right).

3.3 Model Validation and Adaptation Test in Three Case Study Areas

[Subsection 3.3.1](#) presents the three case studies including the available data with respect to the inundation area, validation data and data for the adaptation test. [Subsection 3.3.2](#) describes the method for updating the loss model to adapt it to local settings. The following sections present the results and discussion of the validation of the general BN-FLEMOps model ([subsection 3.3.3](#)), and the result of the adaptation test, that is, application of the loss model with two update steps in the three case study areas ([subsection 3.3.4](#)).

3.3.1 Case Study Descriptions

The BN-FLEMOps model is validated on the meso-scale in three different case studies. We use the European proxy data for building area and flood experience as inputs to the model. Flood intensity information including water depth and return period are available from local data sets. Loss estimates are compared with official loss figures from past flood events.

3.3.1.1 Case Study Caldogno, Italy

The Veneto Region (Italy) was affected by persistent rain between 31 October and 2 November 2010, particularly in pre-Alpine and foothill areas. In some locations, accumulated rainfall exceeded 500 mm in 72 one of the most intense events of the last 50 years (Regione del Veneto, 2011a). Various rivers overflowed, flooding an area of around 140 km² causing three fatalities, and 3,500 evacuees. Flood losses among residential, commercial and public assets amounted to a total of 426 million €. In Caldogno, a municipality in province of Vicenza, losses to those sectors reached 25.7 million € (Regione del Veneto, 2011b).

Official Loss Figure for Validation: Losses to residential buildings amounted to 7.55 million € in the city Caldogno, which is used for validation in this study. This official loss figure is provided by the municipality of Caldogno and corresponds to restoration costs that were collected and verified in the scope of the loss compensation process after the event (Scorzini & Frank, 2017).

Flood Map: Water depth and flow velocity maps were estimated using a coupled 1-D/2-D model of the area between the municipalities of Caldogno and Vicenza. These flood intensity metrics were originally computed on 5 m x 5 m spatial grid (Scorzini & Frank, 2017; Figueiredo et al., 2018) and have here been resampled 10 m x 10 m.

Empirical Microscale Loss Data for Model Updating: For 295 damaged residential buildings, building characteristics at the microscale were obtained through surveys. These characteristics include building type, structural type, number of floors, quality, and year of construction. Building areas were derived from the region of Veneto's cadastral map, and building values were estimated from data provided by the Chamber of Commerce of Vicenza (Figueiredo et al., 2018). Data about the inundation duration and water depth at the building were taken from the flood map described above. Thus, in this case study, the flood intensity information used for model updating is the same as for the model application.

3.3.1.2 Case Study Lech, Austria

The municipalities in the mountain basin of Reutte (catchment size of 1,000 km) in the Lech catchment (Austria) were affected by floods several times in the recent past, with severe flooding in 1999 and 2005 (Cammerer et al., 2013). The flood in 2005 caused direct loss of about 410 million € (with 61 % of loss in the private sector) in the entire federal state of Tyrol (Amt der Tiroler Landesregierung, 2006). In the case study area, the flood in August 2005 had a peak discharge of 943 m³/s which corresponds to an estimated return period of 330 years at the gauge Lechaschau. Particularly, the municipalities of Pflach and Höfen were strongly affected, due to overtopping and breaches of embankments. The flood type can be regarded as an interaction of static and dynamic flooding in the case study area (Cammerer et al., 2013).

Official loss figure for validation: Estimation of structural loss to residential buildings in the study area is 1.9 million €, as provided by Cammerer et al. (2013). This figure is based on official data but is associated with uncertainty, because loss reports did not differentiate well loss to building structure or contents. The study by Cammerer et al. (2013) made assumption to separate the aggregated loss figures, but the true values are not reported.

Flood map: Maximum water depths of the flood event in August 2005 were simulated with hydrodynamic 2-D model “Hydro_AS-2D”. Two dyke failures, which had occurred in the community of Pflach, were considered in the hydraulic simulation. While the simulation and validation were performed on a 1 m x 1 m grid, the used water depths were aggregated on a cell size of 10 m (Cammerer et al., 2013).

Empirical microscale loss data for model updating: To update the Bayesian Network model, local information of 22 households affected during the 2005 flood are available. In the Austrian federal states of Tyrol and Vorarlberg, 218 interviews with private households were carried out in the aftermath of the flood event in 2005 in order to compare different risk transfer systems of three Alpine regions (Raschky & Schwindt, 2009). However, only 72 of all surveyed households were actually affected by the flood in 2005, and only 22 households provided enough information to estimate the relative loss. Additionally, the data contain information about the inundation depth and duration, as well as the building area, building type, flood experience, and precautionary measures. Thus, in this case study, the flood intensity information used for model updating differs from the one used for model application.

3.3.1.3 Case Study Mulde, Germany

The Mulde catchment (7,400 km²) in Saxony (Germany) is prone to recurrent flood hazards. During the last years it has been hit by severe floods in August 2002 and June 2013 (Engel, 2004; Conradt et al., 2012; Schröter et al., 2015). The flood in August 2002 was triggered by extreme precipitation in the headwater areas of the catchment. Record-breaking rainfall amounts of 312 mm were recorded within 24 hours at the station Zinnwald-Georgenfeld of the German Weather Service in the Ore Mountains (Ulbrich et al., 2003). As a result, extreme flash floods hit the upstream parts of the Mulde and its tributaries. The flood wave caused numerous dike breaches in the lower reach of the Mulde River, with 19 municipalities affected. The magnitude of the flood peak discharge was in the order of a 500 year flood, estimated using methods of extreme value statistics based on annual maximum series of mean daily discharge records from the gauge Bad Dueben (Elmer et al., 2010).

Official loss figure for validation: Losses to residential buildings totalled 240.6 million € (Saxon Relief Bank, 2005), for further details refer to (Kreibich et al., 2017b).

Flood map: For this event, an inundation depth map was derived from hydronumeric simulations (Apel et al., 2009) and hydraulic calculations (Grabbert, 2006). This map gives information on maximum inundation depths on a spatial grid with a pixel size of 10 m x 10 m.

Empirical microscale loss data for model updating: Local information from 74 households are available from the post-event computer-aided telephone interview survey for updating the Bayesian Network model (Thieken et al., 2017). The interview data set includes information about the building area, building type, flood experience, and precautionary measures. Also, the data about the inundation duration, and water depth at the building to characterize the flood intensity are available from the survey. Thus, in this case study, the flood information used for model updating differs from the one used for model application.

Table 3.2: Overview of Case Study Characteristics and Available Variables for Model Updating in the Case Studies

Case study characteristics	Caldogno	Lech	Mulde
Reported loss to residential buildings used for validation (M €)	7.5	1.9	240
Number of empirical microscale observations available for updating	295	22	74
Inundated area (km ²)	3.3	3.4	116.5
Event year	2010	2005	2005
Availability of variables in microscale observations			
Water depth	✓	✓	✓
Duration	X	X	X
Return period	✓	✓	✓
Building type	✓	X	X
Building area	✓	X	X
Flood experience	X	✓	✓
Precaution	X	✓	✓
Relative loss	✓	✓	✓

Note. The symbols indicate whether data were available (✓) or unavailable (X)

3.3.2 Updating Method Using Local Data

As described in [subsection 3.2.1](#), the NPTs of the BN-FLEMOps model were trained based on empirical loss data from Germany. Bayesian networks offer the possibility to consistently update the NPTs with additional data, and Wagenaar et al. (2018) showed that such updating improves the spatial transferability of loss models to the microscale. This was explained by the fact, that the additional use of local data better enables the model to cover some effects of implicit assumptions about variables not included in the loss models, like flow velocity (Wagenaar et al., 2018). Thus, we test if the meso-scale application of the BN-FLEMOps model in geographical regions with potentially different socio-economic, building and flood event characteristics can be improved via an updating of the NPTs with empirical data from the target regions. To assess the value of local data on the model performance, we use two update steps with 750 and 1,500 additional observations from the case study regions and compare it to the application of the general BN-FLEMOps model. Given, that the data used to develop the NPTs of BN-FLEMOps consists of 1,522 observations, the update steps correspond to 50 % and 100 % of added data from the case study, respectively. That means, at the update step of 100 %, the added data from the case study has as much weight to describe the damage processes as the German data used to derive the model.

Sampling with replacement is applied to generate the required number of 750 and 1,500 data points in case study regions since in none of the areas enough empirical data is available ([Table 3.1](#), second row). We follow the same approach as for the computation of loss ratios described in [subsection 3.2.2](#), that is, the entire update and validation process is repeated 5,000 times.

In case the empirical data from the case study is not complete, the NPTs of the missing variables can not be updated. In this case, the NPTs of these missing variables reproduce the joint probability distribution of the original model. This means that knowledge about the conditional probabilities between the loss influencing variables is transferred from the meso-scale model to the case study areas.

3.3.3 Result and Discussion of the Validation in the Three Case Study Areas

The results of the loss estimation with BN-FLEMOps on the meso-scale for the three case studies are shown in [Figure 3.9](#) in comparison with the official loss figures. The thick solid line of the empirical cumulative density functions shows the median of the 5,000 repetitions of flood loss estimation. The red vertical solid lines indicate the officially reported loss for each case study. The meso-scale application of BN-FLEMOps yields results that are in the same order of magnitude as the reported loss values for all three case studies. Official loss figures are within the 90 % quantile range of the probabilistic loss estimation. The loss estimation is however associated with large uncertainty manifest in the interquartile range in [Table 3.2](#). For the Mulde case study, the median loss estimate shows an underestimation of about 104 million € (43 %). For the Caldogno and Lech case studies, we see an overestimation by the BN-FLEMOps model by 2.1 million € (28 %) and 5.8 million € (305 %), respectively. We observe that the uncertainty of the upper tails of loss distributions is larger than for the lower tails, which is shown by the variability between the loss estimation repetitions above the 75 % quantile ([Figure 3.9](#)). This uncertainty is explained by multimodal marginal distributions for combinations of observations that lead to high loss ratios. Those multimodal marginal distributions are indicated by groups of deviating empirical cumulative density distributions model replications in the ([Figure 3.9](#)). In addition, the model shows a strong sensitivity to the contribution of the highest loss class. This can be attributed to the discretization of the variable relative building loss. The highest loss class covers a very large range, from 0.274 to 1. We use the median as the representative value for every class to compute the absolute damage (assuming a uniform distribution within each class). The representative value for the highest loss class is 0.637 whereas the representative value for the second-highest loss class is 0.226, which is a difference of more than 40 percent relative building loss. This effect stems from the discretization of the continuous variables and is a known issue and subject to ongoing research.

Using the median of the 5,000 repetitions of the loss estimations (thick solid line in [Figure 3.9](#)), the interquartile range (0.25 to 0.75) of the simulation covers the official loss figure for the Mulde and Caldogno cases, but not for the Lech case study where the overestimation is too high ([Table 3.3](#)). The underestimation of BN-FLEMOps in the Mulde case study may partly be attributed to uncertainties in the temporal transfer of the asset values to the year 2002 as discussed in [subsection 3.2.3](#). The BN-FLEMOps shows the poorest performance in the Lech case study and the best performance in the Caldogno case study. The meso-scale loss estimation approach developed for the whole of Europe seems to be less suitable for data-scarce applications and low-impact events like the Lech case study, where damage to only 22 objects was reported.

Table 3.3: Comparison of Average Loss Estimate From BN-FLEMOps and Official Loss Information

Case studies	Median loss estimate + (IQR) in million €	Official loss in million €
Caldogno (Italy)	9.6 (3.8; 18.8)	7.5
Lech (Austria)	7.7 (3.6; 13.3)	1.9
Mulde (Germany)	136.0 (89.0; 248.2)	240.0

Note. IQG = InterQuartile Range.

The loss estimates with the meso-scale application of BN-FLEMOps for the Lech case study are high also in comparison with the results achieved by Cammerer et al. (2013).

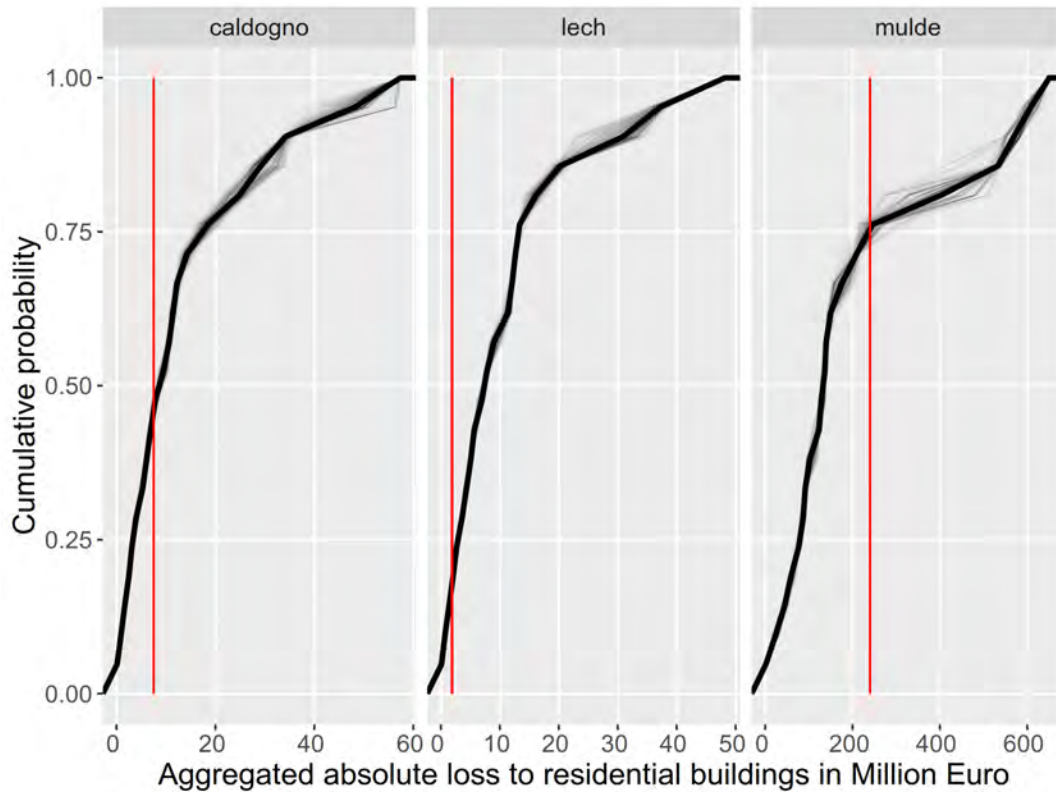


Figure 3.9: Empirical cumulative density function (ecdf) of the absolute residential building loss for the three case studies. The thick solid lines represent the median ecdf, while the grey lines represent the 1,000 model replications and red lines represent the official loss figure.

They tested several international flood loss models in a spatial transfer study, which resulted in a broad range of loss estimates from 0.842 million € (based on model MURL, 2000) to 9.094 million € (based on model Hydrotec, 2002) when applied to the same inundation area as used in this study. However, Cammerer et al. (2013) concluded, that models based on neighbouring regions with similar building and flood event characteristics yield fundamentally better results than models based on loss data from spatially different regions and dissimilar flood events. Given this hypothesis, they identified the most suitable models, estimating loss between 1.836 million and 2.854 million €.

For the 2010 flood in the Caldogno case study, the meso-scale application of BN-FLEMOps achieves equally good results as previous studies. Scorzini and Frank (2017) show a comprehensive overview of different model applications and flood loss estimates for the city of Caldogno ranging from 5.96 million € (model based on Debo, 1982) to 13.49 million € (model based on Dutta et al., 2003).

The Mulde flood event was studied by Kreibich et al. (2017b), who developed a meso-scale model that was applied and validated in the different municipalities affected by the flood event. Their aggregated loss for all municipalities sums up to 238.6 million € and is thus very close to the reported loss of 240 million €. However, also this study reported a high uncertainty in loss estimation with values between 44 million and 400 million € and used a different data source for asset values (Kreibich et al., 2017b).

3.3.4 Result and Discussion of Model Adaptation Test

Previous studies have shown that spatial transfer of flood loss models may entail a decline in model predictive performance and an increase in model uncertainty (e.g., Cammerer et al., 2013; Schröter et al., 2014). Wagenaar et al. (2018) report that BN-FLEMOps can be applied in regions different from its origin while maintaining a stable performance. To further test the spatial transferability of the BN-FLEMOps and compare performance on the microscale and meso-scale, the model is adapted to the different case studies via updating of the NPTs with empirical microscale loss data from each case study. Figure 3.10 shows the median and the 90 % quantile range for the application of the meso-scale BN-FLEMOps model (0 data points used for NPT update) and the two update steps (750 and 1,500 data points used for NPT update). The officially reported loss is depicted as a red line. In the Caldogno and Mulde case studies, the loss estimates improve with updating, which is reflected by slightly decreasing the 90 % quantile ranges and an approximation of the median loss estimate to the official loss figure. In the Caldogno case study, the median of the simulated loss is 8.7 million € for the second updating step (1500 data points added), which reduced the overestimation from 28 % to 16 %. For the Mulde case study, the median simulated loss increased with the first and second update step to 161.5 million and 199.7 million €, respectively. That reduced the underestimation from the initial 43 % to 33 % for the first update step and to 19 % for the second one. The results for the Lech case study do not improve with updating since the estimation for the simulated loss is even increasing with updating and thus the overestimation is increasing.

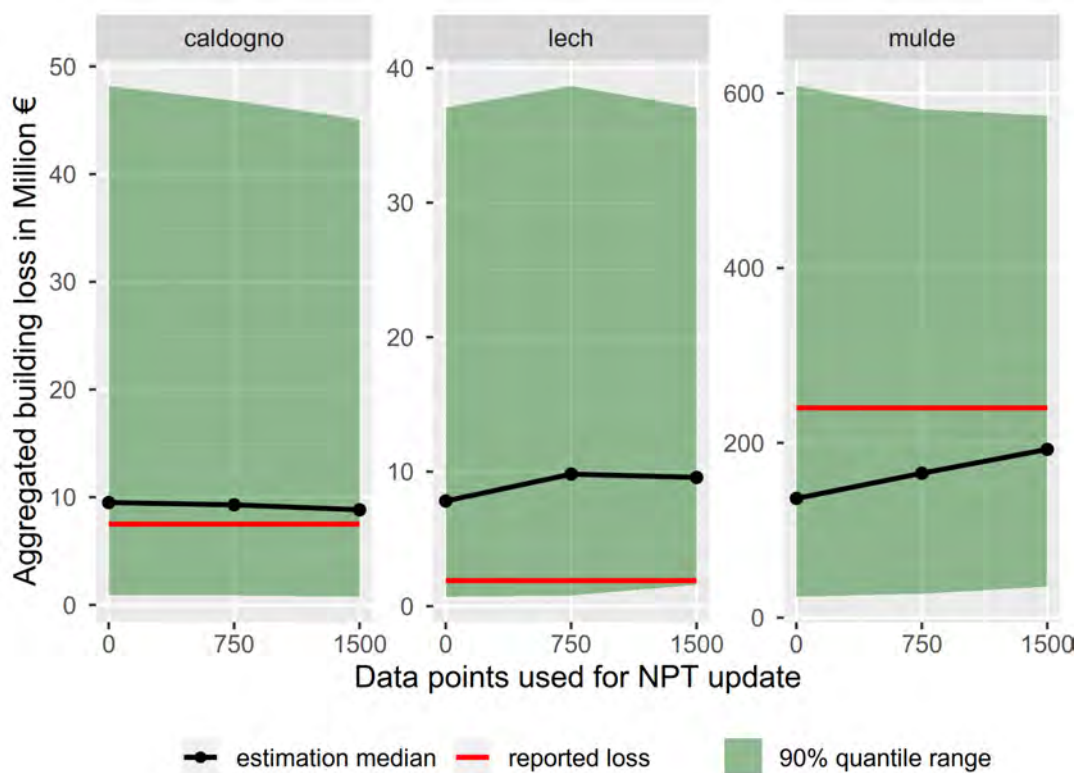


Figure 3.10: Comparison of reported loss and estimated loss for the application of BN-FLEMOps with update.

We attribute the failure of the updating process mainly to the very limited number

of empirical loss data available in the Lech case study. With only 22 loss data points as the base population, the information content about local damage processes is rather limited. Further tests are necessary to identify the boundary conditions for a successful model. The performance of the BN-FLEMOps model in these case studies is comparable to other models, except for the Lech. For the Caldogno case study, the average of seven microscale models yields a 10 % overestimation of reported loss (Scorzini & Frank, 2017). The application of six other meso-scale loss models to the Mulde case study resulted in an average underestimation of 16 % (Kreibich et al., 2017b). For the Lech case study, the site-specific meso-scale model FLEMO_{AT+} covers the observed loss within the 90 % confidence interval, which is also clearly more narrow (1.872–2.343 million €; Cammerer et al., 2013) than for the BN-FLEMOps model. These numbers illustrate that the BN-FLEMOps model provides comparable outcomes. A closely related question concerns the effort of transferring the BN-FLEMOps and other models to other contexts. While the site specific implementation of flood loss models requires the acquisitions and compilation of various input data sets, for the BN-FLEMOps model the building resistance-related meso-scale proxy variables flood experience and building area have been consistently derived for the whole of Europe and will be accessible via online data repositories. This will reduce the effort to apply the BN-FLEMOps model throughout Europe.

3.4 Conclusion

The developed approach for a consistent flood loss modelling in Europe, which is based on the probabilistic, multi-variable BN-FLEMOps is suitable for loss estimation of residential building structures in the whole of Europe. This novel approach advances flood loss modelling since it enables a consistent estimation of losses on a continental scale, with the resulting probabilistic loss estimates inherently providing uncertainty information. Such uncertainty information can significantly improve the quality of decisions, for example, about different risk management strategies based on cost-benefit analyses since the uncertainty information enables decision makers to consider the whole range of possible outcomes. Depending on risk-averse or risk-neutral attitudes of the decision makers, different alternatives might be preferred. Additionally, the credibility and trust in risk analyses can be preserved, since the impression of certain estimates is avoided and uncertainties in the results are communicated clearly.

BN-FLEMOps can be adapted to individual regions in Europe with an updating approach using empirical data from the target region. This approach is validated in three case studies of varying spatial scale in Germany, Italy, and Austria. In two of the three case studies, the performance of the model improved further with this updating. However, results indicate that the meso-scale approach may be less suitable for data-scarce, low-impact events like in the Austrian Lech case study, where the damage to only 22 objects amounts to a comparably small loss figure of 1.9 million €. Given the shortage of detailed empirical flood loss data for model updating, further tests with respect to adapting the model to different European regions are recommended to better understand and characterize the circumstances under which the updating approach improves model performance. Nevertheless, the proposed approach for probabilistic flood loss modelling in Europe can provide estimates of flood loss that reliably cover reported figures in the 90 % quantile range.

Acknowledgements

This research was supported by the European Union’s Horizon 2020 research and innovation program, through the IMPREX project (Grant Agreement 641811). Parts of this work have been supported by funding for the OASIS1 Future Danube: Multi Hazard and Risk Model demonstrator project through the Climate Knowledge and Innovation Community (Climate-KIC) by the European Institute of Innovation and Technology (EIT) and from the European Union’s Horizon 2020 research and innovation program H2020 Insurance (Grant Agreement 730381). In addition, this work was supported by Climate-KIC through project “SAFERPLACES — Improved assessment of pluvial, fluvial and coastal flood hazards and risks in European cities as a mean to build safer and resilient communities”, Task ID TC2018B_4.7.3-SAFERPL_P430-1A KAVA2 4.7.3. We are very grateful to alp-S, Innsbruck, Austria (<https://alp-s.at>), particularly Matthias Huttenlau and Holger Cammerer, for providing the data for the Lech case study. We thank Meike Müller from Deutsche Rückversicherung AG for providing the estimation of the residential building values to complement the empirical flood loss data from Germany and Austria. The empirical loss data of the Lech catchment is available via the German flood damage database HOWAS21 (<https://howas21.gfz-potsdam.de/howas21/>). The survey to collect the empirical damage data in the Mulde catchment in Germany was supported by the German Research Network Natural Disasters (German Ministry of Education and Research (BMBF), 01SFR9969/5), and by a joint venture between the German Research Centre for Geosciences GFZ and the Deutsche Rückversicherung AG (<https://deutscherueck.de>). The empirical data from the Mulde catchment may be obtained upon request. We thank Hans de Moel from the Institute for Environmental Studies (IVM) of VU Amsterdam for his support with the creation of the European Asset Map. Our thanks go to the JRC for providing the flood hazard map for Europe.

Kapitel 4

Das probabilistische Hochwasserschadensmodell für Wohngebäude – BN-FLEMOps

Übersicht

Hochwasserrisikoanalysen und insbesondere die Schätzungen von Hochwasserschäden sind oft mit hohen Unsicherheiten verbunden. Diese Unsicherheiten können mit deterministischen Modellen nicht oder nur eingeschränkt dargestellt werden. In dieser Studie wird das probabilistische Hochwasserschadensmodell **BN-FLEMOps** vorgestellt. Am Beispiel einer Schadensberechnung für das Hochwasser 2002 an der Mulde wird BN-FLEMOps mit einem Modellensemble aus etablierten Hochwasserschadensmodellen verglichen und die Vorteile probabilistischer Modelle für die Entscheidungsfindung unter Unsicherheit diskutiert. Zudem wird die einfache Anwendung des Modells durch seine Implementierung in die standardisierte Modellierungsplattform **OASIS loss modelling framework (OASIS-lmf)** beschrieben. Die Validierung an der Mulde zeigt, dass die von BN-FLEMOps geschätzten mittleren Schadenswerte in 12 von 19 Gemeinden näher am beobachteten Schaden liegen, als die Mediane des Modellensembles. Darüber hinaus, deckt sich die Streuung der von BN-FLEMOps generierten Wahrscheinlichkeitsdichtefunktion in der Mehrheit der Gemeinden mit der Spannweite des Ensembles. BN-FLEMOps ist durch die probabilistische Methodik gut geeignet, Unsicherheiten in der Hochwasserschadensmodellierung zuverlässig abzubilden, die beispielsweise auch in Kosten-Nutzen-Analysen integriert werden können und Entscheidungsträgern somit eine umfassende Informationsgrundlage bieten.

Publiziert als:

Steinhausen, M., Lüdtkke, S., Schröter, K., Figueiredo, R., & Kreibich, H. (2020): Das probabilistische Hochwasserschadensmodell für Wohngebäude – BN-FLEMOps. *Hydrologie und Wasserbewirtschaftung*, 64(4), 188–199. https://doi.org/10.5675/HYWA_2020.4_2

The probabilistic flood damage model for residential buildings – BN-FLEMOps

Abstract

Flood risk analyses and particularly estimates of flood loss are often associated with a high degree of uncertainty. Deterministic models are unable to reflect and account for these uncertainties, or they are only able to do this to a very limited extent. This study presents the probabilistic flood loss estimation model for the private sector BN-FLEMOps remove. Taking the 2002 flood at the Mulde river as an example, the BN-FLEMOps is compared with a model ensemble of established flood loss models and the advantages of probabilistic models for decision-making under uncertainty are discussed. Furthermore, the implementation of the model in the OASIS Loss Modelling Framework (LMF) is described, which enables its easy application. The validation at the Mulde river shows that the median loss estimates by the BN-FLEMOps are closer to the observed loss in 12 out of 19 municipalities in comparison with the ensemble median. The spread of the BN-FLEMOps generated probability density function coincides with the range of the ensemble member estimates in most municipalities. The BN-FLEMOps with its probabilistic approach is therefore well-suited to reliably quantify uncertainties in flood loss modelling and to serve decision-makers as a comprehensive source of information.

Published as:

Steinhausen, M., Lüdtkke, S., Schröter, K., Figueiredo, R., & Kreibich, H. (2020): *The probabilistic flood damage model for residential buildings – BN-FLEMOps*. *Hydrologie und Wasserbewirtschaftung*, 64(4), 188–199. https://doi.org/10.5675/HYWA_2020.4_2

4.1 Einleitung

Im Zeitraum von 1998 bis 2009 haben Hochwasserereignisse in Europa zu einem Verlust von über 1.000 Menschenleben und ökonomischen Schäden von rund 52 Milliarden € geführt (European Environment Agency, 2010). Mit zunehmender Klimaveränderung und fortschreitendem globalen Wandel, durch einen Wachstum der Vermögenswerte (Exposition) in überschwemmungsgefährdeten Gebieten, werden ein Anstieg der hochwasserbedingten Todesfälle sowie höhere ökonomische Schäden erwartet (Dottori et al., 2018). Um diesen Entwicklungen entgegenzuwirken, müssen erhebliche Anstrengungen im Management zur Reduzierung des Hochwasserrisikos unternommen werden. Die Grundlage eines effizienten Hochwasserrisikomanagements sind verlässliche Risikoanalysen auf verschiedenen räumlichen Skalen unter expliziter Betrachtung damit verbundener Unsicherheiten (Apel et al., 2004; Merz et al., 2004; Merz et al., 2010b; de Moel et al., 2015). Für eine Risikoanalyse muss neben einer Simulation der Gefährdung auch eine Abschätzung der zu erwartenden Schäden erfolgen. Schadensmodellierungen liefern quantitative Abschätzungen der zu erwartenden Hochwasserschäden (LfULG, 2012), werden aber bislang bei der Erstellung von Hochwasserrisikokarten und Hochwasserrisikomanagementplänen nicht durchgängig eingesetzt, stattdessen werden die nachteiligen Folgen auf verschiedene wirtschaftliche Tätigkeiten qualitativ erfasst und dargestellt (LAWA, 2018).

Die am weitesten verbreiteten Modelle zur Abschätzung des monetären Schadens sind Wasserstands-Schadens-Funktionen, die den Hochwasserschaden auf Basis der exponierten Objekte (z. B. Wohngebäude) und der Überflutungstiefe abschätzen (Merz et al., 2010b; Meyer et al., 2013). Da die Hochwasserschadensschätzungen mit relativ hohen Unsicherheiten behaftet sind (Merz et al., 2004; Merz et al., 2010b), wurden multivariable Hochwasserschadensmodelle entwickelt, die zusätzlich zum Wasserstand andere Größen bei der Schadensschätzung berücksichtigen. Solche komplexeren Modelle können die Schadensprozesse bei Hochwasser besser abbilden und Unsicherheiten in der Schätzungen reduzieren (Gerl et al., 2016). Verwendete Modellierungsansätze sind beispielsweise regelbasierte Modelle (Penning-Rowse et al., 2005; Zhai et al., 2005; Elmer et al., 2010) oder auf Regressionsbäumen basierende Ansätze (Kreibich et al., 2017b; Sieg et al., 2017). Zur Entwicklung von probabilistischen Schadensmodellen wurden Bayessche Netze verwendet (Schröter et al., 2014; Wagenaar et al., 2018), welche den großen Vorteil haben, dass neben der Schadensschätzung gleichzeitig auch eine Quantifizierung der Unsicherheiten erfolgt und einem geschätzten Schadenswert eine Eintrittswahrscheinlichkeit zugeordnet wird. Die zusätzliche Quantifizierung und Bereitstellung von Unsicherheitsinformationen in Bezug auf die Schadens- und Risikoabschätzung sind besonders wichtig. Unsicherheitsabschätzungen verbessern die Risikoanalysen, unterstützen deren Plausibilisierung und sind essenziell für eine fundierte und robuste Entscheidungsfindung im Risikomanagement (Pappenberger & Beven, 2006; Merz et al., 2008; de Brito & Evers, 2016). Im Vergleich zu konkreten quantitativen Unsicherheitsaussagen ist eine einfache Schätzung des Schadens mit einem deterministischen Modell eine sehr eingeschränkte Information, die zu einer falschen Risikobewertung und somit falschen Investitionsentscheidungen führen kann (Wagenaar et al., 2016). Die Anwendung von mehreren Modellen in einem Ensemble kann Abhilfe schaffen (Figueiredo et al., 2018), erfordert aber erheblichen zusätzlichen Aufwand für die Auswahl der für das Untersuchungsgebiet und die Fragestellung geeigneter Modelle (Gerl et al., 2016) in der Beschaffung von Daten sowie der Implementierung und Anwendung der Modelle. Probabilistische Modelle hingegen liefern die Unsicherheitsinformation als

Bestandteil der Modellergebnisse. Zu den Schadenswerten wird demnach auch eine Eintrittswahrscheinlichkeit vom Modell ausgegeben.

Ein Nachteil der Mehrparametermodelle ist jedoch die aufwendigere Anwendung im Vergleich zu den Wasserstands-Schadens Funktionen, da mehr Eingangsparameter benötigt und komplexere Modellstrukturen verwendet werden. Dieser Umstand hat bisher ihre Verbreitung in der wasserwirtschaftlichen Praxis erschwert. Die OASIS-Initiative (<https://oasislmf.org>), ein OpenSource-Projekt der Versicherungswirtschaft zur Risikomodellierung für Naturgefahren, stellt eine Modellierungsplattform bereit, mit der Daten und Modelle entsprechend einem vorgegebenen Standard verarbeitet werden können. Der OASIS-Standard erlaubt eine flexible und transparente Implementierung, gerade auch von probabilistischen Modellierungsansätzen. Die Anwendung und Kombination unterschiedlicher Modelle wird dadurch deutlich erleichtert. Zudem ermöglichen die Standardisierung und Verwendung freier Software und freier Daten einen iterativen Informationsaustausch zwischen Wissenschaft und Praxis. Dieser ist bei der Entwicklung neuer Ansätze und der Übertragung wissenschaftlicher Erkenntnisse in die praktische Anwendung unabdingbar, auch um die unterschiedlichen Erfahrungen, Bedürfnisse und Ideen zu verstehen und gesellschaftlichen Nutzen zu generieren (Morss et al., 2005; Pappenberger & Beven, 2006).

Neben Forschung und kommerzieller Anwendung können großräumiger Hochwasserschadensmodellierungen auch einen Beitrag zur Unterstützung politischer Entscheidungsfindung liefern. Europaweite Hochwasserrisikoanalysen sind für die Mitgliedstaaten der Europäischen Union (EU) wichtig, um beispielsweise die Entwicklung ihrer Klimaanpassungsstrategien zu unterstützen (van Renssen, 2013), den EU-Solidaritätsfonds zu verwalten (Hochrainer et al., 2010), und um Hochwasserrisikomanagementpläne zu erstellen, die durch die Hochwassermanagementrichtlinie gefordert werden (European Union, 2007). Nationale bis regionale Risikoanalysen werden für die Entscheidungsunterstützung im Risikomanagement und die Erstellung von Risikokarten benötigt (Thielen et al., 2016). Die in Deutschland für das Hochwasserrisikomanagement zuständigen Bundesländer haben Risikoanalysen durchgeführt, um ihre Investitionen in den Hochwasserschutz zu priorisieren und um Gefahren- und Risikokarten für die Raumplanung und die Risikokommunikation zu erstellen. Es sind außerdem bundesländerübergreifende Analysen erforderlich, zum einen für das Nationale Hochwasserschutzprogramm (LAWA, 2014), aber auch zur Unterstützung übergreifender Kooperationen, z. B. von Ober- und Unterliegern, im Hochwasserrisikomanagement und im Katastrophenschutz.

In diesem Beitrag wird das probabilistische Hochwasserschadensmodell BN-FLEMOps vorgestellt und der Nutzen von Unsicherheitsinformationen für die Unterstützung bei der Entscheidungsfindung in der Risikomanagementplanung diskutiert. Das Modell wird beispielhaft an der Mulde angewendet und anhand offizieller Schadensdaten für das Hochwasser im August 2002 validiert. Ein Vergleich mit anderen Schadensmodellen dient der Plausibilisierung der Unsicherheiten in den Schadensschätzungen. Weiterhin werden die Anwendungsmöglichkeiten des Modells und die Implementierung im OASIS Loss Modelling Framework (OASIS-lmf) diskutiert.

4.2 Methoden und Daten

4.2.1 BN-FLEMOps Entwicklung und Modellstruktur

Das Modell BN-FLEMOps (Bayesian Network Flood Loss Estimation MOdel for the private sector) wurde für die Berechnung von Hochwasserschäden an privaten Wohngebäuden auf der Mikroskala entwickelt (Wagenaar et al., 2018). Das probabilistische Modell schätzt den relativen, an der Bausubstanz des Gebäudes durch Hochwasser verursachten Schaden und quantifiziert die Unsicherheiten der Schadensschätzung. Dieser relative Schadenswert zwischen 0 und 1 berechnet sich aus dem tatsächlichen Hochwasserschaden in Euro im Verhältnis zu den Wiederaufbaukosten des Gebäudes. In diesem Wert beinhaltet sind alle Kosten an Material und Lohn zur Reparatur des durch das Hochwasser verursachten Schadens.

