Development of a physical vulnerability model for floods in data-scarce regions: a case study of Nigeria

Inauguraldissertation

der Philosophisch-naturwissenschaftlichen Fakultät

der Universität Bern

vorgelegt von

Mark, Bawa Malgwi

von Adamawa, Nigeria.

LeiterIn der Arbeit:

Prof. Dr. Margreth Keiler

Geographisches Institut

Universität Bern

Prof. Dr. Andreas Zischg

Geographisches Institut

Universität Bern

Development of a physical vulnerability model for floods in data-scarce regions: a case study of Nigeria

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von
Mark, Bawa Malgwi
von Adamawa, Nigeria.

LeiterIn der Arbeit:
Prof. Dr. Margreth Keiler
Geographisches Institut
Universität Bern

Prof. Dr. Andreas Zischg
Geographisches Institut
Universität Bern

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen.

Bern, Prüfungsdatum

Der Dekan/Die Dekanin
Prof. Dr. Name

Creative Commons Attribution (CC BY 4.0)

This thesis is dedicated to the loving memory of my mother, Florence Malgwi, whose passion for higher learning inspired me; and to my father, Bawa Malgwi, whose book shelf opened my mind to different possibilities.

This is my prayer in the desert
When all that's within me feels dry
This is my prayer in my hunger and need
My God is the God who provides

And this is my prayer in the fire
In weakness or trial or pain
There is a faith proved
Of more worth than gold
So refine me, Lord, through the flame

And this is my prayer in the battle
When triumph is still on its way
I am a conqueror and co-heir with Christ
So firm on His promise I'll stand

This is my prayer in the harvest
When favor and providence flow
I know I'm filled to be emptied again
The seed I've received I will sow

Writer: Brooke Ligertwood

(Extract: Psalms 66)

Abstract

With the observed increase in frequency and magnitude of flood hazard globally, knowledge on characterizing and predicting flood impacts is key for disaster risk reduction. Physical vulnerability assessment is identified as a significant component for assessing the impacts of floods on the built environment. Physical vulnerability assessment is usually carried out with a physical vulnerability model (PVM), which typically requires a large sample size of empirical data describing damage and/or flood characteristics. PVM is used to estimate the relative vulnerability between buildings (vulnerability index) or to predict damage grades or monetary loss (stage-damage curves, multivariate methods). Hence, they show the relationship between flood hazard intensity and/or building characteristics with respect to flood damage. Although PVMs have been applied in several regions, the unavailability of key input variables, namely flood hazard and empirical damage data, has restricted its application in many data-scarce regions. As a result, efforts towards disaster risk reduction have been limited. Given that many data-scarce areas are developing countries that have further limited disaster coping capacities, the importance of developing PVMs to address these specific challenges is a vital step towards gaining a better understanding of future flood impacts.

In Africa, for example, floods accounted for more than 60% of natural disasters between 2000-2019. The impacts of floods, in terms of the increase in fatalities, the number of affected people and the amount of severity of damages have increased. Furthermore, increasing urbanization of floodplains coupled with an increase in climate extremes is expected to intensify flood impacts on the continent. Despite identified flood risks, however, there is limited knowledge about the physical vulnerability of buildings in many African countries to date. Building types such as sandcrete block and clay buildings, which are prevalent in many African countries, remain largely under-investigated due to the lack of data for developing PVMs. As a result, studies in such data-scarce locations have been limited to exposure assessments or to the identification of vulnerability indicators without a systematic linkage established between hazard intensity, building characteristics and damage.

In Nigeria, annually occurring floods persist. The presence of large and small rivers with floodplains inhabited by smaller communities or even urban centres have become a cause for concern. In many cases, buildings in these areas are of low construction quality and are, therefore, even more, susceptible to floods. For instance, in 2012, more than 3.8 million people were displaced after floods affected 28 out of Nigeria's 36 states. Post-disaster assessment showed that in 12 states, over 1.3 million buildings were either partially or completely damaged. In these regions, developing methods to characterize flood scenarios to identify buildings susceptible to significant damages provide an important first step for reduction of future risks.

In light of these challenges and motivation, the objective of this thesis is to develop and test a new PVM that is characterized by a reduced data requirement and the ability to characterize and predict flood impacts on buildings. Two study regions were selected from Nigeria to test the applicability of the developed method. Both regions are located in the central part of the country, where small and large rivers are present and fluvial

flooding is common. The overall objective was further divided into three sub-objectives: i) to review existing methods in order to conceptualize a PVM approach that can be tailored for regional situations in typical data-scarce areas, ii) to develop and test a hydrodynamic modelling approach to reconstruct a past flood scenario in a data-scare location so that the use of modelled flood characteristics may be possible in a PVM, and iii) to test the applicability and performance of the new PVM approach by comparing the results to those of existing methods.

To meet the project objectives, firstly, a review of existing methods for physical vulnerability assessment was carried out from which a new concept for PVM was developed. The concept combines approaches with reduced data requirements to allow for application in data-scarce regions. The concept systematically combines vulnerability indicators, damage grades and synthetic what-if analysis, such that it can be fully implemented using knowledge from regional experts. Secondly, a method for characterizing flood hazard at building locations was developed to increase sample sizes of flood depths required for developing PVMs. The method combines hydrodynamic modelling and interview data to reconstruct a plausible scenario of past floods. The result is the extraction of usable flood depths, which can be extrapolated beyond the spatial extent of collected observations. The method does not require hydrological data, given that these data are usually unavailable in data-scarce locations. Rather, flood depths and duration are utilized through a four-round simulation procedure to minimize the root mean square error (RMSE) between observed and modelled flood depths. The method was tested using 300 spatially distributed flood depth and duration data. An overall RMSE of 0.61 m was achieved, which falls within similar studies using hydrodynamic models and demonstrates the applicability of the method in data-scarce areas. Furthermore, the new PVM (expert-based approach) was implemented and compared with an existing data-driven PVM. The expert-based approach systematically combines the vulnerability indicator method and synthetic what-if analysis based on the knowledge of regional experts. The data-driven method uses a random forest model implemented using empirical data. The comparison was based on variables identified as important drivers of building damage and predictive accuracy. Data from two study regions were used to evaluate model performance. Both methods showed that distance to channel, building material, building condition and building quality are important damage drivers for sandcrete and clay building types. A 30% accuracy that was associated with the expert-based approach was considered satisfactory, given identified data challenges and demonstrates the potential of the method.

The thesis has contributed to knowledge about the physical vulnerability of buildings to floods in typical data-scarce regions. The contributions range from i) developing a PVM tailored for typical data-scarce areas with a transferable and updatable framework to ii) developing a method for increasing the sample size to support physical vulnerability assessments. The thesis presents one of the first findings for sandcrete block and clay buildings in terms of damage grades classification, identification of main damage drivers, and a model for predicting probable damage. The findings will support decision makers with regards to planning for effective disaster risk reduction solutions, such as mitigation and emergency planning in areas that are likely to sustain

significant flood damage. The workflow in the developed methods are transferable and can be updated when new data are available.

Acknowledgments

Firstly, I would like to express my deepest appreciation to Margreth Keiler, my primary supervisor, for her unwavering support and guidance from the concept development of the study to its completion. Thank you for the numerous discussions and feedbacks that have shaped the study to its final form. I am extremely grateful to Andreas Zischg, my co-supervisor, for his ingenious insights, firm support, and encouragement throughout the study period. Moreso, I greatly acknowledge the Swiss Government Excellence Scholarship (ESKAS) who funded the PhD. I sincerely acknowledge all my co-authors, without whom the completion of my thesis would not have been possible. Thanks to Margreth Keiler, Andreas Zischg, Sven Fuchs, Jorge Ramirez, Markus Zimmermann, Stephan Schürmann, and Matthias Schlögl. I deeply appreciate your invaluable contributions, experience, constructive advice, and patience throughout the different research milestones.

A significant part of this research relied on field data collected over a cumulative period of four months through interviews. As a result, I am very grateful to the people of Suleja/Tafa and Wurro Jebbe who warmly welcomed me into their homes and provided me with all the information I needed. These interactions have left me with, in addition to other facts, a deeper realization of the impacts of hazards and the urgency to curtail them. In addition, I wish to thank my field guides, Bala Inuwa and Ambrose, who persevered through very sunny, and sometimes rainy days, to ensure that we realize our daily targets on the field. You guys are the best! I would also like to extend my sincere thanks to all the experts who volunteered to take part in the expert assessment. Thanks to Dr. A. A. Komolafe, Dr. M. K. Kawu, Dr. O. S. Rafiu, Dr. A. Richard, Dr. O. Samsideen, Prof. O. D. Jimoh, Prof. I. O. Adelekan, Dr. O. P. Uchenna, Dr. T. Akukwe, Dr. A. B. Ismail. I am grateful for your time, effort, and research on physical vulnerability, which laid the necessary basis for this study.

Furthermore, I am deeply indebted to my colleagues (now friends) whose technical and moral support for my work has been profound and deserve recognition. I very much appreciate Jorge Ramirez, for valuable contributions, unrelenting support, and guidance in ways too numerous to mention. I sincerely thank Candace Chow whose insightful suggestions, through numerous discussions, were instrumental in conceptualizing the study. Also, I am grateful to Mirjam Mertin, Stephan Schürmann, and Mattia Brughelli for making the office such a pleasant place to be. I am especially thankful to Alexander Groos for insightful feedbacks on different manuscripts and wonderful discussions on varying subjects. I also wish to especially thank Eric Sauvageat for supporting with R programming language on different occasions.

Finally, I am extremely grateful to my family for their love, support, and prayers through the study period. Thanks to my Father, Bawa Malgwi, and my siblings, Kucheli, Christy, and Godiya for always being there. A very special thanks to my Fiancée, Sarah, whose love, patience, and encouragement further allowed me to complete my research. For many other colleagues, acquaintances, and family whom I may not be able to mention, but have supported me in one way or the other, I say a profound thank you. Finally, all glory to God,

who gave me breath, health, strength, and wisdom to pursue this dream up to its accomplishment, I remain in complete awe of you.

Table of contents

Chapter 1: Introduction

1.1 Physical vulnerability to floods

Floods are generally characterized by excess water that flows overland. Floods can be i) coastal, which flows inland from oceans or seas, ii) fluvial, which occurs as a result of overflows from rivers, and iii) pluvial, which is characterized by surface runoff where the infiltration capacity of the soil has been exceeded. The increase in the frequency of floods, commonly attributed to an increase in precipitation as a result of climate change (Hoegh-Guldberg et al., 2018) or increase in urbanization especially in flood risk areas, continues to be a cause for concern. The interaction between floods and the built environment (e.g., residential buildings) represents a potential risk with significant adverse human and economic losses worldwide (CRED, 2019). Flood risk is particularly high for regions with limited capacities to cope with hazard consequences. Limited coping capacities, in these cases, may be related in part or in combination with low-quality housing, lack of structural mitigation, or flood protection measures. Generally, such areas with limited coping capacities are defined as vulnerable.

Vulnerability has been long observed to play an instrumental role in producing differential hazard consequences on exposed communities (James and Hall, 1986). UNISDR (2009) defined vulnerability as the condition that makes communities susceptible to hazards. Although vulnerability has physical, social, economic, environmental dimensions, physical vulnerability has been particularly identified as an initial trigger for other vulnerability dimensions (Fuchs, 2009; WHO, 2009; Papathoma-Köhle et al., 2011). Physical vulnerability relates to i) the pre-conditions that can influence the extent of hazard consequences or ii) the expected probable damage to the built environment when impacted by hazards.

Generally, physical vulnerability assessment is carried out using a physical vulnerability model (PVM). PVMs have been used in three different ways. Firstly, they characterize pre-conditions that influence the vulnerability of a building and provide guidance on identifying relative building vulnerabilities (e.g., using vulnerability indicators). Secondly, they can be used to predict the post-event condition of a building in terms of monetary loss or damage pattern (stage-damage curves). Thirdly, they can be used to characterize pre-conditions that influence damage and predict post-event conditions (multivariate methods). Figure 1 shows the characteristics of different PVMs commonly used. Applications of PVMs include mitigation and emergency planning (Walliman et al., 2011), estimating economic losses (Jongman et al., 2012), cost-benefit analysis of flood protection measures (Holub and Fuchs, 2008), and risk assessment for future system scenarios (Mazzorana et al., 2012).

PVMs used for predicting flood damage, either in terms of monetary losses or damage grades, are generally referred to as flood damage models. Flood damage models either show the relationship between damage and water depths (stage-damage curves) or additionally consider other hazard and/or building characteristics (multivariate models). The ability to predict building damage is an important step towards disaster risk reduction (Merz et al., 2010), and it has received increased attention (Ke et al., 2012).