Bayessche Netze, wie auch in BN-FLEMOps angewendet, sind graphische Modelle, die die gemeinsame Wahrscheinlichkeitsverteilung (P) von Variablen (X_i) unter Berücksichtigung von bedingten Abhängigkeiten beschreiben und dem Aufbau eines gerichteten azyklischen Graphen folgen. Die Variablen werden als Knoten dargestellt und ihre Verbindungen über Kanten repräsentieren die bedingten Abhängigkeiten zwischen den Variablen. Es wird zwischen „Eltern-Knoten“, von denen eine Kante wegführt, und „Kinder-Knoten“, auf die eine Verbindung gerichtet ist, unterschieden. Die Richtung der Kanten darf jedoch bei einem aus Daten gelernten Netz nicht zwangsläufig als kausaler Zusammenhang interpretiert werden (Vogel et al., 2018). Für die Berechnung des Gebäudeschadens ist die Richtung der Kanten nicht entscheidend. In einem Bayesschen Netze kann Inferenz in beide Richtungen der Kanten erfolgen (Fenton & Neil, 2013). Alle in das Netz eingehenden Daten werden als unsicherheitsbehaftete Zufallsvariable betrachtet, deren probabilistische Abhängigkeit im Bayesschen Netz abgebildet wird. Die gemeinsame Wahrscheinlichkeitsverteilung der Variablen beinhaltet somit auch die in den Eingangsdaten enthaltenen Unsicherheiten. An jedem Knoten des Netzes ist eine bedingte Wahrscheinlichkeitsfunktion durch die Variablen der „Eltern-Knoten“ definiert und kann mit der folgenden Gleichung berechnet werden (Fenton & Neil, 2013):

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i \mid \text{Eltern}(X_i)) \quad (4.1)$$

Die Netzstruktur des Modells BN-FLEMOps ([Abbildung 4.1](#)) basiert auf der statistischen Auswertung einer Hochwasserschadensdatenbank, in der Informationen aus Umfragekampagnen zu Hochwasserereignissen aus den Jahren 2002, 2005, 2006, 2010, 2011 und 2013 enthalten sind. Eine detaillierte Übersicht und Analyse dieser Daten und deren Einsatz bei der Entwicklung von Hochwasserschadensmodellen bieten Merz et al. (2013) und Schröter et al. (2016). Aus dieser Datenbank wurden 1.522 vollständige Datensätze von Gebäudeschäden verwendet, um das Modell zu trainieren. Die Auswahl der Variablen und ihre Verbindung im Netz wurden aus den statistischen Zusammenhängen abgeleitet und mit Expertenwissen angepasst. Die Netzstruktur wurde unter Verwendung von drei im R-paket "bnlearn" (Scutari, 2010) implementierten Lernalgorithmen, Fast Incremental Association ([Fast-IAMB](#)) (Yaramakala & Margaritis, 2005), Interleaved Incremental Association ([Inter-IAMB](#)) (Tsamardinos et al., 2013) und einem Bayesian Dirichlet equivalence ([BDe](#)) Hill-Climbing-Optimierungsansatz (Heckerman et al., 1995) aufgebaut. Jeder der Algorithmen wurde 500-mal angewendet, um insgesamt 1.500 Netze aus den empirischen

Daten zu lernen. In diesen Wiederholungen beim Bau der Netzwerkstruktur kam jeweils nur eine zufällig ausgewählte Teilmenge (950) der Interviewdaten zum Einsatz, um eine Überanpassung des Modells zu vermeiden. Die in 80 % aller Netze vorkommenden Verbindungen aus Knoten und Kanten wurden, in Kombination mit Überlegungen zur Datenverfügbarkeit und Forschungszielen, für die Entwicklung der in [Abbildung 4.1](#) gezeigten Struktur von BN-FLEMOps verwendet (Wagenaar et al., 2018).

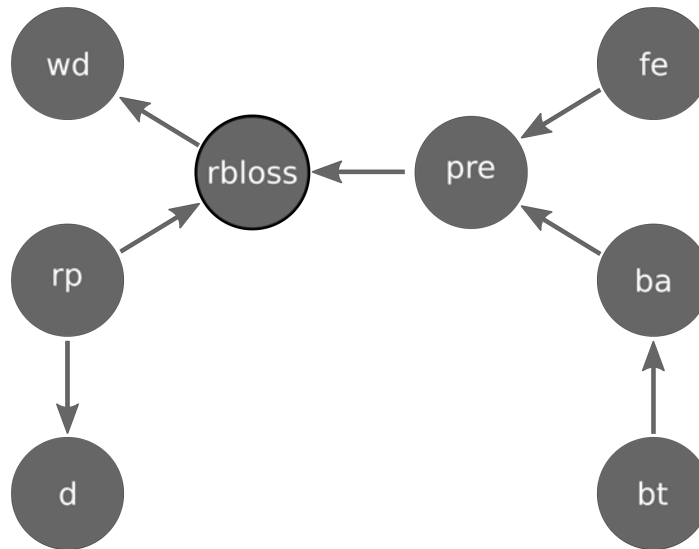


Abbildung 4.1: Modellstruktur des BN-FLEMOps mit Variablennamen (angepasst aus Wagenaar et al. (2018)). rbloss = relativer Gebäudeschaden, wd = Wasserstand, rp = Hochwasserjährlichkeit, d = Dauer der Überflutung, pre = private Vorsorgemaßnahmen, fe = Hochwassererfahrung, ba = Gebäudegrundfläche, bt = Gebäudetyp.

Model structure of the BN-FLEMOps with variable names (adapted from Wagenaar et al. (2018)). rbloss = relative building loss, wd = water depth, rp = return period, d = inundation duration, pre = precautionary measures, fe = flood experience, ba = footprint area of the building, bt = building type.

Informationen zum Wasserstand (wd), der Hochwasserjährlichkeit (rp) und der Überflutungsdauer (d) beschreiben im Modell die Intensität der Überflutung. Die Widerstandsfähigkeit der exponierten Objekte gegenüber Hochwasser wird durch die Variablen Gebäudetyp (bt), Gebäudefläche (ba), Hochwassererfahrung (fe) und private Hochwasser-Vorsorgemaßnahmen (pre) beschrieben. Um den relativen Gebäudeschaden (rbloss) zu schätzen, werden nur die direkt mit dem Knoten "rbloss" verbundenen Variablen benötigt. Das heißt, ist beispielsweise die Jährlichkeit des Hochwasserereignisses bekannt, so spielt die Dauer der Überflutung keine Rolle mehr für die Schadensschätzung. Sind jedoch die privaten Vorsorgemaßnahmen unbekannt, erfolgt die Schadensschätzung über die im Netz hinter diesem Knoten stehenden Variablen Gebäudefläche und Hochwassererfahrung. Dabei wird jede mögliche Ausprägung der Variable (pre) im Netzwerk berücksichtigt. So ist eine Schadensschätzung trotz unvollständiger Eingangsdaten möglich. Eine fehlende Variableninformation kann jedoch zu einer höheren Unsicherheit in den Modellergebnissen führen.

Die oben genannten Variablen gehen diskretisiert in das Modell BN-FLEMOps ein. Der geschätzte relative Schaden sowie der Wasserstand besitzen beispielsweise 10, die Hochwasserjährlichkeit 5 Klassen ([Tabelle 4.1](#)). Die Einteilung der Klassenbreiten

erfolgte, basierend auf den 1.522 Eingangsdatensätzen, nach dem Prinzip gleichmäßig gefüllter Klassen. Die Variablen mit der höchsten Auflösung und dem größten Beitrag zur Erklärung des Schadens besitzen eine höhere Anzahl von Klassen, z. B. Wasserstand. Eine detaillierte Beschreibung der Methodik zur Herleitung des Bayesschen Netzes und aller Variablen ist in Wagenaar et al. (2018) publiziert. Durch die Verwendung europaweit verfügbarer Proxydaten für die Variablen konnte BN-FLEMOps von der Mikroskala (Objekt genau) auf die Mesoskala (flächenaggregiert) übertragen werden (Unterabschnitt 4.2.2). Hierzu wurden die punktuell verfügbaren Umfragedaten mit flächendeckenden Datensätzen statistisch verglichen und durch diese ersetzt. Das so auf die Mesoskala übertragene Modell wurde anhand von drei europäischen Fallstudien validiert (Lüdtke et al., 2019b).

4.2.2 Datengrundlage und -verarbeitung

Im Folgenden werden die für die Validierung an der Mulde verwendeten Datensätze und die aus ihnen extrahierten Modellvariablen (Wasserstand, Jährlichkeit, Hochwassererfahrung und Gebäudegrundfläche) beschrieben sowie die notwendigen Verarbeitungsschritte zur Hochwasserschadensberechnung mit dem Modell BN-FLEMOps dargelegt. Tabelle 4.1 zeigt eine Übersicht der im Modell repräsentierten Variablen. Diese Übersicht schließt auch die Variablen des Netzes ein, welche nicht für die Anwendung an der Mulde verwendet wurden. Diese Variablen können bei Anwendungen mit anderer Datenlage verwendet werden.

Tabelle 4.1: Übersicht der Modellvariablen von BN-FLEMOps.
Overview of the model variables of BN-FLEMOps.

Variable	Abkürzung	Einheit	Klassen
Relativer Gebäudeschaden	rbloss	Relative (0 bis 1)	10
Wasserstand	wd	Meter	10
Hochwasserjährlichkeit	rp	Jahre	5
Dauer der Überflutung	d	Stunden	5
Private Vorsorgemaßnahmen	pre	Rang	3
Hochwassererfahrung	fe	Rang	6
Gebäudegrundfläche	ba	Quadratmeter	3
Gebäudetyp	bt	Index	3

Ereignishydraulik und -jährlichkeit

Die Überflutungsflächen und -tiefen für das Hochwasser im August 2002 an der Mulde (Abbildung 4.2) wurden durch eine hydrodynamische 1D/2D-Simulationsrechnungen im Stadtgebiet von Eilenburg (Apel et al., 2009) in Kombination mit einer linearen Interpolation von Wasserständen entlang der Freiburger und Vereinigten Mulde ermittelt (Grabbert, 2006). Das verwendete hydrodynamische Modell ist das LISFLOOD-FP-Modell von Bates und De Roo (2000), für das als obere Randbedingung die Abflussganglinie des nächsten oberstrom liegenden Pegels Golzern verwendet wurde. Das hydrodynamische Modell wurde auf Grundlage eines digitalen Geländemodells (25 m-Auflösung) (BKG, 2007) durch Anpassung der Oberflächenrauigkeit kalibriert. Hierfür wurden außerdem eine auf Fernerkundungsdaten basierende Flutmaske sowie 380 Hochwassermarken im Stadtgebiet von Eilenburg mit einbezogen. Die lineare Interpolation der Wasserstände entlang der Mulde liefert eine gleichmäßig abfallende Wasserspiegellage, die mit den Höheninformationen des digitalen Höhenmodells verschnitten wurde. Alle Bereiche, die niedriger liegen als diese interpolierte Hochwasserspiegellage, werden als überflutet angesehen. Die Überflutungstiefe resultiert aus der Differenz zwischen Wasserspiegeloberfläche und Geländehöhe. Aus dem Raster der Überflutungstiefen wird die Variable Wasserstand (wd) extrahiert.

Die Hochwasserjährlichkeit (rp) wurde am Pegel Bad Dueben, nördlich von Laußig, auf ca. 500 Jahre geschätzt (Elmer et al., 2010). Die Eingangsdaten für die Variablen Wasserstand und Hochwasserjährlichkeit entsprechen den in Figueiredo et al. (2018) verwendeten Daten. So wird eine höchst mögliche Vergleichbarkeit der Modelle des Ensembles mit BN-FLEMOps erreicht.

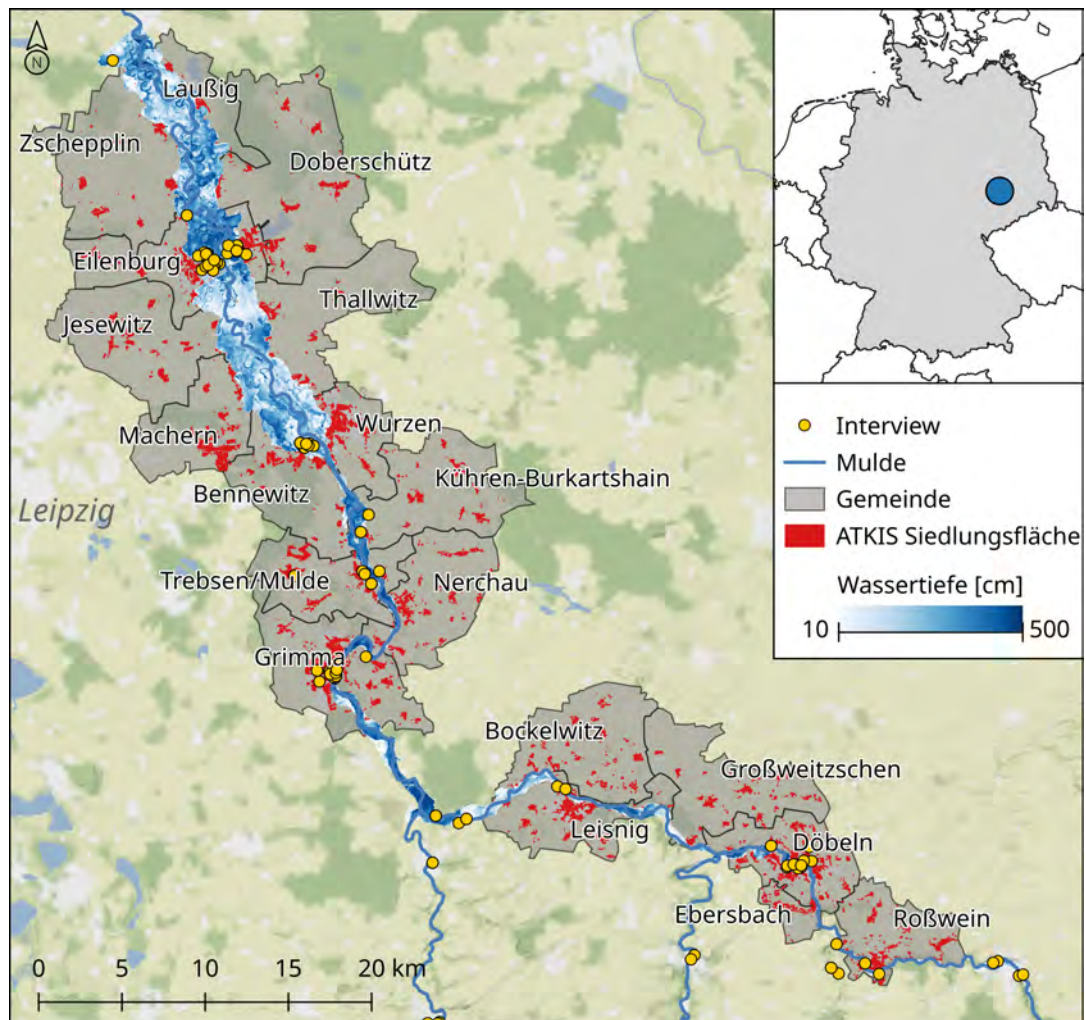


Abbildung 4.2: Übersichtskarte des Untersuchungsgebietes an der Mulde mit Gemeinden mit offiziellen Schadenswerten.

Overview map of the study area at the Mulde river with municipalities with official reported loss figures.

Gebäudegrundflächen und Hochwassererfahrung

Die Modell-Variablen Gebäudegrundfläche (ba) und Hochwassererfahrung (fe) charakterisieren die Widerstandsfähigkeit des Wohngebäudes gegenüber Hochwasser. Zur Übertragung des Modells BN-FLEMOps auf die Mesoskala wurden europaweit verfügbare Datensätze für die Eingangsvariablen der Gebäudefläche und der Hochwassererfahrung auf ihre Anwendbarkeit geprüft. Hierzu wurden die europaweiten Datensätze mit den Umfragedaten der Hochwasserdatenbank statistisch verglichen. In diesem Vergleich wurden die OpenStreetMap (OSM) (OSM contributors, 2018) und der Hochwasserskataloge des Dartmouth Flood Observatory (DFO) (Brakenridge, 2018) als Datengrundlage auf der Mesoskala identifiziert.

Die OSM bietet eine Geodatenbank mit Gebäudegeometrien, die als zuverlässige Quelle für zivile Anwendungen frei genutzt werden kann (Barrington-Leigh & Millard-Ball, 2017). Um die Grundflächen der Wohngebäude zu berechnen, wurden gezielt nur Gebäude mit Wohnnutzung aus dem Datensatz extrahiert. Hierzu wurden die im Hintergrund liegenden Landnutzungsinformationen aus OSM verwendet, um den Gebäuden eine Nutzungsform zuzuschreiben und so Wohngebäude von gewerblicher oder industrieller Nutzung zu unterscheiden. Die OSM-Daten zeigen im Mittel eine gute Übereinstimmung mit den Gebäudeflächen aus den Umfragedaten. Da OSM eine von Nutzern erstellte und kontrollierte Datenbank ist, kann es zu Ungenauigkeiten und Unvollständigkeiten in den Daten kommen. Hecht et al. (2013) stellen in ihrer Untersuchung fest, dass OSM-Daten in Deutschland noch unvollständig sind, aber in den nächsten Jahren durch von staatlichen Stellen und Nutzern hinzugefügte Daten an Vollständigkeit gewinnen werden. In einem statistischen Vergleich der Gebäudegrundflächen zwischen den aus Umfragen nach Hochwasserereignissen gewonnenen und aus OSM extrahierten Daten, stellt Lüdtker et al. (2019b) fest, dass die OSM-Geometrien, im Mittel über 25 NUTS-3-Regionen (Nomenclature des Unités Territoriales Statistiques; Nomenklatur Territorialer Einheiten für Statistiken der EU), 18 m² kleiner sind.

Frühere Studien haben bereits den Zusammenhang zwischen der Erfahrung der Bevölkerung mit Hochwasser und der Höhe von Hochwasserschäden gezeigt und als Variable zur Schätzung in Schadensmodellen verwendet (Thieken et al., 2005; Merz et al., 2013; Schröter et al., 2014). Zur Beschreibung der Hochwassererfahrung auf der Mesoskala wird die Anzahl Hochwasserereignisse der Region in den letzten 25 Jahren verwendet. Diese Zahl wird auf der Datenbasis des DFO berechnet. Diese Datenbank enthält Informationen über die räumliche Ausdehnung, die Hochwasserursache, den Startzeitpunkt etc. von vergangenen Hochwasserereignissen seit dem Jahr 1985. Die Geometrien der Hochwasserereignisse im DFO-Katalog neigen dazu, die Ausdehnung des Hochwassers undifferenziert und zu groß abzubilden (Brakenridge, 2018) und so die Anzahl der vergangenen Ereignisse in jedem Punkt zu überschätzen. Dieser Bias in den DFO-Daten wurde korrigiert, indem die gezählten Ereignisse um 1 reduziert wurden. Durch diese Bias-Korrektur wurde eine gute Übereinstimmung mit den Umfragedaten erzielt. Für die Variablen Gebäudetyp und private Vorsorgemaßnahmen konnten keine geeigneten Datensätze gefunden werden. Die Funktionsweise und Struktur des Bayesschen Netztes macht die Variable des Gebäudetyps (bt) für den in dieser Studie verwendeten Datensatz überflüssig und erlaubt es, mit den beiden „Eltern Knoten“ Hochwassererfahrung (fe) und Gebäudegrundfläche (ba) den relativen Schaden zu schätzen. Das Fehlen der Informationen für die Variable private Vorsorgemaßnahmen (pre), führt zur Berücksichtigung aller Ausprägungen der Variable bei der Schadensschätzung und vergrößert die Unsicherheit der Schätzung (Unterabschnitt 4.2.1).

Gebäudewerte

Die Wohngebäudewerte pro Gemeinde, basierend auf den Wiederherstellungskosten für das Jahr 2000, wurden mithilfe des Baupreisindex auf das Ereignisjahr 2002 hochgerechnet (Kleist et al., 2006), und auf das digitale Landschaftsmodell ATKIS disaggregiert (Abbildung 4.2) (Wünsch et al., 2009). Offizielle Schadensdaten zum Hochwasser im August 2002 auf Gemeindeebene wurden von der Sächsischen Aufbaubank (Saxon Relief Bank, 2005) erfasst und zur Verfügung gestellt. Der gesamte Hochwasserschaden an Wohngebäuden bezogen auf das Hochwasser im August

2002 beträgt aufsummiert über alle Gemeinden im Einzugsgebiet der Mulde 240,6 Millionen € (Kreibich et al., 2017c).

Datenverschneidung und Schadensschätzung

Um die Eingangsdaten für das Modell BN-FLEMOps vorzubereiten, werden die Datensätze in einem Geoinformationssystem räumlich miteinander verschneidet. Durch die Verschneidung entstehen Teilflächen mit allen für das Modell notwendigen Variablen. Jede Teilfläche trägt Informationen zum Wasserstand (wd), der Jährlichkeit (rp), der Gebäudegrundfläche (ba) und der Hochwassererfahrung (fe), auf deren Basis der relative Hochwasserschaden für jede dieser Teilflächen modelliert wird. Eine durch die Verschneidung entstandene Teilfläche kann beispielsweise die folgende Kombination von Eingangsdaten (wd = 120 cm, rp = 500 a, ba = 100 m², fe = 2), einem Gebäudewert von 625.000 € und einem Gemeinamen bestehen. BN-FLEMOps errechnet dann eine in 10 Klassen diskretisierte Schadensverteilung mit entsprechender Eintrittswahrscheinlichkeit pro Schadensklasse. Es werden 10.000 Realisationen des Modells gerechnet und die Mediane der Klassen zur weiteren Berechnung der absoluten Schäden verwendet. Die Gebäudewerte werden mit den relativen Schadenswerten multipliziert, um den absoluten Schaden in Euro zu erhalten. Anhand der ebenfalls enthaltenen Gemeindekennung erfolgt eine Aggregation der geschätzten Schäden auf Gemeindeebene. Weitere Details und eine Abbildung zur Datenverschneidung können Lüdtko et al. (2019b) entnommen werden.

4.2.3 Schadensmodelle des Ensembles

Vergleichsschätzungen für die direkten Wohngebäudeschäden wurden mit einem Ensemble von Hochwasserschadensmodellen durchgeführt. Diese Studie greift dafür auf Ergebnisse von Figueiredo et al. (2018) für das gleiche Untersuchungsgebiet und Hochwasserereignis zurück. Als Vergleichsmodelle wurden zunächst alle Modelle für direkte Schäden an Wohngebäuden aus dem Katalog von Gerl et al. (2016) ausgewählt und hinsichtlich ihrer Eignung für die Gegebenheiten im Muldegebiet geprüft (Figueiredo et al., 2018). Insgesamt wurden zwanzig deterministische Hochwasserschadensmodelle ausgewählt und im Untersuchungsgebiet für das Hochwasser im August 2002 angewendet.

Die Modelle sind in [Tabelle 4.2](#) zusammen mit den Informationen zu wichtigen Eigenschaften und Quellen, in denen die Modelle detailliert beschrieben sind, aufgeführt. Dieses Ensemble von Modellen bildet eine Vielfalt weltweit verfügbarer und geeigneter Modelle ab und liefert somit eine realistische Streuung möglicher Schadensschätzungen. Die Mehrzahl der Modelle wurde in einem europäischen Kontext entwickelt. Die Modelle IKSE und FLEMOps+r wurden auf der Basis von Daten aus dem Elbeeinzugsgebiet entwickelt und stehen daher in enger räumlicher Beziehung zum Untersuchungsgebiet der Mulde. Im Ensemble werden sowohl einfache Wasserstands-Schadens-Funktionen mit nur einer Intensitätsvariablen, als auch komplexere multivariable Modelle berücksichtigt. Die Mehrheit von fünfzehn Modellen generiert relative Schadenszahlen zwischen 0 und 1. Die übrigen 5 der 20 Modelle berechnen absolute Schadenswerte (ANUFlood, HOWAS, MCM, TyrOL, Vojnovic).

Tabelle 4.2: Übersicht der Charakteristika der Hochwasserschadensmodelle im Ensemble und des probabilistischen Modells BN-FLEMOps (nach Figueiredo et al. (2018)).
Overview of the ensemble flood loss model characteristics and the probabilistic model BN-FLEMOps (acc. to Figueiredo et al. (2018)).

Modellname	Ereignis variablen*	Expositions variablen*	Land	Referenzen
BN-FLEMOps	wd, rp, d	pre, fe, ba, bt	Deutschland	Wagenaar et al. (2018) Lüdtke et al. (2019a)
ANUFlood	wd	fa	Australien	NR&M (2002)
Budiyono	wd	bt	Indonesien	Budiyono et al. (2015)
DSM	wd	bt	Niederlande	Klijn et al. (2007)
Dutta	wd	str	Japan	Dutta et al. (2003)
FLEMOps+r	wd, con, rp	bt, bq, pre	Deutschland	Elmer et al. (2010)
HAZUS-MH	wd	bt, nf, bas	USA	Scawthorn et al. (2006)
HOWAS	wd	bt, bas	Deutschland	Buck und Merkel (1999)
HWS-GIS	wd	–	Deutschland	Hydrotec (2002)
ICPR	wd	–	Schweiz, Deutschland, Frankreich, Niederlande	ICPR (2001)
IKSE	wd	–	Deutschland	IKSE (2003)
Luino	wd	–	Italien	Luino et al. (2009)
MCM	wd, id	bt	England, Wales	Penning-Rowse et al. (2005)
MERK	wd	nf, bas	Deutschland	Rees et al. (2003)
Pistrika & Jonkman	wd, fv	–	USA	Pistrika und Jonkman (2010)
Riha & Marcikova	wd, id	bt, oth	Tschechische Republik	Riha und Marcikova (2009)
Toth	wd	bt, str, nf	Ungarn	Tóth et al. (2008)
TYROL	wd	–	Österreich	Huttenlau et al. (2010)
Vanneuville	wd	bt	Belgien	Vanneuville et al. (2006)
Vojinovic	wd	fa	Sint Maarten	Vojinovic et al. (2008)
Yazdi & Neyshabouri	wd	–	Iran	Yazdi und Neyshabouri (2012)

* wd = Wasserstand, rp = Hochwasserjährlichkeit, d = Dauer der Überflutung, pre = private Vorsorgemaßnahmen, fe = Hochwassererfahrung, ba = Gebäudegrundfläche, bt = Gebäudetyp, con = Kontamination, bq = Gebäudequalität, nf = Stockwekanzahl, bas = Keller, str = Gebäudestruktur, fa = Geschossfläche, fv = Fließgeschwindigkeit, id = Dauer der Überflutung

4.3 Ergebnisse und Diskussion

4.3.1 Schadensschätzung und Modellvergleich

Die Ergebnisse der Schadensschätzung mit dem probabilistischen Modell BN-FLEMOps im Vergleich zu den Modellen des Ensembles und den offiziellen Schadensangaben pro Gemeinde (Saxon Relief Bank, 2005) sind in [Abbildung 4.3](#) dargestellt. Die Spannweite des Modellensembles wird in einem Box-Plot den Ergebnissen von BN-FLEMOps, in einem sogenannten Violin-Plot gegenübergestellt. Der Violin-Plot bildet die auf Gemeindeebene aggregierten Schadensschätzungen aller Teilflächen in einer an der y-Achse gespiegelten Dichteverteilung ab. Die waagerechten Linien zeigen die Mediane der Schadensverteilungen im Box-Plot und Violin-Plot an. Die individuellen Ergebnisse der Ensemble-Modelle sind als blaue Punkte im Box-Plot dargestellt. Der offizielle in den Gemeinden erfasste Schaden (Saxon Relief Bank, 2005) ist als Referenzpunkt rot eingetragen.

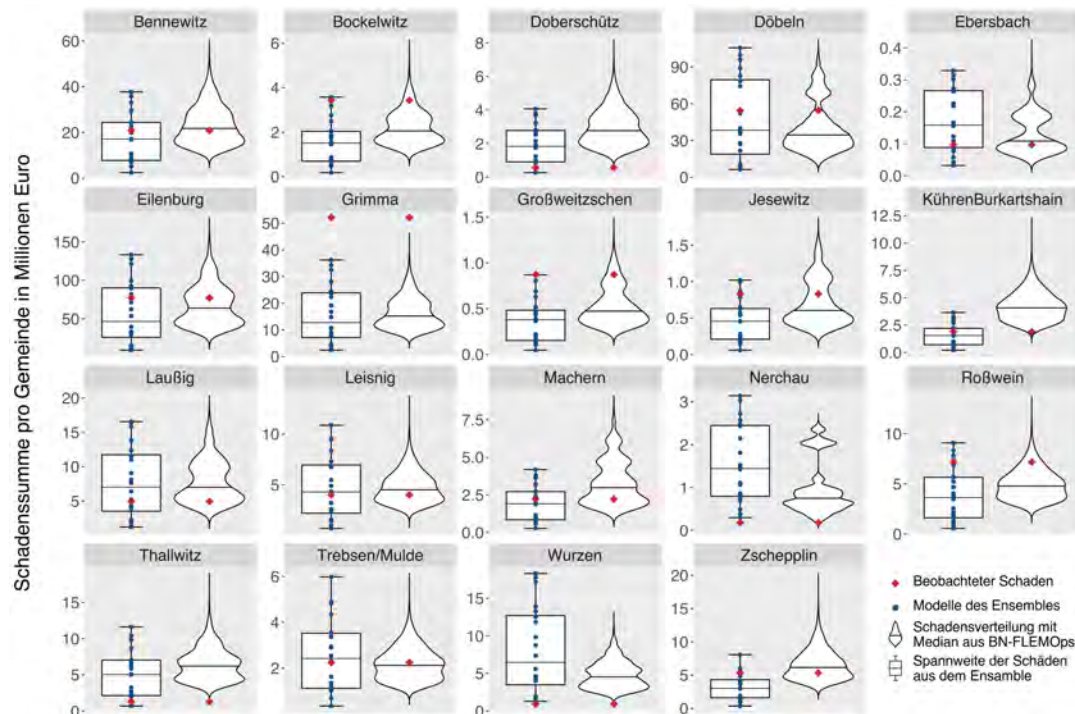


Abbildung 4.3: Ergebnisse der probabilistischen Schadensschätzung mit BN-FLEMOps im Vergleich zu den Schadensmodellen im Ensemble und den offiziellen Schadensangaben vom Hochwasser 2002 (Saxon Relief Bank, 2005) in 19 Gemeinden an der Mulde.

Results of the probabilistic flood loss estimation with BN-FLEMOps in comparison with ensemble models and official loss figures (Saxon Relief Bank, 2005) in 19 municipalities along the Mulde river.

In 12 der 19 Gemeinden liegt der Median der Schadensverteilung von BN-FLEMOps im Vergleich zum Median des Ensembles näher am offiziell beobachteten Schaden. In nahezu allen Gemeinden zeigt sich, dass die Verteilung des mit BN-FLEMOps geschätzten Schadens den beobachteten Schadenswert sowie die gesamte Spannweite der Ensemble-Modelle abdeckt. Stark unterschätzt wird der Schaden durch BN-FLEMOps sowie das Modellensemble nur in der Gemeinde Grimma. Diese Unterschätzung des Schadens liegt mit hoher Wahrscheinlichkeit begründet in den zu

niedrig simulierten Wasserständen und vergleichsweise niedrig geschätzten Gebäudewerten in der Gemeinde Grimma. Die verwendete Hochwassertiefenkarte, welche teilweise auf linearer Interpolation und einem gering aufgelösten Geländemodell basiert, führt hier zur Unterschätzung der Hochwasserschäden. Apel et al. (2009) zeigt, dass komplexere Verfahren zur Wasserstandsmodellierung die Ergebnisse von Schadensmodellen verbessern können. Der Autor stellt jedoch auch fest, dass der Einfluss unterschiedlicher Hazardmodelle auf die Schätzung des Hochwasserschadens deutlich unter der Variabilität verschiedener Schadensmodelle liegt.

Im Vergleich aller Modelle liegt die Performance von BN-FLEMOps, gemessen anhand der Indikatoren „Wurzel der mittleren quadratische Abweichung“ (rmse = 10,25) und der „mittleren typischen Abweichung“ (mbe = -2,93) über alle Gemeinden, an 5. Stelle von 21 Modellen (Tabelle 4.3). Die mittlere typische Abweichung über alle Gemeinden beträgt weniger als 3 Millionen € (-23,26 %). Das Modell BN-FLEMOps tendiert demnach zu einer leichten Unterschätzung des Hochwasserschadens. Diese rührt maßgeblich aus der Unterschätzung der vergleichsweise hohen Schäden in den Gemeinden Grimma und Döbeln. Die in OSM leicht unterrepräsentierte Gebäudegrundfläche hat ebenfalls einen Einfluss auf die negative mittlere typische Abweichung der Modellergebnisse.

Tabelle 4.3: Performance-Indikatoren für alle Modelle des Ensembles und BN-FLEMOps, geordnet nach der „Wurzel der mittleren quadratische Abweichung“ (rmse), alle Werte in Millionen Euro. Für BN-FLEMOps sind die 25 %, 50 % und 75 % Quantile abgebildet.

Performance indicators for the ensemble models and BN-FLEMOps, in ascending order by the root mean square error (rmse). All values are in million Euro. The 25 %, 50 % and 75 % quantiles are presented for BN-FLEMOps.

Modell	rmse	mbe	Gesamtschaden
Luino	8,14	-1,23	217,21
IKSE	9,16	-2,43	194,35
Dutta	9,18	1,87	276,12
DSM	9,47	1,36	266,39
FLEMOps+r	10,92	-3,85	167,42
HAZUS-MH	10,96	4,00	316,55
Riha & Marcikova	11,45	2,99	297,31
Vanneuville	13,61	-5,30	49,93
Toth	13,91	-6,05	125,63
MCM	14,41	-4,27	131,54
HWS-GIS	15,80	-7,24	103,08
ICPR	15,89	-7,20	103,75
MERK	16,50	-7,66	95,12
Pistrika & Jonkman	16,88	8,24	397,05
Yazdi & Neyshabouri	17,17	7,40	381,14
Budiyono	18,26	8,19	396,19
Vojinovic	19,10	-8,67	49,93
HOWAS	20,98	-9,86	50,50
TYROL	21,16	-9,98	54,12
ANUFlood	21,56	-10,27	35,53
BN-FLEMOps	10,25	-2,93	Q ₇₅ = 246,62 Q ₅₀ = 184,84 Q ₂₅ = 137,82

Im Vergleich der Modelle des Ensembles schneiden solche Modelle (Luino, IKSE, DSM), welche für eine dem Untersuchungsgebiet ähnliche Geografie und Ökonomie entwickelt wurden, besonders gut ab. Die Performance der Modelle lässt keine besonderen Muster bezüglich der Komplexität und Variablenauswahl erkennen. Sowohl einfache Wasserstands-Schadens-Funktionen (Luino, IKSE) als auch komplexere

Multi-variable Modelle (FLEMOps+r, HAZUS-MH) liefern akkurate Schadensschätzungen.

Das Beispiel der Gemeinde Leisnig zeigt eine präzise Schadensschätzung mit geringer Unsicherheit. Der Median der Verteilung liegt hier dicht am beobachteten Schaden und die Streuung der Verteilung ist gering. Die Ergebnisse von BN-FLEMOps können in dieser Gemeinde mit hoher Sicherheit als zuverlässig angesehen werden. Für die Gemeinde Machern besitzt die Verteilung eine deutlich größere Streuung, die Fläche innerhalb des Violin-Plots ist zwischen 2 – 8 Millionen € Schaden aufgespannt. Dies resultiert in einer Überschätzung des Schadens im Vergleich zum Ensemble, sowie der Beobachtung und offenbart eine hohe Unsicherheit der Schätzung.

Der von der SAB erfasste Gesamtschaden der 19 Gemeinden an der Mulde liegt bei 240,58 Millionen €. BN-FLEMOps schätzt diesen Wert im Median auf 184,84 Millionen €, während der Ensemblemedian 149,48 Millionen € beträgt. Das probabilistische Modell zeigt demnach im Mittel ein genaueres Ergebnis als das Ensemble und bildet den beobachteten Schadenswert innerhalb des 50 %-Quantils der Schadensschätzung ab.

4.3.2 Probabilistische Modelle in der Entscheidungsfindung

In der klassischen Entscheidungstheorie steht primär die Optimierung des „Nutzens“ unter Berücksichtigung divergierender Interessen und limitierender Faktoren im Vordergrund. Im Bereich des Hochwasserrisikomanagements müssen die „Nutzen“ verschiedener Maßnahmen für unterschiedlichste Akteure vereint und der Nutzen immer im Verhältnis zu den Kosten einer Maßnahme analysiert werden. Probabilistische Schadensmodelle können im Gegensatz zu deterministischen Modellen zusätzlich die Eintrittswahrscheinlichkeiten des geschätzten Schadens abbilden. Dies kann im Risikomanagement bei der Entscheidungsfindung unter Unsicherheit zur Optimierung von Kosten-Nutzen-Funktion verwendet werden (Polasky et al., 2011). Ist beispielsweise eine Entscheidung zwischen zwei alternativen Hochwasserschutzmaßnahmen zu treffen, könnte ein risikoaverser Entscheidungsträger die Maßnahme mit der geringeren Unsicherheit bevorzugen, da er eine ineffiziente Lösung mit (negativem) Nutzen-Kosten-Verhältnis unter null vermeiden will, selbst wenn die alternative Maßnahme im Mittel ein besseres Nutzen-Kosten-Verhältnis aufweist, aber mit größerer Unsicherheit behaftet ist. Demgegenüber würde ein risikofreudiger Entscheidungsträger die Alternative mit dem höheren mittleren Nutzen-Kosten-Verhältnis bevorzugen, obwohl die Wahrscheinlichkeit eines negativen (mehr Kosten als Nutzen) Nutzen-Kosten-Verhältnisses deutlich vorhanden ist (Merz et al., 2008).

Bei risikobehafteten Entscheidungen können beispielsweise anhand des Erwartungswert-Varianz-Prinzips (μ - σ -Regel) (Markowitz, 1952) Alternativen abgewogen werden. Eine Entscheidung basiert hierbei nicht mehr nur auf einem Mittelwert, sondern auch auf der Varianz der Wahrscheinlichkeitsverteilung aus den Modellergebnissen. Unsicherheiten können auf diesem Weg im Entscheidungsprozess berücksichtigt werden. Risikoaverse oder risikofreudige Akteure können ihrer Entscheidungen so besser fundieren und begründen (Knoke & Wurm, 2006). In einer Literaturstudie zur Entscheidungsfindung im Hochwasserrisikomanagement, in der über hundert aktuelle Forschungsarbeiten berücksichtigt wurden, identifizieren de Brito und Evers (2016) die Berücksichtigung von Unsicherheiten als die Herausforderung mit dem größten Forschungsdefizit. Die Autorinnen erkennen Bayessche-Netzwerk-Methoden als die Ansätze mit dem größten Potenzial, die Unsicherheiten in multifaktoriellen

Entscheidungsanalysen besser zu berücksichtigen. Der hohe Informationsgehalt probabilistischer Modelle ermöglicht die Überprüfung verschiedener Entscheidungsalternativen auf der Grundlage einer einzelnen Modellanwendung. Die beispielhafte Hochwasserschadensberechnung mit BN-FLEMOps an der Mulde zeigt, dass ein ganzes Modellensemble durch ein probabilistisches Modell ersetzt werden konnte. Unsicherheitsinformationen in den Ergebnissen probabilistischer Modelle ermöglichen Prozesse der Entscheidungsfindung entlang wahrscheinlicher, optimistischer oder pessimistischer Szenarien (Polasky et al., 2011).

4.3.3 Anwendbarkeit

Das Modell BN-FLEMOps kann auf unterschiedlichen räumlichen Skalen angewendet werden. Auf der Mikroskala können Schäden für individuelle Gebäude geschätzt werden (Wagenaar et al., 2018). Hierzu werden objektspezifische Daten zum Antreiben des Modells benötigt. Auf der Mesoskala erfolgt die Schätzung von Hochwasserschäden für Landnutzungsflächen mit Wohnbebauung. Eine solche Anwendung, wie sie in dieser Studie gezeigt wird, beruht auf flächendeckenden und räumlich aggregierten Datensätzen. Details zur Übertragung von BN-FLEMOps, welches auf der Basis von mikroskaligen Daten entwickelt wurde, auf die Mesoskala können Lüdtko et al. (2019a) entnommen werden.

BN-FLEMOps eignet sich aufgrund der verwendeten Variablen und der Bayesschen-Netzwerk-Methodik zur Anwendung in verschiedenen Untersuchungsgebieten. Das Modell wurde bereits erfolgreich in Untersuchungsgebieten in den Niederlanden, Deutschland, Österreich und Italien zur Nachberechnung historischer Ereignisse verwendet (Wagenaar et al., 2018; Lüdtko et al., 2019a). Zudem zeigt eine Risikoanalyse unter Verwendung von konsistenten, europaweiten Datensätzen in Lüdtko et al. (2019b) die kontinentale Anwendbarkeit des Modells für Europa. Ebenso ist eine deutschlandweite Anwendung mit detaillierterer Datengrundlage möglich. Hierzu können die bebaute Fläche, Gebäudetypen und -werte auf [ATKIS](#)-Datenbasis verwendet werden (Thieken et al., 2006a). Eine detaillierte Repräsentation der exponierten Werte erhöht die Präzision und Zuverlässigkeit von Schadensmodellierungen (Wünsch et al., 2009).

Hattermann et al. (2018) beschreibt die Einbindung von BN-FLEMOps mit Hochwasser-Zukunftsprojektionen in eine Modellkette zur Analyse von Auswirkungen von Klimaveränderungen auf Hochwasserschäden an der Donau. In einer solchen Kette von Modellen kann das Schadensmodell die Auswirkungen von Veränderungen in der Vulnerabilität auf den Hochwasserschaden abbilden (Metin et al., 2018). Das Modell BN-FLEMOps eignet sich insbesondere zur Untersuchung privater Vorsorgemaßnahmen zur Reduktion von Hochwasserschäden (Sairam et al., 2019a). So können Strategien zur Klimaanpassung entwickelt werden, die über rein technische Maßnahmen hinausgehen. Im Kontext schneller Hochwasserereignisanalysen können mit BN-FLEMOps aufgrund der räumlich-zeitlichen Übertragbarkeit und Flexibilität bei den benötigten Eingangsdaten auch schnelle Schadensschätzungen analog zu dem in Dottori et al. (2017) diskutierten Ansatz vorgenommen werden. Die Anwendung von BN-FLEMOps und Kombination mit anderen Modellen wird, wie im folgenden Abschnitt erläutert, durch die Implementierung im OASIS-LMF-Format erleichtert.

4.3.4 Implementierung im OASIS-LMF

Die Anwendung von fortschrittlichen Simulationsmodellen in der Praxis scheitert häufig an erhöhten Kosten, verursacht durch einen größeren Aufwand und zusätzlichen Datenbedarf. Weiterhin können die in der Wissenschaft entwickelten Algorithmen zumeist nur mit erheblichem zusätzlichem Aufwand in bestehende Arbeitsabläufe und Berechnungsverfahren integriert werden. Das vorgestellte probabilistische Hochwasserschadensmodell BN-FLEMOps sowie die Datensätze für die Eingangsparameter Hochwassererfahrung, Gebäudefläche sowie Wohngebäudewerte stehen der wasserwirtschaftlichen Praxis und der Versicherungswirtschaft entsprechend dem OASIS-LMF-Standard (<https://oasislmf.org>) über den OASIS-HUB (<https://oasishub.co/dataset/european-exposure-data-for-bn-flemo-models-gfz>) zur Verfügung. Der OASIS-HUB macht Modelle, Software, Daten, Werkzeuge und Dienstleistungen rund um Naturgefahren zugänglich. Das OASIS-LMF ist eine Open-Source-Plattform zur Entwicklung und Anwendung von Naturkatastrophen (NatKat)-Modellen, die durch eine Initiative der Versicherungswirtschaft entwickelt wurde. Ziele dieser Initiative waren die Schaffung und Bereitstellung transparenter und bewährter Modellierungsansätze mit einem modularen Aufbau. Dieses Angebot wird von der Versicherungsbranche bereits gut angenommen und mehrere Verbundprojekte arbeiten derzeit daran, das Angebot auch bei Behörden bekannter zu machen. Der OASIS-HUB ermöglicht die Kombination von Gefährdungs- und Schadensmodellen im Sinne eines Plug-and-Play-Konzepts, das auf Grundlage eines festgelegten Standards für die verschiedenen Komponenten von NatKat-Modellen funktioniert. Dieser Standard ermöglicht Modellläufe mit dem Kernel der Open-Source-Plattform OASIS-LMF, der die für die Risikobewertungen benötigten Ergebnisse z. B. Schadens Erwartungswerte inklusive deren Unsicherheiten erzeugt.

4.4 Zusammenfassung

In diesem Artikel wird das probabilistische, multivariable Hochwasserschadensmodell BN-FLEMOps vorgestellt. Die beispielhafte Anwendung von BN-FLEMOps für das Augusthochwasser 2002 an der Mulde zeigt, dass das Modell den Gesamtschaden des Ereignisses im Mittel etwas unterschätzt. Der beobachtete Schaden wird jedoch innerhalb des 50 %-Quantils der Schadensschätzung abgebildet. Im Vergleich zu einem Ensemble aus international publizierten Modellen schätzt BN-FLEMOps den Hochwasserschaden an Wohngebäuden im Mittel genauer. Zudem kann die gesamte Spannweite des Ensembles und damit die Unsicherheit der Schadensmodellierung zuverlässig mit dem probabilistischen Modell dargestellt werden. Zu den geschätzten Schadenswerten wird eine Eintrittswahrscheinlichkeit vom Modell ausgegeben. Dies ermöglicht es, Entscheidungsträgern komplexere Methodiken der Entscheidungstheorie anzuwenden und so besser begründete Entscheidungen zu treffen. Das Modell kann durch die verwendete Methodik der Bayesschen Netze und europaweit verfügbarer Datensätze räumlich und zeitlich transferiert werden. Damit ist sowohl die Anwendung auf der Mikro- und Mesoskala möglich als auch die Verwendung für die Nachberechnung historischer Hochwasser und die Schadensprognose unter der Annahme von Klimaveränderungen. Mit der Implementierung von BN-FLEMOps im Standardformat des OASIS-LMF kann das Modell von Anwendern mit geringem Aufwand genutzt und im Plug-and-Play-Verfahren mit anderen Modellen kombiniert werden. Das Open-Source-Prinzip ermöglicht zudem jedem Nutzer das Überprüfen der statistischen Zusammenhänge des Modells und das transparente Nachvollziehen der Schadensschätzung.

Summary

This article introduces the probabilistic multi-variable flood loss model BN-FLEMOps. The application of BN-FLEMOps for the flood of August 2002 at the Mulde river shows that the model slightly underestimates the total loss of the event on average. However, the observed loss is covered within the 50 % quantile. Compared to an ensemble of internationally published models, the BN-FLEMOps estimates the flood loss for residential buildings on average more accurately. In addition, the entire variance of the ensemble and thus the uncertainty of loss modelling can be reliably represented with the probabilistic model. The model also provides the probability of occurrence in its output. This allows decision makers to use more complex methodologies from the field of decision theory to make more substantiated decisions. The model can be spatially and temporally transferred due to the Bayesian networks methodology and available pan-European data sets. Thus, application is possible both on the micro- and mesoscale as well as the use for the recalculation of historical

Danksagung

Diese Forschung wurde durch das IMPREX-Projekt (Fördernummer 641811) sowie H2020 Insurance (Fördernummer 730381) im Forschungs- und Innovationsprogramm der Europäischen Union Horizon 2020 gefördert. Die Erhebung der empirischen Daten im Muldeinzugsgebiet wurde vom Deutschen Forschungsnetz Naturkatastrophen (Bundesministerium für Bildung und Forschung (BMBF), Nr. 01SFR9969/5) unterstützt und durch eine Zusammenarbeit des Deutschen GeoForschungsZentrums GFZ und der Deutschen Rückversicherungs AG (www.deutscherueck.de) ermöglicht

Chapter 5

Drivers of future fluvial flood risk change for residential buildings in Europe

Abstract

Flooding is the most costly natural hazard in Europe. Climatic and socio-economic change are expected to further increase the amount of loss in the future. To counteract this development, policymaking, and adaptation planning need reliable large-scale risk assessments and an improved understanding of potential risk drivers. In this study, recent datasets for hazard and flood protection standards are combined with high-resolution exposure projections and attributes of vulnerability derived from open data sources. The independent and combined influence of exposure change and climate scenarios rcp45 and rcp85 on fluvial flood risk are evaluated for three future periods centred around 2025, 2055 and 2085. Scenarios with improved and neglected private precaution are examined for their influence on flood risk using a probabilistic, multi-variable flood loss model — BN-FLEMOps — to estimate fluvial flood losses for residential buildings in Europe. The results on NUTS-3 level reveal that urban centres and their surrounding regions are the hotspots of flood risk in Europe. Flood risk is projected to increase in the British Isles and Central Europe throughout the 21st century, while risk in many regions of Scandinavia and the Mediterranean will stagnate or decline. Under the combined effects of exposure change and climate scenarios rcp45, rcp85, fluvial flood risk in Europe is estimated to increase seven-fold and ten-fold respectively until the end of the century. Our results confirm the dominance of socio-economic change over climate change on increasing risk. Improved private precautionary measures would reduce flood risk in Europe on an average by 15 %. The quantification of future flood risk in Europe by integrating climate, socio-economic and private precaution scenarios provides an overview of risk drivers, trends, and hotspots. This large-scale comprehensive assessment at a regional level resolution is valuable for multiscale risk-based adaptation planning.

Submitted as:

Steinhausen, M., Paprotny, D., Dottori, F., Nivedita, S., Mentaschi, L., Alfieri, L., Lüdtkke, S., Kreibich, H., & Schröter, K.: Drivers of future fluvial flood risk change for residential buildings in Europe *Global Environmental Change*. (In review).

5.1 Introduction

Floods affect more people globally than any other natural hazard (CRED, 2020). In recent decades, economic flood impacts sharply increased globally as well as in Europe (Formetta & Feyen, 2019). For the future, flood losses and people affected by river floods in Europe are expected to considerably increase because of the combined effects of climate and socio-economic change (Rojas et al., 2013; Dottori et al., 2018). Thus, improved flood risk management and climate adaptation strategies are needed to reduce the adverse effects of floods and counteract the trend of increasing flood risk (European Union, 2007; European Commission, 2021). An important basis is detailed knowledge about the dynamics and drivers of flood risk. Flood risk is influenced by different drivers that can be organized into the components hazard, exposure, and vulnerability (IPCC, 2012).

Changes in flood risk are essential information for risk pooling (Prettenhaler et al., 2017) as facilitated through the European Union Solidarity Fund, (Hochrainer et al., 2010; Hochrainer-Stigler et al., 2017) as well as for the management of (re-)insurance portfolios (Kron, 2005). Flood risk associated with residential areas has a particular importance; the residential sector accounts for a large share of overall direct losses in Europe (Rojas et al., 2013), and impacts directly translate to the population.

Global warming will intensify the hydrological cycle and increase the magnitude and frequency of intense precipitation. Shifts in political-economic and socio-cultural systems foster human encroachment into floodplains and increase the damage potential, and consequently greater wealth tends to drive up economic costs of natural hazards (Mitchell, 2003). The increase in flood losses during recent decades can be largely explained by increasing exposure, both in Europe (Barredo, 2009; Barredo et al., 2012; Stevens et al., 2016; Paprotny et al., 2018b) and beyond (Pielke & Downton, 2000; Tanoue et al., 2016; McAneney et al., 2019). Besides changing climate, exposure growth remains a strong driver of losses, as shown by Rojas et al. (2013) and Alfieri et al. (2015a).

Vulnerability of a residential building to flooding is influenced by the building characteristics and the level of adaptation (Few, 2003). Besides flood prevention via structural protection, many countries have adopted the concept of managing residual flood risk via building-level private precautionary measures (Kreibich et al., 2015; Bubeck et al., 2018). The implementation of measures is influenced by flood experience (Parker et al., 2007; Bubeck et al., 2013; Atreya et al., 2017), socio-economic and building characteristics (Bubeck et al., 2012; Cumiskey et al., 2018). The role of several private precautionary measures has been analysed for countries in Europe and was found to provide resistance against flood loss thereby, reducing vulnerability of residential buildings (Hudson et al., 2014; Poussin et al., 2015). Vulnerability is often quantified using flood loss estimation models (Gerl et al., 2016).

In the past decade, increasing geospatial data availability and computing resources fostered a sharp increase in modelling frameworks for flood risk assessment at continental and global scales (Jongman et al., 2012b; Merz et al., 2014a; Ward et al., 2020). These often enable a modular structure and flexible use in the implementation of what-if scenarios based on appropriate combinations of hazard, exposure, and vulnerability. Typical applications include the reconstruction of past impacts and projections under future climate scenarios (Alfieri et al., 2015b; Alfieri et al., 2016a). Recent applications coupled continental scale flood simulations with detailed impact models for specific sectors, including railway networks (Bubeck et al., 2019), road

networks (van Ginkel et al., 2021), commercial assets (Paprotny et al., 2020a) and ports (Izaguirre et al., 2021), showing the benefits of advanced modelling approaches to refine large-scale impact estimates. However, approaches to estimate flood losses for the residential sector at the European scale have seen to date only simplified depth-damage applications.

Flood loss assessments at European scale have a number of shortcomings. Commonly, rather simple deterministic depth-damage functions are used for flood loss assessments (Jongman et al., 2014; Alfieri et al., 2015b; Alfieri et al., 2016b; Dottori et al., 2017). These models, which use only the water depth to estimate loss, are unable to adequately describe complex damage processes (Meyer et al., 2013; Schröter et al., 2014), so that associated uncertainties might be high but are unknown because of a lack of validation and missing uncertainty quantification. Particularly, resistance factors, such as the presence of precautionary measures, are rarely taken into account by current loss models, but are a precondition for the evaluation and development of effective risk mitigation strategies (Kreibich et al., 2015).

Formetta and Feyen (2019) identify the need for spatially explicit exposure and vulnerability information to better understand the impacts of natural hazards. This study contributes spatially detailed quantifications of different flood risk drivers on the European domain. The probabilistic, multi-variable flood loss model BN-FLEMOps (Wagenaar et al., 2018) makes it possible to quantify the effect of private precaution. In combination with novel input data for exposure, vulnerability, and recent data for hazard and flood protection, this enables the quantification of uncertainties for all components of flood risk. With these innovations, we aim to address the following questions:

- How does flood risk for residential buildings develop in Europe under conditions of changing climate, exposure, and vulnerability?
- What are the unique contributions of flood risk drivers to the changes?
- How do scenarios of adapted private precaution influence flood risk?
- Where are regional hot spots of flood risk and flood risk change in Europe?

Our assessment of flood risk is focused on direct tangible loss to residential buildings caused by river floods. A period between 1981–2010 centred around the year 1995 serves as “baseline” for the analysis of future flood risk. The drivers of flood risk are evaluated in independent and combined scenarios. Results for the historic baseline are compared against reported losses at country level.

5.2 Data and Methods

5.2.1 Hazard

Data and modelling framework developed for the European research project [PESETA VI](#) (Dottori et al., 2020) was used for this study. The occurrence and intensity of flood processes under present and future climate conditions were simulated using the hydrological model LISFLOOD, as described in (Mentaschi et al., 2020). Then a dataset of flood hazard maps based on hydrodynamic modelling to reproduce inundation extent and water depth (Dottori et al., 2021a) was created. Moreover, a spatial dataset of flood protection standards to assess the frequency and magnitude of floods causing impacts was applied (Dottori et al., 2021b). The next sections describe in more detail all the hazard components.

5.2.1.1 River flow projections

Projections of river streamflow derived with the hydrological model LISFLOOD (Knijff et al., 2010; Bubeck et al., 2013) implemented for the European continent with a resolution of 5 km, over the time horizon 1981–2100 were utilized. LISFLOOD is a distributed, physically-based rainfall-run-off model combined with a routing module for river channels that allows to reproduce all the relevant catchment processes. The simulations are based on the latest calibrated version of LISFLOOD (Dottori et al., 2021a) released as open-source software and available at <https://ec-jrc.github.io/lisflood/>. The forcing data of the streamflow projections (temperature, precipitation, radiative forcing, wind, and vapour pressure) were provided by 11 EURO-CORDEX projections under scenarios [rcp85](#) and [rcp45](#) (Jacob et al., 2014; Casanueva et al., 2016). The simulations were carried out using present-day estimations of population, anthropogenic land use and water demand provided by the Land Use-based Integrated Sustainability Assessment of the Joint Research Center ([JRC LUISA](#)) territorial modelling platform (Silva et al., 2013). The extremes of river discharge were studied using the non-stationary approach for Extreme Value Analysis (EVA) proposed by Mentaschi et al. (2016). For more details, the reader is referred to Mentaschi et al. (2020).

5.2.1.2 Inundation modelling

Flood extent and flood depths used in this work are taken from the dataset of hazard maps for river flooding developed by Dottori et al. (2021a). The extent of the dataset spans across geographical Europe (with the exclusion of the Volga river basin), and includes all the rivers entering the Mediterranean Sea and the Black Sea with upstream basin area above 500 km². The maps represent inundation depth at 100 m resolution along each section of the river network for six different flood return periods (10, 20, 50, 100, 200 and 500 years). The hydrological input is based on a long-term hydrological simulation for the period 1990–2016, run using the same version of the LISFLOOD model used for streamflow projections. This simulation provides the hydrological input for the flood simulations performed with the hydrodynamic model LISFLOOD-FP (Bates et al., 2010). Future flood hazard is then assessed, considering changes in flood frequency of the synthetic flood events with associated flood extent magnitude.

5.2.1.3 Structural flood defences

The novel dataset of flood protections described by Dottori et al. (2021b) was utilized here. In the dataset, the level of protection at any river location is defined by the return period (in years) of the maximum design flood that does not produce overflow. Flood protection data were derived at the administrative level and for this study have been downscaled to a 100 m grid to match the other gridded datasets. The dataset is built combining different sources of information on protection levels from technical reports and scientific publications, using modelled and observed flood losses to select the most plausible protection levels for all countries in geographical Europe (excluding Russia, Belarus, Ukraine, and countries in the Caucasus). Where design protection levels were unavailable, the level of flood defence was determined through an inverse modelling approach for each country, identifying the protection values that provide the closest match between modelled and reported losses. To this end, the database considered two datasets of protection standards (Jongman et al., 2014; Scussolini et al., 2016) and a range of uniform protection values at country scale. Future levels of flood protection are adjusted to changes in flood frequency. Specifically, streamflow projections are used to calculate the change in frequency of present-day design flood events under future climate conditions. For instance, a 1-in-100-year flood protection under present conditions may reduce to 1-in-50-year under future conditions because of the increased frequency of the flood event with the same magnitude.

5.2.2 Exposure

The baseline exposure for the year 1995 was obtained from HANZE database (Paprotny et al., 2018a). The exposure layer has a resolution of 100 m and indicates the value of the gross stock of residential buildings in Euro (€) in 2015 prices and exchange rates. It was generated by disaggregating the national stock of dwellings to regions (NUTS-3 level) according to regional gross domestic product (GDP) and then further disaggregating the estimated regional stock proportionally to the population per grid cell. The gridded population in 1995 was computed using a land-use change model. This model redistributes the estimated gridded population in the baseline map for 2011 through modelling land-use transitions (especially between urban and agricultural or natural land). Detailed information on the underlying methodology is provided by Paprotny et al. (2018a). Here, we recalculated the original map for 1995 from that study by including only residential buildings instead of all types of fixed assets and adjusted it for inflation (of GDP) and change in exchange rates up to year 2015.

Exposure in 2025, 2055, 2085 was calculated by adjusting the 1995 map according to the following formula:

$$E_{c,t} = E_{c,1995} \times \frac{P_{r,t}}{P_{r,1995}} \times \frac{G_{r,t}}{G_{r,1995}} \times \frac{W_{c,t}}{W_{c,1995}} \quad (5.1)$$

where:

$E_{c,t}$ is the exposure in grid cell and year

$P_{r,t}$ is the population in NUTS-3 region in which grid cell is located, and year

$G_{r,t}$ is the regional GDP in NUTS-3 region in which grid cell is located, and year

$W_{n,t}$ is the ratio of the value of residential buildings to GDP (“wealth-to-income” ratio) in country in which grid cell is located, and year.

Estimates of future exposure require computing three components: population, GDP and the wealth-to-income ratio. Each was obtained separately with different methods, which are explained below. However, because of the resolution of available demographic and economic data, the analysis operates on regional level and therefore neglects sub-regional population and land-use change after 1995 (see subsection B.1.1). Because of changes in NUTS regional classification over the years, growth for years 1995–2010 was taken directly from the HANZE database by intersecting the NUTS-3 (version 2013) polygon layer with the gridded residential exposure layers generated for 1995 and 2010. Afterwards, we use our own compilation of historical data: regional population (2010–2015), regional GDP per capita (from 2010 to the most recent year available) and national-level wealth-to-income ratio (2010–2018).

In the period up to 2085, the following projections were used:

- population from EUROPOP2019 projections for NUTS-3 regions by Eurostat (2021a), supplemented by projections prepared by statistical agencies of each constituent country of the UK; uncertainty ranges were added based on probabilistic national population projections by the United Nations (2019b).
- GDP computed through a Markov Chain-Monte Carlo probabilistic simulation of future regional economic growth based on empirically-observed “Beta” convergence (Barro & Sala-i-Martin, 1992; Monfort, 2008) within European NUTS-3 regions during 2000–2018.
- wealth-to-income ratio through extrapolation (with uncertainty bounds) of the 1950–2018 pan-European upward trend of this component (Paprotny et al., 2018b; Paprotny & Schröter, 2020).

Detailed information on the derivation of the projections is provided in the appendix.

In the process described above, an uncertainty distribution of population, GDP per capita and wealth-to-income ratio in 2025, 2055, 2085 for each NUTS-3 region was derived. Those distributions were combined into one, assuming independence of each component through sampling each distribution 10,000 times. Thus, a combined exposure uncertainty distribution for each region and time-step was achieved. We selected two scenarios: the first and third quartiles of the distribution, which can be translated as low and high exposure growth scenarios by historical standards. The projected growth in residential exposure since 1995 is presented in the appendix.

5.2.3 Vulnerability

In this section, we describe the drivers influencing flood vulnerability of residential buildings in relation to the predictors and model structure of the Bayesian Network — Flood Loss Estimation MOdel for the private sector (BN-FLEMOps; Wagenaar et al., 2018). BN-FLEMOps captures the interaction between flood hazard, and vulnerability for individual residential buildings. The model estimates relative flood loss (rloss) as the relation between the absolute building loss and the replacement costs of the building. All costs (e.g., material, wages) that are associated with repair of damage and reconstruction of the building are included in the loss calculation to estimate

the direct tangible damage to the building structure. This does not include movable household contents in the buildings.

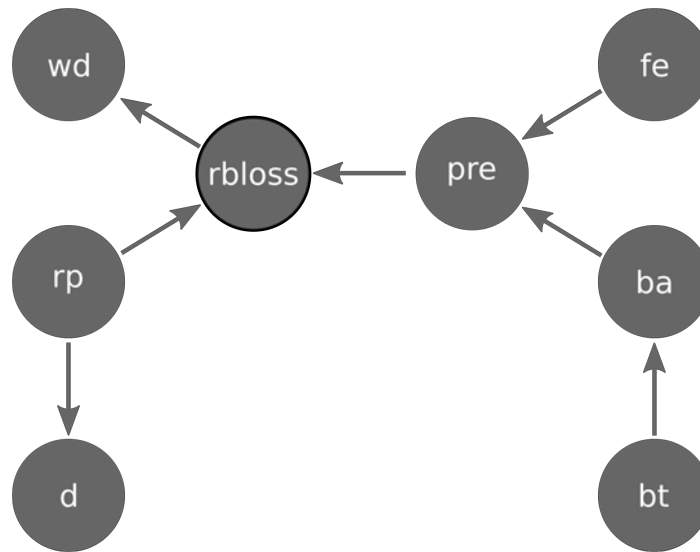


Figure 5.1: Directed acyclic graph structure of BN-FLEMOps with the variables: rbloss = relative building loss; wd = water depth; rp = return period; d = inundation duration; pre = precautionary measures; fe = flood experience; ba = footprint area of the building; bt = building type (adapted from Wagenaar et al., 2018)

The Bayesian network approach represents the statistical dependency of multiple input variables in a directed acyclic graph [Figure 5.1](#). Variables describing the hazard characteristics are water depth (wd), inundation duration (d) and return period (rp). Vulnerability is represented by private precautionary measures (pre) in terms of different levels of precaution, previous flood experience of the population (fe), building footprint area (ba) and building type (bt) (see [Table B.1](#) in the appendix).

BN-FLEMOps was developed with empirical microscale damage data collected via surveys with over 5000 flood affected households after large flood events in Germany between 2002 and 2013. The structure of the model network was learned from a set of 1522 complete survey records in combination with expert knowledge. For further details on model development and the empirical data basis, we refer to Wagenaar et al. (2018).

BN-FLEMOps can be transferred to the meso-scale and applied to areas where no empirical microscale information is available (Lüdtke et al., 2019a). The meso-scale application of BN-FLEMOps relies on Europe-wide consistent proxy data for model input. A description of all model input data and other relevant datasets for this study, such as asset data and structural flood protection, is given in [subsection 5.2.1 – 5.2.3](#). The proxy data approach has been validated against reported loss figures in three case studies in Germany, Austria, and Italy, showing high predictive performance for larger events with adequate empirical data. In a comparison with results from a model ensemble, BN-FLEMOps achieves equal or better performance, estimating the loss from the 2002 flood in the Mulde basin. As a probabilistic approach, the model inherently provides uncertainty information for the flood loss estimates (Steinhausen et al., 2020). Further, it captures the loss-reducing effect of private precaution (Sairam et al., 2019a). Detailed and consistent information on levels of precaution (pre) does not exist on the continental scale. Therefore, we infer levels of private precaution for Europe from the loss model using evidence of flood experience (fe) and building

area (ba), i.e., two variables in the model which are connected with precaution (pre). Individuals and communities learn from past flood experiences. The memory of past flood impacts motivates households to update and improve private precaution (Bubeck et al., 2012). Flood experience (fe) is indicated by the number of floods in each NUTS-3 region in the 25 years preceding the baseline (1971–1995) from the HANZE database (Paprotny et al., 2018a). The uptake and effectiveness of private precaution is also influenced by the characteristics of the residential building. Building footprint area (ba) is the size of the building exposed to floods. It provides an indication about building types. For example, detached single family homes have footprint areas that on average are smaller than those of apartments. Building area sizes for residential buildings are calculated from OSM building geometries within residential land use (Lüdtke et al., 2019a). For the baseline scenario, the level of private precaution is derived based on flood experience of the regions in 1995 and building characteristics obtained from OpenStreetMap (OSM contributors, 2020). The level of private precaution reflects the number of measures implemented to reduce negative flood impacts. These measures may include dry and wet proofing, adapted use and flood barriers.

5.2.4 Risk

Flood risk for a given area is expressed in monetary terms as accumulated direct tangible loss in Euro per time period and event probability, as well as Expected Annual Damage (EAD) in Euro. EAD is calculated over all return periods (10–500) with the inclusion of residual risk for areas protected against 500 year flood events or higher. EAD is a measure for risk that is widely used in large-scale modelling studies, e.g., (Alfieri et al., 2015b; Dottori et al., 2021b; van Ginkel et al., 2021).

EAD is calculated as the probability weighted sum of damage for flood scenarios with different return periods.

$$EAD = \sum_{i=1}^n (\Delta P_i \times \Sigma D_i) + R_{res} \quad (5.2)$$

Where:

The damage D_i is calculated as: $\Sigma D_i = D_i + D_{i+1}$ per exceedance probability interval (inverse return period RP_i): $\Delta P = \frac{1}{RP_i} - \frac{1}{RP_{i+1}}$ Note that based on the level of flood protection, EAD is calculated excluding the damages caused by floods with return periods below the protection standard. For protected scenarios, where the flood protection level FP is greater than the maximum return period considered in this study (rp 500), the residual risk is added as: $R_{res} = \frac{D_{max}}{FP}$

For the estimation of residual risk, the damage caused by flood events with a return period greater than 500 years is assumed not to increase beyond the damage of the 500-year event. Residual risk in this context is defined as the risk remaining despite structural flood protection, for example because of dike over-topping. Flood risk is estimated for the 27 member states of the European Union with the addition of the UK, Iceland, Norway, and Switzerland. Malta is excluded from the analysis as it doesn't include any major rivers.

5.2.5 Scenarios

Scenarios for a historic baseline period and for three future periods are created for Europe. These scenarios represent a range of possible future developments to explore the influence of climate and exposure changes on flood losses, as well as effects of private precaution on risk reduction. We use the median EAD to present our findings, and refer to the appendix for further uncertainty quantification in terms of the interquartile range (IQR). To compare the different levels of uncertainty in the flood risk drivers and in the flood loss model itself, we present a separate analysis of each component and discuss sources of uncertainty. An overview of all scenarios is depicted in [Figure 5.2](#).

5.2.5.1 Baseline scenarios

The baseline period is centred around the year 1995 and serves as reference for the comparison to the future scenarios. Hazard characteristics for the baseline were derived from the Europe-wide flood hazard maps ([subsubsection 5.2.1.2](#)) that include the inundation areas and depth for flood events of six return periods (10, 20, 50, 100, 200 and 500 years). For each of these hazard maps the return period is uniform throughout the entire domain with different water depth depending on local topography and river discharge. The hazard maps for the baseline represent the hydrology of a 30-year period from 1981 to 2010. Hazard scenarios with homogeneous return period flood discharges as they are used in this study are unrealistic. When assuming full dependence, i.e., simultaneous occurrence of the same return period flood at every location/spatial unit, the total loss integrated over the entire domain is strongly overestimated (Lamb et al., 2010; Metin et al., 2020; Nguyen et al., 2020). Considering or disregarding spatial dependence does not influence the EAD, but biases the risk curve (Metin et al., 2020). Vulnerability is characterized by the level of precaution derived from flood experience of the population (HANZE), counting the number of floods in each NUTS-3 region in the 25 years prior to the baseline and the building footprint area derived from a 2020 Version of OSM (OSM contributors, 2020). Exposure values for residential buildings in the baseline period are based on the building stock data and population density for the year 1995. Flood protection information is derived from the gridded [PESETA VI](#) layer by calculating the statistical mode of flood protection levels for each NUTS-3 region. In regions where the flood protection level is equal or higher than the return period of the flood events found in the hazard maps, no loss is considered to occur. For comparison, a baseline without flood protection and a baseline with flood protection are generated.

5.2.5.2 Future scenarios

The future periods considered in this study each cover 30 years and are centred around 2025, 2055 and 2085, thus represent the beginning, middle and end of the 21st century. To evaluate future flood risk under scenarios of climate change, the median, first, and third quartile of the return period shifts are derived as described in [subsubsection 5.3.2.1](#). EAD calculations are based on these “shifted” return periods and in accordance with the statistical mode of the “shifted” flood protection levels per NUTS-3 region. Change in exposure values (median, first, and third quartile) based on economic development and population growth are used to scale the baseline loss estimations for the independent analysis of the impact of exposure changes on risk. To examine the combined effects of climate and exposure change, results of the

climate scenarios rcp45 and rcp85 are scaled by the changes in exposure for the three future period.

5.2.5.3 Risk reduction scenarios

To explore the effects of structural flood protection and private precaution on risk reduction, multiple scenarios for the baseline and the future are created. For the baseline period, a scenario with current levels of private precaution, but without structural flood protection is compared to the baseline with current levels of flood protection and private precaution. Two further scenarios with flood protection, but improved precaution and neglected precaution, are created. In the scenario with improved precaution, the level of private precaution is increased by one for all areas where precaution is not at the maximum level. On the contrary, precaution levels are reduced by one in areas where they are greater than zero to simulate precaution neglect in Europe. For future climate and exposure conditions, all three scenarios of private precaution (baseline levels, improvement, and neglect) are calculated to examine the possible effects for risk reduction. The levels of private precaution in the three scenarios are kept constant throughout all future time periods.

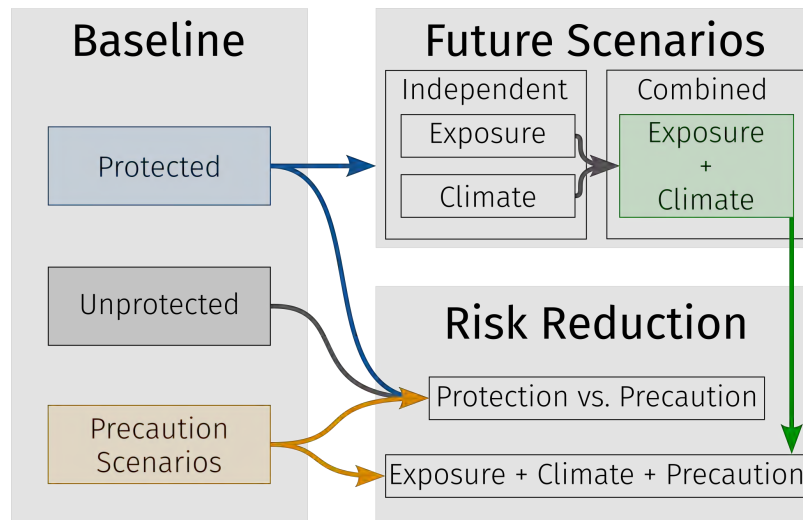


Figure 5.2: Overview of the scenarios for flood risk estimation. The protected and unprotected scenarios assume baseline levels of private precaution. Precaution scenarios utilize flood protection standards. Future scenarios for 2025, 2055 and 2085 are based on the protected baseline. The comparison of risk reduction by protection and private precaution uses all baseline scenarios. The influence of precaution on flood risk is assessed for the combined future scenario.