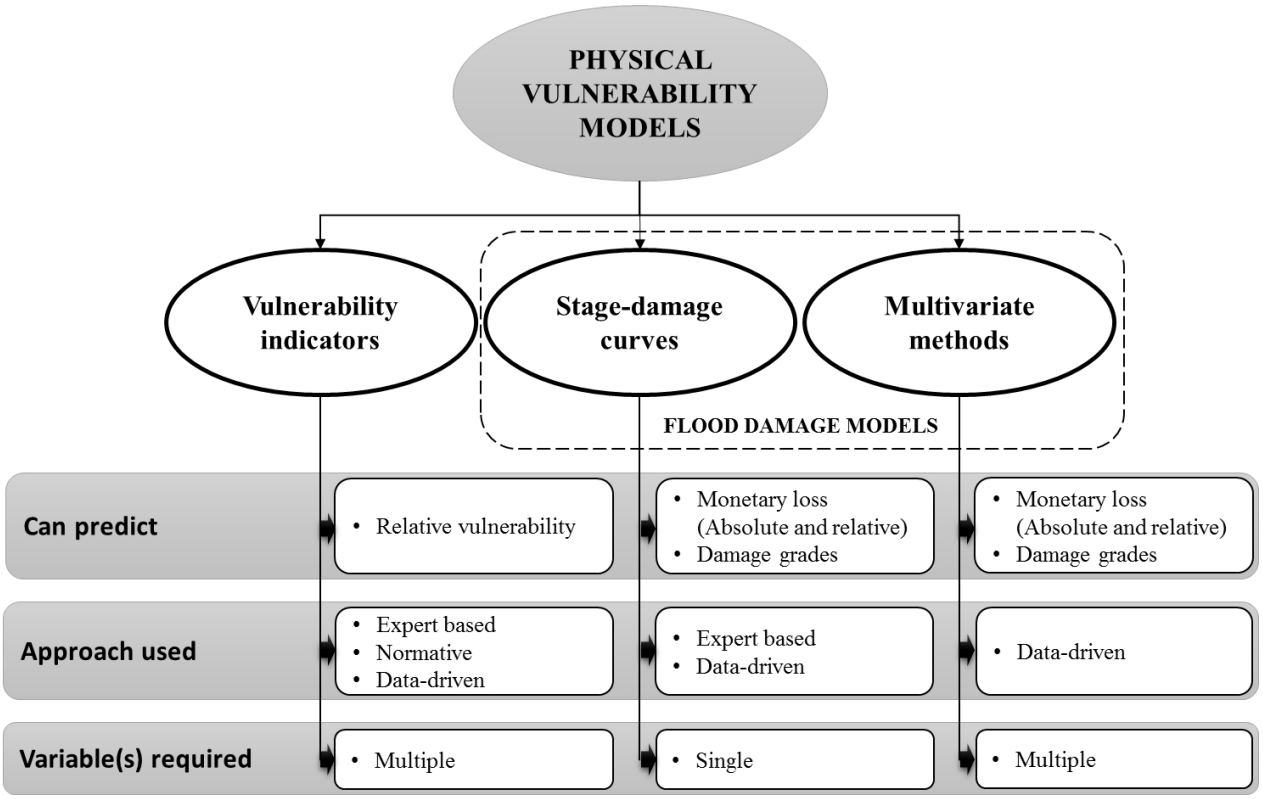


Figure 1: Overview of physical vulnerability models

Flood damage models require a systematic combination of damage influencing variables: these variables have been categorized as impact (action) and resistance variables (Thieken et al., 2005). The impact variable is a primary input for PVM required to characterize hazard (e.g. flood depth and velocity) at building locations: they can be acquired through hydrodynamic modelling. Resistance variables relate to characteristics of the building or the immediate environment which can influence the degree of damage. Resistance variables can be obtained from census data (e.g. building characteristics) or using satellite data (exposure). Furthermore, the application of flood damage models require empirical damage data for performance evaluation: these data are used to quantify the degree of hazard impact on a building. Empirical damage data are often in form of monetary loss (absolute and relative) or damage grades, which are obtained from insurance companies or government compensation reports. While these three data (impact variables, resistance variables, and empirical damage data) are available in some regions, they are either unavailable or very limited in other countries consequently limiting disaster risk reduction efforts.

Within the context of this thesis, data-scarce areas generally refer to regions that lack the required data for developing PVMs. Scarcity in this case pertains to either or a combination of the following three criteria: i) unavailability of data (no records are documented), ii) insufficient data (data are too sparse to be representative or to be used to infer significant statistical deductions), or iii) limited accessibility (data may be available but expensive to acquire). In particular, data that are usually scarce in many regions pertain to

systematically documented records of past flood characteristics and empirical damage data at the micro-(building) scale, both of which are key for developing PVMs.

Over the past decades, while the reported occurrences and impacts of various flood-related hazards in Africa have been on the rise (EM-DAT 2013), the actual number of flood events may have been grossly underestimated in the past (UNISDR, 2011). An increase in the urbanization of flood prone areas and the presence of low-quality housing and infrastructure in many African countries (UNISDR, 2009, Habitat, 2011) have led to an increase in the number and severity of hazard impacts. Consequently, Adelekan et al. (2015) identified populations and assets in African cities to be among the most globally vulnerable. Many of the African countries are data-scarce, as a result, the application of PVMs remains difficult. For example, in a detailed review of flood risk in Nigeria, Komolafe et al. (2015) reported that a comprehensive database for documenting flood hazard data and damage at the building level remains unavailable, either by government or private agencies.

Adelekan et al. (2015) emphasized the need for further research to investigate the vulnerability of communities to floods in Africa. In a project titled "Climate Change and Vulnerability of African Cities (CLUVA)" (Gasparini, 2013), the need for developing physical vulnerability models specifically tailored for representative buildings in Africa (e.g., sandcrete blocks and clay building materials) was emphasized, given increasing flood risk therein. Gasparini (2013) noted that although sandcrete blocks and clay have been identified as common building materials in most African countries, research on damage drivers and flood induced damage patterns remain largely under-investigated.

Nigeria is one of the countries most affected by disasters in Africa; between 1990-2019, Nigeria ranked third in the number of fatalities from natural disasters in the continent (CRED, 2019). In Nigeria, flood disasters are majorly fluvial. The presence of small and large rivers with floodplains that are inhabited by many small communities or larger urban centers have constituted a cause for serious concern. In many of these regions, floods have become yearly occurrences resulting in fatalities and damage (Komolafe et al., 2015). Flood risk in Nigeria, in particular fatalities and losses or building damage, are expected to rise. This is primarily attributed to a rapidly increasing population that is generally linked to urbanization of flood prone areas, in addition to increase in climate extremes. Furthermore, Nigeria is typically data-scarce; a situation observed in many African countries (Gasparini, 2013). Records of past floods and its consequences (losses or damages) at the micro-scale are unavailable. Komolafe et al. (2015) noted that the lack of data is related to the fact that flood insurance is uncommon. In particular, government-issued compensations after the occurrence of disasters is flawed; consequently, the mistrust and lack of dependency on government-led measures have led to more citizen-driven initiatives that take on the full responsibility of repairing affected buildings immediately after flood occurrence (Komolafe et al., 2015).

Given the expected increase in flood risk and unavailability of relevant data to characterize physical vulnerability, Nigeria is selected as a study region for this thesis. In the next section, further details are provided on flood risk in Nigeria and study locations.

1.2 Flood impacts in Nigeria

Although floods have recurred annually in Nigeria, the human and economic losses incurred from the floods in 2012 were the worst in the country's recent history (FGN, 2013; NIHSA, 2013). These floods affected 28 out of Nigeria's 36 states consequently leading to 363 fatalities (NIHSA, 2013) and displaced over 6 million people (accounting for about 4 percent of the country's population in 2012). As a result of these consequences, Nigeria experienced the second-highest number of disaster-induced displacements worldwide in 2012 (Yonetani and Morris, 2013). After the 2012 floods, the federal government of Nigeria, in partnership with several international agencies (e.g. EU, World bank), carried out a detailed post-disaster assessment of incurred losses in 11 of the most affected states within the country. The assessment showed that about 700,000 buildings were totally destroyed and over 400,000 buildings were partially destroyed (FGN, 2013). Consequently, the housing sector accounted for over 44% of overall economic losses which was estimated at 2.6 trillion Naira (US $16.9 billion) (FGN, 2013). Further assessment into the cause of the damage revealed that more than 60% of households acquire their houses through private resources and initiatives and only a few use the services of formal institutions. As a result, many structures were constructed without complying with building codes, which exacerbated the overall damage from the floods.

The 2012 post-flood assessment revealed a rather alarmingly high awareness of potential flood damage and emphasized a need for increased efforts dedicated to risk reduction, particularly targeted to minimize impacts in built environments. In particular, both stakeholders and researchers in Nigeria have stressed the need for a systematic approach to evaluate both the preconditions of buildings before floods and the potential for structural or non-structural damage after floods by Komolafe et al. (2015). These efforts are even more urgently needed, given the expected increase in climate extremes and urbanization of flood-prone regions in Nigeria (CRED, 2019). In the sub-sections, general i) geographical overview of Nigeria and study locations and ii) common building types (sandcrete block and clay) are provided.

1.2.1 Geographical setting and study regions

Nigeria is located in West Africa with the southern part of the country bounded by the Atlantic Ocean (Figure 2). Several inland rivers flow within the country, of which the Benue and Niger rivers are the most extensive. Both major rivers, with sources outside of the country, flow through the north-central part of the country and downstream through the south to the Atlantic Ocean. The floodplains of the inland and two major rivers have become permanent settlements of several urban areas and smaller communities. River floods are regular and occur almost yearly in some locations (Komolafe et al., 2015). Two study regions selected for this study are located in the north-central and north-eastern parts of the country, both of which are part of the guinea

savannah and share similar climatic conditions (Ayanlade, 2009). Study region 1 is located in Suleja and Tafa areas in Niger state and study region 2 is in Wuro-Jebbe, in Adamawa state. These regions are located in the central part of the country, mostly characterized by many rivers; a region that is among the most affected by floods in

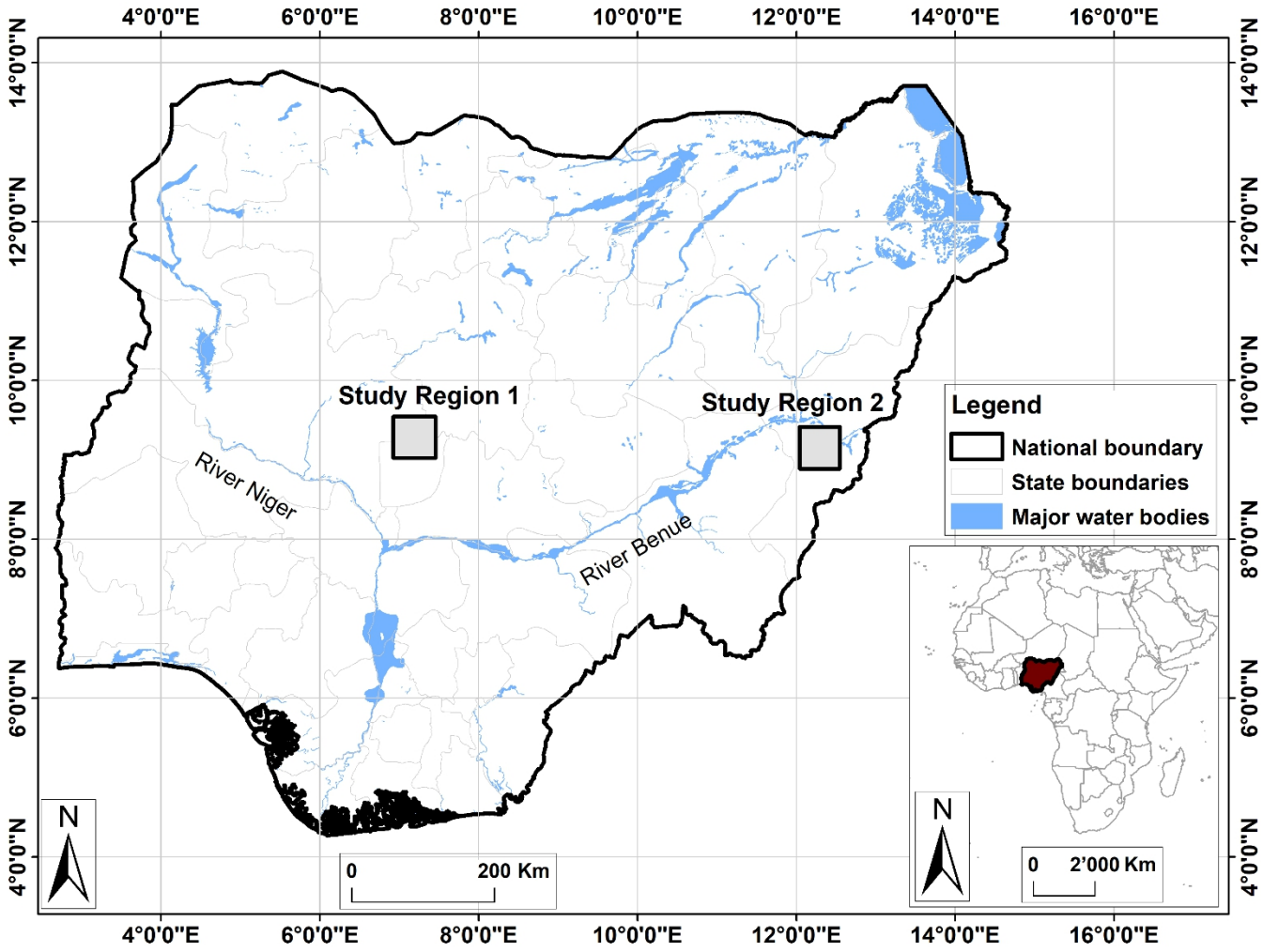


Figure 2: Overview map of Nigeria. The inset map shows the location of Nigeria in Africa.

Nigeria (Mohammed, 2014). The northern part of Nigeria is characterized by two seasons: the rainy season (March to August) and the dry season (November to February) (Ayanlade, 2009). Floods mostly occur towards the end of the rainy season between July – August. Both regions are data-scarce, with no available records of past flood or incurred damage at the micro-(building) scale.

1.2.2 Sandcrete and clay buildings

Common building types in Nigeria can be categorized into three classes: informal, traditional, and modern housing. This classification is primarily based on the wall unit used for construction (FGN, 2013). Informal buildings (e.g. corrugated iron sheets, thatched leaves) mostly serve as temporary shelters and are non-engineered, characterized by standardized load transfer mechanism, so are not discussed further.

Traditional buildings are commonly constructed as either clay or mixed buildings (with sandcrete blocks) by local technicians. They are characterized by a wall thickness of around 15 cm, generally with lightweight materials as roofing elements (e.g. corrugated iron sheets, thatched leaves) and are predominantly one-story structures (i.e with only a ground floor). Clay buildings, especially unburnt clay bricks, are vulnerable to

floods since the material can lose its comprehensive strength when inundated: for instance, Thakur et al. (2012) found a threshold of around 3 days of inundation for clay buildings to collapse. Traditional buildings are predominantly found in rural and suburban areas and makeup about 40% of the building stock in Nigeria (NBS, 2012).

The sandcrete block building is the most prevalent modern building type in Nigeria and is constructed according to the Nigerian building code (NBC, 2006). Sandcrete buildings are made from hollow blocks (i.e. a mixture of cement and sand) with either 15 or 24 cm wide walls. They are predominantly built as single-story structures in Nigeria but can also occur as multiple stories high. Sandcrete blocks make up about 50 % of Nigeria's building stock (NBS, 2012).