5.3 Results

The distribution of flood risk in Europe for the baseline (1995) is described first. Subsequently, the independent and combined effects of climate and exposure change on future flood risk are reported. Finally, scenarios for flood risk reduction are evaluated.

5.3.1 Baseline flood risk

For the baseline period the EAD estimations with structural flood defences show Germany (314 million €), France (203 million €), Italy (97 million €) and Spain (52 million €) as the European countries with the highest flood risk for residential buildings (Figure 5.3). The United Kingdom (UK) is estimated to have similarly high losses as the alpine countries of Austria and Switzerland. However, the uncertainty around EAD estimates for the UK is much greater than for other countries with similar EAD. This becomes apparent when comparing the IQR of the UK EAD estimates (13.7–65.5 million €) with that of Austria (15.0–58.1 million €). Scandinavian countries Norway, Sweden, and Finland fall into the middle range of flood risk in Europe with EADs of 4.6, 9.0 and 12.3 million € respectively. The Netherlands because of extremely high flood protection standards of up to 10,000 years exhibit relative low losses of 6.5 million € per year. Low annual losses under 1 million € are estimated for Estonia, Denmark, Bulgaria, Iceland, and Cyprus.

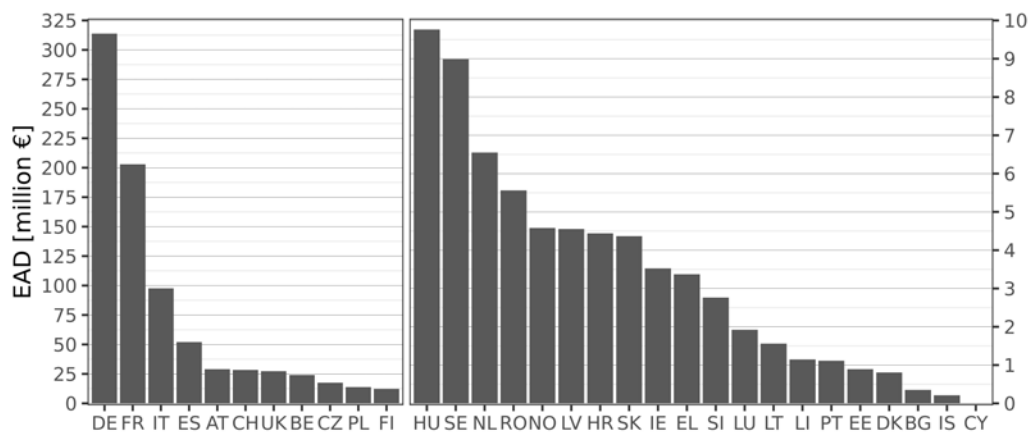


Figure 5.3: Expected Annual Damage per country for the baseline scenario. Note the change of scale along the y-axis of the right panel for better readability.

At regional (NUTS-3) level, the highest losses in Europe are metropolitan areas situated along major rivers (Figure 5.4). These areas combine the hazard of fluvial flooding with a high concentration of exposed assets vulnerable to damage. Hamburg at the mouth of Elbe River, Paris at the confluence of Seine and Marne, Florence, Zaragoza, London (combining multiple NUTS-3 regions), Geneva, Ghent, and Linz are the regions with the overall highest losses in the baseline period.

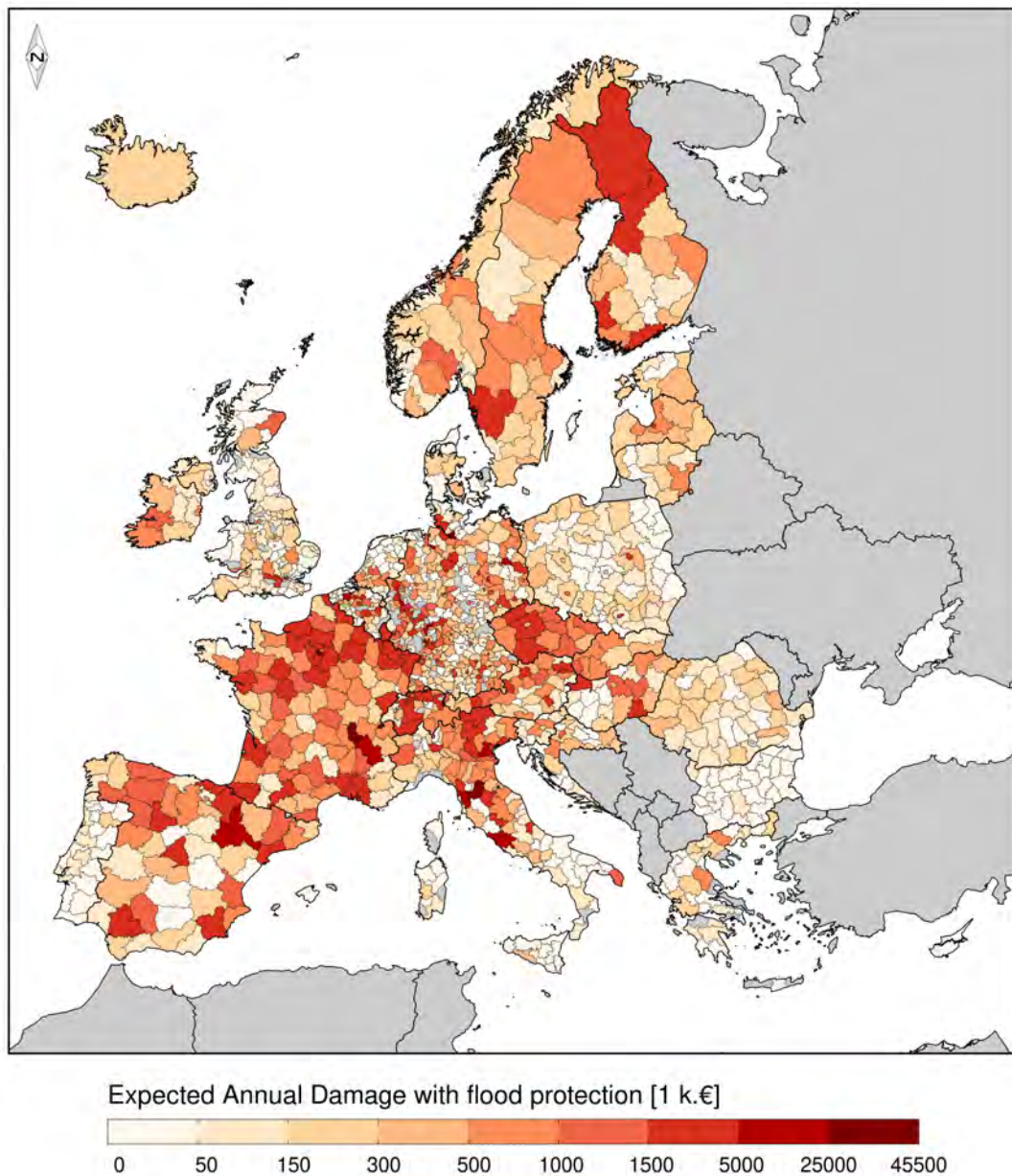


Figure 5.4: Expected Annual Damage per NUTS-3 region for the baseline scenario. Grey NUTS-3 regions have no risk estimation (no hazard and/or no exposure). Grey countries were not included in the analysis. EAD values are not normalized to region area size. Map projection EPSG3035. Shape files are provided in the appendix.

Besides these regions of the highest risk, areas of high flood risk (EAD over 2.5 million €) are found in Germany in the city of Dresden in the Elbe catchment as well

as Cologne and Frankfurt located along the Rhine, and Main River respectively. The Po valley in northern Italy with the city of Padua, Pisa by the Arno River in Tuscany, and the capital region of Rome by the Tiber River. High-risk regions in France are in the Rhone catchment in the region of Lyon, Grenoble, and Avignon as well as Genève (Switzerland). The Danube and its tributary Inn generate the highest losses in Austria in the cities of Linz and Innsbruck. Also, the region of Warsaw on the Vistula River falls into the high-risk category in Poland.

Regions of low risk (EAD below 40,000 €) are predominantly remote and rural without major river systems and or areas with high flood protection levels. Most low-risk regions are located in eastern and Southern Europe. All but one NUTS-3 region of Bulgaria has an EAD below 40,000 €. Likewise, low losses are expected in central and eastern Poland, as well as in the region surrounding lake Balaton in Hungary and northern Croatia. In the Mediterranean, low-risk areas are in southern Italy and the islands of Sicily and Cyprus. Correspondingly, most regions in Portugal are estimated to have low losses in the baseline period. Regions in the east and south of England, Wales, and the north of Scotland show particularly low flood risk. The same can be seen for most regions in the northern part of the Netherlands. High levels of flood protection in northern Scotland (200 years) and in the north of The Netherlands (1250–10,000 years) combined with moderate exposure explain the low risk in these regions.

5.3.2 Drivers of future fluvial flood risk

The two most important drivers, climate, and exposure change are explored separately to understand the different degree and spatial patterns in which they contribute to future flood risk in Europe.

5.3.2.1 Climate change

Little overall differentiation of climate driven risk change in Europe is visible between rcp45 and rcp85, the beginning of the century (2025). Both scenarios entail a median risk increase of around 70 % compared to the baseline, with IQR from 25 % to 150 % EAD increase (Figure 5.6). In the middle and end of the century, the two climate scenarios become more distinct as the loss more than doubles from 2025 to 2055 under rcp85. Also, the uncertainty, indicated by the IQR around rcp85 estimates, substantially increases and is skewed towards higher change rates compared to the more optimistic rcp45 scenario. For the end of the century (2085) an increase in risk driven by climate change is estimated at 196 % under rcp85 with an IQR from 70 % to 448 %. The large uncertainty around these estimates reflects the variability within the 11 different EURO-CORDEX projections that drives the hydrological and hydraulic simulations.

Most regions in Europe show a consistent trend of either decreasing or increasing flood risk under the two scenarios of climate change. Regions with decreasing flood risk are located primarily in Scandinavia in northern Norway and parts of Sweden and Finland, Iceland, as well as in north-eastern Poland (Figure 5.5). Furthermore, under rcp45, parts of the Baltic States are projected to have decreasing flood risk. In the end of the century, risk is estimated to decrease in the central regions of Romania under rcp45 and rcp85. Decreasing risk throughout all three future periods and under both rcps is also expected in many regions of Portugal and Spain, with other parts of the Iberian Peninsula showing small increase in expected risk. Especially under

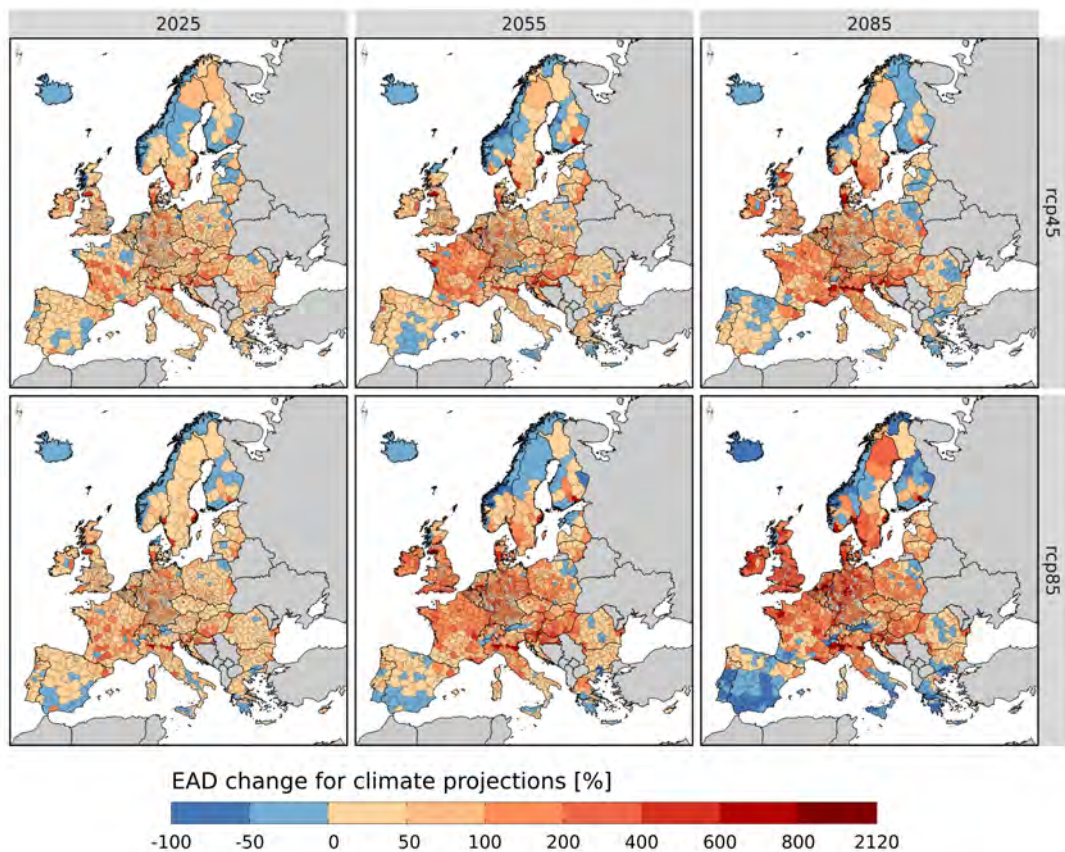


Figure 5.5: Relative change in EAD per NUTS-3 region from the baseline to climate projections in three future periods and for two rcp scenarios. Blue colours indicate lower risk and red colours higher risk compared to the baseline. Map projection EPSG3035.

rcp85, decreasing risk is estimated for southern Spain and Portugal at the end of the century. Sicily and southern Greece are regions in the Mediterranean with decreasing risk.

Overall increasing risk is expected in almost all regions in the UK and Central Europe from France to Hungary and from Denmark to northern Italy with the exception for the western part of Austria which shows a decrease in flood risk. The highest increase in risk in Northern Europe are projected for the British Isles, Denmark and southern Sweden. In Southern Europe, the Po valley in Italy and the regions along the Sava River in Croatia show very high rates of risk increase. The projected regional patterns of climate driven flood risk change align well with and are a continuation of the observed trends in river flood discharges for Europe in the period 1960-2010 (Blöschl et al., 2019).

Under the less extreme rcp45, flood risk in Europe would only increase by 125 % at the end of the century, instead of almost 200 % under rcp85. A future with lower global emissions reduces flood risk and also entails lower uncertainties around climate predictions, making the estimations of future flood risk more reliable. Additionally, uncertainties around the median estimates are significantly lower under rcp45 in comparison, especially with the very high uncertainty for rcp85 at the end of the century. Climate adaptation plans to reduce flood risk can be based on more certain estimations under rcp45.

5.3.2.2 Exposure change

At the beginning of the century, flood risk is expected to double, solely based on the increase in value of exposed buildings (Figure 5). The rate of increasing risk because of exposure change, follows a linear trend and is estimated to reach nearly 200 % in the middle of the century with an IQR from 150 % to 244 %. In the end of the century, this exposure driven increase will reach nearly 300 %.

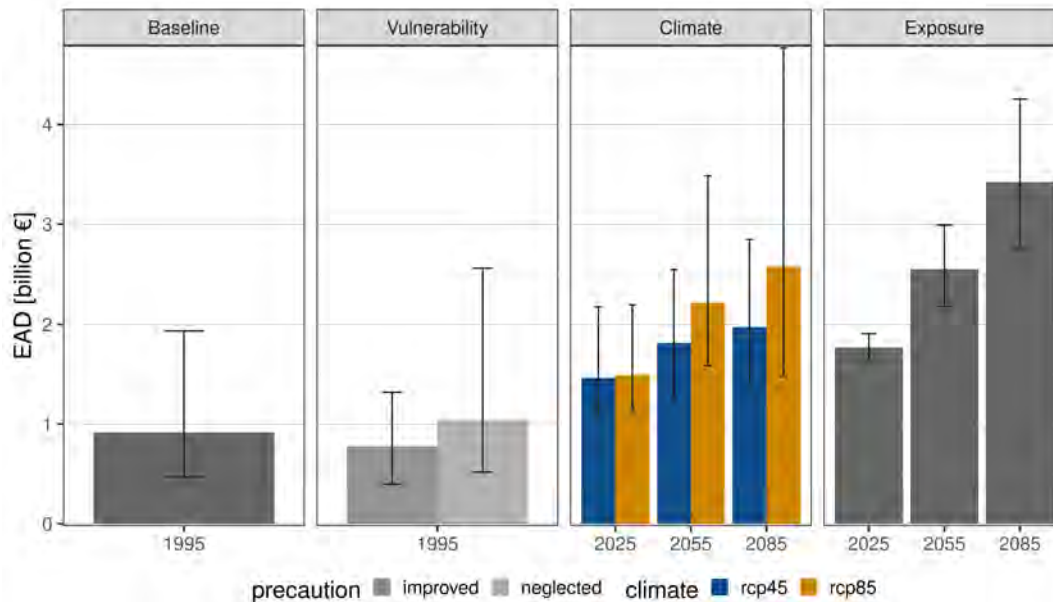


Figure 5.6: EAD in Europe for the baseline and scenarios of future flood risk change. The first panel displays the EAD for all of Europe in the baseline period. In the second panel, scenarios of private precaution show possible changes in risk attributed to vulnerability. The third panel shows the effects of climate change under two *rcp* scenarios with baseline precaution in three future time periods. Baseline exposure values are used for the first three panels. The fourth panel uses baseline climate and precaution, but includes changes in the exposure component. Uncertainty for each risk component separately is displayed by the IQR around the bar plot of the median estimate.

Regional differences in flood risk change attributed to the increase in building values are smaller in comparison with climate driven changes. Most European regions experience a steady increase in risk from 1995. Exceptions to the continuous risk increase are the County of Sisak-Moslavina in eastern central Croatia, a few regions in the south and east of Poland and in north-eastern Lithuania and Estonia. Note that the building value decrease in these areas does not exceed 10 % and only persist in beginning of the century projections. Low increase over all future periods is projected for northern Spain, Portugal, central France, rural regions of Poland and northern Lithuania.

South-Eastern Europe is expected to experience the highest increase in risk. For some regions in Bulgaria and Romania, more than 10-fold increase in flood risk is predicted until the end of the century (Figure 5.7). Furthermore, the British Isles and southern Sweden are regions with substantial increases in exposure driven risk. Strong increases in risk are expected in particular in and around urban centres. Examples for this spatial pattern can be found all over Europe in the regions surrounding Warsaw, Poznań, and Wrocław (Poland); the regions surrounding Prague (Czech Republic),

Budapest (Hungary), Bucharest (Romania), Sofia (Bulgaria) as well as Stockholm and Malmö (Sweden).

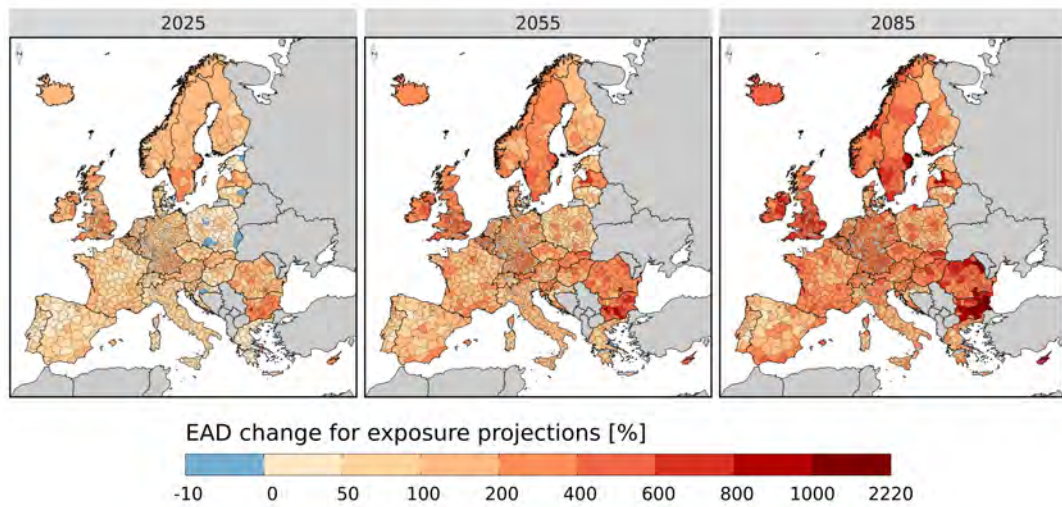


Figure 5.7: Relative change in EAD per NUTS-3 region from the baseline to exposure projections in three future periods. Blue colour indicates lower risk and red colour higher risk compared to the baseline. Map projection EPSG3035.

5.3.2.3 Combined effects

After attributing the role of flood risk drivers, the combination of climate and exposure change is simulated to create more realistic future scenarios of flood risk in Europe. The effects of climate and exposure change on EAD, per country, are presented in [Figure 5.8](#). [Figure 5.9](#) shows the relative changes in EAD compared to the baseline, per NUTS-3 region.

As presented in [subsection 5.3.2.1](#) the effects of climate change on flood risk in Europe are spatially heterogeneous. While most countries and regions must expect increases in risk, some regions, especially in northern and Southern Europe, will see stable or slightly decreasing risk in the future. Changes in exposed building values will lead to increasing risk in all European regions in the middle and end of the century, with the largest changes occurring in Eastern Europe and the British Isles. In combination, these drivers can counteract or build up their respective effects. Nevertheless, compared to the baseline, all European countries are expected to face increased flood risk for residential buildings. For the baseline period, an EAD of 885 million € for residential buildings in Europe was estimated. Under future climate conditions and exposure, the loss is projected to increase to 2.0–2.1 billion € in 2025 and 4.3–5.5 billion € in 2055 under rcp45 and rcp85, respectively. By the end of the 21st century the EAD will increase seven-fold to 6.8 billion € under rcp45 or 10-fold to 9.3 billion € under rcp85. These results assume current levels of private precaution.

Under rcp45, Spain sees a steady increase in flood risk throughout the century because the risk increase based on socio-economic development surpasses the slight decrease driven by climate change. Only under rcp85 will Spain's flood risk stagnate because of the stronger effects of climate change at the end of the century. Under rcp45, Finland will see a decrease in risk from 2055 to the end of the century, as opposed

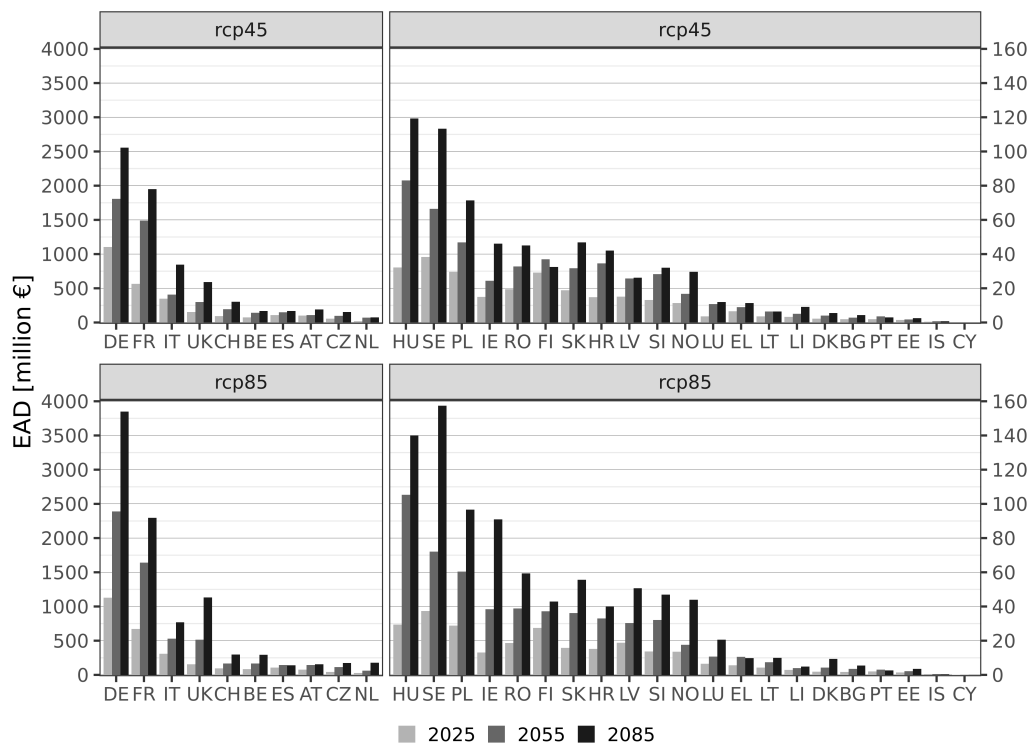


Figure 5.8: Combined effects of climate and exposure change on EAD per country. Note the change of scale along the y-axis of the right panels for better readability.

to an increase under rcp85. The socio-economic development in Scandinavia and the Baltic States surpasses the risk reducing effects of climate change, leading to an overall increase in flood risk (Figure 5.9).

The highest relative EAD changes are observed in the United Kingdom, Ireland, The Netherlands, Sweden, Slovenia, and Bulgaria at the end of the century under rcp85. In contrast, the lowest relative changes in risk are observed in the Baltic States, Portugal, Spain, and Greece. Portugal, the country with the lowest relative changes in EAD, is expected to experience a 71 % (rcp45) to 85 % (rcp85) increase in risk at the beginning of the century, 193 % (rcp85) to 219 % (rcp45) by the middle of the century and 129 % (rcp85) to 164 % (rcp45) in the end of the century. Under rcp85, the risk increase at the end of the century is most pronounced. Countries such as Finland and Greece where risk would stagnate from middle to end of the century under rcp45, show higher EAD under rcp85 at the end of the century.

5.3.3 Risk reduction

The exceedance probability curves in Figure 5.10 show the total loss for uniform return period flood events in Europe. A comparison of baseline flood risk (green graph) and a scenario that removes all structural flood protection in Europe while retaining the current level of private precaution (black graph) reveals the effect of risk reduction for flood events of the considered return periods. Examining a Europe-wide 100-year flood as an example, losses with the current level of structural flood protection and private precaution sum up to 21 billion €. Without flood protection, the losses could be as high as 137 billion €. The risk reduction that structural flood protections offer in Europe is especially substantial for events under the 100-year

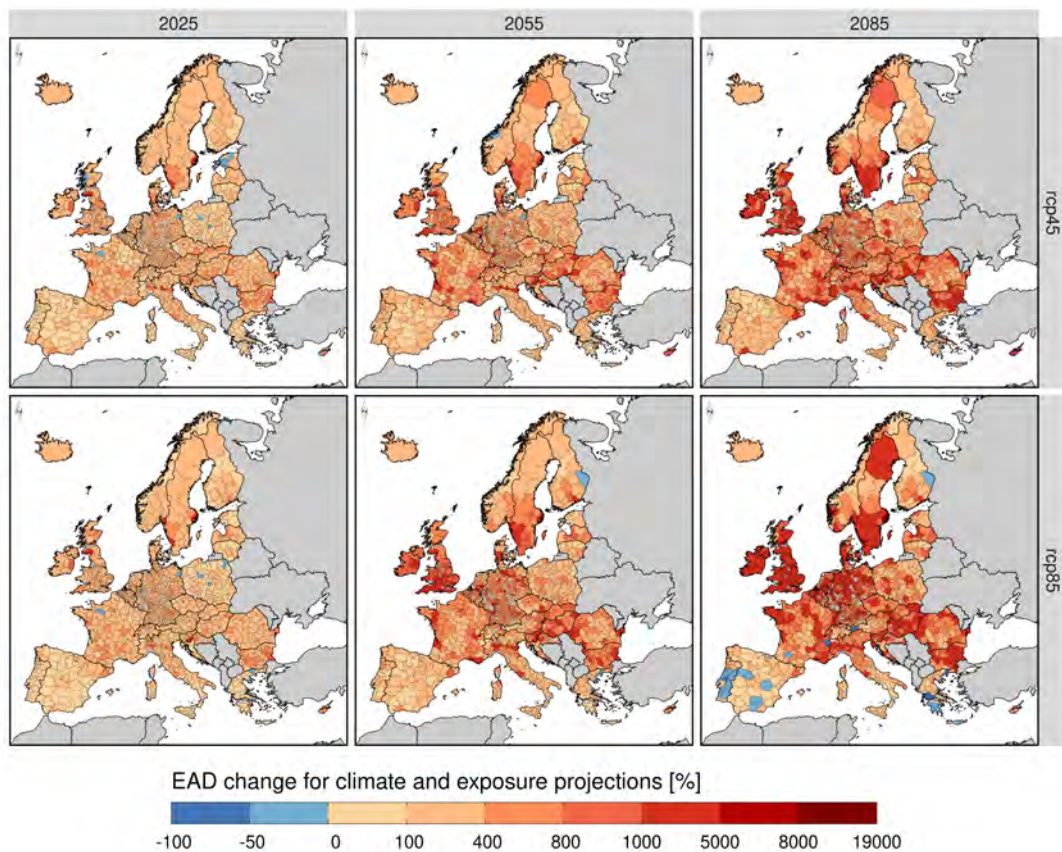


Figure 5.9: Relative change in EAD per NUTS-3 region from the baseline to climate projections, combined with exposure projections in three future periods and for two rcp scenarios. Blue colours indicate lower risk and red colours higher risk compared to the baseline. Map projection EPSG3035.

return period. For a flood event with a return period of 200 years, the losses for residential buildings in Europe are drastically higher and amounts to 126 billion €. Flood protection reduces the losses for 200 and 500 year events only by 31 and 26 billion €, respectively. Risk reduction by structural flood protection for 500 year events is a result of the high protection levels in the Netherlands (rp 1250–10,000), London (rp 1000) and Budapest (rp 1000). “Existing” flood protection levels reduce the total EAD in Europe by 4.7 billion € for the baseline period.

A scenario with improved private precaution levels and current flood protection standards (blue graph) results in a reduction in losses for high return period events (200 and 500 years) of 17 % (-21 billion €) and 12 % (-17 billion €). These results suggest that improvements in private precautionary measures would significantly reduce the residual flood risk left by missing or insufficient structural flood protection. In a pessimistic scenario in which private precaution is neglected in Europe, the losses for a 500-year event is of the same magnitude as in the scenario without structural flood protection. The scenario with improved precaution for the same return period, however, reveals the potential of private precautionary measures to reduce residual flood risk also for extreme events. Improved precautionary measures can reduce EAD in Europe by 15 % (137 million €) in the baseline period. Under future climate and exposure conditions the highest risk reduction through improved private precaution can be achieved in Luxembourg, Norway, Slovenia, Greece, UK, and Latvia with over

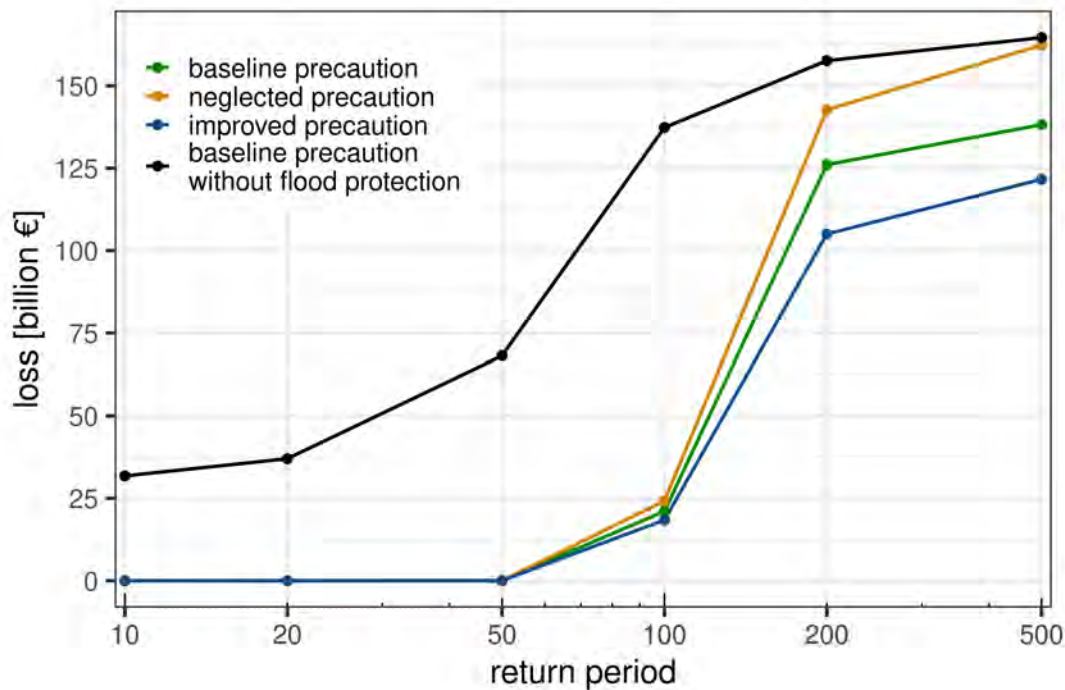


Figure 5.10: Exceedance probability curve of flood losses for Europe. The green curve shows the baseline as defined in [subsubsection 5.2.5.1](#). The black curve shows losses for a baseline without flood protection. A baseline scenario with increased and decreased private precaution is displayed in blue and orange, respectively.

20 % lower EAD. Lithuania, Iceland, and Portugal show the lowest potential for risk reduction with EAD decreases below 10 % ([Figure B.2](#) and [Figure B.3](#)).

Neglecting private precaution would cause increased risk in Iceland by 24 % and in France by up to 23 % depending on the future period and climate scenarios. The uncertainty range around the effects of private precaution on EAD can be considerable, especially for countries with low EAD. Bulgaria, for instance, may see adverse developments in risk because of precaution neglect between 6 % to 21 % depending on climate scenario and future period.

5.4 Discussion

5.4.1 Comparisons with previous studies

Comparing our results to Rojas et al. (2013) reveal similarities in spatial patterns of climate related flood risk change. Strong increases are projected for the UK and Ireland and northern Italy, while decrease or stagnation are expected in regions of Portugal, Spain, and Greece. Our results do not show a decrease in risk in North-Eastern Germany and Poland, as presented in Rojas et al. (2013). Our projections more closely agree with the spatial patterns of change reported in Dottori et al. (2021b) where decreases in flood risk are shown in parts of Sweden and Finland as well as in Portugal, Spain, and Greece. Meanwhile, Central Europe, including Germany and Poland, are expected to have greater flood risk at the end of the century. Germany, Italy, and France are identified as the three European countries with the highest EAD. Alfieri et al. (2015b) supports these findings, but reports significantly higher risks in Italy compared to Germany. This deviation is most likely because of

lower flood protection levels in the Po River catchment in northern Italy, assumed by Jongman et al. (2014). We identify further agreement in the strong risk increase and high uncertainty around results for the end of the 21st century under rcp85. All comparisons drawn to these previous studies are qualitative in their nature since simpler depth-damage functions for flood loss estimation were used but assets from all sectors (e.g., residential, commercial, industrial etc.) were included without differentiation in the loss figures.

Willner et al. (2018) identify the need to increase flood protection levels in Europe especially in southern Sweden, northern Germany, Poland and in France along the Rhone, Loire, and Seine to mitigate increasing flood risk in the middle of the 21st century. Our findings agree with most of the European regions identified to be under increasing risk in the future. We show that improvements in private precaution can reduce residual risk by 15 % in Europe and provide an additional mitigation strategy next to flood protection level increases. These findings agree with Dottori et al. (2021b) who found that flood proofing of buildings may cause an average damage reduction in Europe by 12 %. The recommendations given by Alfieri et al. (2016a) to combine dykes, local retention, and relocation in a strategy for risk reduction in Europe remain valid and should be extended by the improvement of private precautionary measures. The development of new buildings in flood prone areas has to be avoided to not increase exposure unnecessarily. Growing uncertainties in the middle and end of the century in flood risk projections call for iterative approaches for risk reducing strategies to avoid under or over adaptation. Private precaution is locally implemented at the object level, and can thus be used as a more targeted flood risk reduction strategy in places where large-scale measures may not be cost-effective.

5.4.2 Comparison with reported losses

Flood losses estimated by the model for the baseline historical period can be compared with reported losses in the same years. A multi-source compilation of historical flood loss data was extracted from HANZE database (Paprotny et al., 2018a). Several adjustments had to be made to the reported losses to be comparable with this study. Only river floods were selected, as defined in the HANZE database, without compound or flash floods. Losses were normalized to 1995 exposure levels using the HANZE gridded exposure estimates for different time periods, and adjusted to 2015 prices and exchange rates. Further, economic losses for flood events that were missing this information, but had other quantitative impact data, were gap-filled based on the dependency of relative economic loss to relative fatalities/people affected, exposure at the time of the event and estimated footprint of the event (same methodology as the gap-filling in Paprotny et al. (2018a)). Finally, the economic losses were converted to residential building loss using the relative share of housing in the total stock of fixed assets in a given country and year of event (also drawn from the HANZE database). The final estimate of residential flood losses in our study area based on reported data is 1443 million € annually for 1981–2010. This value is well within the 50 % confidence interval of our modelled protected losses, which is 454–1933 million €, and considerably lower than the unprotected loss estimates, which is expected given that flood protection prevents a large share of flood losses in practice. On a country level, a high correlation is achieved ($R^2 = 0.81$) (see Table B.2 in the appendix), with a very accurate prediction for Germany, predictions within the 50 % confidence interval for France, Hungary, or the Netherlands, but also underestimation of losses for Italy, Poland, Spain, and the UK, suggesting that the flood protection estimates for those

countries might be too high (see [Figure B.1](#) in the supplement). Nonetheless, there are variations in completeness of reported data, as well as uncertainty in gap-filling, normalizing and converting loss data to housing losses only. Additionally, the definition of river floods in HANZE database does not fully match the criterion of 500+ km² catchment size used in LISFLOOD. The modelled flood losses in the baseline scenario are clearly in the correct order of magnitude, giving greater confidence in the absolute values of losses in the projections.

5.4.3 Uncertainties

The results derived from the analysis of the future scenarios should not be understood as deterministic predictions of future flood risk, but rather as attempts to study the effects of independent and combined drivers on the possible changes of flood risk. All efforts of modelling possible future conditions in the Earth system are naturally uncertain. A systematic distinction can be made between epistemic uncertainty (because of lack of knowledge) and aleatory uncertainty (because of randomness). These uncertainties are part of all datasets and the imperfect relations of model variables, thus exercising an efficacious influence on flood risk estimates presented in this study.

Our projections of exposure growth neglect the effect of land use change. This is caused by the lack of any useful projections of this factor for Europe. Recent gridded population projections by Jones and O'Neill (2016) indicate a very strong urbanization, contrary to processes of sub- or even de-urbanization taking place in Europe, (Salvati et al., 2018; Shaw et al., 2020) and is caused by the global setting of that study as intense urbanization is featured throughout the developing world (J. Liu et al., 2007; Güneralp et al., 2017). When aggregated to NUTS-3 level, the datasets fail to reproduce regional demographic growth patterns projected until 2050 by Eurostat for almost all countries in the study area. Another set of projections by Murakami and Yamagata (2019) also strongly misrepresents regional population growth in Europe. Further, land-use changes are multi-layered, very complex and to a large extent policy-driven, making any projections deeply uncertain (Antrop, 2004; J. Liu et al., 2007; Lambin & Meyfroidt, 2010; Bryan et al., 2016). Finally, Paprotny et al. (2018b) show that the share of population, GDP, and wealth within the 100-year flood zone in Europe has varied little over the past 150 years. We therefore concentrated on projecting demographic- and economic-driven increase in the value of residential assets and did not consider land-use change.

The spread between the CORDEX models is a major source of uncertainty for our projections of flood impact. It has several causes, e.g., the limited resolution, the lack of understanding and parametric representation of some processes, dissimilarities among the models, use of different metrics and criteria to calibrate different models, lack of consensus on the assessment of the quality of future projections, internal variability and others. As climate projections are interdependent and forced by hypothetical emission scenarios, the spread among the models does not represent an estimation of the statistical uncertainty, but rather a range of opportunity, and should not stop future decisions (Knutti et al., 2013). The relative uncertainty U_r (IQR/rl_{chng}) tends to increase with the year, and with the scenario's emissions. The uncertainty is also not homogeneous in space, with higher values of U_r generally found in Eastern and Southern Europe ([Figure B.4](#) in appendix).

In the present work, we did not explicitly consider the uncertainty associated to flood protection data. In Europe, information on actual protection standards can be found only for a few countries and urban areas (Scussolini et al., 2016; Dottori et al., 2017). Recent studies tried different approaches to estimate protection standards in areas with no reported information, using modelled flood risk (Jongman et al., 2014), empirical functions related to gross domestic product (Scussolini et al., 2016), and recorded loss data (Dottori et al., 2021b). In several countries in Europe, these datasets propose substantially different protection levels, meaning that the overall uncertainty is considerable and difficult to quantify (Dottori et al., 2021b).

BN-FLEMOps is a fully probabilistic model trained based on a combination of expert knowledge and micro-level empirical data from multiple events in Germany. The model accounts for uncertainties in flood loss predictions based on 1) the ability of the model to capture damage processes given the model structure and predictors and 2) availability of relevant empirical data concerning the predictors for model application. The model accounts for these two sources of uncertainties — represented by the NPTs and provide an aggregated uncertainty range for the loss predictions.

5.5 Conclusions

Our study is the first high-resolution assessment of future flood risk in Europe, considering changes in all three components of risk: hazard, exposure and vulnerability. It reveals regional patterns of future flood risk change for residential buildings in Europe under scenarios of climate change and exposure development. Changes in the value of exposed buildings will increase flood risk all over Europe throughout the 21st century with higher rates of increase in Southern and Eastern Europe. Climate driven risk changes are more heterogeneously distributed in Europe. Climate change will contribute to increasing risk in the UK and Central Europe, while stagnating or decreasing risk is projected for parts of Scandinavia and the Mediterranean. Results show that in particular, urban centres and their surrounding regions along major European rivers will experience the highest rates of risk increase in the future.

In all future periods exposure has a greater influence on risk change than climate driven hazard. Climatic changes alone will increase risk in 2085 by 116 % (rcp45) and 181 % (rcp85), while exposure increases the risk by 275 % by 2085. The combination of Exposure and climate change amplifies the increase of flood risk in Europe: under rcp45 and rcp85 scenarios, risk is estimated to increase by 670 % and 950 % respectively by the end of the century. Projections under the rcp85 scenario are more uncertain, with a tendency towards higher losses within the interquartile range. Even though rcp85 should be regarded as a rather extreme scenario (see Hausfather and Peters, 2020), these results reveal the threat of insufficient climate mitigation. In comparison, the expected annual damage under rcp45 at the end of the century is 500 million € lower and less uncertain. Striving towards the less extreme climate scenario, rcp45 facilitates more reliable planning of adaptation to reduce future flood risk in Europe.

Existing flood protection levels reduce the expected annual damage for residential buildings in Europe by 4.7 billion € in the baseline period. Besides structural flood protection, private precautionary measures can reduce the risk by 15 % (137 million €) based on expected annual damage estimates for Europe under a scenario of improved private precaution. For regions where flood protection standards are low or where climate driven changes in flood hazard will increase the need for protection, private

precaution can significantly reduce risk by up to 20 % of the expected annual damage. Improvements in private precaution should be part of adaptation strategies alongside other risk reduction measures such as increased dyke heights, higher local retention and relocation of buildings and dykes.

Acknowledgements:

Dominik Paprotny was supported by the German Research Foundation through project Decomposition of flood losses by environmental and economic drivers, project no. 44917597. Nivedita Sairam is supported by the German Federal Ministry of Education and Research (BMBF) project DECIDER, project no. 01LZ1703A-H.

Chapter 6

Discussion, conclusions, and outlook

6.1 Summary of key findings

This thesis advances flood risk modelling for residential buildings in Europe and thereby improve our understanding of flood risk under current and future climatic and socio-economic conditions. Probabilistic, multi-variable models represent damage processes in detail and offer transparent uncertainty information. To fulfil the data requirements for these approaches, I identified and integrated new data sources into flood loss modelling. This enables the transfer of the microscale model to the meso-scale, where it is validated against reported flood loss in case studies and performance tested. Based on the identified vulnerability data and the most recent data sets for hazard and exposure, future flood risk change is modelled for residential buildings in Europe. The key findings from Chapters 2 – 5 are presented in bullet points below, and the research questions guiding this thesis are addressed in detail in the subsequent paragraphs:



Key findings

- Geometric characteristics of building footprints are useful proxy variables to represent building resistance to damage in flood loss models.
- The integration of open data sources such as OSM, DFO, and HANZE improves the applicability and transferability of multi-variable models in Europe.
- The newly developed probabilistic, multi-variable flood loss model BN-FLEMOps accurately estimates loss on the micro- and meso-scale and provides useful uncertainty information for decision-making.
- Fluvial flood risk for residential buildings in Europe will increase up to 10-fold until the end of the 21st century. Climate and exposure change contribute to the risk increase in a ratio of two to three.
- Urban regions along major European rivers will suffer the highest risk increase.
- Improving the level of private precaution could reduce flood risk in Europe by 15 % on average.

1. Which new data sources can advance multi-variable approaches for flood loss modelling?

In [Chapter 2](#) the question “Are OpenStreetMap building data useful for flood vulnerability modelling?” is addressed. To tackle this question, Random Forest models based on water depth and several of numerical spatial measures derived from **OpenStreetMap (OSM) building footprint geometries** were built. In comparison with a simple stage-damage curve, these **OSM-based** models have significantly higher prediction precision and deliver more reliable loss estimates. Compared with a detailed multi-variable model based on empirical damage data, the OSM models show similar prediction accuracy and reliability, with slightly lower precision. Building area is identified as one of the best predictors of flood loss at the microscale (object level) that can be derived from the OSM data. On the meso-scale (regional level), OSM derived building area is a good representation of building area information collected via surveys. Besides OSM, other **data sources that provide building footprints** can also serve as data basis for this approach.

The number of floods in recent history in Europe can be derived from [DFO](#) and [HANZE flood event databases](#) to serve as proxy data for flood experience of the local population. These new data sets can be integrated into [BN-FLEMOps](#) as predictors for regions where detailed local survey data is not available. In cases where **detailed local data** is available, it can be used to update the model and improve its performance.

These findings improve the performance of multi-variable loss models for the microscale without the need for household surveys, which are often resource intense. Furthermore, these new and **open data sources** improve transferability of existing multi-variable models from the micro- to the meso-scale.

2. What is required to improve the transferability of loss models with included uncertainty quantification?

[Chapter 2](#) shows that **models based on spatial measures** derived from OSM building geometries are well suited for the transfer across regions where footprint data is available. These loss models **require water depth information and building footprint geometries**. In transfer applications at the microscale, OSM based models outperform conventional stage-damage models and often perform as well as multi-variable models with micro-level predictors. The building data should be complete and precise to ensure model accuracy. In regions where building characteristics are significantly different from the training area, the performance, especially the reliability, of OSM-based models are lower than models using detailed survey data.

To train these detailed **multi-variable models, extensive empirical data sets** are required. For model transfers, **heterogeneous data** from multiple events in different regions are best suited. BN-FLEMOps is a Bayesian Network model that meets these requirements and can be updated with local data to deliver more accurate and reliable loss estimates. The transfer of BN-FLEMOps from the microscale to the meso-scale is accomplished by the integration of **widely available, consistent data sets from OSM, DFO and HANZE**. These open datasets, together with the implementation of BN-FLEMOps in the open source, [OASIS-lmf](#) provide better accessibility to complex multi-variable loss models. The **Bayesian Network approach** for loss modelling enables **applications with incomplete input data**, and thus increases the transferability to

more data scarce regions. Incomplete input data is transparently reflected through higher uncertainties in the probabilistic predictions.

3. How does a probabilistic, multi-variable model, built on novel data, perform in comparison with traditional approaches?

The probabilistic multi-variable model BN-FLEMOps provides, compared to an ensemble of flood loss models, a **higher average prediction performance**. Compared to the individual members of the ensemble, BN-FLEMOps results in the **fifth most accurate and least biased estimation**. While the median of the model prediction slightly underestimates the overall loss, the interquartile range of predictions covers the reported loss.

Probabilistic predictions for each municipality show that BN-FLEMOps adequately covers the model ensemble variance in its probabilistic results. This means that the probabilistic multi-variable model **provide detailed information about estimation uncertainties** without the need to implement and run an entire model ensemble. This uncertainty information is valuable for decision-making and supports flood risk management.

4. How will flood risk for residential buildings in Europe change throughout the 21st century?

In the historic baseline period, Germany, France, and Italy are the three European countries with the overall highest fluvial flood risk. Climate change effects are estimated to **double (116 %) flood risk under rcp45 and almost triple (181 %) the risk under rcp85** until the end of the 21st century. The later periods of the century 2055 and 2085 show a growing difference between the climate scenarios. Under rcp85, the risk increase overall will be stronger, while uncertainties in the estimates become larger (see [Figure B.5](#)). Climate change will lead to **increasing** flood risk in most of **Central Europe and the British Isles**. Parts of **Scandinavia and the Mediterranean** are estimated to experience **stagnating** or even **declining** flood risk.

The influence of changing **exposure will almost quadruple (275 %) the flood risk** until 2085. **South-Eastern Europe** and parts of the British Isles as well as southern Sweden are projected to have stronger increases in risk, driven by exposure change, than the rest of Europe. The **combined effect** of climate and exposure change is estimated to increase the flood risk by 670 % under rcp45 and 950 % under rcp85 until the end of the century. Overall, a divide between higher **risk increase in urban regions** and lower increase in rural regions can be seen in the estimates.

Current structural flood protection standards are especially high in the Netherlands, London, and Budapest and reduce the EAD in Europe by 4.7 billion €. In order to reduce the impacts of changing climate and increasing exposure, vulnerability reduction via adaptation measures is crucial. In this respect, the flood loss model BN-FLEMOps is the first to directly include private precautionary measures as a predictor of flood loss in Europe. It is estimated that, on average, the implementation of improved **private precaution is capable of reducing flood risk by 15 %**.

6.2 Discussion

In this chapter, several aspects of the research performed in this thesis are discussed, along with opportunities for future exploration. The discussion follows the structure of the thesis as presented in [Figure 1.4](#), by addressing the subject of new data sources for flood loss models, the issue of uncertainty in flood risk and the implications of flood risk change in Europe. The discussion includes all components of flood risk, but focuses on the vulnerability component of flood loss estimation.

6.2.1 New data sources and model transfer

Flood loss models built only using water depth and numerical spatial measures (OSM-based) as described in [Chapter 2](#) greatly reduce the effort of data collection and processing. The comparison shows that OSM based models can achieve similar predictive performance as multi-variable models which are based on detailed object level data collected in post flood surveys. OSM is an openly available data source and offers great potential for automation of data extraction and calculation of spatial measures to build, tests and transfer loss models.

However, OSM does not provide all relevant data to comprehensively represent flood damage processes – for example, information about building characteristics are limited in OSM. Building material, number of storeys, or building age are useful to explain building vulnerability (Thieken et al., 2005), but are rarely available from OSM. The OSM database can be extended with additional features such as building height and floor space based on population density (Paprotny & Schröter, 2020), and building age (Schorlemmer et al., 2020). With the addition of new features, OSM could support the implementation of more detailed loss models beyond the ones developed in this thesis. Algorithmic correction of building locations (Vargas-Muñoz et al., 2019), prediction of land use (Srivastava et al., 2018) and automated error detection (Basiri et al., 2016) show promising results that will further improve the quality of OSM data. In continuation, OSM based flood loss models should be developed and tested for more regions in Europe and globally to examine their potential outside the geographical context of the tested case study Central Europe (Germany). In this respect, the degree to which building data quality influence loss model performance should be determined in future studies.

OSM is not the only available data source for high resolution building information. Microsoft published building footprint data under the Open Data Commons Open Database Licence (ODbL) for the USA, Canada, Tanzania, and Uganda. The datasets were generated using a classification algorithm from Bing Maps to classify to delineate building footprints in satellite and aerial imagery (Microsoft, 2021a, 2021b, 2021c). Open Building footprints for the African continent were extracted from high resolution satellite data by Google Research (Sirko et al., 2021). Other approaches for building footprint classification with remote sensing data show prospects in urban and data-scarce regions (Schuegraf & Bittner, 2019; Li et al., 2020), where they could provide important information such as the distance of buildings to channels (Malgwi et al., 2021). These data sources and approaches also offer the potential to build flood loss models for regions where OSM data is less reliable or unavailable. Moreover, datasets created by corporate contributors via automated classification have in several cases already been integrated into OSM as Anderson et al. (2019) reports.

Similar loss models for other asset categories could be build, analogous to the concepts set out in [Chapter 2](#) and [3](#). Paprotny et al. (2020a) for example, can provide up-to-date

exposure values for commercial assets. Schoppa et al. (2020) provides several viable multi-variable models that can be transferred and applied on the European scale for flood loss estimation for the commercial sector. Koks et al. (2019) and van Ginkel et al. (2021) demonstrate the usefulness of OSM data for modelling flood risk to road and rail infrastructure. In addition to these advancements, the potential of extracting relevant vulnerability information from building geometry has not yet been fully explored and offers opportunity for future research.

The thesis demonstrates how empirical flood loss models can be transferred between scales and geographical regions, and it is confirmed that multi-variable models built on heterogeneous data are well suited for model transfer applications, as was previously suggested by Wagenaar et al. (2018). Wagenaar et al. (2018) report that the German empirical damage data set is very diverse and covers many possible damage cases, and BN-FLEMOps is therefore more suited than other models to be transferred in space and time. Nevertheless, the comparison with reported losses in [Figure B.1](#) indicates that losses in Germany are especially well represented, while deviations for other countries are larger. There are good reasons to assume that expanding the empirical database for the training of BN-FLEMOps with flood damage data from multiple European countries would improve estimation accuracy and reduce uncertainties. A Europe-wide effort to collect flood damage data in a standardized form, similar to the National Flood Insurance Program of FEMA (2021) in the USA, would benefit national and continental flood risk assessments. Larger consistent damage databases would also open up opportunities for the development of loss models based on more data intense machine learning methodologies such as artificial neural networks (Opella & Hernandez, 2019; Chen et al., 2020; Tabbussum & Dar, 2020; Lazin et al., 2021).

The use of open data and software makes the approaches presented in this thesis more accessible and reproducible. Many components of the risk modelling approach used for the Europe-wide study in [Chapter 5](#) such as the climate data (WCRP, 2021), the hydrological model (JRC, 2021) and the vulnerability data sources (Paprotny et al., 2018a; OSM contributors, 2020) are openly accessible. Combining the different components of the risk modelling chain is often time and resource expensive. The implementation of BN-FLEMOps in the open-source loss model framework OASIS makes the model more accessible to users and more compatible with other components in the risk modelling chain (Hattermann et al., 2018; Schröter et al., 2021). [OASIS-lmf](#) is an open platform for loss modelling with many developers and users from the research and (re-)insurance sector. In research and practice, many flood risk model components are still closed source or use proprietary software. Well documented and openly accessible data and software contribute to the advancement of flood risk modelling.

The experience of households with floods is a relevant predictor for flood loss Vogel et al. (2018) and Mohor et al. (2021). However, vulnerability information such as the flood experience and precautionary measures installed to protect a building can usually only be quantified through on-site assessment or remote surveys (Kreibich et al., 2005; Poussin et al., 2015). Through the integration of historic flood event counts from [DFO](#) and [HANZE](#) database, it was possible to represent these less tangible vulnerability information in the Europe-wide application of BN-FLEMOps. Nonetheless, flood experience and private precaution are defined in broad classes in BN-FLEMOps (see [Figure 3.1](#)) and therefore only general conclusions about their

influence on risk reduction can be drawn in this thesis. Furthermore, data in large-scale event database suffer from poor spatial resolution. The DFO catalogue only roughly traces the outlines of flood events, which can lead to an overestimation of flood experience, as discussed in [Chapter 3](#). HANZE data with its flood counts on NUTS-3 level offers higher resolution for most flood events and in addition dates back until the beginning of the 20th century. HANZE data was therefore chosen for the European-wide flood risk assessment in [Chapter 5](#). Future risk assessments on all scales may explore the effects of individual measures, flood awareness or warning lead time in more detail.

6.2.2 Uncertainty in flood loss estimation

All components of risk; hazard, exposure, and vulnerability are associated with uncertainties. When uncertainty information is reported transparently, it supports risk-based decision-making (Aven, 2004) and helps to prevent biases in adaptation decisions (Oakley et al., 2020). Vulnerability modelling and the assumptions involved introduce the largest uncertainty to risk estimation (Winter et al., 2018). Combined uncertainties from vulnerability models and exposure data can account for variations up to a factor of four in loss estimates at the meso-scale (de Moel & Aerts, 2011). Several sources of uncertainty are discussed in this section, with a focus on vulnerability and the risk assessment for residential buildings in Europe.

Vulnerability

One of the sources of uncertainty in vulnerability models arise from the model structure and parametrization. In the BN-FLEMOps model, the discretization of damage classes influences the level of uncertainty. The model distinguishes between ten classes of relative damage with very granular separation in lower classes, but the highest class covers a large range from 0.226 to 0.637. This is because the empirical database does not include enough cases of severely damaged or completely destroyed buildings to provide evidence for a finer resolution with respect to high damage classes. Thus, the uncertainty is higher for larger damages. [Chapter 3](#) demonstrates that model updating with detailed local data can improve the prediction performance and reduce uncertainties in loss estimates. But the example for the River Lech shows that when local data is of insufficient quantity (sample size) or quality (reliability), updating may not successfully improve the predictions. In this respect, it is important to acknowledge that detailed empirical damage data is still rare and will often not suffice for training, validation and updating of models. More efforts to collect post event damage data in a standardized form is called for in [subsection 6.2.1](#). Data collection for small local events can add to the empirical data basis for model training and validation, as was demonstrated by Rözer et al. (2016), Laudan et al. (2017), and Molinari et al. (2017).

The probabilistic results presented in [Chapter 4](#) capture the ensemble variance and demonstrate the usefulness of uncertainty representation with only one model. Despite the high accuracy of the modelling of the reported loss overall, some municipalities were not well represented. The same regions that were also not well modelled by the ensemble are only marginally better captured by BN-FLEMOps. On average, BN-FLEMOps is more accurate, but still makes similar “errors” as the ensemble. The fact that this miss match of flood reported and modelled loss exists with the ensemble and BN-FLEMOps suggests that there is a systematic error in the models or that the reported losses are incorrect. In loss comparisons, it has to be considered that

also the reported loss data may be incomplete and contain uncertainties (Molinari et al., 2020). The probabilistic graph-based approach of BN-FLEMOps makes these modelling challenges transparent and quantifies the uncertainties for more informed decision-making.

Exposure

According to, Koivumäki et al. (2010) uncertainties in exposure modelling for buildings are often disregarded in risk analysis, while the hazard and vulnerability component receive greater attention. Significant uncertainties are associated with the asset value of exposed elements that can influence the variation of loss estimates by a factor of two (de Moel & Aerts, 2011). To quantify uncertainties in exposure projections for the Europe-wide risk assessment, the Inter Quartile Range **IQR** is reported. **Figure 5.6** reveals that the relative variation around exposure projections is smaller than compared to the hazard (climate) and vulnerability components. This suggests that exposure projections are more reliable or that not all uncertainties are adequately represented. Exposure projections in **Chapter 5** for example could not include land use change projections, because of the unreliability of current data sets for Europe. Furthermore, they do not consider possible impacts of climate change on the socio-economic development and the influence on exposure values in Europe. Climate change is projected to impact many economic sectors in Europe (Harrison et al., 2015; Holtermann & Rische, 2020). Changes in these sectors may influence reconstruction costs of buildings. Policy initiatives like the “European Green Deal” that will allocate substantial public investment into retrofitting of the European building stock (Wolf et al., 2021) could influence asset values. These coupled effects should be analysed in future studies to further improve exposure models for flood loss estimation and to better inform climate adaptation strategies.

Hazard

Uncertainties in the flood hazard data can originate from the resolution, parametrization, and structure of the hydrological and hydrodynamic models and their underlying data used to predict flood water depth (Feyen et al., 2007; de Moel, 2012; Z. Liu & Merwade, 2018).

In the Europe-wide application, for example, smaller catchments (<500 km²) are excluded from the hazard modelling because they are not accounted for in the hydrological and hydraulic simulations. Furthermore, the resolution of the hydrological model is 5 x 5 km (Dottori et al., 2021a). Therefore, flood events of small extent and the contributions from small catchments are not included in the Europe-wide risk assessment. This contributes to the overall uncertainty of risk estimates, but the exact contribution is difficult to quantify. Alfieri et al. (2018) identifies the lack of high-resolution Digital Elevation Models (**DEM**) for Europe as one of the main contributors to hazard uncertainty. The current best DEM for Europe has a resolution of 25 by 25 meters and many small, but hydrodynamically relevant features can not be represented, resulting in inaccuracies in flood extent and depth estimations. In a comparison of large-scale hazard models in the USA, Devitt et al. (2021) found that models exhibit regional differences in performance and especially high variation in mountainous areas. The authors concluded that model ensembles should be used to better represent these uncertainties. An ensemble setup of LISFLOOD-FP shows that model averaging is also a viable approach to represent uncertainties in hydraulic simulations for flood extent and water depth (Z. Liu & Merwade, 2018). The hazard

data for the European flood risk assessment uses an ensemble approach, and an analysis of its performance and the involved uncertainties is reported in Dottori et al. (2021a).

Flood protection standards for Europe are a source of large uncertainty for risk estimation. The dataset for protection levels used in the European-wide risk assessment in [Chapter 5](#) recognizes the high variability in data quality throughout Europe. The developers Dottori et al. (2021b) report a lack of information in Eastern Europe in particular. Actual design levels are only available for few countries and urban areas (Scussolini et al., 2016; Dottori et al., 2017). Much of the protection information used is therefore derived from policy recommendations or via comparison of modelled losses against reported loss figures. The creation of a Europe-wide dataset including the true location and height of protection infrastructure has the potential to greatly reduce uncertainties in continental risk assessments. Dike location and crest heights can for example be derived from high resolution ($>10 \text{ m}^2$) DEM's based on [LiDAR](#) data (Wing et al., 2019). Such high-resolution elevation data would also improve the accuracy of hydrodynamic models in regions without flood protection.

Another important source of uncertainty are the climate projections used to predict changes in the return periods of flood events and flood protection standard in Europe under two rcp scenarios. An ensemble of 11 EURO-CORDEX models is used to estimate the climatic changes in Europe forced by hypothetical emission scenarios that heavily depend on societal and technological developments (van Vuuren et al., 2011). The different parametrization of these models results in a spread between estimations, as can be seen in [Figure B.4](#). The regional patterns of climate model variance can be traced in the risk projections for Europe. [Figure B.5](#) depicts the variance of expected [EAD](#) values around the median estimate. The variation in the [IQR](#) of the estimated [EAD](#) for Europe grows towards the end of the century and is higher under rcp85. Regional hotspots of high [qcv](#) are located in northern Scandinavia, western part of the Alps and in Greece. Risk estimates in these regions have to be understood as more uncertain and should be critically compared to regional studies based on detailed local data. For example, in a nation-wide study for Finland Veijalainen et al. (2010) estimated decreasing flood risk for most regions of the country under climate change comparable to the results of [Chapter 5](#). The variation between climate models, though, was found to be higher than between emission scenarios. To reduce and better quantify the uncertainties related to climate projections, further research on the parametrization and downscaling methodologies will be required.

6.2.3 Future flood risk in Europe

The results of this thesis concerning fluvial flood risk in Europe align with findings of the IPCC (2021) report. Climate change will have heterogeneous effects on the various regions of Europe. The British Isles and Central Europe will experience higher flood risk, while in Northern and Southern Europe stagnating or decreasing flood risk is expected. Lower flood risk in Scandinavia and the Mediterranean may be a positive outlook for the future, but the IPCC (2021) projects increased aridity and risk of droughts in the Mediterranean at $2 \text{ }^\circ\text{C}$ global warming. Besides fluvial floods and droughts, pluvial floods pose an increasing risk for Europe (Kaspersen et al., 2017; Prokic et al., 2019; IPCC, 2021). In contrast to the findings of Hosseinzadehtalaei et al. (2020) that climate change is a stronger driver of pluvial flood risk in comparison to socio-economic development in Europe, [Chapter 5](#) on the drivers of fluvial flood risk projects that increase in exposure has a dominant impact on fluvial flood risk

change in all European countries. Development of exposure can be influenced in the near-term and more directly than changes in the climate system. Societies can use the tools of spatial planning to limit new constructions in flood plains and adapt to the climate driven changes in flood risk (Wilson, 2006). The protection of buildings that already exist can be improved through flood protection infrastructure and private precaution. The new EU strategy on adaptation to climate change recognizes climate change and exposure development as risk drivers and explores political options for adaptation (European Commission, 2021). Current levels of flood protection as reported in Chapter 5 reduce annual fluvial flood risk to residential buildings by 4.7 billion € in Europe. To preserve the current state of protection under climate change, many regions in Europe will need to increase their protection level by the middle of the century (Willner et al., 2018). Nevertheless, It is not economically viable to build and maintain dyke structures in all risk areas. Dottori et al. (2021b) for example, shows that strengthening dykes is not cost-effective in many regions of the Mediterranean.

Chapter 5 shows that, improved private precaution can reduce flood losses by 15 % on average and up to 20 % in European regions. With the currently available data and models, however, it was not possible to give detailed recommendations for which exact measures should be implemented to achieve the best damage reduction. A broader data basis of building characteristics and further research on the effectiveness of measures would be required to transfer the knowledge generated in case studies by Kreibich et al. (2015), Aerts (2018), and Du et al. (2020) for implementation in a Europe-wide multi-variable model. Dottori et al. (2021b) analysed the effects and benefit-cost ratio of flood proofing measures in general and based on literature values found a potential of 12 % EAD reduction for Europe. Other effective adaptation strategies include increased local retention and relocation of exposed assets and population.

6.3 Conclusion and outlook

This thesis has advanced flood loss modelling for residential buildings in Europe by developing new modelling approaches and adapting existing probabilistic multi-variable models through the integration of new data sets.

Information derived from OSM building footprints were found to be valuable for developing performant loss models. This approach enables transfers of flood loss models to regions where building geometries are available, but detailed vulnerability data is lacking. OSM-based models perform significantly better than conventional stage-damage curves and deliver comparable results to detailed empirical data-based multi-variable models. Data-scarce regions can therefore profit from more accurate loss estimations by OSM based models, without the need for resource intense data collection. Since the OSM based approach does not solve all modelling challenges concerning the characterization of building vulnerability and the transfer of models, extending open building databases by additional features such as building level, age, and construction material will address some of its current shortcomings. The advancements in automated classification and growing popularity of the OSM project hold promise for rapid improvements in the development of consistent building data sets for loss modelling. Through the use of open data and software, this thesis makes loss modelling approaches more transparent and accessible, while also offering great potential for automation.

The transfer of multi-variable loss models from the micro- to the meso-scale is possible through the integration of large consistent data sets describing building characteristics (OSM) and flood experience of the local population (DFO, HANZE). The results of this thesis that model development and transfer, especially for continental applications, could be further improved through the extension of the empirical database. Standardized post event damage data collection on all scales should be a priority to promote for future advancements in loss modelling. The finding, that updating existing graph-based models with new local data improves model performance, further emphasizes this conclusion.

Loss estimates from probabilistic multi-variable approaches are shown to be more accurate than those of a model ensemble, and the probabilistic results cover the variability of the entire ensemble. The probabilistic approach therefore delivers useful uncertainty information for decision-making requiring the application of only one model. In flood risk management, this detailed uncertainty information can be used by individuals or organizations to support the decision process. BN-FLEMOps is the first probabilistic multi-variable flood loss model for residential buildings in Europe. Combined with the most recent hazard and exposure projections, BN-FLEMOps can estimate the contribution of each risk component to flood risk change individually and in combination. It is shown that exposure change increases risk throughout all regions of Europe, but has the strongest effects in South-East Europe and the British Isles. Avoiding unnecessary new construction of buildings in flood prone areas is a clear conclusion of these results. Climate change as a risk driver has a more heterogeneous impact on Europe. Central-Europe and the British Isles will experience higher levels of flood risk. On the contrary, a stagnation, or decline in fluvial flood risk is projected for Northern Scandinavia and most regions in the Mediterranean. The two diverging trends of climate driven risk change propagate throughout the century and lead to more extreme changes under rcp85 than under rcp45. The uncertainties around the climate driven risk increase are the largest in these regions of northern and Southern Europe, and uncertainties are higher under rcp85. Limiting climate change as much as still possible promises lower flood risk in most of Europe and higher certainty in risk projections.

The detailed spatial resolution of this approach reveals that flood risk increase is expected to be strongest in urban centres and their surrounding regions located along major European rivers. Therefore, suggesting that European flood risk management should focus on these regions. By integrating new vulnerability data into the multi-variable model the possibility to reduce flood risk for residential buildings by improving private precautionary measures is quantified at 15 % on average for Europe. Especially in areas where large scale flood protection structures are unviable, incentivize or requiring higher levels of private precaution can help to reduce flood risk significantly.

In conclusion, flood loss estimation for residential buildings can be advanced through the integration of new data sets representing building characteristics and flood experience. This allows for model development in data-scarce areas and the transfer in space and scale. The consistent data basis and transparent graph-based model structure of the European-wide loss estimation enables comparisons of risk change at the regional level. Finally, probabilistic multi-variable approaches provide new information about vulnerability and uncertainty for flood risk management and can contribute to building more resilient societies.

Appendix A

Appendix to Chapter 2

A.1 Numerical spatial measures





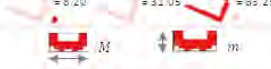
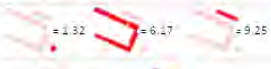


Variables	Equation	Meaning	Range and example values
1 Area (Area)	$Area$	Area of the building in square meter	$[0 \text{ m}^2 \quad \infty$
2 Perimeter (Perimeter)	$Perimeter$	Perimeter of the building in meter	$[0 \text{ m} \quad \infty$
3 Degree of compactness (DegrComp)	$\frac{Area \cdot 4\pi}{Perimeter^2}$	Compactness of the building shape, relative vicinity of the internal points, normalized to a circle	$[0 \quad 1]$ 
4 Perimeter-area ratio (FARatio)	$Perimeter / Area$	Simple measure of shape complexity, biased by building size	$[0 \quad \infty$ 
5 Shape index (ShapeIndex)	$\frac{Perimeter / 4}{\sqrt{Area}}$	Shape complexity, adjusted to building size, normalized to a square	$[1 \quad \infty$ 
6 Fractal dimension index (FracDimInd)	$\frac{2 \ln(Perimeter/4)}{\ln Area}$	Shape complexity, scaled from 1 to 2, adjusted to the size, norm to a square	$[1 \quad 2]$ 
7 Radius of gyration (RadGyras)	$\sum_{i=1}^n \frac{dist(vertex_i - centroid)}{num \ of \ vertices}$	Building extent and compactness in meter	$[0 \text{ m} \quad \infty$ 
	$Major, \ minor$	Major and minor axes of the Minimum bounding rectangle	
8 Linear segment indicator (LinSegInd)	$Major / minor$	Elongation of the polygon, normalized to a square	$[1 \quad \infty$ 
9 Ratio of bounding rectangle area (BoundRatio)	$\frac{Bound \ rectangle \ Area}{Area}$	Shape complexity, normalized to the hypothetical simplest polygon	$[1 \quad \infty$ 

Figure A.1: Definition and examples for numerical spatial measures.

A.2 Code and data availability

Flood damage data of the 2005, 2006, 2010, 2011, and 2013 events along with instructions on how to access the data are available via the German flood damage database, HOWAS21 (<http://howas21.gfz-potsdam.de/howas21/>). Flood damage data of the 2002 event was partly funded by the reinsurance company Deutsche Rückversicherung (<https://deutscherueck.de>) and may be obtained upon request. The surveys were supported by the German Research Network Natural Disasters (German Ministry of Education and Research (BMBF), 01SFR9969/5), the MEDIS project (BMBF; 0330688) the project "Hochwasser 2013" (BMBF; 13N13017), and by a joint venture between the German Research Centre for Geosciences GFZ, the University of Potsdam, and the Deutsche Rückversicherung AG, Düsseldorf. OSM is an open data project and the cartographic information can be downloaded, altered and redistributed under the Open Data Commons Open Database Licence (ODbL)

(OSM contributors, 2020). In the presented study, the geographic data were processed in PostgreSQL 12.2 with PostGIS 3.0.1 extension and R version 3.6.3 (2020-02-29) (R Core Team, 2020). The spatial measures were calculated in PostgreSQL and imported in to R for further processing. The RandomForest model was built and applied in R with the use of the following packages: randomForest 4.6-14 (Liaw & Wiener, 2002), sf 0.6-3 (Pebesma, 2018), reshape2_1.4.3 (Wickham, 2007), gdalUtilities_1.1.0 (O'Brien, 2021), rpostgis_1.4.3 (Bucklin & Basille, 2018), rgdal_1.4-8 (Bivand et al., 2021), raster_3.0-7 (Hijmans et al., 2021), RPostgreSQL_0.6-2 (Conway et al., 2021), tidyverse_1.3.0 (Wickham et al., 2019).