1.3 Challenges of PVMs in data-scarce areas

Although some efforts have already been made to support the development of PVMs in data-scarce areas (highlighted in preceding sub-sections and Chapter 1), a number of challenges persist. The development of PVMs relies on sizable and accurate documentation of past flood hazard events and their consequences from which usable deductions can be made. This type of data is generally limited or unavailable in data-scarce regions and constitute a primary challenge that needs to be addressed. Only then can physical vulnerability assessments in these regions be adequately performed. These challenges are briefly elaborated on in the following sub-sections.

1.3.1 Scarcity of empirical data

Firstly, the development of PVMs for predicting building damage (i.e via flood damage modelling) requires that a sizable amount of empirical data be available at the building level. In particular, this data describes damage influencing variables and building damage or monetary loss. However, empirical damage data are usually unavailable in many data-scarce regions. In place of direct data sources, past studies have applied a viable alternative for data-scarce areas by applying what-if analysis, such as synthetic stage-damage curves (Naumann et al., 2009; Neubert et al., 2008; Penning-Rowsell et al., 2005). This approach uses assumed water depths in combination with knowledge from regional experts, to characterize likely flood damage. While the application of this approach is suitable for data-scarce areas, the use of water depth as the only damage driver has been observed to be associated with large uncertainties and reduced prediction accuracies (Merz et al., 2004, Pistrika et al., 2014; Schröter et al., 2014; Fuchs et al., 2013; 2019). A new approach that improves upon the synthetic method would consider multiple damage drivers, while systematically integrating methods with minimal data requirements; the development of such an approach may provide a first step for the use of more accurate results from PVMs in data-scarce areas. While Jongman et al. (2012) and Mancini et al. (2017) have already called for such new approaches that apply best practice from existing methods and are adaptable to regional situations, Papathoma-Köhle et al. (2017) recommended a combination of existing methods to optimize their respective advantages. In effect, a PVM adapted to data-scarce areas will have to balance

empirical data requirements and model uncertainty. In particular, the model should i) allow for the consideration of multiple regional damage influencing variables, in line with multivariate methods and ii) have minimal empirical data requirements. The conceptualization of such a method presents the first challenge in data-scarce areas.

1.3.2 Scarcity of flood hazard data

Furthermore, another basic requirement for PVMs is the integration of reliable estimates of flood characteristics at building locations, where a statistically significant sample size of data is available. The most common variable used to characterize floods is flood depth, which can be acquired through field interviews. However, field interviews may require significant resources such as time and effort to obtain a sufficiently large sample size. The use of hydrodynamic modelling to reconstruct scenarios of past flood events to obtain flood data for PVMs was demonstrated by some studies (e.g. Kreibich and Thieken, 2008; Wagenaar et al., 2017). This approach is more efficient in terms of i) obtaining distributed flood characteristics (flood depth and additionally velocity and duration) with good spatial coverage and ii) offers the possibility to model future scenarios. However, limitations to applying hydrodynamic models persist in data-scarce regions due to the unavailability of i) high-resolution DEMs to accurately characterize floodplain properties and channel bathymetry and ii) hydrological data (precipitation or river discharge measurements). While the former can be obtained at a high cost from satellite companies, the latter require a comprehensive network of stations to install and maintain which are costly, rendering both options to be unsuitable for many data-scarce regions (Sanyal et al., 2013; Yan et al., 2014; Schumann et al., 2015). As a result, another important challenge is to develop an approach to facilitate the extraction of flood characteristics at building locations. The approach should not require hydrological data and be able to use a globally available DEM.

1.4 Research questions and objective

1.4.1 Research questions

In light of the identified challenges for developing PVMs in data-scarce areas, three research questions have been formulated:

1. How can existing PVMs be utilized to develop a new model based on i) reduced requirements for empirical data and ii) that considers multiple damage influencing variables?
2. How can a hydrodynamic model be utilized to reconstruct plausible scenarios of past flood events in areas without hydrological data, such that modelled results have an acceptable accuracy and can be extrapolated to further develop PVMs?
3. How does the new PVM perform against a traditional (existing) PVM using damage prediction accuracy and identified damage drivers as a set of evaluation metrics?

1.4.2 Objective

Based on the known challenges, motivation, and problem formulation, the main objective of the thesis is to develop and test a new PVM for improving the characterization and prediction of floods impacts and that can be applied in typical data-scarce areas. In particular, the new PVM aims to support physical vulnerability assessment in data-scarce areas by evaluating both pre-event conditions and expected post-event building damage states. With this overall objective and research questions in mind, three sub-objectives were developed to serve as project milestones:

1. Review existing methods in order to conceptualize a PVM approach that can be tailored for regional situations in typical data-scarce areas (Chapter 2)
2. Develop and test a hydrodynamic modelling approach to reconstruct a past flood scenario in a data-scare location to allow for the extrapolation of flood characteristics for use in a PVM (Chapter 3)
3. Test the applicability and performance of the new PVM approach by comparing with an existing PVM (Chapter 4)

1.5 Structure of the thesis

The thesis is organized into five chapters. An introduction is presented in Chapter 1 followed by three research articles in Chapters 2, 3, and 4, followed by a final synthesis of the project in Chapter 5. A broader overview of the individual chapters are as follows:

- Chapter 1 presents a general overview of the thesis. It provides a background on physical vulnerability models and physical vulnerability to floods in Nigeria. In addition, it discusses challenges and existing gaps for the application of flood damage models, particularly in data-scarce areas and the specific objectives of the thesis.
- Chapter 2 is a published research article titled “A generic physical vulnerability model for floods: Review and concept for data-scarce regions” (Malgwi et al., 2020). The study reviews existing physical vulnerability assessment methods from which a new conceptual framework was developed for data-scarce regions. The new concept combines approaches with reduced data requirements, such as vulnerability indicators, damage grades, and synthetic what-if analysis. By integrating the vulnerability indicator method, the new framework innovatively provides a multi-variable approach through the recommendation of an aggregated vulnerability index. The vulnerability index facilitated a building vulnerability classification where synthetic what-if analysis results are better targeted to specific vulnerability classes and damage grades than with baseline methods.
- Chapter 3 is a research article under review titled “Flood scenario reconstruction using field interview data and hydrodynamic modelling: A method for data-scarce regions”. The study presents a method that utilizes hydrodynamic modelling and interview data to reconstruct a plausible scenario of past

flood events in a typical data-scarce region. The developed method does not require hydrological data, rather, it utilizes observed distributed flood depths and durations at building locations to implement a four-step simulation procedure that aims to minimize the root mean square error (RMSE) between observed and simulated flood depths. The method was tested using 300 data points collected from study region 1 (see Figure 2). Using this method, flood depths can be extrapolated beyond the spatial extent of the observed (building) locations to increase the sample size for further development of PVMs in data-scarce areas.

- Chapter 4 is a submitted research article titled "Expert-based versus data-driven flood damage models: a comparative evaluation for data-scarce regions". The newly developed PVM (Chapter 1) is here referred to as an expert-based approach. The performance of the expert-based method and a multivariate (data-driven) method using random forest was compared. The comparison between the two methods is to accurately i) predict expected damage grades and ii) identify main damage drivers. Data from both study regions 1 and 2 (see Figure 2) were used for the analysis.
- Chapter 5 presents the synthesis of the thesis. The main findings were discussed with respect to the defined research questions in the first chapter. A conclusion is provided to address how the sub- and overall thesis objectives were addressed. Finally, an outlook for future research is presented.

Acknowledgements

Many thanks to Candace Chow for constructive reviews on the document. Also, special thanks to Jorge Ramirez and Sabrina Erlwein for providing insightful feedback.

References

Ayanlade, A.: Seasonal rainfall variability in Guinea Savanna part of Nigeria: A GIS approach, Int. J. Clim. Chang. Strateg. Manag., 1(3), 282–296, doi:10.1108/17568690910977492, 2009.

CRED, (Centre for Research on the Epidemiology of Disasters): Disasters in Africa: 20 Year Review (2000-2019)., 2019.

FGN, (Federal Government of Nigeria): Nigeria: Post-Disaster needs assessment - 2012 Floods. [online] Available from: https://www.gfdrr.org/sites/gfdrr/files/NIGERIA_PDNA_PRINT_05_29_2013_WEB.pdf (Accessed 1 January 2019), 2013.

Fuchs, S.: Susceptibility versus resilience to mountain hazards in Austria-paradigms of vulnerability revisited., Nat. Hazards Earth Syst. Sci., 9(2), 337–352, doi:10.5194/nhess-9-337-2009, 2009.

Fuchs, S., Keiler, M., Ortlepp, R., Schinke, R. and Papathoma-Köhle, M.: Recent advances in vulnerability assessment for the built environment exposed to torrential hazards: challenges and the way forward, J. Hydrol., 575, 587–595, doi:10.1016/j.jhydrol.2019.05.067, 2019.

Gasparini, P.: Analysis and monitoring of environmental risk: CLUVA Final Report, Clim. Chang. Urban Vulnerability Africa, 1–26 [online] Available from: http://cordis.europa.eu/docs/results/265137/final1-cluva-final-publishable-summary-report.pdf (Accessed 1 January 2019), 2013.

Habitat, U. N.: Global report on human settlements 2011: Cities and climate change, United Nations Hum. Settlements Program, Earthscan, 2011.

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K. L., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R. and Zhou, G.: Special Report on Global Warming of 1.5 °C - Chapter 3: Impacts of 1.5º C global warming on natural and human systems, Glob. Warm. 1.5°C. An IPCC Spec. Rep. impacts Glob. Warm. 1.5°C above pre-industrial levels Relat. Glob. Greenh. gas Emiss. pathways, Context Strength. Glob. response to Threat Clim. Chang., 175–311, doi:10.1002/ejoc.201200111, 2018.

Holub, M. and Fuchs, S.: Benefits of local structural protection to mitigate torrent-related hazards, WIT Trans. Inf. Commun. Technol., 39, 401–411, 2008.

James, L. D. and Hall, B.: Risk information for floodplain management, J. Water Resour. Plan. Manag., 112(4), 485–499, 1986.

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., Gericke, A., Neal, J., Aerts, J. C. J. H. and Ward, P. J.: Comparative flood damage model assessment: Towards a European approach, Nat. Hazards Earth Syst. Sci., 12(12), 3733–3752, doi:10.5194/nhess-12-3733-2012, 2012.

Ke, Q., Jonkman, S. N., Rijcken, T. and Gelder, P. V: Flood damage estimate for downtown shanghai citysensitivity analysis, in Proceedings of the Poster Presentation, the 3rd Conference of the International Society for Integrated Disaster Risk Management, Sept, pp. 7–9., 2012.

Komolafe, A. A., Adegboyega, S. A. A. and Akinluyi, F. O.: A review of flood risk analysis in Nigeria, Am. J. Environ. Sci., 11(3), 157–166, doi:10.3844/ajessp.2015.157.166, 2015.

Kreibich, H. and Thieken, A. H.: Assessment of damage caused by high groundwater inundation, Water Resour. Res., 44(9), 1–14, doi:10.1029/2007WR006621, 2008.

Malgwi, M. B., Fuchs, S. and Keiler, M.: A generic physical vulnerability model for floods: review and concept for data-scarce regions, Nat. Hazards Earth Syst. Sci., 20(7), 2067–2090, doi:10.5194/nhess-20-2067-2020, 2020.

Mancini, M., Lombardi, G., Mattia, S., Oppio, A. and Torrieri, F.: An Integrated Model for Ex-ante Evaluation of Flood Damage to Residential Building BT - Appraisal: From Theory to Practice: Results of SIEV 2015, edited by S. Stanghellini, P. Morano, M. Bottero, and A. Oppio, pp. 157–170, Springer International Publishing, Cham., 2017.

Mazzorana, B., Levaggi, L., Keiler, M. and Fuchs, S.: Towards dynamics in flood risk assessment, Nat. Hazards Earth Syst. Sci., 12(11), 3571–3587, doi:10.5194/nhess-12-3571-2012, 2012.

Merz, B., Kreibich, H., Thieken, A. and Schmidtke, R.: Estimation uncertainty of direct monetary flood damage to buildings, Nat. Hazards Earth Syst. Sci., 4(1), 153–163, doi:10.5194/nhess-4-153-2004, 2004.

Merz, B., Kreibich, H., Schwarze, R. and Thieken, A.: Review article "assessment of economic flood damage," Nat. Hazards Earth Syst. Sci., 10(8), 1697–1724, doi:10.5194/nhess-10-1697-2010, 2010.

Merz, B., Kreibich, H. and Lall, U.: Multi-variate flood damage assessment: A tree-based data-mining approach, Nat. Hazards Earth Syst. Sci., 13(1), 53–64, doi:10.5194/nhess-13-53-2013, 2013.

Mohammed, A.: Rural Hazards and Vulnerability Assessment in the Downstream Sector of Shiroro Dam , Nigeria, Planet@ Risk, 2(6), 370–375, 2014.

Naumann, T., Nikolowski, J. and Sebastian, G.: Synthetic depth-damage functions – A detailed tool for analysing flood resilience of building types, Road Map Towar. a Flood Resilient Urban Environ. Proc. Final Conf. COST Action C, (November), doi:10.1093/CERCOR/8.4.321, 2009.

NBC, (Nigerian Building Code): National Building Code, Federal Republic of Nigeria, 2006.

NBS, (National Bureau of Statistics): Annual Abstract of Statistics , Federal Republic of Nigeria., 2012.

Neubert, M., Naumann, T. and Deilmann, C.: Synthetic water level building damage relationships for GIS-supported flood vulnerability modeling of residential properties, in Flood Risk Management: Research and Practice. Proceedings of the European Conference on Flood Risk Management Research into Practice, FLOODrisk, p. 294., 2008.

NIHSA, 2013: Nigeria Hydrological Services Agency, 2013 Flood Outlook., 2013.

Papathoma-Köhle, M., Kappes, M., Keiler, M. and Glade, T.: Physical vulnerability assessment for alpine hazards: State of the art and future needs, Nat. Hazards, 58(2), 645–680, doi:10.1007/s11069-010-9632-4, 2011.