A.3 Competing interests

Authors Heidi Kreibich and Kai Schröter are members of the editorial board of Natural Hazards and Earth System Sciences

Appendix B

Appendix to Chapter 5

B.1 Exposure projections

File available for download (`exposure_1995_2085.csv`).

B.1.1 Population growth

Population growth Projections of regional population are based on recent EUROPOP 2019 projections for NUTS-3 regions by Eurostat (2021a) for years 2019-2100. As the Eurostat projections include only scenario-based projections, we added country-level uncertainty ranges from probabilistic population projections by the United Nations (2019b). The Bayesian probabilistic projections provided by the agency based on the methodology developed by Raftery et al. (2012) present the median and two confidence intervals (60% and 90%). We assumed that the Eurostat baseline scenario corresponds to the United Nations' median scenario. The uncertainty distributions of United Nations projections evaluated at five given percentiles are almost exactly uniform for all countries, hence we fit the data points to a uniform distribution to obtain population numbers for other percentiles.

The Eurostat projections cover all countries in this study except the United Kingdom. National projections for England, Northern Ireland, Scotland, Wales were gathered from the statistical offices of the different components of the UK. As the projections are only until year 2043, the final growth rates were extrapolated assuming convergence with national growth rate by year 2100 (a manifestation of "Sigma" convergence; Sala-i-Martin, 1996; Monfort, 2008). Population growth-rate for the UK was taken from the United Nations projections. For Northern Ireland, an older set of demographic projections until 2041 were used because of changed in administrative divisions. Also, the regional projections from Eurostat are provided for NUTS version 2016, therefore they were adjusted to NUTS version 2013 by linking the corresponding regions in both classifications.

B.1.2 Economic growth

The analysis of economic growth started with compiling a database of GDP per capita in constant prices for all NUTS-3 regions in the study area between 2000 and 2018, using data from (United Nations, 2019a; OECD, 2020; Eurostat, 2021b) and national statistical institutes. The growth rate of GDP per capita is negatively correlated with the level of GDP per capita in the preceding year in this dataset (rank correlation of -0.16). This finding is consistent with the "Beta" convergence (Barro & Sala-i-Martin, 1992; Monfort, 2008), which states that poorer regions or

countries will grow faster than richer ones. This convergence in turn stems from Solow (1956) neoclassical growth theory, which postulates (in simplification) that long-term economic growth will be driven by the diminishing return of increases in factors of production, particularly capital. Here, we also find that the dependency structure resembles most closely a Clayton copula, i.e., growth rates of poorer regions are more concentrated in the probability space than growth rates of richer regions. We use this information to perform a Markov Chain Monte Carlo (MCMC) simulation. In this calculation, GDP per capita growth rate in year t depends only on the level of GDP per capita in year $t-1$. The Clayton copula quantified with our GDP dataset is sampled at each step to obtain a prediction of growth rate based on the previous year's level of GDP per capita. The procedure is carried out for each region starting with the GDP level in the latest year available from historical data in yearly time steps up to 2085. This random walk is repeated 10,000 times to obtain the empirical uncertainty distribution of GDP per capita in every region for all years until 2085. The starting year is mostly 2017 or 2018, but for a minority of regions it is within 2014–2016. An out-of-sample validation of the method was carried out for regions with complete data for 2000–2018. Half of the (randomly selected) regions quantified the Clayton copula, and the other half was used in the MCMC simulation starting in 2000. The results, show that 85% and 60% of observed GDP per capita levels were within the 90% and 60% confidence intervals of the simulation in 2018.

B.1.3 Wealth-to-income ratio

The problem of wealth-to-income ratio has become very prominent since Piketty and Goldhammer (2014) prediction that wealth will continue to outpace income growth in the 21st century. Empirical data show that from 1950 to 2015 the gross stock of residential buildings in 37 European countries has grown from 150% to almost 220% of GDP (Paprotny et al., 2018b). Without new construction of dwellings, the wealth-to-income ratio would have declined in our study area from 184% to 141% between 2000 and 2018, according to Paprotny and Schröter (2020). Instead, it increased to 221%. Decomposition of this difference shows that almost half can be attributed to the increase in demand for dwellings: population growth, decline in the average number of persons per household, increase in average dwelling size and increase in the number of second homes. The other half results from a strong increase in dwelling construction costs (not to be mistaken with market prices), which have continuously outpaced the price inflation in the general economy (as measured by the GDP deflator). We therefore assume that the upward trend will continue in the future Europe-wide, according to historical data (1950–1995 from Paprotny et al. (2018a) and 2000–2018 from Paprotny and Schröter (2020)). This pan-European trend was applied to modify national ratios recorded in 2018 to estimate their values in 2025, 2055, 2085. The standard error in fitting the trend was computed and used to generate normally-distributed confidence intervals.

B.2 Flood loss model variables

Table B.1: Overview of BN-FLEMOps variables

Abbreviation	Variable	Unit	Classes
rbloss	Relative building loss of residential buildings	Relative 0 to 1	10
wd	Water depth relative to ground level	Meters	10
rp*	Return period of the flood event calculated for peak flood discharges with extreme value statistics on annual maximum series of discharge	Years	3
d**	Inundation duration at the affected building	Hours	5
pre	Precautionary measures — indicator (0 = no, 1 = good, 2 = very good precaution) considering the number and type of private precautionary measures undertaken	Score	3
fe	Flood experience — number of floods experienced before the respective damaging flood event	Score	6
ba	Footprint area of residential buildings	Square meter	3
bt**	Building type — 1 = single-family houses, 2 = (semi-)detached houses, 3 = multifamily houses	Index	3

* For this application of BN-FLEMOps the number of classes for the return period variable was reduced from three to five bins. This alteration was necessary because the empirical survey data did not populate five classes with sufficiently large sample sizes.

** Variable not used in this study

B.3 Baseline period

(baseline_protected.shp; baseline_unprotected.csv)

The shape file linked here includes the baseline [EAD](#) estimations per NUTS-3 region with flood protection and without flood protection. The 2013 version of NUTS-3 borders and IDs was used. It is projected in EPSG3035 and the results in S4 – S7 can be joined to this spatial dataset via the column nuts_id.

B.4 Climate change

File available for download (climate_change.csv).

B.5 Exposure change

File available for download (exposure_change.csv).

B.6 Combined effects

File available for download (combined_effects.csv).

B.7 Private precaution effects

File available for download (private_precaution_effects.csv).

B.8 Comparison with reported loss

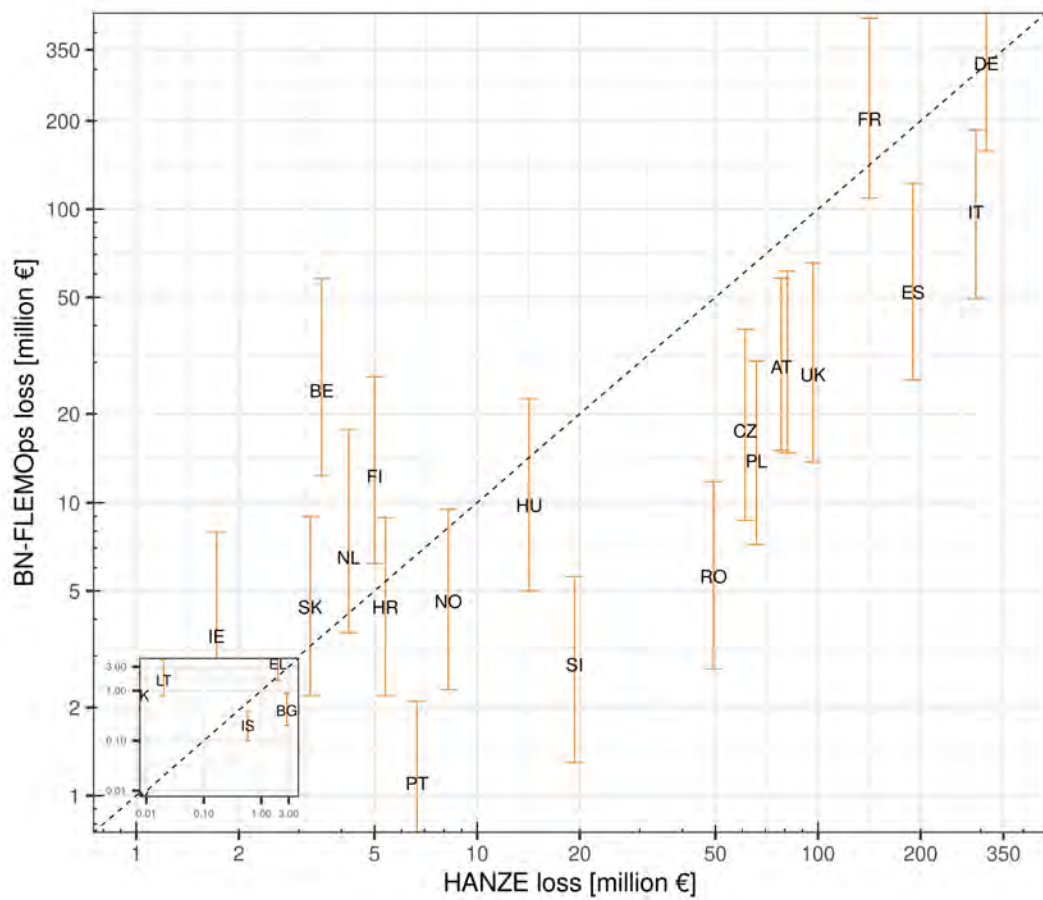


Figure B.1: Comparison of baseline losses with the HANZE database EAD [million €] for the period from 1981 to 2010 on the national level. The location of CH is shifted to avoid overlapping. Its true location is given by the orange point. Missing from the plot are the countries: LV, LI, EE, CY, DK because EAD values are out of range.

Table B.2: Performance indicators for the comparison of baseline losses with the HANZE database EAD [million €] for the period from 1981 to 2010 on the national level

Model spread	BN-FLEMOps	HANZE	Pearson	Spearman	RMSE	MBE
1. Quartile	453.50	/	0.82	0.84	64.69	-30.93
2. Quartile (median)	885.00	1443.40	0.81	0.84	49.20	-17.45
3. Quartile	1932.70	/	0.81	0.83	90.29	15.29

B.9 Private precaution scenarios under future conditions



Figure B.2: Future EAD per country under rcp45, combining the effects of climate and exposure change with three levels of private precautionary measures. Note the change of scale along the y-axis of the right panel for better readability.

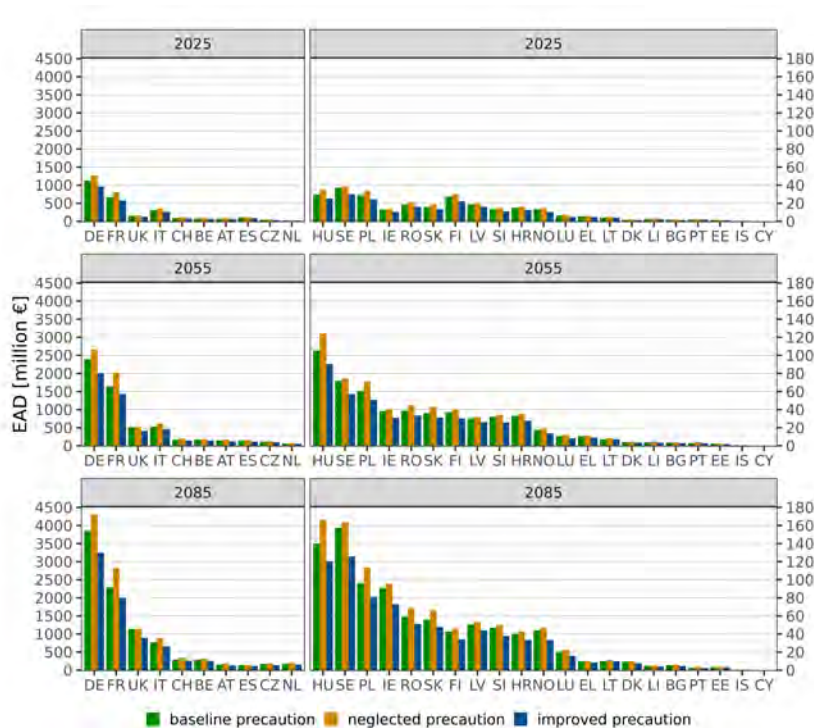


Figure B.3: Future EAD per country under rcp85 combining the effects of climate and exposure change with three levels of private precautionary measures. Note the change of scale along the y axis of the right panel for better readability.

B.10 Uncertainty of climate model ensemble

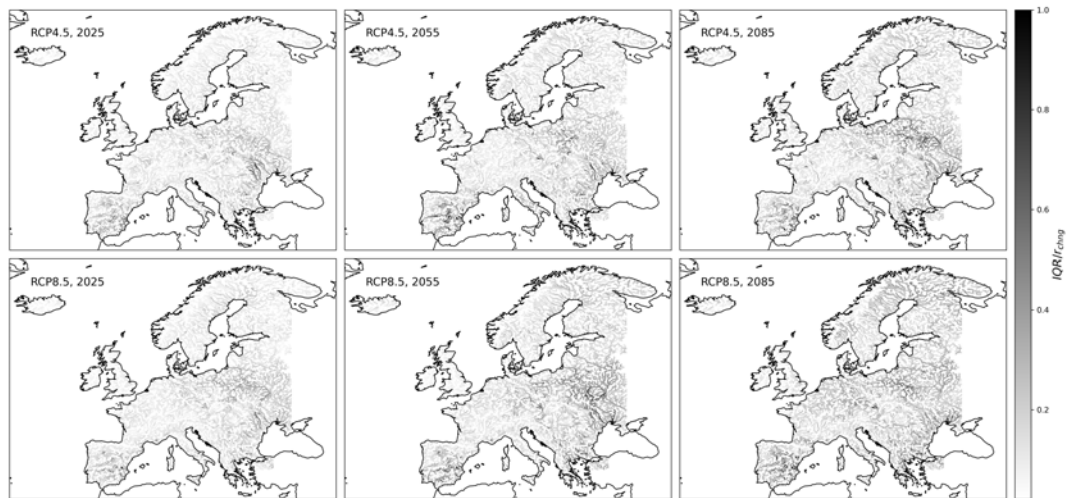


Figure B.4: Relative uncertainty of the model ensemble for discharge change projections under rcp45 and rcp85 scenarios. Darker colour indicates higher uncertainty.

B.11 Uncertainty of loss estimates under climate scenarios

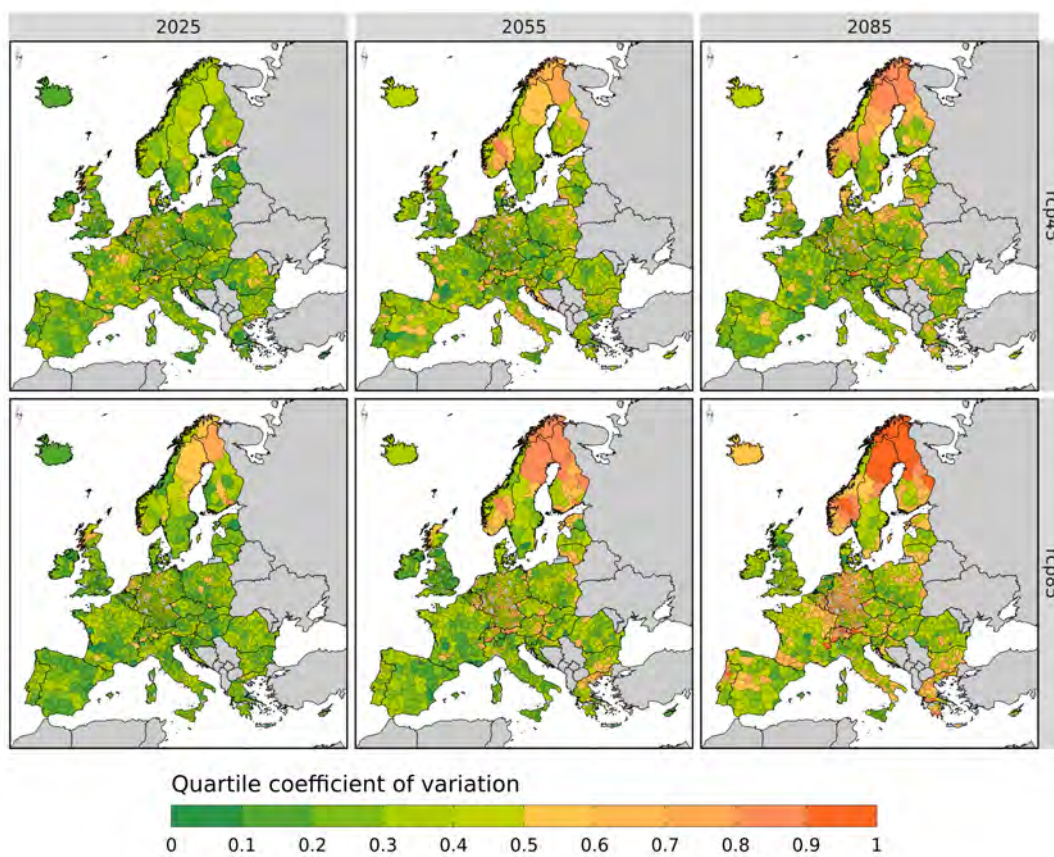


Figure B.5: Quartile coefficient of variation for Expected Annual Damage in three future periods and under two rcp scenarios. Calculated as: $qcv = (Q_3 - Q_1) / (Q_3 + Q_1)$. Green colours indicate lower uncertainty and orange colours higher uncertainty. Map projection EPSG3035. (Amended after paper submission.)

Bibliography

- Aerts, J. C. J. H. (2018). A review of cost estimates for flood adaptation. *Water*, 10(11), 1646. <https://doi.org/10.3390/w10111646>
- Aerts, J. C. J. H., Botzen, W. J., Clarke, K. C., Cutter, S. L., Hall, J. W., Merz, B., Michel-Kerjan, E., Mysiak, J., Surminski, S., & Kunreuther, H. (2018). Integrating human behaviour dynamics into flood disaster risk assessment. *Nature Climate Change*, 8(3), 193–199. <https://doi.org/10.1038/s41558-018-0085-1>
- Alfieri, L., Burek, P., Feyen, L., & Forzieri, G. (2015a). Global warming increases the frequency of river floods in europe. *Hydrology and Earth System Sciences*, 19(5), 2247–2260. <https://doi.org/10.5194/hess-19-2247-2015>
- Alfieri, L., Dottori, F., Betts, R., Salamon, P., & Feyen, L. (2018). Multi-model projections of river flood risk in europe under global warming. *Climate*, 6(1), 6–6. <https://doi.org/10.3390/cli6010006>
- Alfieri, L., Feyen, L., & Di Baldassarre, G. (2016a). Increasing flood risk under climate change: A pan-european assessment of the benefits of four adaptation strategies. *Climatic Change*, 136(3), 507–521. <https://doi.org/10.1007/s10584-016-1641-1>
- Alfieri, L., Feyen, L., Dottori, F., & Bianchi, A. (2015b). Ensemble flood risk assessment in europe under high end climate scenarios. *Global Environmental Change*, 35, 199–212. <https://doi.org/10.1016/j.gloenvcha.2015.09.004>
- Alfieri, L., Feyen, L., Salamon, P., Thielen, J., Bianchi, A., Dottori, F., & Burek, P. (2016b). Modelling the socio-economic impact of river floods in europe. *Natural Hazards and Earth System Sciences*, 16(6), 1401–1411. <https://doi.org/https://doi.org/10.5194/nhess-16-1401-2016>
- Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., & Feyen, L. (2014). Advances in pan-european flood hazard mapping. *Hydrological Processes*, 28(13), 4067–4077. <https://doi.org/10.1002/hyp.9947>
- Amadio, M., Scorzini, A. R., Carisi, F., Essenfelder, A. H., Domeneghetti, A., Mysiak, J., & Castellarin, A. (2019). Testing empirical and synthetic flood damage models: The case of italy. *Natural Hazards and Earth System Sciences*, 19(3), 661–678. <https://doi.org/10.5194/nhess-19-661-2019>
- Amirebrahimi, S., Rajabifard, A., Mendis, P., & Ngo, T. (2016). A framework for a microscale flood damage assessment and visualization for a building using BIM–GIS integration. *International Journal of Digital Earth*, 9(4), 363–386. <https://doi.org/10.1080/17538947.2015.1034201>
- Amt der Tiroler Landesregierung. (2006). *Rechnungsabschluss 2005 des landes tirol*. Landesrechnungshof. Innsbruck.
- Anderson, J., Sarkar, D., & Palen, L. (2019). Corporate editors in the evolving landscape of OpenStreetMap. *ISPRS International Journal of Geo-Information*, 8(5), 232. <https://doi.org/10.3390/ijgi8050232>
- Antrop, M. (2004). Landscape change and the urbanization process in europe. *Landscape and Urban Planning*, 67(1), 9–26. [https://doi.org/10.1016/S0169-2046\(03\)00026-4](https://doi.org/10.1016/S0169-2046(03)00026-4)
- Apel, H., Aronica, G. T., Kreibich, H., & Thielen, A. H. (2009). Flood risk analyses—how detailed do we need to be? *Natural Hazards*, 49(1), 79–98. <https://doi.org/10.1007/s11069-008-9277-8>
- Apel, H., Thielen, A. H., Merz, B., & Blöschl, G. (2004). Flood risk assessment and associated uncertainty. *Natural Hazards and Earth System Science*, 4(2), 295–308. <https://doi.org/10.5194/nhess-4-295-2004>

- Atreya, A., Czajkowski, J., Botzen, W., Bustamante, G., Campbell, K., Collier, B., Ianni, F., Kunreuther, H., Michel-Kerjan, E., & Montgomery, M. (2017). Adoption of flood preparedness actions: A household level study in rural communities in tabasco, mexico. *International Journal of Disaster Risk Reduction*, 24, 428–438. <https://doi.org/10.1016/j.ijdrr.2017.05.025>
- Aven, T. (2004). On how to approach risk and uncertainty to support decision-making. *Risk Management*, 6(4), 27–39. <https://doi.org/10.1057/palgrave.rm.8240196>
- Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253(1), 1–13. <https://doi.org/10.1016/j.ejor.2015.12.023>
- Barredo, J. I. (2009). Normalised flood losses in europe: 1970-2006. *Natural Hazards and Earth System Science*, 9(1), 97–104. <https://doi.org/10.5194/nhess-9-97-2009>
- Barredo, J. I., Saurí, D., & Llasat, M. C. (2012). Assessing trends in insured losses from floods in spain 1971–2008. *Natural Hazards and Earth System Sciences*, 12(5), 1723–1729. <https://doi.org/10.5194/nhess-12-1723-2012>
- Barrington-Leigh, C., & Millard-Ball, A. (2017). The world's user-generated road map is more than 80% complete. *PLoS ONE*, 12(8), 1–20. <https://doi.org/10.1371/journal.pone.0180698>
- Barrington-Leigh, C., & Millard-Ball, A. (2019). Correction: The world's user-generated road map is more than 80% complete. *PLOS ONE*, 14(10), e0224742. <https://doi.org/10.1371/journal.pone.0224742>
- Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2), 223–251. <https://doi.org/10.1086/261816>
- Basiri, A., Jackson, M., Amirian, P., Pourabdollah, A., Sester, M., Winstanley, A., Moore, T., & Zhang, L. (2016). Quality assessment of OpenStreetMap data using trajectory mining. *Geo-spatial Information Science*, 19(1), 56–68. <https://doi.org/10.1080/10095020.2016.1151213>
- Basu, S., Kumbier, K., Brown, J. B., & Yu, B. (2018). Iterative random forests to discover predictive and stable high-order interactions. *Proceedings of the National Academy of Sciences*, 115(8), 1943–1948. <https://doi.org/10.1073/pnas.1711236115>
- Bates, P. D., & De Roo, A. P. J. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, 236(1), 54–77. [https://doi.org/10.1016/S0022-1694\(00\)00278-X](https://doi.org/10.1016/S0022-1694(00)00278-X)
- Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, 387(1), 33–45. <https://doi.org/10.1016/j.jhydrol.2010.03.027>
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Sumner, M., Hijmans, R., Baston, D., Rouault, E., Warmerdam, F., Ooms, J., & Rundel, C. (2021). *Rgdal: Bindings for the 'geospatial' data abstraction library* (Version 1.5-23). Retrieved August 4, 2021, from <https://CRAN.R-project.org/package=rgdal>
- BKG. (2007). Digitales geländemodelle für deutschland DGM-d (DGM25).
- Blanco-Vogt, A., & Schanze, J. (2014). Assessment of the physical flood susceptibility of buildings on a large scale – conceptual and methodological frameworks. *Natural Hazards and Earth System Sciences*, 14(8), 2105–2117. <https://doi.org/10.5194/nhess-14-2105-2014>
- Blöschl, G., Nester, T., Komma, J., Parajka, J., & Perdigão, R. A. P. (2013). The june 2013 flood in the upper danube basin, and comparisons with the 2002, 1954 and 1899 floods. *Hydrology and Earth System Sciences*, 17(12), 5197–5212. <https://doi.org/10.5194/hess-17-5197-2013>
- Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G. T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Frolova, N., Ganora, D., ... Živković, N. (2019). Changing climate both increases and decreases european

- river floods. *Nature*, 573(7772), 108–111. <https://doi.org/10.1038/s41586-019-1495-6>
- Brakenridge, G. (2018). *Global active archive of large flood events* [Dartmouth flood observatory, university of colorado]. Retrieved January 21, 2019, from <http://floodobservatory.colorado.edu/Archives/index.html>
- Breiman, L. (1998). *Classification and regression trees* (1st ed.). Chapman & Hall/CRC.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brovelli, M., & Zamboni, G. (2018). A new method for the assessment of spatial accuracy and completeness of OpenStreetMap building footprints. *ISPRS International Journal of Geo-Information*, 7(8), 289. <https://doi.org/10.3390/ijgi7080289>
- Brückner, J., Schott, M., Zipf, A., & Lautenbach, S. (2021). Assessing shop completeness in OpenStreetMap for two federal states in germany. *AGILE: GIScience Series*, 2, 1–7. <https://doi.org/10.5194/agile-giss-2-20-2021>
- Bryan, B. A., Nolan, M., McKellar, L., Connor, J. D., Newth, D., Harwood, T., King, D., Navarro, J., Cai, Y., Gao, L., Grundy, M., Graham, P., Ernst, A., Dunstall, S., Stock, F., Brinsmead, T., Harman, I., Grigg, N. J., Battaglia, M., ... Hatfield-Dodds, S. (2016). Land-use and sustainability under intersecting global change and domestic policy scenarios: Trajectories for australia to 2050. *Global Environmental Change*, 38, 130–152. <https://doi.org/10.1016/j.gloenvcha.2016.03.002>
- Bubeck, P., Botzen, W. J. W., Kreibich, H., & Aerts, J. C. J. H. (2012). Long-term development and effectiveness of private flood mitigation measures: An analysis for the german part of the river rhine. *Natural Hazards and Earth System Sciences*, 12(11), 3507–3518. <https://doi.org/10.5194/nhess-12-3507-2012>
- Bubeck, P., Botzen, W. J. W., Kreibich, H., & Aerts, J. C. J. H. (2013). Detailed insights into the influence of flood-coping appraisals on mitigation behaviour. *Global Environmental Change*, 23(5), 1327–1338. <https://doi.org/10.1016/j.gloenvcha.2013.05.009>
- Bubeck, P., Botzen, W. J. W., Laudan, J., Aerts, J. C. J. H., & Thieken, A. H. (2018). Insights into flood-coping appraisals of protection motivation theory: Empirical evidence from germany and france. *Risk Analysis*, 38(6), 1239–1257. <https://doi.org/10.1111/risa.12938>
- Bubeck, P., Dillenardt, L., Alfieri, L., Feyen, L., Thieken, A. H., & Kellermann, P. (2019). Global warming to increase flood risk on european railways. *Climatic Change*, 155(1), 19–36. <https://doi.org/10.1007/s10584-019-02434-5>
- Buck, W., & Merkel, U. (1999). *Auswertung der HOWAS-schadendatenbank. analysis of the HOWAS data base* (HY98/15). Institut für Wasserwirtschaft und Kulturtechnik der Universität Karlsruhe.
- Bucklin, D., & Basille, M. (2018). Rpostgis: Linking r with a PostGIS spatial database. *The R Journal*, 10(1), 251. <https://doi.org/10.32614/RJ-2018-025>
- Budiyono, Y., Aerts, J., Brinkman, J., Marfai, M. A., & Ward, P. (2015). Flood risk assessment for delta mega-cities: A case study of jakarta. *Natural Hazards*, 75(1), 389–413. <https://doi.org/10.1007/s11069-014-1327-9>
- Bui, Q.-T., Nguyen, Q.-H., Nguyen, X. L., Pham, V. D., Nguyen, H. D., & Pham, V.-M. (2020). Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. *Journal of Hydrology*, 581, 124379. <https://doi.org/10.1016/j.jhydrol.2019.124379>
- Burton, H., Rabito, F., Danielson, L., & Takaro, T. K. (2016). Health effects of flooding in canada: A 2015 review and description of gaps in research. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, 41(1), 238–249. <https://doi.org/10.1080/07011784.2015.1128854>
- Cammerer, H., Thieken, A. H., & Lammel, J. (2013). Adaptability and transferability of flood loss functions in residential areas. *Natural Hazards and Earth System Sciences*, 13(11), 3063–3081. <https://doi.org/10.5194/nhess-13-3063-2013>
- Carisi, F., Schröter, K., Domeneghetti, A., Kreibich, H., & Castellarin, A. (2018). Development and assessment of uni- and multivariable flood loss models for

- emilia-romagna (italy). *Natural Hazards and Earth System Sciences*, 18(7), 2057–2079. <https://doi.org/https://doi.org/10.5194/nhess-18-2057-2018>
- Casanueva, A., Kotlarski, S., Herrera, S., Fernández, J., Gutiérrez, J. M., Boberg, F., Colette, A., Christensen, O. B., Goergen, K., Jacob, D., Keuler, K., Nikulin, G., Teichmann, C., & Vautard, R. (2016). Daily precipitation statistics in a EURO-CORDEX RCM ensemble: Added value of raw and bias-corrected high-resolution simulations. *Climate Dynamics*, 47(3), 719–737. <https://doi.org/10.1007/s00382-015-2865-x>
- Changnon, S. A. (2003). Shifting economic impacts from weather extremes in the united states: A result of societal changes, not global warming. *Natural Hazards*, 29(2), 273–290. <https://doi.org/10.1023/A:1023642131794>
- Chen, J., Li, Q., Wang, H., & Deng, M. (2020). A machine learning ensemble approach based on random forest and radial basis function neural network for risk evaluation of regional flood disaster: A case study of the yangtze river delta, china. *International Journal of Environmental Research and Public Health*, 17(1), 49. <https://doi.org/10.3390/ijerph17010049>
- Chinh, D., Gain, A., Dung, N., Haase, D., & Kreibich, H. (2015). Multi-variate analyses of flood loss in can tho city, mekong delta. *Water*, 8(1), 6. <https://doi.org/10.3390/w8010006>
- Choryński, A., Pińskwar, I., Kron, W., Brakenridge, G. R., & Kundzewicz, Z. W. (2012). Catalogue of large floods in europe in the 20th century. *Changes in flood risk in europe* (1st Edition, p. 28). CRC Press. <https://doi.org/10.1201/b12348>
- Conradt, T., Roers, M., Schröter, K., Elmer, F., Hoffmann, P., Koch, H., Hattermann, F. F., & Wechsung, F. (2012). Vergleich der extremhochwässer 2002 und 2013 im deutschen teil des elbegebiets und deren abflusssimulation durch SWIM-live. *Hydrologie und Wasserbewirtschaftung*, 56(3), 241–245. https://doi.org/10.5675/HyWa_2013_5_4
- Conway, J., Eddelbuettel, D., Nishiyama, T., Prayaga, S. K., & Tiffin, N. (2021). *RPostgreSQL: R interface to the 'PostgreSQL' database system* (Version 0.7). Retrieved August 4, 2021, from <https://CRAN.R-project.org/package=RPostgreSQL>
- Cortès, M., Turco, M., Llasat-Botija, M., & Llasat, M. C. (2018). The relationship between precipitation and insurance data for floods in a mediterranean region (northeast spain). *Natural Hazards and Earth System Sciences*, 18(3), 857–868. <https://doi.org/10.5194/nhess-18-857-2018>
- CRED. (2020). *Human cost of disasters: An overview of the last 20 years (2000-2019)*. Centre for Research on the Epidemiology of Disasters (CRED). Brussels. Retrieved December 4, 2020, from <https://reliefweb.int/sites/reliefweb.int/files/resources/Human%20Cost%20of%20Disasters%202000-2019%20Report%20-%20UN%20Office%20for%20Disaster%20Risk%20Reduction.pdf>
- Cullen, A. C., Frey, H. C., & Frey, C. H. (1999). *Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs*. Springer Science & Business Media.
- Cumiskey, L., Priest, S., Valchev, N., Viavattene, C., Costas, S., & Clarke, J. (2018). A framework to include the (inter)dependencies of disaster risk reduction measures in coastal risk assessment. *Coastal Engineering*, 134, 81–92. <https://doi.org/10.1016/j.coastaleng.2017.08.009>
- Debo, T. N. (1982). Urban flood damage estimating curves [Publisher: American Society of Civil Engineers]. *Journal of the Hydraulics Division*, 108(10), 1059–1069. <https://doi.org/10.1061/JYCEAJ.0005906>
- de Brito, M. M., & Evers, M. (2016). Multi-criteria decision-making for flood risk management: A survey of the current state of the art. *Natural Hazards and Earth System Sciences*, 16(4), 1019–1033. <https://doi.org/10.5194/nhess-16-1019-2016>
- de Moel, H. (2012). *Uncertainty in flood risk* (Doctoral dissertation) [ISBN: 9789462030558 OCLC: 810066546]. S.l. s.n.
- de Moel, H., & Aerts, J. C. (2011). Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates. *Natural Hazards*, 58(1), 407–425. <https://doi.org/10.1007/s11069-010-9675-6>

- de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E. C., & Ward, P. J. (2015). Flood risk assessments at different spatial scales. *Mitigation and Adaptation Strategies for Global Change*, 20(6), 865–890. <https://doi.org/10.1007/s11027-015-9654-z>
- de Moel, H., van Vliet, M., & Aerts, J. C. J. H. (2013). Evaluating the effect of flood damage-reducing measures: A case study of the unembanked area of rotterdam, the netherlands. *Regional Environmental Change*. <https://doi.org/10.1007/s10113-013-0420-z>
- Devitt, L., Neal, J., Wagener, T., & Coxon, G. (2021). Uncertainty in the extreme flood magnitude estimates of large-scale flood hazard models. *Environmental Research Letters*, 16(6), 064013. <https://doi.org/10.1088/1748-9326/abfac4>
- Dietz, H., Fischer, S., & Gierschek, C. (2014, December 16). *Wohngebäudeversicherung: Kommentar* (3rd edition). VVW GmbH.
- Dorn, H., Törnros, T., & Zipf, A. (2015). Quality evaluation of VGI using authoritative data—a comparison with land use data in southern germany. *ISPRS International Journal of Geo-Information*, 4(3), 1657–1671. <https://doi.org/10.3390/ijgi4031657>
- Dottori, F., Alfieri, L., Bianchi, A., Skoien, J., & Salamon, P. (2021a). A new dataset of river flood hazard maps for europe and the mediterranean basin region [[Preprint]]. *Earth System Science Data Discussions*, 1–35. <https://doi.org/10.5194/essd-2020-313>
- Dottori, F., Figueiredo, R., Martina, M. L., Molinari, D., & Scorzini, A. R. (2016). INSYDE: A synthetic, probabilistic flood damage model based on explicit cost analysis. *Natural Hazards and Earth System Sciences*, 16(12), 2577–2591. <https://doi.org/10.5194/nhess-16-2577-2016>
- Dottori, F., Kalas, M., Salamon, P., Bianchi, A., Alfieri, L., & Feyen, L. (2017). An operational procedure for rapid flood risk assessment in europe. *Natural Hazards and Earth System Sciences*, 17(7), 1111–1126. <https://doi.org/https://doi.org/10.5194/nhess-17-1111-2017>
- Dottori, F., Mentaschi, L., Bianchi, A., Alfieri, L., & Feyen, L. (2020). *Adapting to rising river flood risk in the EU under climate change: JRC PESETA IV project : Task 5*. (EUR 29955 EN JRC118425) [ISBN 978-92-76-12946-2]. Publications Office of the European Union. Luxembourg. Retrieved May 4, 2021, from <https://data.europa.eu/doi/10.2760/14505>
- Dottori, F., Mentaschi, L., Bianchi, A., Alfieri, L., & Feyen, L. (2021b). Adaptation is cost-effective to offset rising river flood risk in europe [Preprint]. *Nature Portfolio*. <https://doi.org/10.21203/rs.3.rs-519118/v1>
- Dottori, F., Szewczyk, W., Ciscar, J.-C., Zhao, F., Alfieri, L., Hirabayashi, Y., Bianchi, A., Mongelli, I., Frieler, K., Betts, R. A., & Feyen, L. (2018). Increased human and economic losses from river flooding with anthropogenic warming. *Nature Climate Change*, 8. <https://doi.org/10.1038/s41558-018-0257-z>
- Du, S., Scussolini, P., Ward, P. J., Zhang, M., Wen, J., Wang, L., Koks, E., Diaz-Loaiza, A., Gao, J., Ke, Q., & Aerts, J. C. J. H. (2020). Hard or soft flood adaptation? advantages of a hybrid strategy for shanghai. *Global Environmental Change*, 61, 102037. <https://doi.org/10.1016/j.gloenvcha.2020.102037>
- Dutta, D., Herath, S., & Musiakke, K. (2003). A mathematical model for flood loss estimation. *Journal of Hydrology*, 277(1), 24–49. [https://doi.org/10.1016/S0022-1694\(03\)00084-2](https://doi.org/10.1016/S0022-1694(03)00084-2)
- El Adlouni, S., Bobée, B., & Ouarda, T. B. M. J. (2008). On the tails of extreme event distributions in hydrology. *Journal of Hydrology*, 355(1), 16–33. <https://doi.org/10.1016/j.jhydrol.2008.02.011>
- Elmer, F., Thielen, A. H., Pech, I., & Kreibich, H. (2010). Influence of flood frequency on residential building losses. *Natural Hazards and Earth System Science*, 10(10), 2145–2159. <https://doi.org/10.5194/nhess-10-2145-2010>
- EM-DAT. (2019). *CRED*. Retrieved May 30, 2019, from <http://www.emdat.be>
- Engel, H. (2004). The flood event 2002 in the elbe river basin, causes of the flood, its course, statistical assessment and flood damages. *La Houille Blanche*, 6, 33–36.