Papathoma-Köhle, M., Gems, B., Sturm, M. and Fuchs, S.: Matrices, curves and indicators: A review of approaches to assess physical vulnerability to debris flows, Earth-Science Rev., 171, 272–288, doi:10.1016/j.earscirev.2017.06.007, 2017.

Penning-Rowsell, E., Tunstall, S., Tapsell, S., Morris, J., Chatterton, J. and Colin, G.: The benefits of flood and coastal risk management: a manual of assessment techniques, ISBN 19047., 2005.

Pistrika, A., Tsakiris, G. and Nalbantis, I.: Flood Depth-Damage Functions for Built Environment, Environ. Process., 1(4), 553–572, doi:10.1007/s40710-014-0038-2, 2014.

Sanyal, J., Carbonneau, P. and Densmore, A. L.: Hydraulic routing of extreme floods in a large ungauged river and the estimation of associated uncertainties: a case study of the Damodar River, India, Nat. Hazards, 66, 1153–1177, doi:10.1007/s11069-012-0540-7, 2013.

Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F. and Merz, B.: How useful are complex flood damage models?, Water Resour. Res., 50(4), 3378–3395, 2014.

Schumann, G. J. P., Bates, P. D., Neal, J. C. and Andreadis, K. M.: Measuring and Mapping Flood Processes, in Hydro-Meteorological Hazards, Risks, and Disasters, edited by F. S. John, D. B. Giuliano, and P. Paolo, pp. 35–64, Elsevier Inc., 2015.

Thakur, P. K., Maiti, S., Kingma, N. C., Hari Prasad, V., Aggarwal, S. P. and Bhardwaj, A.: Estimation of structural vulnerability for flooding using geospatial tools in the rural area of Orissa, India, Nat. Hazards, 61(2), 501–520, doi:10.1007/s11069-011-9932-3, 2012.

Thieken, A. H., Müller, M., Kreibich, H. and Merz, B.: Flood damage and influencing factors: New insights from the August 2002 flood in Germany, Water Resour. Res., 41(12), 1–16, doi:10.1029/2005WR004177, 2005.

UNISDR, (United Nations International Strategy for Disaster Reduction): Terminology on disaster risk reduction, United Nations International Strategy for Disaster Reduction Geneva, Geneva, Switzerland. [online] Available from: https://www.undrr.org/publication/2009-unisdr-terminology-disaster-risk-reduction (Accessed 15 March 2020), 2009.

UNISDR, (United Nations International Strategy for Disaster Reduction): Revealing Risk, Redefining Development, 2011 Glob. Assess. Rep. Disaster Risk Reduct., i–x [online] Available from: http://www.preventionweb.net/english/hyogo/gar/2011/en/bgdocs/GAR-2011/GAR2011_Report_Prelims.pdf, 2011.

Wagenaar, D., De Jong, J. and Bouwer, L. M.: Multi-variable flood damage modelling with limited data using supervised learning approaches, Nat. Hazards Earth Syst. Sci., 17(9), 1683–1696, doi:10.5194/nhess-17-1683-2017, 2017.

Walliman, N., Ogden, R., Baiche, B., Tagg, A. and Escarameia, M.: Development of a tool to estimate individual building vulnerability to floods, WIT Trans. Ecol. Environ., 155, 1005–1016, doi:10.2495/SC120842, 2011.

WHO, (World Health Orgaonization): WHO Guidelines for indoor air quality: dampness and mould, Copenhagen, Denmark. [online] Available from: https://www.who.int/airpollution/guidelines/dampness-mould/en/ (Accessed 26 February 2020), 2009.

Yan, K., Pappenberger, F. and Solomatine, D. P.: Regional versus physically-based methods for flood inundation modelling in data scarce areas: an application to the Blue Nile, Int. Conf. Hydroinformatics 8-1-2014, 2014.

Yonetani, M. and Morris, T.: Global Estimates 2012: people displaced by disasters, IDMC., 2013.

Chapter 2: A generic physical vulnerability model for floods: review and concept for data-scarce regions

Mark Bawa Malgwi[1,2], Sven Fuchs[3], Margreth Keiler[1,2,4]

[1] University of Bern, Institute of Geography, Hallerstrasse 12, 3012 Bern, Switzerland

[2] University of Bern, Oeschger Centre for Climate Change Research, Hochschulstrasse 6, 3012 Bern, Switzerland

[3] University of Natural Resources and Life Sciences, Institute of Mountain Risk Engineering, Peter-Jordan-Str. 82, 1190 Vienna, Austria

[4]University of Bern, Mobiliar Lab for Natural Risks, Hallerstrasse 12, 3012 Bern, Switzerland

Author contributions: MBM designed the study with the support of MK. MBM was responsible for data collection, analysis, and literature review; MK and SF supported with the literature review and analysis. All authors were jointly involved in manuscript preparation and editing.

Published in *Natural Hazards and Earth System Sciences*

Published on: 31 July, 2020.

DOI: https://doi.org/10.5194/nhess-20-2067-2020

Abstract

The use of different methods for physical flood vulnerability assessment has evolved over time, from traditional single-parameter stage–damage curves to multi-parameter approaches such as multivariate or indicator-based models. However, despite the extensive implementation of these models in flood risk assessment globally, a considerable gap remains in their applicability to data-scarce regions. Considering that these regions are mostly areas with a limited capacity to cope with disasters, there is an essential need for assessing the physical vulnerability of the built environment and contributing to an improvement of flood risk reduction. To close this gap, we propose linking approaches with reduced data requirements, such as vulnerability indicators (integrating major damage drivers) and damage grades (integrating frequently observed damage patterns). First, we present a review of current studies of physical vulnerability indicators and flood damage models comprised of stage–damage curves and the multivariate methods that have been applied to predict damage grades. Second, we propose a new conceptual framework for assessing the physical vulnerability of buildings exposed to flood hazards that has been specifically tailored for use in data-scarce regions. This framework is operationalized in three steps: (i) developing a vulnerability index, (ii) identifying regional damage grades, and (iii) linking resulting index classes with damage patterns, utilizing a synthetic "what-if" analysis. The new framework is a first step for enhancing flood damage prediction to support risk reduction in data-scarce regions. It addresses selected gaps in the literature by extending the application of the vulnerability index for damage grade prediction through the use of a synthetic multi-parameter approach. The framework can be adapted to different data-scarce regions and allows for integrating possible modifications to damage drivers and damage grades.

Keywords: Data-scarce regions, vulnerability indicator, damage grade, flood, building, disaster risk reduction

1 Introduction

The magnitude and frequency of floods and their impact on elements at risk have increased globally (Quevauviller, 2014). Risks associated with floods are especially high for communities with a limited capacity to resist their impacts. Communities with a low resistance to impacts of hazards are often referred to as vulnerable. Although the definition of vulnerability varies in different fields of study, efforts to understand and reduce vulnerability are regarded as important steps for disaster risk reduction (UNISDR, 2015). UNISDR (2009) defined vulnerability as the conditions that make communities susceptible to the impact of hazards. These conditions may be linked to limited access to resources, missing risk transfer mechanisms, and/or poor housing quality if elements at risk are considered. Focusing on the latter, poor housing conditions have been shown to be a key factor if different regions exposed to the same hazard level are compared (Papathoma et al., 2003; Keiler et al., 2006). Although the vulnerability of a community has social, economic, physical, environmental, institutional, and cultural dimensions (Birkmann et al., 2013), these dimensions are all interconnected (Mazzorana et al., 2014). Fuchs (2009) and Papathoma-Köhle et al. (2011) identified physical vulnerability as a primer for other vulnerability dimensions. WHO (2009) also highlighted that there is a strong connection between physical vulnerability and other vulnerability dimensions, pointing out that the disruption of physical elements directly affects social and economic activities within a society. Physical vulnerability assessment supports evaluation of economic losses (Blanco-Vogt and Schanze, 2014), analysis of physical resilience (Papathoma-Köhle et al., 2011), cost–benefit analysis (Holub and Fuchs, 2008), risk assessment for future system scenarios (Mazzorana et al., 2012), and decision-making by stakeholders responsible for hazard protection through, e.g. resource allocation (Fuchs, 2009).

Common approaches used for assessing physical vulnerability to flood hazards include stage–damage curves (vulnerability curves), vulnerability matrices, vulnerability indicators (Papathoma-Köhle et al., 2017), and, more recently, multivariate methods. Stage–damage curves show the relationship between flood depths and the degree of impact (e.g. damage grades, relative or absolute monetary loss). These curves are developed using empirical data or expert knowledge (Merz et al., 2010). The empirical method requires data on flood depths and related building damage patterns or monetary losses after a flood event (Totschnig and Fuchs, 2013). These data allow for searching for suitable curves to correlate flood depths to damage or losses. Synthetic methods are based on a "what-if" analysis derived from expert knowledge to determine expected damage for selected intervals of flood depths (Naumann et al., 2009; Merz et al., 2010; Romali et al., 2015). Multivariate methods statistically deduce relationships between empirical building damage or loss data and multiple damage-influencing parameters.

Generally, both stage–damage curves and multivariate methods are used to predict flood damage. This ability to predict damage is increasingly seen as an important step towards disaster risk reduction (Merz et al., 2010). Stage–damage curves and multivariate methods used for damage prediction are commonly referred to as flood damage models. Most flood damage models are based on empirical damage or monetary loss data (see

reviews by Merz et al., 2010; Jongman et al., 2012; Hammond and Chen, 2015; Gerl et al., 2016). However, due to the scarcity of such data in data-scarce regions, limitations exist in developing these models, and this consequently hinders efforts to reduce disaster risk (Niang et al., 2015). More recently, Englhardt et al. (2019) reemphasized data scarcity as the limiting factor for physical vulnerability assessment in developing countries. A few flood damage models have been developed using a synthetic and expert-based what-if analysis (e.g. Penning-Rowsell et al., 2005; Neubert et al., 2008; Naumann et al., 2009) that aim to reduce the dependency on empirical damage and loss data. However, synthetic approaches often use flood depth as the only damage-influencing parameter, leading to increased uncertainty in damage prediction (Pistrika et al., 2014; Schröter et al., 2014).

Flood damage models have been applied to predict damage grades (e.g. Maiwald and Schwarz, 2015; Ettinger et al., 2016) or the monetary value of such damage (e.g. Thieken et al., 2008; Merz et al., 2013; Fuchs et al., 2019b). Damage grades are more suitable for data-scarce regions, as they represent qualitative descriptions of frequently observed damage patterns within a region (for floods: moisture defects, cracks on supporting walls). As they are not dependent on information about monetary loss (e.g. insurance data), damage grades provide a good basis for damage estimation and enhance the comparability of flood impacts between different flood events, regions, and buildings types (Blong, 2003a). Besides, since damage grades are comparable for similar building types (Maiwald and Schwarz, 2015), they improve the transferability of flood damage models (Wagenaar et al., 2017).

Another approach increasingly used to assess physical vulnerability is based on vulnerability indicators (Barroca et al., 2006; Barnett et al., 2008; Papathoma-Köhle et al., 2017). Several studies have re-emphasized the importance of identifying and understanding vulnerability indicators as a fundamental step in disaster risk reduction (e.g. UNISDR, 2015; Zimmermann and Keiler, 2015; Klein et al., 2019). Vulnerability indicators are based on aggregated variables to communicate the state of a system (e.g. the resistance of a building) and to provide insights in the level to which this system will be impacted by a certain hazard level (Birkmann, 2006). Since the vulnerability indicator approach has a low requirement for empirical damage or loss data, the method has gained increasing popularity in data-scarce regions. In addition, vulnerability indicators supplement the use of stage–damage curves in a way that the overall picture of flood vulnerability becomes clearer. This clarity is achieved by an integration of multiple drivers of vulnerability providing a more holistic perspective of vulnerability-contributing factors.

Papathoma-Köhle et al. (2017) recommended a combination of physical vulnerability assessment methods to take advantage of their individual strengths while minimizing their weaknesses. A combination of methods here refers to the integration of approaches (or techniques) from two different physical vulnerability assessment methods into one method (or model). Such a combination of methods that utilize expert-based approaches in place of data-driven methods might provide a desirable compromise for data-scarce regions. For example, Godfrey et al. (2015), using Romania as a case study, combined an approach based on vulnerability

indicators and an approach based on stage–damage curves to develop an expert-based model for data-scarce regions. However, wider applications of the method have shown to be restricted to regions where stage–damage curves for specific building types already exist. In addition, because of a limited sample size used to test the method, results may be biased (Godfrey et al., 2015).

Only little is known at this point about the flood vulnerability and damage mechanisms of buildings that are exposed in developing countries, such as in Africa. Adelekan et al. (2015) identified population and assets in African cities as being among the most vulnerable globally. Consequently, with climate change, the number of extreme events and catastrophic impacts in these regions is expected to increase (Mirza, 2003). In Africa in particular, the need to develop a systematic approach for evaluating preconditions of buildings and their impact by flood hazards has been stressed by stakeholders and researchers (Komolafe et al., 2015). Although sandcrete block and clay buildings are the most predominant building types in many African countries (Gasparini, 2013), flood damage models remain underdeveloped for such building types (Komolafe et al., 2015). Commonly, exposure and vulnerability are mainly assessed in a regional context based on very coarse data and aggregated land use classes, resulting in considerable uncertainties, especially in a rural context (de Moel et al., 2015). Thus, along with recent studies addressing flood exposure and vulnerability in data-scarce areas, there is a strong need to refine approaches for vulnerability and risk assessments in such regions.

Approaches using damage grades and/or vulnerability indicators are in general more suitable for data-scarce areas; however, so far there is a gap in systematically linking them. This paper aims to develop a conceptual framework for assessing the vulnerability of the built environment to floods in data-scarce areas. To do this, we first provide a review of physical vulnerability indicators for flood hazards, as well as an overview of flood damage models. Second, we develop a conceptual framework that links physical vulnerability indicators and flood damage grades by utilizing local expert knowledge.

This paper is structured as follows: Sect. 2 provides an overview of available information on vulnerability indicators, including indicator selection, aggregation, and weighting, and unveils challenges and gaps of using this method. In Sect. 3, a brief review of flood damage models is presented with a particular focus on the use of damage grades and associated challenges. While Sect. 4 addresses the need for linking vulnerability indicators and damage grades, Sect. 5 introduces the conceptual framework for such linkages, as well as the steps for operationalizing the framework. Concluding remarks are presented in Sect. 6.

2 Review of indicators for physical vulnerability to floods

In this section, we present an overview of different studies using indicators to assess the vulnerability of buildings to flood hazards (for details, see Table A1 in the Appendix).

2.1 Background

A vulnerability index is obtained by selecting, weighting, and aggregating vulnerability indicators. Generally, a vulnerability indicator is a parameter (or variable) that can influence and(or) communicate the degree of damage (or loss) of a system (e.g. a building). The indicator approach aims to simplify a concept through the use of an index (Heink and Kowarik, 2010; Hinkel, 2011). Before establishing an index, a framework should be developed to address how major components of the indicator fit together (Birkmann, 2006; JRC and OECD, 2008). Moreover, the framework of such an index should allow adaption to possible future system changes such that it can be used to analyse potential disaster risk. Such adaptation may include possible changes in selected indicators or indicator weights. The framework includes a variety of elements (we refer to these as indicator elements), which helps to clearly outline the extent of applicability and validity of the derived index. Basic elements defining the framework of a vulnerability index include the aim, the vulnerability dimension, the spatial scale, and the region of application (see Table A1).

A first step in developing a framework for indicators is to define the aim, including the different vulnerability dimensions to be assessed, so that the indicators and the final derived index fit into the overall risk assessment framework. Although some studies focus on one specific dimension of vulnerability (e.g. Dall'Osso et al., 2009), other studies examine multiple dimensions of vulnerability (e.g. Kienberger et al., 2009). The interaction between different vulnerability dimensions generates challenges for assessing vulnerability, as well as the use of a high number of indicators in multidimensional studies (Cutter and Finch, 2008). Birkmann (2006) noted that choosing a multidimensional study design is only worth the effort if data are available in a certain quality and quantity, which in turn have to meet the scale requirements of the study (Birkmann, 2007; Aubrecht et al., 2013; Fuchs et al., 2013; Kundzewicz et al., 2019). Consequently, the spatial scale for applying a vulnerability indicator approach varies depending on the availability of data (de Ruiter et al., 2017) and the aim of the assessment. Spatial scales for assessing vulnerability can be on micro, meso or macro level. Microscale assessment is usually challenging in terms of data collection (Günther, 2006), in particular in developing countries with missing metadata on land use, exposure, and population. Microscale assessments can provide an overview of vulnerability (hotspot assessment) over a larger area; hence, decision-makers can use them in allocating resources for emergency responses or risk mitigation. Other indicators operate on a larger scale, i.e. mesoscale (regional to national) and macroscale (international). Moreover, as vulnerability indicators are adaptive to a regional context, a set of indicators selected for a particular region may not necessarily be transferable to another region (Papathoma-Köhle et al., 2017, 2019).

2.2 Application of physical vulnerability indicators

Commonly applied steps, corresponding outputs, and methods for constructing a physical vulnerability index are presented in Fig. 1. Different methods used in deriving the index include deductive (based on theories and/or basic assumptions), inductive (based on empirical data), and normative (based on value judgement) approaches. In physical vulnerability assessments for flood hazards using vulnerability indicators, the

deductive approach is the most commonly applied method, relying on expert judgement and information provided in the relevant scientific literature without any further empirical data. It is also common to use a combination of inductive and deductive approaches either during the indicator selection or during indicator weighting and aggregation. Table 1 shows different studies that derived a physical vulnerability index to assess flood hazards and the various methods employed. Since our attention is on data-scare regions, further discussions in this section will be focused on the deductive and normative approaches, since they do not rely on empirical data.

2.2.1 Indicator selection

Before a variable is qualified as an indicator, certain criteria have to be met to allow for consistency and methodical soundness. Important criteria for selecting a variable as an indicator include measurability, relevance, analytical and statistical soundness, etc. (see Birkmann, 2006, and JRC and OECD, 2008, for a complete list of criteria for indicator selection). The selection of vulnerability indicators can be categorized into two steps (cf. Table 1). In a preliminary step, an initial selection of a range of identified variables is carried out. This serves to identify all possible parameters that influence vulnerability within a region. As shown in Fig. 1, the preliminary selection is commonly carried out either using a deductive or normative approach. In the final step, the number of variables to be used for weighting or aggregation is reduced. The final selection can be based on data availability, statistical analysis, expert opinion, or other evaluation procedures. For example, Kienberger et al. (2009) reported a spatial vulnerability assessment tool using the indicator approach. In their study, expert knowledge was used for the preliminary selection of indicators. Thereafter, based on structured rounds of questionnaire evaluation, a final selection was made based on a Delphi approach. The Delphi approach utilizes several indicator suggestions from different experts and combines the suggestions after a consensus is reached through several rounds of questionnaire exchange. During the Delphi process, preselected indicators that are identified to be less relevant are removed in order to arrive at a set of more effective indicators.

2.2.2 Indicator weighting

After the selection of indicators, weights are assigned to allocate the extent to which each indicator is relevant with respect to the targeted vulnerability assessment. Prior to assigning weights to different indicators, a scoring is assigned for subcategories of indicators, for example, “building type” as an indicator can have “reinforced concrete”, “masonry” and “wooden” buildings as subcategories: we refer to these subcategories as sub-indicators. The scoring of these sub-indicators, which is a form of internal weighting, results in information of the vulnerability of the individual indicator. Both the scoring of sub-indicators and the weighting of indicators can be carried out using (i) deductive and (ii) normative approaches. Other approaches for weighting vulnerability indicators are the inductive methods based on empirical data (e.g. principal component analysis, PCA), which are excluded due to the data requirements in data-scarce regions. More details of the PCA are given in JRC and OECD (2008).

Table 1: Applications of physical vulnerability indicators, including methods used for variable selection, weighting, and aggregation, and parameters needed for assessing physical (building) vulnerability. AHP stands for analytical hierarchy process, and PCA stands for principal component analysis.

Author(s)	**Variable selection**		**Variable weighting**		**Vulnerability aggregation**	**Parameters considered (pertaining to building vulnerability)**
	Preliminary	**Final (used in model or equation)**	**Approach**	**Consideration for scoring/weighting**		
Balica et al. (2009)	Literature	Experts	Equal (no) weights	Conditions that induce flood damage	Direct additive method	Flood depth, duration, velocity and return period, proximity to the river, land use, topography (slope), building codes
Kienberger et al. (2009)	Experts	Experts (Delphi approach)	AHP	Relative importance and contribution to the vulnerability of people	Weighted additive method	Buildings, infrastructure (transportation system), land cover
Müller et al. (2011)	Literature, field survey, experts	Experts	Expert knowledge, household surveys	Relevance of selected variables with respect to flood risk	Weighted additive method	Material for roof, walls, and floor, the position of building in relative to the street level, the proportion of green spaces per building block, flood protection measures
Kappes et al. (2012)	Literature	Experts	Expert appraisals	Ability of the building to withstand the impact of the process	Weighted additive method	Building type, building use, building condition (using age and maintenance), building material, number of floors, row towards the river, trees towards the river
Thouret et al. (2014)	Literature, experts	Experts	Equal weights, experts, PCA	Weakness relative to a given hazard magnitude	Direct additive method	Heterogeneity of city block (using building size and use), building type (height and number of story, construction material, roof type, and building condition), the shape of the city block, building density
Blanco-Vogt and Schanze (2014)	Literature, experts	Literature, experts	Literature, experts	General resistance characteristics after flooding (biological, chemical and material)	Weighted additive method	Building height, size, elongatedness (height/width ratio), building compactness, adjacency, roof, slabs, external fenestration, external wall, floor

Godfrey et al. (2015)	Literature, experts	Experts	AHP	Based on hazard impact	Normalized weighted linear combination	Floor height, number of floors, structural type, building size, wall material, presence of basement, number of openings, quality of construction, building maintenance, protection wall
Behanzin et al. (2015)	Literature	Experts	Equal (no) weights	-	Direct additive method	Building material, roof material, floor material, land cover around the building
Akukwe and Ogbodo (2015)	Literature	PCA	PCA	Significance in explaining the variance in indicator data set	Weighted additive method	Building material, proximity to water, flood depth, flood frequency
Fernandez et al. (2016)	Literature	PCA	No weights and PCA	Significance in explaining the variance in the data set	Direct and weighted additive method	Building density, number of floors, construction period, building material
Ntajal et al. (2016)	Literature, experts	Experts	Equal (no) weights	-	Direct additive method	Distance to the river, flood depth, flood duration, building and roof material, land cover (the area around the building)
Krellenberg and Welz (2017)	Literature, experts	Experts	Equal (no) weights	Probability to be exposed under certain socio-environmental conditions	Direct additive method	Building quality, building structure, protection wall, trees in foreyard, roof form, land cover, housing condition
Sadeghi-Pouya et al. (2017)	Literature, experts	Experts	Experts (scoring)	Variable influence on vulnerability	Direct additive method	Building quality (material), building age, number of floors, land use
Carlier et al. (2018)	Literature	Literature, experts	Experts	Total consequence of a natural hazard on an element at risk	Weighted additive method	Building material, building condition, building age, building function, opening in hazard direction, building in the area affected by flood (recurrence interval), land cover

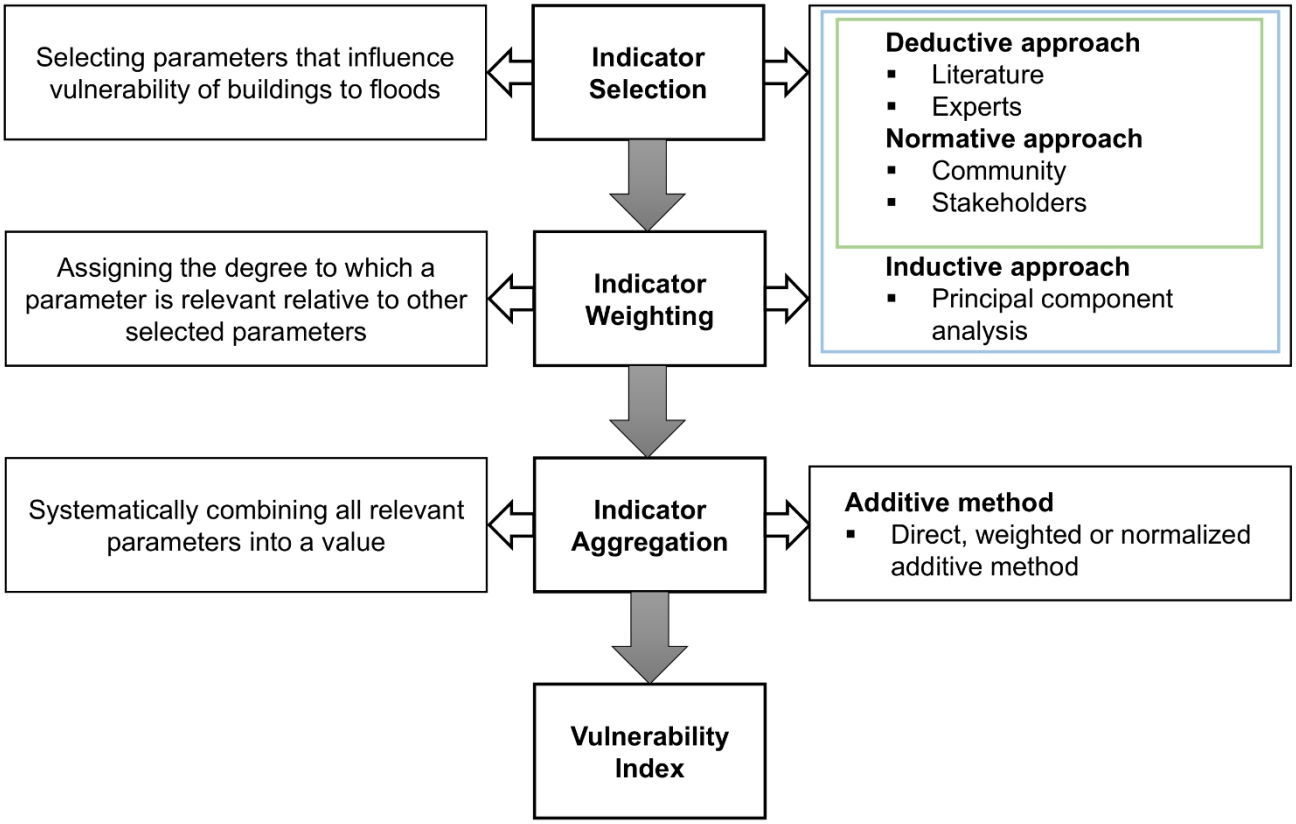


Figure 1: Steps and commonly applied methods for developing a physical flood vulnerability index. Steps include the indicator selection, the indicator weighting, and the indicator aggregation. The green box applies to initial indicator selection, and the blue box applies to final indicator selection.

The deductive approach is based on research-based knowledge and conclusions of previous studies. The weighting is based on deduction or inference from frameworks, a set of concepts, or theories on vulnerability (Hinkel, 2011). Commonly applied deductive weighting includes direct expert weights, expert weights in combination with literature analysis and the application of an analytical hierarchy process (AHP) from expert knowledge.

Direct expert weights refer to weights assigned to indicators using the knowledge of experts either through questionnaires or interviews. A scheme of standardized weights (e.g. from 0 to 10) is provided for the weighting in order to maintain a comparable scale of weights from different experts. Some vulnerability studies therefore used weights from the literature in combination with expert knowledge to formulate new indicator weights. However, this is only possible if (i) the vulnerability of the region of interest has been previously studied or (ii) the region of interest is comparable (in building and hazard characteristics) to a previously studied region. Another commonly applied weighting method for physical vulnerability assessments is based on the analytical hierarchy process (AHP), a multi-criteria decision tool utilizing a pairwise comparison system (Saaty, 1980). The AHP assigns weights between pairs of indicators instead of evaluating each indicator relative to all other indicators. The pairwise comparison evaluates which indicator in

every pair is more important than the other using a scale of 1 (equal importance) to 9 (extreme importance) (Chen et al., 2012). The decision on which indicator is more important can be evaluated from analysing data or expert knowledge. To ensure minimal subjectivity in a pairwise comparison, the consistency ratio (CR) is computed. The CR checks if the subjectivity of pairwise comparisons are within an allowable limit. If the condition of the CR is not fulfilled, a repetition of the process has to be carried out (Golz, 2016). Depending on the total number of indicators, the AHP can be computationally demanding.