- European Commission. (2021). *Forging a climate-resilient europe - the new EU strategy on adaptation to climate change* (SWD(2021) 25 final). Brussels. Retrieved September 15, 2021, from https://ec.europa.eu/clima/sites/clima/files/adaptation/what/docs/swd_2021_25_en.pdf
- European Environment Agency. (2010). *Mapping the impacts of recent natural disasters and technological accidents in europe - an overview of the last decade* (No. 978-92-9213-168-5). <https://doi.org/10.2800/62638>
- European Environment Agency. (2016). Corine land cover (CLC) 2012, version 18.5.1, seamless vector data. <https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012?tab=download>
- European Union. (2007). Directive 2007/60/EC of the european parliament and of the council of 23 october 2007 on the assessment and management of flood risks.
- Eurostat. (2018). *Gross domestic product (GDP) at current market prices by NUTS 3 regions*. Retrieved November 30, 2018, from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_3gdp&lang=en
- Eurostat. (2021a). EUROPOP2019 - regional projections [data file]. Retrieved June 14, 2020, from https://ec.europa.eu/eurostat/databrowser/view/proj_19rp3/default/table?lang=en
- Eurostat. (2021b). Eurostat - database. Retrieved July 13, 2021, from <https://ec.europa.eu/eurostat/data/database>
- Fan, H., Zipf, A., Fu, Q., & Neis, P. (2014). Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science*, 28(4), 700–719. <https://doi.org/10.1080/13658816.2013.867495>
- Felder, G., Gómez-Navarro, J. J., Zischg, A. P., Raible, C. C., Röthlisberger, V., Bozhinova, D., Martius, O., & Weingartner, R. (2018). From global circulation to local flood loss: Coupling models across the scales. *Science of the Total Environment*, 635, 1225–1239. <https://doi.org/10.1016/j.scitotenv.2018.04.170>
- FEMA. (2021). *OpenFEMA data sets* | FEMA.gov. Retrieved September 13, 2021, from <https://www.fema.gov/about/openfema/data-sets>
- Fenton, N., & Neil, M. (2013). *Risk assessment and decision analysis with bayesian networks*. CRC Press.
- Few, R. (2003). Flooding, vulnerability and coping strategies: Local responses to a global threat. *Progress in Development Studies*, 3(1), 43–58. <https://doi.org/10.1191/1464993403ps049ra>
- Feyen, L., Vrugt, J. A., Nualláin, B. Ó., van der Knijff, J., & De Roo, A. (2007). Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model. *Journal of Hydrology*, 332(3), 276–289. <https://doi.org/10.1016/j.jhydrol.2006.07.004>
- Figueiredo, R., & Martina, M. (2016). Using open building data in the development of exposure data sets for catastrophe risk modelling. *Natural Hazards and Earth System Sciences*, 16(2), 417–429. <https://doi.org/10.5194/nhess-16-417-2016>
- Figueiredo, R., Schröter, K., Weiss-Motz, A., Martina, M. L. V., & Kreibich, H. (2018). Multi-model ensembles for assessment of flood losses and associated uncertainty. *Natural Hazards and Earth System Sciences*, 18(5), 1297–1314. <https://doi.org/10.5194/nhess-18-1297-2018>
- Formetta, G., & Feyen, L. (2019). Empirical evidence of declining global vulnerability to climate-related hazards. *Global Environmental Change*, 57, 101920. <https://doi.org/10.1016/j.gloenvcha.2019.05.004>
- Förster, S., Kneis, D., Gocht, M., & Bronstert, A. (2005). Flood risk reduction by the use of retention areas at the elbe river. *International Journal of River Basin Management*, 3(1), 21–29. <https://doi.org/10.1080/15715124.2005.9635242>
- Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236. <https://doi.org/10.1016/j.patrec.2010.03.014>
- Gerl, T., Kreibich, H., Franco, G., Marechal, D., & Schröter, K. (2016). A review of flood loss models as basis for harmonization and benchmarking (G. J.-P. Schumann, Ed.). *PLOS ONE*, 11(7), 1–22. <https://doi.org/10.1371/journal.pone.0159791>

- Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378. <https://doi.org/10.1198/016214506000001437>
- Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4), 211–221. <https://doi.org/10.1007/s10708-007-9111-y>
- Grabbert, J.-H. (2006). *Analyse der schadensbeeinflussenden faktoren des hochwassers 2002 und ableitung eines mesoskaligen abschätzungsmodells für wohngebäudeschäden* (Thesis). University of Potsdam. Potsdam.
- Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, 27(3), 659–678. <https://doi.org/10.1007/s11222-016-9646-1>
- Güneralp, B., Lwasa, S., Masundire, H., Parnell, S., & Seto, K. C. (2017). Urbanization in africa: Challenges and opportunities for conservation. *Environmental Research Letters*, 13(1), 015002. <https://doi.org/10.1088/1748-9326/aa94fe>
- Hajat, S., Ebi, K. L., Kovats, R. S., Menne, B., Edwards, S., & Haines, A. (2005). The human health consequences of flooding in europe: A review. In W. Kirch, R. Bertollini, & B. Menne (Eds.), *Extreme weather events and public health responses* (pp. 185–196). Springer. https://doi.org/10.1007/3-540-28862-7_18
- Hall, J. W., Meadowcroft, I. C., Sayers, P. B., & Bramley, M. E. (2003). Integrated flood risk management in england and wales. *Natural Hazards Review*, 4(3), 126–135. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2003\)4:3\(126\)](https://doi.org/10.1061/(ASCE)1527-6988(2003)4:3(126))
- Hanea, A. M., Kurowicka, D., & Cooke, R. M. (2006). Hybrid method for quantifying and analyzing bayesian belief nets. *Quality and Reliability Engineering International*, 22(6), 709–729. <https://doi.org/10.1002/qre.808>
- Hanea, A. M., Morales Napoles, O., & Ababei, D. (2015). Non-parametric bayesian networks: Improving theory and reviewing applications. *Reliability Engineering & System Safety*, 144, 265–284. <https://doi.org/10.1016/j.res.2015.07.027>
- Harrison, P. A., Dunford, R., Savin, C., Rounsevell, M. D. A., Holman, I. P., Kebede, A. S., & Stuch, B. (2015). Cross-sectoral impacts of climate change and socio-economic change for multiple, european land- and water-based sectors. *Climatic Change*, 128(3), 279–292. <https://doi.org/10.1007/s10584-014-1239-4>
- Hasanzadeh Nafari, R., Ngo, T., & Lehman, W. (2016a). Calibration and validation of FLFArs-a new flood loss function for australian residential structures. *Natural Hazards and Earth System Sciences*, 16(1), 15–27. <https://doi.org/10.5194/nhess-16-15-2016>
- Hasanzadeh Nafari, R., Ngo, T., & Lehman, W. (2016b). Development and evaluation of FLFAcs- a new flood loss function for australian commercial structures. *International Journal of Disaster Risk Reduction*, 17, 13–23. <https://doi.org/10.1016/j.ijdrr.2016.03.007>
- Hasanzadeh Nafari, R., Ngo, T., & Mendis, P. (2016c). An assessment of the effectiveness of tree-based models for multi-variate flood damage assessment in australia. *Water*, 8(7), 282. <https://doi.org/10.3390/w8070282>
- Hattermann, F. F., Wortmann, M., Liersch, S., Toumi, R., Sparks, N., Genillard, C., Schröter, K., Steinhausen, M., Gyalai-Korpos, M., Máté, K., Hayes, B., del Rocio Rivas López, M., Rácz, T., Nielsen, M. R., Kaspersen, P. S., & Drews, M. (2018). Simulation of flood hazard and risk in the danube basin with the future danube model. *Climate Services*, 12, 14–26. <https://doi.org/10.1016/j.cliser.2018.07.001>
- Hausfather, Z., & Peters, G. P. (2020). Emissions – the ‘business as usual’ story is misleading. *Nature*, 577(7792), 618–620. <https://doi.org/10.1038/d41586-020-00177-3>
- Hecht, R., Kunze, C., & Hahmann, S. (2013). Measuring completeness of building footprints in OpenStreetMap over space and time. *ISPRS International Journal of Geo-Information*, 2(4), 1066–1091. <https://doi.org/10.3390/ijgi2041066>
- Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20(3), 197–243. <https://doi.org/10.1007/BF00994016>
- Hijmans, R. J., Etten, J. v., Sumner, M., Cheng, J., Baston, D., Bevan, A., Bivand, R., Busetto, L., Canty, M., Fasoli, B., Forrest, D., Ghosh, A., Golicher, D., Gray, J.,

- Greenberg, J. A., Hiemstra, P., Hingee, K., Geosciences, I. f. M. A., Karney, C., . . . Wueest, R. (2021). *Raster: Geographic data analysis and modeling* (Version 3.4-13). Retrieved August 4, 2021, from <https://CRAN.R-project.org/package=raster>
- Hochrainer, S., Linnerooth-Bayer, J., & Mechler, R. (2010). The european union solidarity fund. *Mitigation and Adaptation Strategies for Global Change*, 15(7), 797–810. <https://doi.org/10.1007/s11027-009-9209-2>
- Hochrainer-Stigler, S., Linnerooth-Bayer, J., & Lorant, A. (2017). The european union solidarity fund: An assessment of its recent reforms. *Mitigation and Adaptation Strategies for Global Change*, 22(4), 547–563. <https://doi.org/10.1007/s11027-015-9687-3>
- Hoeppel, P. (2016). Trends in weather related disasters – consequences for insurers and society. *Weather and Climate Extremes*, 11, 70–79. <https://doi.org/10.1016/j.wace.2015.10.002>
- Holtermann, L., & Rische, M.-C. (2020). The subnational effect of temperature on economic production: A disaggregated analysis in european regions. *MPRA*, (104606), 64. <https://mpa.ub.uni-muenchen.de/104606/>
- Hosseinzadehtalaei, P., Tabari, H., & Willems, P. (2020). Satellite-based data driven quantification of pluvial floods over europe under future climatic and socioeconomic changes. *Science of The Total Environment*, 721, 137688. <https://doi.org/10.1016/j.scitotenv.2020.137688>
- Huang, B. F., & Boutros, P. C. (2016). The parameter sensitivity of random forests. *BMC Bioinformatics*, 17(1), 331. <https://doi.org/10.1186/s12859-016-1228-x>
- Hudson, P., Botzen, W. J. W., Kreibich, H., Bubeck, P., & Aerts, J. C. J. H. (2014). Evaluating the effectiveness of flood damage mitigation measures by the application of propensity score matching. *Natural Hazards and Earth System Sciences*, 14(7), 1731–1747. <https://doi.org/10.5194/nhess-14-1731-2014>
- Huizinga, J. (2007). *Flood damage functions for EU member states*. HKV Lijn in water. Lelystad, the Netherlands.
- Huizinga, J., Moel, H. D., & Szewczyk, W. (2017). *Global flood depth-damage functions database with guidelines*. <https://doi.org/10.2760/16510>
- Huttenlau, M., Stötter, J., & Stiefelmeyer, H. (2010). Risk-based damage potential and loss estimation of extreme flooding scenarios in the austrian federal province of tyrol. *Natural Hazards and Earth System Science*, 10(12), 2451–2473. <https://doi.org/10.5194/nhess-10-2451-2010>
- Hydrotec. (2002). *Hochwasser-aktionsplan lippe – grundlagen, überflutungsgebiete, schadenspotenzial, defizite und maßnahmen*. Aachen.
- ICPR. (2001). *Atlas on the risk of flooding and potential damage due to extreme floods of the rhine*. International Commission for the Protection of the Rhine. Koblenz.
- IKSE. (2003). *Aktionsplan hochwasserschutz elbe*. Internationale Kommission zum Schutz der Elbe. Magdeburg.
- IPCC. (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation. a special report of working groups I and II of the intergovernmental panel on climate change* (C. B. Field, V. Barros, T. F. Stocker, Q. Dahe, D. Dokken, K. Ebi, M. Mastrandrea, K. Mach, G.-K. Plattner, S. Allen, M. Tignor, & P. Midgley, Eds.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139177245>
- IPCC. (2021). Summary for policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate change 2021: The physical science basis. contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*. Cambridge University Press. Retrieved August 13, 2021, from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM.pdf
- Irwin, A. (2018). No PhDs needed: How citizen science is transforming research. *Nature*, 562(7728), 480–482. <https://doi.org/10.1038/d41586-018-07106-5>

- Izaguirre, C., Losada, I. J., Camus, P., Vigh, J. L., & Stenek, V. (2021). Climate change risk to global port operations. *Nature Climate Change*, 11(1), 14–20. <https://doi.org/10.1038/s41558-020-00937-z>
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., ... Yiou, P. (2014). EURO-CORDEX: New high-resolution climate change projections for european impact research. *Regional Environmental Change*, 14(2), 563–578. <https://doi.org/10.1007/s10113-013-0499-2>
- Jäger, W. S., Christie, E. K., Hanea, A. M., den Heijer, C., & Spencer, T. (2018). A bayesian network approach for coastal risk analysis and decision making. *Coastal Engineering*, 134, 48–61. <https://doi.org/10.1016/j.coastaleng.2017.05.004>
- Ji, Z., Li, N., Xie, W., Wu, J., & Zhou, Y. (2013). Comprehensive assessment of flood risk using the classification and regression tree method. *Stochastic Environmental Research and Risk Assessment*, 27(8), 1815–1828. <https://doi.org/10.1007/s00477-013-0716-z>
- Jones, B., & O'Neill, B. C. (2016). Spatially explicit global population scenarios consistent with the shared socioeconomic pathways. *Environmental Research Letters*, 11(8), 084003. <https://doi.org/10.1088/1748-9326/11/8/084003>
- Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., Gericke, A., Neal, J., Aerts, J. C., & Ward, P. J. (2012a). Comparative flood damage model assessment: Towards a european approach. *Natural Hazards and Earth System Science*, 12(12), 3733–3752. <https://doi.org/10.5194/nhess-12-3733-2012>
- Jongman, B. (2018). Effective adaptation to rising flood risk. *Nature Communications*, 9(1), 1986. <https://doi.org/10.1038/s41467-018-04396-1>
- Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J. C. J. H., Mechler, R., Botzen, W. J. W., Bouwer, L. M., Pflug, G., Rojas, R., & Ward, P. J. (2014). Increasing stress on disaster-risk finance due to large floods. *Nature Climate Change*, 4(4), 264–268. <https://doi.org/10.1038/nclimate2124>
- Jongman, B., Ward, P. J., & Aerts, J. C. (2012b). Global exposure to river and coastal flooding: Long term trends and changes. *Global Environmental Change*, 22(4), 823–835. <https://doi.org/10.1016/j.gloenvcha.2012.07.004>
- Jonkman, S. N. (2007). *Loss of life estimation in flood risk assessment; theory and applications* (Doctoral dissertation) [ISBN:978-0-9021950-9]. TU Delft. Delft. Retrieved August 19, 2021, from <https://repository.tudelft.nl/islandora/object/uuid%3Aabc4fb945-55ef-4079-a606-ac4fa8009426>
- JRC. (2021). *Lisflood OS*. European Commission, Joint Research Centre (JRC). Retrieved September 13, 2021, from <https://github.com/ec-jrc/lisflood-code>
- Jung, M. (2016). LecoS — a python plugin for automated landscape ecology analysis. *Ecological Informatics*, 31, 18–21. <https://doi.org/10.1016/j.ecoinf.2015.11.006>
- Kaspersen, P. S., Ravn, N. H., Arnbjerg-Nielsen, K., Madsen, H., & Drews, M. (2017). Comparison of the impacts of urban development and climate change on exposing european cities to pluvial flooding. *Hydrology and Earth System Sciences*, 21(8), 4131–4147. <https://doi.org/10.5194/hess-21-4131-2017>
- Kellermann, P., Schröter, K., Thielen, A. H., Haubrock, S.-N., & Kreibich, H. (2020). The object-specific flood damage database HOWAS 21. *Natural Hazards and Earth System Sciences*, 20(9), 2503–2519. <https://doi.org/10.5194/nhess-20-2503-2020>
- Kienzler, S., Pech, I., Kreibich, H., Müller, M., & Thielen, A. H. (2015). After the extreme flood in 2002: Changes in preparedness, response and recovery of flood-affected residents in germany between 2005 and 2011. *Natural Hazards and Earth System Sciences*, 15(3), 505–526. <https://doi.org/10.5194/nhess-15-505-2015>
- Kleist, L., Thielen, A. H., Köhler, P., Müller, M., Seifert, I., Borst, D., & Werner, U. (2006). Estimation of the regional stock of residential buildings as a basis for a comparative risk assessment in germany. *Natural Hazards and Earth System*

- Science*, 6(4), 541–552. Retrieved October 22, 2019, from <https://hal.archives-ouvertes.fr/hal-00299328>
- Klijn, F., Baan, P., de Bruijn, K., & Kwadijk, J. (2007). *Overstromingsrisico's in nederland in een veranderd klimaat. verwachtingen, schattingen en berekeningen voor het project nederland later* (Q4290.00). Milieu- en Natuurplanbureau (MNP). WL | delft hydraulics. Retrieved October 7, 2019, from https://www.researchgate.net/publication/268809909_Overstromingsrisico's_in_Nederland_in_een_veranderd_klimaat_Verwachtingen_schattingen_en_berekeningen_voor_het_project_Nederland_Later
- Knijff, J. M. V. D., Younis, J., & Roo, A. P. J. D. (2010). LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, 24(2), 189–212. <https://doi.org/10.1080/13658810802549154>
- Knoke, T., & Wurm, J. (2006). Mixed forests and a flexible harvest policy: A problem for conventional risk analysis? *European Journal of Forest Research*, 125(3), 303–315. <https://doi.org/10.1007/s10342-006-0119-5>
- Knutti, R., Masson, D., & Gettelman, A. (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, 40(6), 1194–1199. <https://doi.org/10.1002/grl.50256>
- Koivumäki, L., Alho, P., Lotsari, E., Käyhkö, J., Saari, A., & Hyyppä, H. (2010). Uncertainties in flood risk mapping: A case study on estimating building damages for a river flood in finland: Uncertainties in flood risk mapping. *Journal of Flood Risk Management*, 3(2), 166–183. <https://doi.org/10.1111/j.1753-318X.2010.01064.x>
- Koks, E. E., Rozenberg, J., Zorn, C., Tariverdi, M., Vousdoukas, M., Fraser, S. A., Hall, J. W., & Hallegatte, S. (2019). A global multi-hazard risk analysis of road and railway infrastructure assets. *Nature Communications*, 10(1), 2677. <https://doi.org/10.1038/s41467-019-10442-3>
- Kreibich, H., Christenberger, S., & Schwarze, R. (2012). Corrigendum to "economic motivation of households to undertake private precautionary measures against floods" published in *nat. hazards earth syst. sci.*, 11, 309–321, 2011. *Natural Hazards and Earth System Sciences*, 12(2), 391–392. <https://doi.org/10.5194/nhess-12-391-2012>
- Kreibich, H., Thielen, A. H., Petrow, T., Müller, M., & Merz, B. (2005). Flood loss reduction of private households due to building precautionary measures – lessons learned from the elbe flood in august 2002 [Publisher: Copernicus GmbH]. *Natural Hazards and Earth System Sciences*, 5(1), 117–126. <https://doi.org/10.5194/nhess-5-117-2005>
- Kreibich, H., Baldassarre, G. D., Vorogushyn, S., Aerts, J. C. J. H., Apel, H., Aronica, G. T., Arnbjerg-Nielsen, K., Bouwer, L. M., Bubeck, P., Caloiero, T., Chinh, D. T., Cortès, M., Gain, A. K., Giampá, V., Kuhlicke, C., Kundzewicz, Z. W., Llasat, M. C., Mård, J., Matczak, P., ... Merz, B. (2017a). Adaptation to flood risk: Results of international paired flood event studies. *Earth's Future*, 5(10), 953–965. <https://doi.org/https://doi.org/10.1002/2017EF000606>
- Kreibich, H., Botto, A., Merz, B., & Schröter, K. (2017b). Probabilistic, multivariable flood loss modeling on the mesoscale with BT-FLEMO. *Risk Analysis*, 37(4), 774–787. <https://doi.org/10.1111/risa.12650>
- Kreibich, H., Bubeck, P., Van Vliet, M., & De Moel, H. (2015). A review of damage-reducing measures to manage fluvial flood risks in a changing climate. *Mitigation and Adaptation Strategies for Global Change*, 20(6), 967–989. <https://doi.org/10.1007/s11027-014-9629-5>
- Kreibich, H., Schröter, K., & Merz, B. (2016). Up-scaling of multi-variable flood loss models from objects to land use units at the meso-scale. *Proceedings of the International Association of Hydrological Sciences*, 373, 179–182. <https://doi.org/10.5194/piahs-373-179-2016>
- Kreibich, H., Seifert, I., Merz, B., & Thielen, A. H. (2010). Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector. *Hydrological Sciences Journal*, 55(8), 1302–1314. <https://doi.org/10.1080/02626667.2010.529815>

- Kreibich, H., & Thieken, A. H. (2009). Coping with floods in the city of dresden, germany. *Natural Hazards*, 51(3), 423. <https://doi.org/10.1007/s11069-007-9200-8>
- Kreibich, H., Thieken, A. H., Haubrock, S.-N., & Schröter, K. (2017c). HOWAS21, the german flood damage database. In D. Molinari, S. Menoni, & F. Ballio (Eds.), *Flood damage survey and assessment: New insights from research and practice* (1st ed., pp. 65–75). American Geophysical Union. <https://www.wiley.com/en-gb/Flood+Damage+Survey+and+Assessment%3A+New+Insights+from+Research+and+Practice-p-9781119217923>
- Kron, W. (2005). Flood risk = hazard • values • vulnerability. *Water International*, 30(1), 58–68. <https://doi.org/10.1080/02508060508691837>
- Kunreuther, H., Heal, G., Allen, M., Edenhofer, O., Field, C. B., & Yohe, G. (2013). Risk management and climate change. *Nature Climate Change*, 3(5), 447–450. <https://doi.org/10.1038/nclimate1740>
- Lamb, R., Keef, C., Tawn, J., Laeger, S., Meadowcroft, I., Surendran, S., Dunning, P., & Batstone, C. (2010). A new method to assess the risk of local and widespread flooding on rivers and coasts. *Journal of Flood Risk Management*, 3(4), 323–336. <https://doi.org/10.1111/j.1753-318X.2010.01081.x>
- Lambin, E. F., & Meyfroidt, P. (2010). Land use transitions: Socio-ecological feedback versus socio-economic change. *Land Use Policy*, 27(2), 108–118. <https://doi.org/10.1016/j.landusepol.2009.09.003>
- Lang, S., & Tiede, D. (2003). vLATE extension für ArcGIS – vektorbasiertes tool zur quantitativen landschaftsstrukturanalyse.
- Laudan, J., Rözer, V., Sieg, T., Vogel, K., & Thieken, A. H. (2017). Damage assessment in braunsbach 2016: Data collection and analysis for an improved understanding of damaging processes during flash floods. *Natural Hazards and Earth System Sciences*, 17(12), 2163–2179. <https://doi.org/10.5194/nhess-17-2163-2017>
- Lauritzen, S. L. (1996). *Graphical models*. Oxford University Press.
- LAWA. (2014). *Nationales Hochwasserschutzprogramm Kriterien und Bewertungsmaßstäbe für die Identifikation und Priorisierung von wirksamen Maßnahmen sowie ein Vorschlag für die Liste der prioritären Maßnahmen zur Verbesserung des präventiven Hochwasserschutzes*. Bund/Länderarbeitsgemeinschaft Wasser (LAWA). Kiel. https://www.bmu.de/fileadmin/Daten_BMU/Download_PDF/Binnengewasser/hochwasserschutzprogramm_bericht_bf.pdf
- LAWA. (2018). *Empfehlungen zur Aufstellung von Hochwassergefahrenkarten und Hochwasserisikokarten*. Bund/Länderarbeitsgemeinschaft Wasser (LAWA). Weimar. https://www.lawa.de/documents/lawa-empfehlungen_aufstellung_hw-gefahrenkarten_und_hw-risikokarten_2_1552298996.pdf
- Lazin, R., Shen, X., & Anagnostou, E. (2021). Estimation of flood-damaged cropland area using a convolutional neural network. *Environmental Research Letters*, 16(5), 054011. <https://doi.org/10.1088/1748-9326/abeba0>
- LfULG. (2012). *Umsetzung der EU hochwasserrisikomanagementrichtlinie im rahmen des INTERREG IV b projektes LABEL - grenzüberschreitender hochwasserrisikomanagementplan weiße elster* - (Kurzbericht No. 14801432). Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie. Dresden.
- Li, Q., Shi, Y., Huang, X., & Zhu, X. X. (2020). Building footprint generation by integrating convolution neural network with feature pairwise conditional random field (FPCRF). *IEEE Transactions on Geoscience and Remote Sensing*, 58(11), 7502–7519. <https://doi.org/10.1109/TGRS.2020.2973720>
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22. <https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf>
- Liu, J., Dietz, T., Carpenter, S. R., Alberti, M., Folke, C., Moran, E., Pell, A. N., Deadman, P., Kratz, T., Lubchenco, J., Ostrom, E., Ouyang, Z., Provencher, W., Redman, C. L., Schneider, S. H., & Taylor, W. W. (2007). Complexity of coupled human and natural systems. *Science*, 317(5844), 1513–1516. <https://doi.org/10.1126/science.1144004>

- Liu, Z., & Merwade, V. (2018). Accounting for model structure, parameter and input forcing uncertainty in flood inundation modeling using bayesian model averaging. *Journal of Hydrology*, 565, 138–149. <https://doi.org/10.1016/j.jhydrol.2018.08.009>
- Lüdtke, S., Schröter, K., Steinhausen, M., Weise, L., Figueiredo, R., & Kreibich, H. (2019a). A consistent approach for probabilistic residential flood loss modeling in europe. *Water Resources Research*, 55(12), 10616–10635. <https://doi.org/https://doi.org/10.1029/2019WR026213>
- Lüdtke, S., Steinhausen, M., Schröter, K., Figueiredo, R., & Kreibich, H. (2019b). Pan-european probabilistic flood loss data for residential buildings. <http://doi.org/10.5880/GFZ.4.4.2019.002>
- Lugeri, N., Kundzewicz, Z. W., Genovese, E., Hochrainer, S., & Radziejewski, M. (2010). River flood risk and adaptation in europe-assessment of the present status. *Mitigation and Adaptation Strategies for Global Change*, 15(7), 621–639. <https://doi.org/10.1007/s11027-009-9211-8>
- Luino, F., Cirio, C. G., Biddoccu, M., Agangi, A., Giulietto, W., Godone, F., & Nigrelli, G. (2009). Application of a model to the evaluation of flood damage. *GeoInformatica*, 13(3), 339–353. <https://doi.org/10.1007/s10707-008-0070-3>
- Macdonald, N., Chester, D., Sangster, H., Todd, B., & Hooke, J. (2012). The significance of gilbert f. white's 1945 paper 'human adjustment to floods' in the development of risk and hazard management. *Progress in Physical Geography: Earth and Environment*, 36(1), 125–133. <https://doi.org/10.1177/0309133311414607>
- Malgwi, M. B., Schlögl, M., & Keiler, M. (2021). Expert-based versus data-driven flood damage models: A comparative evaluation for data-scarce regions. *International Journal of Disaster Risk Reduction*, 57, 102148. <https://doi.org/10.1016/j.ijdrr.2021.102148>
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- McAneney, J., Sandercock, B., Crompton, R., Mortlock, T., Musulin, R., Jr, R. P., & Gissing, A. (2019). Normalised insurance losses from australian natural disasters: 1966–2017. *Environmental Hazards*, 18(5), 414–433. <https://doi.org/10.1080/17477891.2019.1609406>
- Mentaschi, L., Alfieri, L., Dottori, F., Cammalleri, C., Bisselink, B., Roo, A. D., & Feyen, L. (2020). Independence of future changes of river runoff in europe from the pathway to global warming. *Climate*, 8(2), 22. <https://doi.org/10.3390/cli8020022>
- Mentaschi, L., Vousdoukas, M., Voukouvalas, E., Sartini, L., Feyen, L., Besio, G., & Alfieri, L. (2016). The transformed-stationary approach: A generic and simplified methodology for non-stationary extreme value analysis. *Hydrology and Earth System Sciences*, 20(9), 3527–3547. <https://doi.org/10.5194/hess-20-3527-2016>
- Merz, B., Aerts, J., Arnbjerg-Nielsen, K., Baldi, M., Becker, A., Bichet, A., Blöschl, G., Bouwer, L. M., Brauer, A., Cioffi, F., Delgado, J. M., Gocht, M., Guzzetti, F., Harrigan, S., Hirschboeck, K., Kilsby, C., Kron, W., Kwon, H.-H., Lall, U., ... Nied, M. (2014a). Floods and climate: Emerging perspectives for flood risk assessment and management. *Natural Hazards and Earth System Sciences*, 14(7), 1921–1942. <https://doi.org/10.5194/nhess-14-1921-2014>
- Merz, B., Hall, J., Disse, M., & Schumann, A. (2010a). Fluvial flood risk management in a changing world. *Natural Hazards and Earth System Sciences*, 10, 509–527. <https://doi.org/10.5194/nhess-10-509-2010>
- Merz, B., Kreibich, H., & Apel, H. (2008). Flood risk analysis: Uncertainties and validation. *Österreichische Wasser- und Abfallwirtschaft*, 60(5), 89–94. <https://doi.org/10.1007/s00506-008-0001-4>
- Merz, B., Kreibich, H., & Lall, U. (2013). Multi-variate flood damage assessment: A tree-based data-mining approach. *Natural Hazards and Earth System Science*, 13(1), 53–64. <https://doi.org/10.5194/nhess-13-53-2013>

- Merz, B., Kreibich, H., Schwarze, R., & Thieken, A. H. (2010b). Review article: Assessment of economic flood damage. *Natural Hazards and Earth System Science*, 10(8), 1697–1724. <https://doi.org/10.5194/nhess-10-1697-2010>
- Merz, B., Kreibich, H., Thieken, A. H., & Schmidtke, R. (2004). Estimation uncertainty of direct monetary flood damage to buildings. *Natural Hazards and Earth System Science*, 4(1), 153–163. <https://doi.org/10.5194/nhess-4-153-2004>
- Merz, B., Elmer, F., Kunz, M., Mühr, B., Schröter, K., & Uhlemann-Elmer, S. (2014b). The extreme flood in june 2013 in germany. *La Houille Blanche*, 100(1), 5–10. <https://doi.org/10.1051/lhb/2014001>
- Merz, B., & Thieken, A. H. (2005). Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, 309(1), 114–132. <https://doi.org/10.1016/j.jhydrol.2004.11.015>
- Metin, A. D., Dung, N. V., Schröter, K., Guse, B., Apel, H., Kreibich, H., Vorogushyn, S., & Merz, B. (2018). How do changes along the risk chain affect flood risk? *Natural Hazards and Earth System Sciences*, 18(11), 3089–3108. <https://doi.org/10.5194/nhess-18-3089-2018>
- Metin, A. D., Dung, N. V., Schröter, K., Vorogushyn, S., Guse, B., Kreibich, H., & Merz, B. (2020). The role of spatial dependence for large-scale flood risk estimation [Publisher: Copernicus GmbH]. *Natural Hazards and Earth System Sciences*, 20(4), 967–979. <https://doi.org/10.5194/nhess-20-967-2020>
- Meyer, V., Becker, N., Markantonis, V., Schwarze, R., van den Bergh, J. C. J. M., Bouwer, L. M., Bubeck, P., Ciavola, P., Genovese, E., Green, C., Hallegatte, S., Kreibich, H., Lequeux, Q., Logar, I., Papyrakis, E., Pfurtscheller, C., Poussin, J., Przulski, V., Thieken, A. H., & Viavattene, C. (2013). Review article: Assessing the costs of natural hazards – state of the art and knowledge gaps. *Natural Hazards and Earth System Sciences*, 13(5), 1351–1373. <https://doi.org/10.5194/nhess-13-1351-2013>
- Meyer, V., & Messner, F. (2005). *National flood damage evaluation methods: A review of applied methods in england, the netherlands, the czech republic and germany* (Working Paper No. 21/2005). UFZ Discussion Paper. Retrieved August 23, 2021, from <https://www.econstor.eu/handle/10419/45193>
- Microsoft. (2021a). *Computer generated building footprints for canada*. Retrieved September 8, 2021, from <https://github.com/microsoft/CanadianBuildingFootprints>
- Microsoft. (2021b). *Computer generated building footprints for the united states*. Retrieved September 8, 2021, from <https://github.com/microsoft/USBuildingFootprints>
- Microsoft. (2021c). *Open dataset of machine extracted buildings in uganda and tanzania*. Retrieved September 8, 2021, from <https://github.com/microsoft/Uganda-Tanzania-Building-Footprints>
- Middelmann-Fernandes, M. H. (2010). Flood damage estimation beyond stage–damage functions: An australian example. *Journal of Flood Risk Management*, 3(1), 88–96. <https://doi.org/10.1111/j.1753-318X.2009.01058.x>
- Mitchell, J. K. (2003). European river floods in a changing world. *Risk Analysis*, 23(3), 567–574. <https://doi.org/10.1111/1539-6924.00337>
- Mohor, G. S., Thieken, A. H., & Korup, O. (2021). Residential flood loss estimated from bayesian multilevel models. *Natural Hazards and Earth System Sciences*, 21(5), 1599–1614. <https://doi.org/10.5194/nhess-21-1599-2021>
- Molinari, D., Ballio, F., & Menoni, S. (2013). Modelling the benefits of flood emergency management measures in reducing damages: A case study on sondrio, italy. *Natural Hazards and Earth System Sciences*, 13(8), 1913–1927. <https://doi.org/10.5194/nhess-13-1913-2013>
- Molinari, D., Menoni, S., & Ballio, F. (2017). *Flood damage survey and assessment: New insights from research and practice*. John Wiley & Sons.
- Molinari, D., Scorzini, A. R., Arrighi, C., Carisi, F., Castelli, F., Domeneghetti, A., Gallazzi, A., Galliani, M., Grelot, F., Kellermann, P., Kreibich, H., Mohor, G. S., Mosimann, M., Natho, S., Richert, C., Schroeter, K., Thieken, A. H., Zischg, A. P., & Ballio, F. (2020). Are flood damage models converging to reality? lessons learnt from a blind test. *Natural Hazards and Earth System Sciences*, 20, 2997–3017. <https://doi.org/10.5194/nhess-2020-40>