Another form of weighting that is not very common in physical vulnerability assessments is the normative approach. Using the normative approach, weights can be assigned based on value judgement (Hinkel, 2011). The normative approach is based on the priorities of individuals. A common application of the normative approach is the equal-weighting approach. Meaning, based on value judgement, all parameters influencing vulnerability are taken to be equally important (Frazier et al., 2014). Adopting an equal-weights approach is sometimes required in cases where no consensus is reached on a suitable weighting alternative. In studies where multiple dimensions of vulnerability are considered, the equal-weights approach will favour dimensions with a higher number of indicators if an unequal number of indicators are used. However, such irregularities can be corrected by a systematic normalization. Furthermore, Chen et al. (2012) noted that the equal-weighting approach cannot properly handle indicators that are highly correlated because these are double-counted. Another implication of the approach, particularly at the aggregation step, was noted by Hinkel (2011): equal weighting means all indicators are ideal replacements of each other, and low values in one indicator can be compensated by high values in another indicator. Another example of the use of value judgement for weighting indicators was demonstrated by Müller et al. (2011), who focused on weighting preferences of homeowners.

2.2.3 Indicator aggregation

Indicator aggregation refers to a systematic combination (or joining) of indicator weights to create a single value. This value is usually referred to as an index. The index carries information on the extent to which an element can be impacted by a hazard relative to other elements, given the combined influence of selected indicators. Physical vulnerability assessment incorporates different types of indicators with non-uniform units, such as building material (no unit) and distance to the hazard source (metres). Therefore, before aggregating indicators, it is necessary to find a systematic and consistent means of representing the (sub-)indicators while retaining their theoretical range. Achieving a rather objective representation of different indicators is carried out by scaling. Asadzadeh et al. (2017) noted that the scaling of indicators is sensitive to the normalization and aggregation method; hence, it is important to adopt a scaling that fits the data and the overall vulnerability framework. In a physical vulnerability assessment, it is common to adopt the ordinal scale to represent both qualitative or quantitative (sub-)indicators. On the ordinal scale, indicators are represented using an increasing or decreasing categorical order. The order selected is mostly subjective depending on the indicator framework and data property (JRC and OECD, 2008). A good example of the use of the ordinal scale was demonstrated

by Dall'Osso et al. (2009), where five categories were used to transform all (sub-)indicators into an ordinal scale.

Generally, several methods for indicator aggregation exist; however, a commonly applied method for physical vulnerability assessment is the additive method (see Table 1). This method is based on a summation of the product of the weights and scores (or the scaled value) of all selected indicators. The summation can be made directly on scores of the indicators (direct additive method) or after multiplying weights and the scores of the indicators (weighted additive method). The result of the indicator aggregation is influenced by the applied aggregation technique as some approaches allow counterbalancing indicators with low values (compensation). In the additive method, a constant level of compensation for indicators with lower values is allowed (JRC and OECD, 2008).

The last step in aggregating indicators is a normalization that ensures that the output from indicator aggregation lies within defined intervals. These intervals should be suitable to communicate the extent to which an element at risk is vulnerable relative to others. JRC and OECD (2008) pointed out that the choice of a normalization approach should be related to data properties and underlying theoretical frameworks. Although there are several normalization techniques, most studies in physical flood vulnerability assessment apply the minimum–maximum normalization. In the minimum–maximum normalization, index outputs are bound within a fixed range, commonly between 0 (not vulnerable) and 1 (highly vulnerable). The minimum–maximum normalization can increase the range of small-interval indicators or reduce the range of large-interval indicators. Hence, all indicators are allowed a proportionate effect on the aggregated index. Detailed descriptions of different normalization methods can be found in JRC and OECD (2008).

2.3 Challenges and gaps in physical vulnerability indicators and indices

Despite the current success in the development of physical vulnerability indicators, a few challenges persist. We identify these challenges for physical vulnerability indicators by focusing on the potential for developing indicator approaches in data-scarce regions and in order to foster adaptability, transferability, and harmonization of indicators across spatial and temporal scales.

Firstly, for the effective operationalization of an index in the vulnerability concept, there is a need for proper management of the underlying data. In many studies, data transformation methods (e.g. missing data, scaling and normalization) are either not mentioned or only briefly highlighted. Such data operations influence the index or model output considerably, as has already been demonstrated by several studies (e.g. UNDP, 1992; Tate, 2012; Mosimann et al., 2018; Chow et al., 2019), and thus data operations should be carried out using appropriate methods that fit the data type and indicator framework. During the indicator development, the following few points have to be clarified: (i) the relationship between indicators, (ii) scaling and normalization needed, (iii) necessary range of variables, and (iv) data quality and quantity.

Secondly, it is important to understand the sensitivity of the vulnerability index depending on the use of deductive, inductive, and normative approaches. So far, no detailed sensitivity analysis has been carried out focusing on physical vulnerability indicators, except for Fernandez et al. (2016), who have taken the first steps by analysing the sensitivity to different aggregation methods. JRC and OECD (2008), Tate (2012), and Papathoma-Köhle et al. (2019) have stressed the need for such internal validation to assess the robustness of indices and evaluate the influence of each approach on the index stability. Such analysis can convey information on the suitability of different approaches for specific datasets, and hence provide useful guidance for further indicator development.

Furthermore, after developing the vulnerability index, it is important to assess how well the index performs by using hazard impact metrics such as building damage or monetary loss data. However, in physical vulnerability assessment, index performance evaluations have only rarely been carried out (Eriksen and Kelly, 2007; Müller et al., 2011). A performance test will allow robust evaluation of underlying indicator frameworks and basic assumptions (Eddy et al., 2012) and will also identify the suitability of selected indicators with respect to the indicator aim (Birkmann, 2006). Few studies, however, provide a qualitative description (e.g. level of agreement), as is the case for performance analysis using a comparison of the deduced index and observed damage data (e.g. Godfrey et al., 2015; Sadeghi-Pouya et al., 2017) or based on visualizations of the spatial agreement using GIS maps by comparing hotspots and observed damage (e.g. Fernandez et al., 2016). In general, a lack of a performance test might be due to (i) the scarcity of empirical data and (ii) the lack of a systematic linkage between the vulnerability index and building damage or monetary loss.

Furthermore, vulnerability indices have been identified to lack a stand-alone meaning outside a relative comparison of building vulnerability (Tarbotton et al., 2012; Dall'Osso and Dominey-Howes, 2013). This is a major limitation given the quality of information contained in the vulnerability index. Further investigation on the additional applicability of the vulnerability index should be carried out. Papathoma-Köhle et al. (2017) recently recommended a combination of methods to fully explore the potential in individual vulnerability assessment methods. Such a combination is particularly encouraged for data-scarce regions.

3 Review of flood damage models

Flood damage models show the relationship between the extent of building damage and damage-influencing (or vulnerability-influencing) factors. First, we focus on an analysis of background information of flood damage models and the application and used methods. Second, we will identify the challenges and current gaps in the context of data-scarce regions.

3.1 Background

Flood damage models provide the basis for decision-making through multiple applications, such as cost-benefit analyses of mitigation measures (Thieken et al., 2005; Schröter et al., 2014), economic impact

assessments (Jongman et al., 2012), planning and implementation of individual mitigation measures (Walliman et al., 2011), and flood risk mapping (Meyer et al., 2012). In general, developing flood damage models require clear communication of model parameters, e.g. if the model is based on an individual damage parameter (stage–damage curves) or if the model is comprised of multiple damage-influencing parameters (multivariate methods). Further important information includes data source and sample size, method of analysis to extract the significance of variables, the scale of application, damage-influencing parameters, and status of the validation or performance test. The different choice of parameters and methods considered within the flood damage models already sets the conditions regarding the model transferability and guide further model development. In Table 2, we highlight these parameters for several studies.

Stage–damage curves are continuous curves relating to the magnitude of a hazard process (x axis) to the damage state of a building (y axis), usually expressed as the degree of loss (Fuchs et al., 2019a). Individual buildings are represented as points in the x–y-axis system and then the function that ensures the best fit may be chosen (Totschnig et al., 2011). Empirically developed stage–damage curves are widely used for assessing flood hazard risk where the number of affected buildings is large enough to deduce a reliable curve (Fuchs et al., 2019a). The shape of the empirically derived stage–damage function depends on the population and spread of data related to buildings within the inundation area under consideration, as well as the type of function chosen. Synthetic stage–damage curves are based on expert knowledge to describe a relationship between flood damage with flood depth for a specific building or land use type. Synthetic curves can be developed independently (e.g. Penning-Rowsell et al., 2005; Neubert et al., 2008; Naumann et al., 2009) or supported by empirical data (e.g. NRE, 2000). For data-scarce regions, utilizing the synthetic (what-if) analysis can serve as an important first step for establishing flood damage models. More details on the synthetic what-if analysis are given by Penning-Rowsell et al. (2005), Neubert et al. (2008), and Naumann et al. (2009).

Multivariate methods utilize empirical data to relate multiple damage-influencing variables and building damage by applying a variety of statistical methods (see Table 2). Such empirical data can be collected from insurance companies (e.g. Chow et al., 2019), through field surveys (e.g. Ettinger et al., 2016), or via telephone interviews (Thieken et al., 2005; Schwarz and Maiwald, 2008; Maiwald and Schwarz, 2015). As demonstrated by Cervone et al. (2016), empirical data can also be collected using social media accounts. Multivariate models may become more common in the near future since they offer a more comprehensive approach compared to the stage–damage curves. Schröter et al. (2014) evaluated the applicability of flood damage models and showed that models that consider a higher number of damage-influencing variables demonstrated superiority in predictive power both spatially (transfer to other regions) and temporally (different flood events). The multivariate method has been shown to better explain the variability in damage data (Merz et al., 2004) and reduce uncertainty in flood damage prediction (Schröter et al. 2014).

3.2 Application of flood damage models

The applications of both stage–damage curves and multivariate methods vary depending on the user requirements. These user requirements may range from estimating damage grades (e.g. Ettinger et al., 2016), estimating absolute or relative monetary loss (e.g. Thieken et al., 2008), or both (e.g. Maiwald and Schwarz, 2015). In particular, the use of damage grades is especially encouraged for data-scarce regions since it relies

Table 2: Applications of flood damage models, indicating the data source, the approach for evaluating variable significance, the scale of application, the parameters needed for developing the vulnerability function, and, where appropriate, the validation or performance test (PCA stands for principal component analysis).

Author	Case study/ region of application	Study aim	Data source for physical vulnerability indicators	Variable significance	Scale of application	Sample size	Parameters considered for developing the vulnerability function (pertaining to physical -building vulnerability)	Validation or performance test
Thieken et al. (2005)	Germany	Investigation of flood damage and influencing factors	Computer-aided phone interviews	PCA and quantile classification	Micro-scale (individual building)	1697	Flood depth, duration and velocity, contamination, precautionary measures, building type, building size, building quality	
Thieken et al. (2008)	Germany	Develop a model for flood loss (direct monetary) estimation for private sector	Computer-aided phone interviews	(Multi)factor analysis	Micro-scale (individual building) and meso-scale (regional)	1697	Flood depth, building type (occupancy), building quality, precaution, contamination	Using a different data set
Vogel et al. (2012)	Germany	Flood damage assessment of residential buildings	Computer-aided phone interviews	Bayesian network	Micro-scale (individual building)	1135	Flood depth, velocity and duration, contamination, return period, precautionary measures, building type (occupancy), building size (floor space), building value, number of flats in a building	Using a subset of training data (bootstrap samples)
Merz et al. (2013)	Germany	Develop tree-based damage prediction models and compare their performance to established models	Computer-aided phone interviews	Regression trees and bagging decision trees	Micro-scale (individual building)	1103	Flood depth, velocity and duration, contamination, return period, precautionary measures, building type (occupancy), building size (floor space), building quality	Using a subset of training data

Spekkers et al. (2014)	The Netherlands	Investigate damage-influencing factors and their relationships with rainfall-related damage	Insurance data data and data from government agencies	Poisson (decision) trees	Meso-scale (district)		Rainfall (intensity, volume, and duration) related variables, building age, ground floor area, real estate value	Using a subset of training data
Ettinger et al. (2016)	Peru	Analysis of building vulnerability	Field Survey and analysis of high spatial resolution images	Logistic regression	Micro-scale (individual buildings)	898	Distance from the channel, distance from bridge, shape of city block, structural building type (material), building footprint	Using a subset of training data
Maiwald and Schwarz (2015)	Germany	Develop engineering vulnerability-oriented for damage and loss prediction	Questionnaire survey, computer-aided phone interviews, evaluation of damage reports, flood simulation	Tangent hyperbolic (damage grade) and an exponential function (relative loss)	Micro-scale (individual building) and meso-scale (regional)		Flood depth and velocity, specific energy (flood depth, velocity, and acceleration due to gravity), building type, presence of basement, building location with respect to flow direction, number of stories	Using a different data set
Wagenaar et al. (2017)	The Netherlands	Prediction of absolute (monetary value for content and structural) flood damage	Experts, flood simulation, cadastre information	Bagging trees	Micro-scale (individual buildings)	4398	Damage data (content and structural), flood depth, duration and velocity, building footprint, return period, building age, building area (footprint, living), basement, detached house	Using a 'withheld' part of the data set

only on observable damage patterns within a region and expert knowledge. In addition, damage grades are easily understandable by experts and non-experts, making them easy communication tools (Attems et al., 2020a). One of the most prominent damage grades is the European Macroseismic Scale EMS-98 for earthquakes (Grünthal, 1998), which was later used as a basis to develop damage grades for flood hazards by Schwarz and Maiwald (2007).

Developing a damage grade requires data (or knowledge) of regional building damage patterns resulting from flood impact. Damage patterns, which are repeatedly observed within a region, can be categorized into damage grades (Schwarz and Maiwald, 2007). Grünthal et al. (1998) and Maiwald and Schwarz (2015) noted that damage grades should not only consider the physical effects of damage but also the number of buildings that show such effects. Hence, in developing damage grades, the focus should be given to both physical damage features and their corresponding proportion. Damage grades express frequently observed damage patterns as categories on an ordinal scale, whereby numbers are assigned to each damage pattern with higher numbers depicting a higher degree of damage (see Table 3). Damage grades vary from non-structural to structural damage. Non-structural damage refers to damage that does not immediately affect the structural integrity of a building. Examples of non-structural damage by floods include moisture defects or light cracks on building finishes. Structural damage mostly occurs on load-bearing elements of the building, for example, cracks in or collapses of walls, beams, and columns (Milanesi et al., 2018).

Generally, there is a wide range of damage patterns available to describe how buildings respond to flood impact. However, including all of these patterns will lead to unnecessarily complex flood damage models. Nonetheless, damage grades should be detailed enough to capture predominantly observed patterns of damage within a region. In such a way, damage grades serve as a compromise between comprehensiveness and simplicity (Blong 2003b). Grünthal (1993) recommended guidelines for good practice in developing damage grades, including (i) checking a wide range of information sources and considering their value, (ii) focusing more on repetitive damage than on extreme damage patterns, and (iii) additionally considering undamaged buildings. As an additional recommendation, Blong (2003b) suggested that damage models should be flexible enough to allow integration of new damage patterns over time. An example of such flexibility is demonstrated in Maiwald and Schwarz (2015, 2019) when expanding an originally five-category damage grade scheme to a six-category scheme. Damage grades are not affected by temporal changes (increase or decrease) in market value or wages, which can affect relative and absolute losses (Blong, 2003a). Due to this robustness to changes, they are easily transferable to regions with comparable building and hazard characteristics. This transferability is particularly important for data-scarce regions, where resources are limited for comprehensive data collection campaigns. Other characteristics of damage grades include simplicity, clarity, reliability, robustness, and spatial suitability (Blong, 2003b).

3.3 Challenges and gaps in flood damage models

Predicting damage grades using commonly applied stage–damage curves and multivariate methods has some weaknesses. These weaknesses are either manifest in both data-rich and data-scarce regions or specific to the latter. For example, despite the wide usage of stage–damage curves, several studies have highlighted inherent uncertainties particularly regarding damage predictions since they consider flood depth as the only damage-influencing parameter (e.g. Merz et al., 2004, 2013; Vogel et al., 2012; Pistrika et al., 2014; Schröter et al., 2014; Wagenaar et al., 2017; Sturm et al., 2018b, a; Fuchs et al., 2019b). These studies have demonstrated that flood damage is not only influenced by water depth but also by other hazard parameters (e.g. velocity and duration) and building characteristics (e.g. construction type, quality, and material). For instance, Merz et al. (2004) demonstrated the poor explanatory power of flood depth in explaining the variance in a data set. Although applying multivariate methods reduced uncertainties associated with models based on a single damage-influencing parameter, in data-scarce regions a disadvantage of the multivariate method is the lack of empirical data for developing and validating such models.

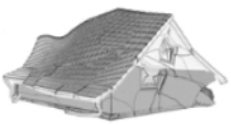
Several other challenges exist in data-scarce regions, which further limits the development of flood damage models. Merz et al. (2010) noted that selecting a method depends on data availability and knowledge of damage mechanisms. The absence of insurance against damage from natural hazards and effective government compensation schemes, typical for many data-scarce regions, contributes to a lack of data to support physical flood vulnerability assessment. For example, Komolafe et al. (2015) reported that no research institute or agency has a central database to document flood damage in many African countries, such as Nigeria. They further pointed out that such scarcity of damage data might be related to the fact that the practice of flood insurance is uncommon and government compensation after flood disasters is flawed. As such, people immediately repair their buildings after a flood event. Additionally, regulatory policies on building standards are less well implemented in many areas. Similar observations were made by Englhardt et al. (2019) in Ethiopia, pointing out a considerable difference in building quality and value, especially in rural areas. Also, in Nigeria, FGN (2013) reported that over 60 % of households acquire their houses through private resources and initiatives; thus, only a few use the services of formal institutions. This often leads to substantial differences in the quality of buildings, consequently increasing the challenges in developing building-type vulnerability assessment schemes. In addition, such differences in building quality further limit the application of flood damage models that use relative or absolute monetary losses due to a high range of replacement costs and property values.

4 The need for linking indicators and damage grades

A combination of damage grades (representing repeatedly observed damage patterns) with vulnerability indicators (capturing important damage-influencing variables within a region) using an expert-based what-if approach offers a convenient and comprehensive method for assessing flood damage. This allows us to tailor

flood damage models to the specific needs of data-scarce regions and simultaneously to take advantage of the strengths of the methods while limiting their individual weaknesses.

Table 3: Damage grades developed by Schwarz and Maiwald (2007) showing structural and non-structural damage to buildings. For each damage grade class, a description and a graphical representation are shown. The grey colour in the graphical representation indicates flood depth.

Damage grade class	**Damage**		**Description**	**Graphical representation**
	Structural	*Non-structural*		
D1	No	Slight	Only penetration and pollution	
D2	No to slight	Moderate	Slight cracks in supporting elements Impressed doors and windows Contamination	
D3	Moderate	Heavy	Major cracks and/or deformations in supporting walls and slabs Settlements	
D4	Heavy	Very heavy	Structural collapse of supporting walls, slabs	
D5	Very heavy	Very heavy	Collapse of the building or of major parts of the building	

Several weaknesses highlighted in Sects. 2.3 and 3.3 have limited the assessment of physical vulnerability. However, specific aspects of these approaches can be utilized for data-scarce regions. Although the vulnerability index has been identified as lacking a stand-alone meaning, its combination with damage grades will extend its applicability for damage grade prediction. Besides, the use of damage grades will help us to evaluate the performance of vulnerability indices. Current flood damage models were identified to be either data-intensive (multivariate methods) or to not consider other damage-influencing variables (stage–damage curves). However, an integration of damage grades with vulnerability indicators can provide a suitable model

to overcome these challenges. This integration can be fostered through utilizing the expert-based synthetic what-if analysis, which has been applied for developing synthetic stage–damage curves.

	i	ii	iii
			
REGION A: Aggregated Index	0.7	0.6	0.9
Vulnerability indicators ▪ Building material ▪ Building condition ▪ Distance to channel ▪ Flood depth	 Sandcrete block Moderate 100 m 1 m	 Sandcrete block Good 50 m 1.2 m	 Clay Poor < 20 m 0.60 m
REGION B: Aggregated Index	0.5	0.4	0.7
Vulnerability indicators ▪ Building age ▪ Building quality ▪ Sheltering effect ▪ Flood depth	 < 10 years Good Complete 1m	 20 years Moderate Partial sheltering 1m	 > 30 years Poor No sheltering 1m
Maiwald and Schwarz (2015) 5-category damage grade Germany	**DG 2** Slight cracks in supporting element	**DG 4** Partial collapse of supporting element	**DG 5** Collapse
Ettinger et al. (2016) 4-category damage grade Peru	**Light** Signs of impact	**Heavy** Partial/Total collapse	**Heavy** Total Collapse

Figure 2: Illustration of the need for linking vulnerability index and damage grades using real damage cases (a, b, and c) documented after a 2017 flood in Suleja and Tafa, Nigeria; hypothetical vulnerability indicators and regions (A and B); and damage grades developed from studies by Maiwald and Schwarz (2015) and Ettinger et al. (2016).

To demonstrate the added value of this linkage, we use a combination of (i) observed flood damage data, (ii) a hypothetical physical vulnerability index for two regions A and B, and (iii) two flood damage models developed for predicting damage grades. The observed damage data (see Fig. 2) was documented from a field survey conducted after the 2017 flood event in Suleja and Tafa, Nigeria. The flood event was caused by prolonged rainfall for about 12 h between 8 and 9 July 2017. The flood event resulted in the loss of lives and damaged hundreds of buildings and a large amount of infrastructure (Adeleye et al., 2019). A field study was conducted in March 2018 in order to document damage to the built environment and to interview affected homeowners. From the documented cases, we use three buildings to illustrate the potential weakness that may occur from only using a vulnerability index approach and the added value of the suggested linkage with damage grades.

The three buildings shown in Fig. 2 are constructed from sandcrete block (Fig. 2, buildings a, b) and clay bricks (Fig. 2, building c). The buildings have different damage patterns, ranging from moisture defects on walls resulting in peeling-off of plaster material and slight cracks (e.g. building a), partial collapse of supporting walls (e.g. building b), and complete collapse (e.g. building c). A hypothetical physical vulnerability index is considered for the two regions A and B (see Fig. 2). In the two regions, hypothetical

vulnerability indicators were assigned as the main damage-influencing parameters. Indicators for region A included building material, building condition, distance to channel, and flood depth. Indicators for region B included building age, building quality, sheltering effect, and flood depth. Vulnerability indices for regions A and B both express relative vulnerability from 0 (low vulnerability) to 1 (high vulnerability). Hypothetical vulnerability indices, after aggregating identified indicators, are given in Fig. 2. We further consider two damage grades presented by Maiwald and Schwarz (2015) for Germany and by Ettinger et al. (2016) for Peru. We use identified damage patterns on the buildings from the field study to assign a damage grade to each building.

From Fig. 2, we see that although we can use the developed index to identify which building is highly or moderately vulnerable within a region, we cannot compare the indices between different regions because they contain aggregated information from different parameters. However, in the case of damage grades, although they were developed in two different regions, qualitative descriptions of these grades can be used to assign damage grade classes for the identified damage patterns in buildings a, b, and c (Fig. 2).

A combination of physical vulnerability indicators and damage grades using the synthetic approach has a number of advantages for data-scarce areas.

i. Employing the synthetic what-if analysis to link damage grades and damage drivers allows us to overcome high empirical data requirements of the multivariate method. Consequently, the linkage will capture multiple damage-influencing variables. Also, using the damage grades will allow us to carry out performance checks on the effectiveness and robustness of selected vulnerability indicators.
ii. The linkage will enable us to compare consequences of flood hazards across spatial and temporal scales in data-scarce regions. Spatial comparability can be achieved through the identification of similar damage characteristics (Fig. 2) between regions with similar building types and hazard characteristics. Temporal comparability can be achieved by relating the severity of observed damage grades between different flood events since damage grades are not readily affected by market values or wages. In addition, using similar hazard scenarios, damage can be estimated and compared between regions while still considering individual damage drivers (Fig. 2).
iii. Since damage grades are physically observable features, the linkage will foster the provision of an easy communication tool for stakeholders and community residents about the consequences of hazards.

5 Conceptual framework

In this section, we present a new conceptual framework that aims to link physical vulnerability indicators and damage grades in order to make use of their individual strengths for data-scarce regions. We first provide

background information on terminologies used within the framework and second present step-by-step details on how to operationalize the framework.

5.1 Background for operationalizing the new framework

Vulnerability indicators are used to capture damage-influencing variables, which include characteristics of flood hazard, the built environment, and its surroundings. Damage grades represent the physical consequences of hazard impacts on a building that depends on both hazard and building characteristics. Figure 3 shows the conceptual framework and the proposed approach for linking physical vulnerability indicators and damage grades, the terminology is given below.

-*Vulnerability.* The degree to which an exposed building will experience damage from flood hazards under certain conditions of exposure, susceptibility, and resilience or local protection (adapted from Balica et al., 2009).

-*Impact (action) and resistance parameters*. The framework considers two major damage-influencing parameters: action (impact) and resistance parameters. The action and resistance parameters have been identified by Thieken et al. (2005) and Schwarz and Maiwald (2007) as the primary classes of damage drivers. Impact (or action) parameters relate to the flood parameters comprising of hazard frequencies and magnitudes (Thieken et al., 2005). Resistance parameters are related to the predisposition of the building to suffer damage, either permanently (e.g. building material) or temporarily (e.g. measures for flood preparedness) (Thieken et al., 2005). In the framework, resistance parameters comprise elements of the building and its surroundings, which are divided into susceptibility, exposure, and local protection parameters.

-*Exposure*. Refers to the extent to which a building is spatially or temporarily affected by a flood event (adapted from Birkmann et al., 2013). Exposure parameters include features of the natural and built environment that either increase or decrease the impact of floods on buildings, such as topography and distance to the flood source.

-*Susceptibility*. Refers to the disposition of a building to be damaged by a flood event (adapted from Birkmann et al., 2013). Susceptibility parameters relate specifically to the structural characteristics of the building at risk, neglecting any effects of local protection measures that may provide flood protection.

-*Local protection*. Refers to deliberate or non-deliberate measures that are put in place and can reduce the impact of floods on a building. These measures can be directly included in the building structure, e.g. elevation of the entrance door, or measures located in the immediate surrounding of a building. While many local structural protection measures may not be primarily constructed as a protection mechanism against floods, they reduce the impact of floods on a building (Holub and Fuchs, 2008; Attems et al., 2020b). In the context of this framework, a fencing wall will be an example of a local protection measure.