- Monfort, P. (2008). *Convergence of EU regions measures and evolution*. European Commission. Retrieved May 6, 2021, from https://ec.europa.eu/regional_policy/sources/docgener/work/200801_convergence.pdf
- Mooney, P., & Minghini, M. (2017). A review of OpenStreetMap data. In Foody, G, See, L, Fritz, S, Mooney, P, Olteanu-Raimond, A-M, Fonte, C C, & Antoniou, V. (Eds.), *Mapping and the citizen sensor* (pp. 37–59). Ubiquity Press. Retrieved August 5, 2021, from <https://doi.org/10.5334/bbf.c>
- Morss, R. E., Wilhelmi, O. V., Downton, M. W., & Gruntfest, E. (2005). Flood risk, uncertainty, and scientific information for decision making: Lessons from an interdisciplinary project. *Bulletin of the American Meteorological Society*, 86(11), 1593–1602. <https://doi.org/10.1175/BAMS-86-11-1593>
- Munich-RE. (2019). *NatCatService*. Retrieved January 12, 2020, from <https://www.munichre.com/en/solutions/for-industry-clients/natcatservice.html>
- Murakami, D., & Yamagata, Y. (2019). Estimation of gridded population and GDP scenarios with spatially explicit statistical downscaling. *Sustainability*, 11(7), 2106. <https://doi.org/10.3390/su11072106>
- MURL. (2000). Potentielle Hochwasserschäden am Rhein in NRW. Retrieved August 5, 2021, from <https://docplayer.org/storage/59/43420963/1628196118/0hmVtYS23OrJjkew4LMbwg/43420963.pdf>
- Murphy, K. (2012). Machine learning: A probabilistic perspective. adaptive computation and machine learning. MIT Press.
- Nguyen, V. D., Metin, A. D., Alfieri, L., Vorogushyn, S., & Merz, B. (2020). Biases in national and continental flood risk assessments by ignoring spatial dependence. *Scientific Reports*, 10(1), 19387. <https://doi.org/10.1038/s41598-020-76523-2>
- NR&M. (2002). *Guidance on the assessment of tangible flood damages*. Department of Natural Resources and Mines, Queensland Government. Queensland, Australia.
- Oakley, M., Mohun Himmelweit, S., Leinster, P., & Casado, M. (2020). Protection motivation theory: A proposed theoretical extension and moving beyond rationality—the case of flooding. *Water*, 12(7), 1848. <https://doi.org/10.3390/w12071848>
- O'Brien, J. (2021). *gdalUtilities: Wrappers for 'GDAL' utilities executables* (Version 1.1.2). Retrieved August 4, 2021, from <https://CRAN.R-project.org/package=gdalUtilities>
- OECD. (2020). OECD.stat [data file]. Retrieved June 15, 2020, from <https://stats.oecd.org/>
- Opella, J. M. A., & Hernandez, A. A. (2019). Developing a flood risk assessment using support vector machine and convolutional neural network: A conceptual framework. *2019 IEEE 15th International Colloquium on Signal Processing & Its Applications (CSPA)*, 260–265. <https://doi.org/10.1109/CSPA.2019.8695980>
- OSM. (2018). *OpenStreetMap - Copyright* [OpenStreetMap]. Retrieved November 30, 2018, from <https://www.openstreetmap.org/copyright/en>
- OSM contributors. (2018). *OpenStreetMap* [OpenStreetMap]. Retrieved November 30, 2018, from <https://www.openstreetmap.org/>
- OSM contributors. (2020). *OpenStreetMap* [OpenStreetMap]. Retrieved October 30, 2020, from <https://www.openstreetmap.org/>
- Papathoma-Köhle, M., Kappes, M., Keiler, M., & Glade, T. (2011). Physical vulnerability assessment for alpine hazards: State of the art and future needs. *Natural Hazards*, 58(2), 645–680. <https://doi.org/10.1007/s11069-010-9632-4>
- Pappenberger, F., & Beven, K. J. (2006). Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, 42(5). <https://doi.org/10.1029/2005WR004820>
- Paprotny, D., Kreibich, H., Morales-Nápoles, O., Castellarin, A., Carisi, F., & Schröter, K. (2020a). Exposure and vulnerability estimation for modelling flood losses to commercial assets in europe. *Science of The Total Environment*, 737, 140011. <https://doi.org/10.1016/j.scitotenv.2020.140011>
- Paprotny, D., Kreibich, H., Morales-Nápoles, O., Terefenko, P., & Schröter, K. (2020b). Estimating exposure of residential assets to natural hazards in europe using

- open data. *Natural Hazards and Earth System Sciences*, 20(1), 323–343. <https://doi.org/https://doi.org/10.5194/nhess-20-323-2020>
- Paprotny, D., Morales-Nápoles, O., & Jonkman, S. N. (2018a). HANZE: A pan-european database of exposure to natural hazards and damaging historical floods since 1870. *Earth System Science Data*, 10(1), 565–581. <https://doi.org/https://doi.org/10.5194/essd-10-565-2018>
- Paprotny, D., & Schröter, K. (2020). Residential exposure to natural hazards in europe, 2000–2018. <https://doi.org/10.5880/GFZ.4.4.2020.003>
- Paprotny, D., Sebastian, A., Morales-Nápoles, O., & Jonkman, S. N. (2018b). Trends in flood losses in europe over the past 150 years. *Nature Communications*, 9(1), 1985. <https://doi.org/10.1038/s41467-018-04253-1>
- Parker, D., Green, C. H., & Thompson, P. M. (1987). *Urban flood protection benefits: A project appraisal guide*. Gower Technical Press.
- Parker, D., Tapsell, S., & McCarthy, S. (2007). Enhancing the human benefits of flood warnings. *Natural Hazards*, 43(3), 397–414. <https://doi.org/10.1007/s11069-007-9137-y>
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann Publishers Inc.
- Pebesma, E. (2018). Simple features for r: Standardized support for spatial vector data. *The R Journal*, 10(1), 439. <https://doi.org/10.32614/RJ-2018-009>
- Penning-Rowsell, E. C., & Chatterton, J. B. (1977). *Benefits of flood alleviation*. Saxon House.
- Penning-Rowsell, E. C., & Green, C. (2000). New insights into the appraisal of flood-alleviation benefits: (1) flood damage and flood loss information. *Water and Environment Journal*, 14(5), 347–353. <https://doi.org/10.1111/j.1747-6593.2000.tb00272.x>
- Penning-Rowsell, E. C., Johnson, C., Tunstall, S., Tapsell, S., Morris, J., Chatterton, J., Green, C., Wilson, T., Koussela, K., & Fernandez-Bilbao, A. (2005). *The benefits of flood and coastal risk management: A manual of assessment techniques (multi-coloured handbook 2005)*. Middlesex University Press.
- Penning-Rowsell, E. C., Priest, S., Parker, D., Morris, J., Tunstall, S., Viavattene, C., Chatterton, J., & Owen, D. (2018). *Flood and coastal erosion risk management: A manual for economic appraisal* (1st ed.). Routledge. <https://doi.org/10.4324/9780203066393>
- Petrucci, O., Aceto, L., Bianchi, C., Bigot, V., Brázdil, R., Pereira, S., Inbar, M., Kahraman, A., Kılıç, Ö., Kotroni, V., Llasat, M., Llasat-Botija, M., Mercuri, M., Pappagiannaki, K., Řehoř, J., Rossello-Geli, J., Salvati, P., Vinet, F., & Zêzere, J. (2020). EUropean flood fatalities (EUFF) database 1980-2018 (updated) [Version Number: 2 Type: dataset]. <https://doi.org/10.4121/UUID:489D8A13-1075-4D2F-ACCB-DB7790E4542F>
- Pielke, R. A., & Downton, M. W. (2000). Precipitation and damaging floods: Trends in the united states, 1932–97. *Journal of Climate*, 13(20), 3625–3637. [https://doi.org/10.1175/1520-0442\(2000\)013<3625:PADFTI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<3625:PADFTI>2.0.CO;2)
- Piketty, T., & Goldhammer, A. (2014). *Capital in the twenty-first century*. The Belknap Press of Harvard University Press.
- Pistrika, A. K., & Jonkman, S. N. (2010). Damage to residential buildings due to flooding of new orleans after hurricane katrina. *Natural Hazards*, 54(2), 413–434. <https://doi.org/10.1007/s11069-009-9476-y>
- Pitchforth, J., & Mengersen, K. (2013). A proposed validation framework for expert elicited bayesian networks. *Expert Systems with Applications*, 40(1), 162–167. <https://doi.org/10.1016/j.eswa.2012.07.026>
- Pittore, M., Wieland, M., & Fleming, K. (2017). Perspectives on global dynamic exposure modelling for geo-risk assessment. *Natural Hazards*, 86, 7–30. <https://doi.org/10.1007/s11069-016-2437-3>
- Plapp, S. T. (2003). *Wahrnehmung von Risiken aus Naturkatastrophen. Eine empirische Untersuchung in sechs gefährdeten Gebieten Süd- und Westdeutschlands* (Doctoral dissertation). Universität Karlsruhe (TH). Karlsruhe. <https://doi.org/10.5445/IR/3542003>

- Polasky, S., Carpenter, S. R., Folke, C., & Keeler, B. (2011). Decision-making under great uncertainty: Environmental management in an era of global change. *Trends in Ecology & Evolution*, 26(8), 398–404. <https://doi.org/10.1016/j.tree.2011.04.007>
- Poussin, J. K., Wouter Botzen, W. J., & Aerts, J. C. J. H. (2015). Effectiveness of flood damage mitigation measures: Empirical evidence from french flood disasters. *Global Environmental Change*, 31, 74–84. <https://doi.org/10.1016/j.gloenvcha.2014.12.007>
- Prettenthaler, F., Albrecher, H., Asadi, P., & Köberl, J. (2017). On flood risk pooling in europe. *Natural Hazards*, 88(1), 1–20. <https://doi.org/10.1007/s11069-016-2616-2>
- Prokic, M. N., Savić, S., & Pavić, D. (2019). Pluvial flooding in urban areas across the european continent. *Geographica Pannonica*, 23(4). <https://doi.org/10.5937/gp23-23508>
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria, R Foundation for Statistical Computing. <https://www.R-project.org/>
- Raftery, A. E., Li, N., Sevcikova, H., Gerland, P., & Heilig, G. K. (2012). Bayesian probabilistic population projections for all countries. *Proceedings of the National Academy of Sciences*, 109(35), 13915–13921. <https://doi.org/10.1073/pnas.1211452109>
- Raschky, P. A., & Schwindt, M. (2009). *Aid, natural disasters and the samaritan's dilemma*. The World Bank. <https://doi.org/10.1596/1813-9450-4952>
- Rees, S., Markau, H.-J., & Sterr, H. (2003). *MERK – mikroskalige evaluation der risiken in überflutungsgefährdeten küstenniederungen*. Bundesministeriums für Bildung und Forschung und des Ministeriums für ländliche Räume, Landesplanung, Landwirtschaft und Tourismus des Landes Schleswig-Holstein. Kiel.
- Regione del Veneto. (2011a). *31 ottobre–2 novembre 2010: L'alluvione dei santi (Rapporto Statistico 2011)*. <http://statistica.regione.veneto.it/Pubblicazioni/RapportoStatistico2011/Capitolo17.html>
- Regione del Veneto. (2011b). *La grande alluvione*. http://www.regione.veneto.it/web/guest/comunicati-stampa/dettaglio-comunicati?_spp_detailId=361297
- Rehan, B. M. (2018). An innovative micro-scale approach for vulnerability and flood risk assessment with the application to property-level protection adoptions. *Natural Hazards*, 91(3), 1039–1057. <https://doi.org/10.1007/s11069-018-3175-5>
- Riha, J., & Marcikova, M. (2009). Classification and estimation of flood losses. *International Symposium on Water Management and Hydraulic Engineering*, 1–5.
- Rojas, R., Feyen, L., & Watkiss, P. (2013). Climate change and river floods in the european union: Socio-economic consequences and the costs and benefits of adaptation. *Global Environmental Change*, 23(6), 1737–1751. <https://doi.org/10.1016/j.gloenvcha.2013.08.006>
- Rözer, V., Kreibich, H., Schröter, K., Müller, M., Sairam, N., Doss-Gollin, J., Lall, U., & Merz, B. (2019). Probabilistic models significantly reduce uncertainty in hurricane harvey pluvial flood loss estimates. *Earth's Future*, 7(4), 384–394. <https://doi.org/10.1029/2018EF001074>
- Rözer, V., Müller, M., Bubeck, P., Kienzler, S., Thielen, A. H., Pech, I., Schröter, K., Buchholz, O., & Kreibich, H. (2016). Coping with pluvial floods by private households. *Water*, 8(7), 304. <https://doi.org/10.3390/w8070304>
- Rusnack, W. (2021). *MinimumBoundingBox*. Retrieved August 4, 2021, from <https://github.com/BebeSparkelSparkel/MinimumBoundingBox>
- Sairam, N., Schröter, K., Lüdtke, S., Merz, B., & Kreibich, H. (2019a). Quantifying flood vulnerability reduction via private precaution. *Earth's Future*, 7(3), 235–249. <https://doi.org/10.1029/2018EF000994>
- Sairam, N., Schröter, K., Rözer, V., Merz, B., & Kreibich, H. (2019b). Hierarchical bayesian approach for modeling spatiotemporal variability in flood damage processes. *Water Resources Research*, 55(10), 8223–8237. <https://doi.org/10.1029/2019WR025068>
- Sala-i-Martin, X. X. (1996). The classical approach to convergence analysis. *The Economic Journal*, 106(437), 1019–1036. <https://doi.org/10.2307/2235375>

- Salvati, L., Zambon, I., Chelli, F. M., & Serra, P. (2018). Do spatial patterns of urbanization and land consumption reflect different socioeconomic contexts in Europe? *Science of The Total Environment*, 625, 722–730. <https://doi.org/10.1016/j.scitotenv.2017.12.341>
- Saxon Relief Bank. (2005). Schäden durch das Hochwasser 2002. Angaben aus der Fördermitteldatenbank [damages of the flood 2002. data from the compensation data bank] [private communication].
- Scawthorn, C., Flores, P., Blais, N., Seligson, H., Tate, E., Chang, S., Mifflin, E., Thomas, W., Murphy, J., Jones, C., & Lawrence, M. (2006). HAZUS-MH flood loss estimation methodology. II. damage and loss assessment. *Natural Hazards Review*, 7(2), 72–81. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2006\)7:2\(72\)](https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(72))
- Schoppa, L., Sieg, T., Vogel, K., Zöller, G., & Kreibich, H. (2020). Probabilistic flood loss models for companies. *Water Resources Research*, 56(9), 1–19. <https://doi.org/https://doi.org/10.1029/2020WR027649>
- Schorlemmer, D., Beutin, T., Cotton, F., Ospina, N. G., Hirata, N., Ma, K.-F., Nievas, C., Prehn, K., & Wyss, M. (2020). Global dynamic exposure and the OpenBuildingMap - a big-data and crowd-sourcing approach to exposure modeling. <https://doi.org/10.5194/egusphere-egu2020-18920>
- Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F., & Merz, B. (2014). How useful are complex flood damage models? *Water Resources Research*, 50(4), 3378–3395. <https://doi.org/10.1002/2013WR014396>
- Schröter, K., Kunz, M., Elmer, F., Mühr, B., & Merz, B. (2015). What made the June 2013 flood in Germany an exceptional event? a hydro-meteorological evaluation. *Hydrology and Earth System Sciences*, 19(1), 309–327. <https://doi.org/10.5194/hess-19-309-2015>
- Schröter, K., Lüdtke, S., Redweik, R., Meier, J., Bochow, M., Ross, L., Nagel, C., & Kreibich, H. (2018). Flood loss estimation using 3D city models and remote sensing data. *Environmental Modelling & Software*, 105, 118–131. <https://doi.org/10.1016/j.envsoft.2018.03.032>
- Schröter, K., Lüdtke, S., Vogel, K., Kreibich, H., & Merz, B. (2016). Tracing the value of data for flood loss modelling (M. Lang, F. Klijn, & P. Samuels, Eds.). *E3S Web of Conferences*, 7, 05005. <https://doi.org/10.1051/e3sconf/20160705005>
- Schröter, K., Steinhausen, M., Wortmann, M., Lüdtke, S., Hayes, B., Drews, M., Hattermann, F., & Kreibich, H. (2021). Current and future flood risk in the Danube region using an open loss modelling framework. <https://doi.org/10.3311/FloodRisk2020.9.4>
- Schuegraf, P., & Bittner, K. (2019). Automatic building footprint extraction from multi-resolution remote sensing images using a hybrid FCN. *ISPRS International Journal of Geo-Information*, 8(4), 191. <https://doi.org/10.3390/ijgi8040191>
- Scorzini, A. R., & Frank, E. (2017). Flood damage curves: New insights from the 2010 flood in Veneto, Italy. *Journal of Flood Risk Management*, 10(3), 381–392. <https://doi.org/10.1111/jfr3.12163>
- Scussolini, P., Aerts, J. C. J. H., Jongman, B., Bouwer, L. M., Winsemius, H. C., de Moel, H., & Ward, P. J. (2016). FLOPROS: An evolving global database of flood protection standards. *Natural Hazards and Earth System Sciences*, 16(5), 1049–1061. <https://doi.org/10.5194/nhess-16-1049-2016>
- Scutari, M. (2010). Learning Bayesian networks with the bnlearn R package. *Journal of Statistical Software*, 35(3), 1–22. <https://doi.org/10.18637/jss.v035.i03>
- Scutari, M., & Denis, J.-B. (2014). *Bayesian networks*. Chapman; Hall/CRC.
- Sedighi, T., Varga, L., Hosseinian-Far, A., & Daneshkhah, A. (2021). Economic evaluation of mental health effects of flooding using Bayesian networks. *International Journal of Environmental Research and Public Health*, 18(14), 7467. <https://doi.org/10.3390/ijerph18147467>
- Senaratne, H., Mobasher, A., Ali, A. L., Capineri, C., & Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1), 139–167. <https://doi.org/10.1080/13658816.2016.1189556>

- Shaw, B. J., van Vliet, J., & Verburg, P. H. (2020). The peri-urbanization of Europe: A systematic review of a multifaceted process. *Landscape and Urban Planning*, 196, 103733. <https://doi.org/10.1016/j.landurbplan.2019.103733>
- Sieg, T., Vogel, K., Merz, B., & Kreibich, H. (2017). Tree-based flood damage modeling of companies: Damage processes and model performance. *Water Resources Research*, 53(7), 6050–6068. <https://doi.org/10.1002/2017WR020784>
- Sieg, T., Vogel, K., Merz, B., & Kreibich, H. (2019). Seamless estimation of hydrometeorological risk across spatial scales. *Earth's Future*, 7(5), 574–581. <https://doi.org/10.1029/2018EF001122>
- Silva, F. B. e., Gallego, J., & Lavallo, C. (2013). A high-resolution population grid map for Europe. *Journal of Maps*, 9(1), 16–28. <https://doi.org/10.1080/17445647.2013.764830>
- Sirko, W., Kashubin, S., Ritter, M., Annkah, A., Bouchareb, Y. S. E., Dauphin, Y., Keyzers, D., Neumann, M., Cisse, M., & Quinn, J. (2021). Continental-scale building detection from high resolution satellite imagery. *arXiv:2107.12283 [cs]*. Retrieved September 11, 2021, from <http://arxiv.org/abs/2107.12283>
- Smith, D. (1994). Flood damage estimation - a review of urban stage-damage curves and loss functions. *Water SA*, 20(3), 231–238.
- Sohn, W., Brody, S. D., Kim, J.-H., & Li, M.-H. (2020). How effective are drainage systems in mitigating flood losses? *Cities*, 107, 102917. <https://doi.org/10.1016/j.cities.2020.102917>
- Solín, L., Madajová, M., & Skubinčan, P. (2018). Mitigating flood consequences: Analysis of private flood insurance in Slovakia. *Journal of Flood Risk Management*, 11, S173–S185. <https://doi.org/10.1111/jfr3.12191>
- Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1), 65–94. <https://doi.org/10.2307/1884513>
- Spekkers, M. H., Clemens, F. H. L. R., & ten Veldhuis, J. a. E. (2015). On the occurrence of rainstorm damage based on home insurance and weather data. *Natural Hazards and Earth System Sciences*, 15(2), 261–272. <https://doi.org/https://doi.org/10.5194/nhess-15-261-2015>
- Srivastava, S., Lobry, S., Tuia, D., & Munoz, J. V. (2018). Land-use characterisation using google street view pictures and OpenStreetMap, 1–5. Retrieved September 11, 2021, from <https://library.wur.nl/WebQuery/wurpubs/545325>
- Stanke, C., Murray, V., Amlôt, R., Nurse, J., & Williams, R. (2012). The effects of flooding on mental health: Outcomes and recommendations from a review of the literature. *PLoS Currents*, 4, e4f9f1fa9c3cae. <https://doi.org/10.1371/4f9f1fa9c3cae>
- Steinhausen, M., Lüdtke, S., Schröter, K., Figueiredo, R., & Kreibich, H. (2020). Das probabilistische Hochwasserschadensmodell für Wohngebäude BN-FLEMOps. The probabilistic flood damage model for residential buildings BN-FLEMOps. *Hydrologie und Wasserbewirtschaftung*, 64(4), 188–199. https://doi.org/10.5675/HYWA_2020.4_2
- Steinhausen, M., Wagner, P. D., Narasimhan, B., & Waske, B. (2018). Combining sentinel-1 and sentinel-2 data for improved land use and land cover mapping of monsoon regions. *International Journal of Applied Earth Observation and Geoinformation*, 73, 595–604. <https://doi.org/10.1016/j.jag.2018.08.011>
- Stevens, A. J., Clarke, D., & Nicholls, R. J. (2016). Trends in reported flooding in the UK: 1884–2013. *Hydrological Sciences Journal*, 61(1), 50–63. <https://doi.org/10.1080/02626667.2014.950581>
- Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., & Feuston, B. P. (2003). Random forest: A classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information and Computer Sciences*, 43(6), 1947–1958. <https://doi.org/10.1021/ci034160g>
- Tabbussum, R., & Dar, A. Q. (2020). Comparative analysis of neural network training algorithms for the flood forecast modelling of an alluvial Himalayan river. *Journal of Flood Risk Management*, 13(4), 12656. <https://doi.org/10.1111/jfr3.12656>

- Tanoue, M., Hirabayashi, Y., & Ikeuchi, H. (2016). Global-scale river flood vulnerability in the last 50 years. *Scientific Reports*, 6(1), 36021. <https://doi.org/10.1038/srep36021>
- Tariq, M. a. U. R., Hoes, O. a. C., & Giesen, N. C. V. d. (2014). Development of a risk-based framework to integrate flood insurance. *Journal of Flood Risk Management*, 7(4), 291–307. <https://doi.org/10.1111/jfr3.12056>
- Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., & Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature*, 596(7870), 80–86. <https://doi.org/10.1038/s41586-021-03695-w>
- Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F. W., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling & Software*, 90, 201–216. <https://doi.org/10.1016/j.envsoft.2017.01.006>
- Teske, D. (2014). Geocoder accuracy ranking [Series Title: Communications in Computer and Information Science]. In A.-L. Lamprecht & T. Margaria (Eds.), *Process design for natural scientists* (pp. 161–174). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45006-2_13
- Thieken, A. H., Bessel, T., Kienzler, S., Kreibich, H., Müller, M., Pisi, S., & Schröter, K. (2016). The flood of june 2013 in germany: How much do we know about its impacts? *Natural Hazards and Earth System Sciences Discussions*, 1–57. <https://doi.org/10.5194/nhess-2015-324>
- Thieken, A. H., Kreibich, H., Müller, M., & Lamond, J. (2017). Data collection for a better understanding of what causes flood damage—experiences with telephone surveys. *Flood damage survey and assessment* (pp. 95–106). American Geophysical Union (AGU). <https://doi.org/10.1002/9781119217930.ch7>
- Thieken, A. H., Kreibich, H., Müller, M., & Merz, B. (2007). Coping with floods: Preparedness, response and recovery of flood-affected residents in germany in 2002. *Hydrological Sciences Journal*, 52(5), 1016–1037. <https://doi.org/10.1623/hysj.52.5.1016>
- Thieken, A. H., Müller, M., Kleist, L., Seifert, I., Borst, D., & Werner, U. (2006a). Regionalisation of asset values for risk analyses. *Natural Hazards and Earth System Sciences*, 6(2), 167–178. <https://doi.org/https://doi.org/10.5194/nhess-6-167-2006>
- Thieken, A. H., Müller, M., Kreibich, H., & Merz, B. (2005). Flood damage and influencing factors: New insights from the august 2002 flood in germany. *Water Resources Research*, 41(12), 1–16. <https://doi.org/10.1029/2005WR004177>
- Thieken, A. H., Petrow, T., Kreibich, H., & Merz, B. (2006b). Insurability and mitigation of flood losses in private households in germany. *Risk Analysis*, 26(2), 383–395. <https://doi.org/10.1111/j.1539-6924.2006.00741.x>
- Tian, Y., Zhou, Q., & Fu, X. (2019). An analysis of the evolution, completeness and spatial patterns of OpenStreetMap building data in china. *ISPRS International Journal of Geo-Information*, 8(1), 35. <https://doi.org/10.3390/ijgi8010035>
- Tóth, S., Kovács, S., & Kummer, L. (2008). *Vulnerability analysis in the körös-corner flood area along the middle-tisza river – pilot study application of general vulnerability analysis techniques* (Deliverable report D22.3 T22-08-03). Vituki Hungary. <https://repository.tudelft.nl/islandora/object/uuid:b813a25e-9675-4329-8a7f-35b6bbe0cbc4/datastream/OBJ/download>
- Tsamardinos, I., Aliferis, C. F., & Statnikov, A. (2013). Algorithms for large scale markov blanket discovery. *AAAI Press*, 376–38.
- Ulbrich, U., Brücher, T., Fink, A. H., Leckebusch, G. C., Krüger, A., & Pinto, J. G. (2003). The central european floods of august 2002: Part 2 -synoptic causes and considerations with respect to climatic change. *Weather*, 58(11), 434–442. <https://doi.org/10.1256/wea.61.03B>
- UNISDR. (2015). *Sendai framework for disaster risk reduction 2015 - 2030*. United Nations Office for Disaster Risk Reduction. Geneva. https://www.preventionweb.net/files/43291_sendaiframeworkfordrren.pdf

- United Nations. (2019a). National accounts - analysis of main aggregates (AMA). Retrieved June 15, 2020, from <https://unstats.un.org/unsd/snaama/>
- United Nations. (2019b). World population prospects - population division - united nations. Retrieved June 14, 2020, from <https://population.un.org/wpp/>
- van der Bles, A. M., van der Linden, S., Freeman, A. L. J., Mitchell, J., Galvao, A. B., Zaval, L., & Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. *Royal Society Open Science*, 6(5), 181870. <https://doi.org/10.1098/rsos.181870>
- van Ginkel, K. C. H., Dottori, F., Alfieri, L., Feyen, L., & Koks, E. E. (2021). Flood risk assessment of the european road network. *Natural Hazards and Earth System Sciences*, 21(3), 1011–1027. <https://doi.org/10.5194/nhess-21-1011-2021>
- Vanneuville, W., Maddens, R., Collard, C., Bogaert, P., De Maeyer, P., & Antrop, M. (2006). *Impact op mens en economie t.g.v. overstromingen bekeken in het licht van wijzigende hydraulische condities, omgevingsfactoren en klimatologische omstandigheden* (MIRA/2006/02). Universiteit Gent. www.milieurapport.be
- van Renssen, S. (2013). EU adaptation policy sputters and starts. *Nature Climate Change*, 3(7), 614–615. <https://doi.org/10.1038/nclimate1943>
- van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., & Rose, S. K. (2011). The representative concentration pathways: An overview. *Climatic Change*, 109(1), 5. <https://doi.org/10.1007/s10584-011-0148-z>
- Vargas-Muñoz, J. E., Lobry, S., Falcão, A. X., & Tuia, D. (2019). Correcting rural building annotations in OpenStreetMap using convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, 283–293. <https://doi.org/10.1016/j.isprsjprs.2018.11.010>
- Veijalainen, N., Lotsari, E., Alho, P., Vehviläinen, B., & Käyhkö, J. (2010). National scale assessment of climate change impacts on flooding in finland. *Journal of Hydrology*, 391(3), 333–350. <https://doi.org/10.1016/j.jhydrol.2010.07.035>
- Vlek, C. (1996). A multi-level, multi-stage and multi-attribute perspective on risk assessment, decision-making and risk control. *Risk Decision and Policy*, 1, 9–31. <https://doi.org/10.1080/135753096348772>
- Vogel, K., Riggelsen, C., Korup, O., & Scherbaum, F. (2014). Bayesian network learning for natural hazard analyses. *Natural Hazards and Earth System Sciences*, 14(9), 2605–2626. <https://doi.org/10.5194/nhess-14-2605-2014>
- Vogel, K., Riggelsen, C., Scherbaum, F., Schröter, K., Kreibich, H., & Merz, B. (2013). Challenges for bayesian network learning in a flood damage assessment application. <https://doi.org/10.1201/b16387-452>
- Vogel, K., Weise, L., Schröter, K., & Thielen, A. H. (2018). Identifying driving factors in flood-damaging processes using graphical models. *Water Resources Research*, 54, 1–26. <https://doi.org/10.1029/2018WR022858>
- Vojinovic, Z., Ediriweera, J., & Fikri, A. (2008). An approach to the model-based spatial assessment of damages caused by urban floods.
- Wagenaar, D., De Jong, J., & Bouwer, L. M. (2017). Multi-variable flood damage modelling with limited data using supervised learning approaches. *Natural Hazards and Earth System Sciences*, 17(5), 1683–1696. <https://doi.org/10.5194/nhess-17-1683-2017>
- Wagenaar, D., de Bruijn, K. M., Bouwer, L. M., & de Moel, H. (2016). Uncertainty in flood damage estimates and its potential effect on investment decisions. *Natural Hazards and Earth System Sciences*, 16(1), 1–14. <https://doi.org/https://doi.org/10.5194/nhess-16-1-2016>
- Wagenaar, D., Lüdtke, S., Schröter, K., Bouwer, L. M., & Kreibich, H. (2018). Regional and temporal transferability of multivariable flood damage models. *Water Resources Research*, 54(5), 3688–3703. <https://doi.org/10.1029/2017WR022233>
- Walczak, Z. (2007). Graphics in LATEX using TikZ. 29, 4. <https://tug.org/TUGboat/tb29-1/tb91walczak.pdf>

- Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S., & Bai, X. (2015). Flood hazard risk assessment model based on random forest. *Journal of Hydrology*, 527, 1130–1141. <https://doi.org/10.1016/j.jhydrol.2015.06.008>
- Ward, P. J., Blauhut, V., Bloemendaal, N., Daniell, J. E., de Ruiter, M. C., Duncan, M. J., Emberson, R., Jenkins, S. F., Kirschbaum, D., Kunz, M., Mohr, S., Muis, S., Riddell, G. A., Schäfer, A., Stanley, T., Veldkamp, T. I. E., & Winsemius, H. C. (2020). Review article: Natural hazard risk assessments at the global scale. *Natural Hazards and Earth System Sciences*, 20(4), 1069–1096. <https://doi.org/10.5194/nhess-20-1069-2020>
- Waske, B., & Braun, M. (2009). Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(5), 450–457. <https://doi.org/10.1016/j.isprsjprs.2009.01.003>
- WCRP. (2021). *Data access - how to access the data* [Cordex]. Retrieved September 13, 2021, from <https://cordex.org/data-access/>
- Weise, K., & Woger, W. (1993). A bayesian theory of measurement uncertainty. *Measurement Science and Technology*, 4(1), 1–11. <https://doi.org/10.1088/0957-0233/4/1/001>
- White, G. F. (1936). The limit of economic justification for flood protection. *The Journal of Land and Public Utility Economics*, 12(2), 133–148. <https://doi.org/10.2307/3158294>
- White, G. F. (1945). Human adjustments to floods : A geographical approach to the flood problem in the united states. *Department of Geography, University of Chicago, Research Paper 29*, 225. https://biotech.law.lsu.edu/climate/docs/Human_Adj_Floods_White.pdf
- Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(1), 1–20. <https://doi.org/10.18637/jss.v021.i12>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Willner, S. N., Levermann, A., Zhao, F., & Frieler, K. (2018). Adaptation required to preserve future high-end river flood risk at present levels. *Science Advances*, 4(1), 1–9. <https://doi.org/10.1126/sciadv.aao1914>
- Willows, R., Reynard, N., Meadowcroft, I., & Connell, R. (2003). *Climate adaptation: Risk, uncertainty and decision-making* (Part 2). UK Climate Impacts Programme. Oxford, UK. Retrieved January 2, 2019, from <http://nora.nerc.ac.uk/id/eprint/2969/1/N002969CR.pdf>
- Wilson, E. (2006). Adapting to climate change at the local level: The spatial planning response. *Local Environment*, 11(6), 609–625. <https://doi.org/10.1080/13549830600853635>
- Wind, H. G., Nierop, T. M., Blois, C. J. d., & Kok, J. L. d. (1999). Analysis of flood damages from the 1993 and 1995 meuse floods. *Water Resources Research*, 35(11), 3459–3465. <https://doi.org/10.1029/1999WR900192>
- Wing, O. E. J., Bates, P. D., Neal, J. C., Sampson, C. C., Smith, A. M., Quinn, N., Shustikova, I., Domeneghetti, A., Gilles, D. W., Goska, R., & Krajewski, W. F. (2019). A new automated method for improved flood defense representation in large-scale hydraulic models. *Water Resources Research*, 55(12), 11007–11034. <https://doi.org/10.1029/2019WR025957>
- Winsemius, H. C., Van Beek, L. P. H., Jongman, B., Ward, P. J., & Bouwman, A. (2013). A framework for global river flood risk assessments. *Hydrology and Earth System Sciences*, 17(5), 1871–1892. <https://doi.org/10.5194/hess-17-1871-2013>
- Winter, B., Schneeberger, K., Huttenlau, M., & Stötter, J. (2018). Sources of uncertainty in a probabilistic flood risk model. *Natural Hazards*, 91(2), 431–446. <https://doi.org/10.1007/s11069-017-3135-5>

- WMO. (2021). *The atlas of mortality and economic losses from weather, climate and water extremes (1970–2019)* (WMO- No. 1267) [ISBN: 978-92-63-11267-5]. World Meteorological Organization (WMO). Geneva, Switzerland. Retrieved September 2, 2021, from https://library.wmo.int/doc_num.php?explnum_id=10769
- Wolf, S., Teitge, J., Mielke, J., Schütze, F., & Jaeger, C. (2021). The european green deal — more than climate neutrality. *Intereconomics*, 56(2), 99–107. <https://doi.org/10.1007/s10272-021-0963-z>
- Wüest, M. (2019). Uncertainty quantification by ensemble application in (re-)insurance [personal communication].
- Wünsch, A., Herrmann, U., Kreibich, H., & Thieken, A. H. (2009). The role of disaggregation of asset values in flood loss estimation: A comparison of different modeling approaches at the mulde river, germany. *Environmental Management*, 44(3), 524–541. <https://doi.org/10.1007/s00267-009-9335-3>
- Yaramakala, S., & Margaritis, D. (2005). Speculative markov blanket discovery for optimal feature selection [ISSN: 2374-8486]. *Fifth IEEE International Conference on Data Mining (ICDM'05)*, 1–4. <https://doi.org/10.1109/ICDM.2005.134>
- Yazdi, J., & Neyshabouri, S. A. A. (2012). Optimal design of flood-control multi-reservoir system on a watershed scale. *Natural Hazards*, 63(2), 629–646. <https://doi.org/10.1007/s11069-012-0169-6>
- Zhai, G., Fukuzono, T., & Ikeda, S. (2005). Modeling flood damage: Case of tokai flood 2000. *Journal of the American Water Resources Association*, 41(1), 77–92. <https://doi.org/10.1111/j.1752-1688.2005.tb03719.x>
- Zhou, Q. (2018). Exploring the relationship between density and completeness of urban building data in OpenStreetMap for quality estimation. *International Journal of Geographical Information Science*, 32(2), 257–281. <https://doi.org/10.1080/13658816.2017.1395883>