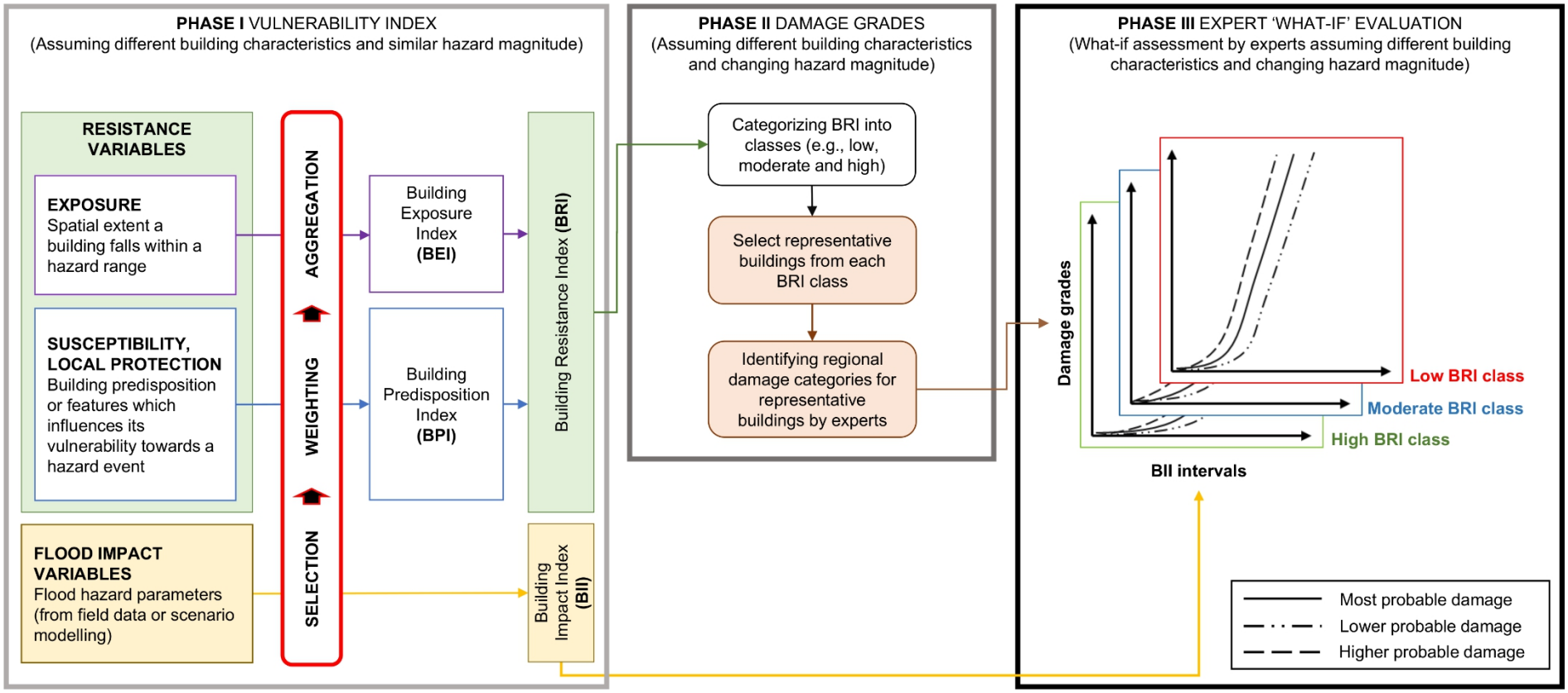


Figure 3: The proposed conceptual framework, linking vulnerability indicators to damage grades so that vulnerability to the built environment can be better assessed in data-scarce regions. The framework consists of three consecutive steps (phases) from the vulnerability index development (assuming different building characteristics but similar hazard magnitudes) to the damage grades (assuming different building characteristics and changing hazard magnitudes) and finally an expert-based "what-if"-evaluation, leading to functions linking damage grades from phase II to building impact indices (BIIs) from phase I for each BRI class.

5.2 Operationalizing the framework

In order to operationalize the new framework, three phases are proposed: (i) developing a vulnerability index, (ii) developing a damage grade classification, and (iii) linking the vulnerability index to the damage grade classification.

5.2.1 Phase 1: developing a vulnerability index

We develop a vulnerability index aimed at systematically integrating damage-influencing parameters. These parameters represent vulnerability indicators or damage drivers adapted for a selected region. As a result, we structure indicators into impact and resistance parameters as shown in Fig. 3 (phase I). In order to allow an evaluation of how different components contribute to damage, we categorize resistance parameters into separate components, exposure, susceptibility, and local protection. Application of the method is aimed at the microscale level, however, it can be applied at meso- or macroscale if data are available. Generally, the selection, weighting, and aggregation of indicators are similar to the procedure discussed in Sect. 2.2. Since our focus is on data-scarce regions, we focus the framework on expert-based approaches.

Indicators are mainly selected using expert surveys. Where possible, experts should include individuals from different disciplines in order to have a wide-ranging assessment. Expert surveys are carried out by conducting standardized interviews using questionnaires. The main focus of the questionnaire is on asking each expert to identify parameters representing damage drivers within a region. A set of indicators can be identified and included in the questionnaire, with the support of a literature review. Experts can then either select from the suggested indicators or propose new ones. All variables suggested by experts at this step serve as preselected indicators.

Indicator (or parameter) weighting is carried out using an expert-based approach. Here, experts are asked to weight how each preselected variable influences damage. The weighting is carried out using a scale of influence table based on Saaty (1980), as shown in Table 4. Because the table by Saaty (1980) was originally used for making a pairwise comparison between two parameters, it was slightly modified so as to be used in weighting preselected parameters with respect to how they influence flood damage. The scale (Table 4) will help to bring consistency and comparability in weighting when using the framework. Using Table 4, experts can assign a certain influence (e.g. slight, strong) for each preselected indicator. For each parameter, a mean value of the assigned weights from all experts is calculated and checked based on Table 4. The mean value here represents the central value used to communicate how all the experts evaluate a parameter based on its influence on damage within a region. The mean weights for each parameter are used for the final selection of indicators. For example, a mean weight of 2 from Table 4 will infer that, on average, experts consider the parameter to have only a slight effect on damage. A decision has to be made on a threshold (e.g. 1, 2, or 3 from Table 4) for parameter inclusion for the final selection. The threshold will depend, however, on the specific need (e.g. level of accuracy) or aim (e.g. identifying major damage-influencing parameters) of the study. Only parameters included in the final

Table 4. Table of influence for indicator weighting, ranging from slight influence of an indicator (1) to extreme influence (9) (modified from Saaty, 1980).

1	2	3	4	5	6	7	8	9
Slight influence	Slight to moderate influence	Moderate influence	Moderate to strong influence	Strong influence	Strong to very strong influence	Very strong influence	Very strong to extreme influence	Extreme influence

selection will be used in the indicator aggregation step. Next, using mean values for each indicator that has passed the final selection, the AHP is implemented to determine indicator weights (see Sect. 2). For detailed information on the procedure for implementing the AHP, we refer the reader to JRC and OECD (2008) and Saaty (1980).

A normalized weighted additive method is used for aggregating indicators. As shown in Fig. 3 (phase 1), selected parameters for exposure are aggregated to derive a building exposure index (BEI). The BEI is a measure of the extent to which a building is likely to be damaged as a result of (i) the spatial location relative to the flood source and (ii) surrounding buildings. Indicators for susceptibility and local protection are aggregated to derive a building predisposition index (BPI). The BPI provides a measure of the extent to which a building is likely to be damaged based on the building characteristics and available protection measures. Both the BEI and BPI are aggregated to derive a building resistance index (BRI). The BRI measures expected resistance a building can offer at a specific degree of impact, given its predisposition and exposure. Hence, given the same degree of hazard impact, a building with a high BRI (high resistance) is expected to experience less damage compared to a building with a low BRI (low resistance). As pointed out earlier, a building-type vulnerability classification can be challenging in data-scarce areas. Therefore, we propose the use of the BRI to classify buildings into different resistance classes (e.g. low, moderate, and high). Such classifications of buildings into vulnerability categories have been shown to facilitate a better understanding of the distribution of damage data (Schwarz and Maiwald, 2008). Elements within the same vulnerability class are expected to experience similar damage when impacted by the same degree of hazard.

The last step in phase 1 is to utilize the additive model to aggregate flood hazard parameters (e.g. depth, duration) in order to derive a building impact index (BII). The BII is used to express the combined effect of hazard parameters on a building structure. The BII is computed using interview data collected after a flood event (Malgwi et al., 2020).

5.2.2 Phase 2: developing the damage grades

We adopt a slightly modified procedure outlined in Naumann et al. (2009) for developing damage grades. Figure 3 (phase II) shows the systematic steps for developing the damage grades using an expert-based approach. The main aim of this step is to identify commonly observed damage patterns within a region and

categorize them into classes. As such, basic outputs of this phase are classes of different damage patterns ordered into damage grades.

Sourcing for damage patterns within a region is carried out by analysing observed damage data or by structured interviews with experts or community residents. Such structured interviews are undertaken using questionnaires in flood-prone communities. Community residents or experts are asked which damage patterns are observed after flood events. They are also asked about how frequently these observed patterns occur after floods. In addition, questions about which damage types are usually repaired (or replaced) after flood events can be asked. From such information, the original damage can be deduced. Other sources of information are literature reviews, reviews of damage reports, news, and social media (videos and images). Such a wide range of information sources is particularly encouraged by Grünthal (1993) in order to have a comprehensive damage grade classification. Attention should also be given to the proportion of buildings observed to exhibit each damage grade (Grünthal, 1993). The damage grades should not focus on isolated (uncommon) damage patterns, instead more attention should be given to frequently observed patterns.

We present an overview of a synthetic method for developing a damage grade as described by Naumann et al. (2009). The necessary steps are as follows.

i. Identification of building types and building representatives. A first step for developing a flood damage model is to assess building types within a region and select building representatives (Walliman et al., 2011; Maiwald and Schwarz, 2015). The assessment of building types can be carried out based on field surveys, expert surveys or remote sensing. Where a large-scale building assessment is required, a method conceptualized by Blanco-Vogt and Schanze (2014) for semi-automatic extraction and classification of buildings can be applied. The representatives should include building types (material, form of construction and quality) that are predominant within a region. Additionally, Naumann et al. (2009) noted other attributes used for classifying buildings, these include the period of construction and the original use, the characteristic formation of buildings, and spatial patterns and geometry. In the framework, we use the BRI for classifying buildings into different categories since it ideally captures parameters that influence damage. A suitable classification for the BRI is a generic categorization into “low BRI”, “moderate BRI”, and “high BRI” classes (Fig. 3, phase II). The class represents buildings that will offer a low (low BRI), medium (moderate BRI), and good (high BRI) resistance if we consider the same impact magnitude. Such a generic building classification, which is not building-type based, is especially suitable in areas with a high variability in building quality. From each BRI class, a representative building is selected. Suitably, these representatives can be selected from different building types and should communicate the typical characteristics of buildings in the BRI class.

ii. Identification and grading of regional damage patterns. Flood damage to buildings can be generally categorized into three major parts, these include water penetration damage (moisture), chemical damage (pollution and contamination), and structural damage (Schwarz and Maiwald, 2007; Walliman et al., 2011). These three general damage categories can serve as a basis for developing further damage classification in regions where such damage assessment was not previously carried out. For each BRI representative, different patterns of damage are identified. Patterns that are repeatedly observed are indications of a damage grade category (Maiwald and Schwarz, 2015). Where the damage patterns for different representatives are the same, a single damage grade scheme can be adopted. However, where the damage patterns are substantially different, the damage grade is adapted for each BRI representative. This step ensures that predominant building and damage types are considered.

In the next step, identified damage patterns are assigned to a scale representing the degree of damage severity. A commonly applied scale for damage grades is the ordinal scale (e.g. Table 3). The ordinal scale provides suitable classes for damage grades since the intervals between different categories are not consistent. For example, in Table 3, the difference in severity between damage grades 1 and 2 is not the same as between 2 and 3. Minimum damage (usually water contact with external walls or water penetration) and maximum damage (complete collapse or washing away of a building) have to be decided. Additionally, a decision has to be made on how many damage grades to consider. As pointed out earlier, a balance has to be set between comprehensiveness and simplicity. Where difficulties exist in deciding which damage grade is of higher or lower severity, local technicians or other persons with building construction expertise can be asked to estimate the repair cost for each damage grade. In this case, a high repair cost will infer a higher damage grade.

5.2.3 Phase 3: expert "what-if" analysis

With a focus on data-scarce regions, we present steps to link damage-influencing variables (from phase 1) and predominant damage patterns (from phase 2). Expert knowledge is utilized to predict damage grade(s) for each representative building type (BRI class) using synthetic flood depths. The synthetic flood depths will represent scenario-based flood depths that are typical for a region. Intervals for synthetic flood depths are integrated using the BII (Fig. 3).

In the what-if analysis, expert knowledge of regional flood damage mechanisms is crucial. Based on a given flood depth, experts propose a probable damage grade for a specific building type. Estimating a single damage grade for a given water depth can result in uncertainties. Therefore, we propose the use of three probable damage states to capture the range of possible damage. Figure 3 (phase III) shows an idealized curve depicting the relationship between damage grades, BII, and BRI. The methodical steps for linking damage grades with the BRI and BII were adopted from and modified following Naumann et al. (2009) and Maiwald and Schwarz (2015). Steps for the linkage include the following processes.

i. To develop suitable intervals for the BII, such as flood depths in steps of 0.5 or 1 m intervals.

ii. For each defined interval of BII, local experts estimate the expected damage for each BRI class. Experts should provide three possible damage grades for each BII interval. The possible damage grades should include (i) most probable damage grades, (ii) lower probable damage grades, and (iii) higher probable damage grades. As an example, if a representative building type (e.g. one-story sandcrete block building) is selected from the BRI category "low resistance", experts will estimate for each BII interval (e.g. 1 m water depth) the damage to be expected. Such damage estimates can be (i) most probable: slight cracks on supporting walls; (ii) lower probable damage: only water penetration; and (iii) higher probable damage: heavy cracks on supporting walls.

iii. For each BRI class, a suitable curve is used to join most probable, lower probable, and higher probable damage for all BII values, as exemplified in Fig. 3 (phase III).

6 Concluding remarks

With increasing magnitudes and frequencies of floods, assessing the physical vulnerability of exposed communities is crucial for reducing risk (UNISDR, 2015). The success of risk reduction methods is even more critical for data-scarce areas, which are mostly developing countries with limited capacity to cope with flood risk (UNDRR, 2019). Mitigating flood risks on the built environment requires knowledge of flood characteristics, building typology, and damage mechanisms so that viable predictions can be carried out for damage grades and monetary loss. Physical vulnerability assessment includes the identification of major damage drivers and the evaluation of possible damage to exposed buildings. For data-scarce regions, such a vulnerability assessment, which can be adapted to regional building types, may serve as a first step in overall risk reduction.

Taking up this challenge, we presented reviews and concepts for assessing the physical vulnerability of buildings exposed to flood hazards. Furthermore, we proposed linking approaches with reduced data requirements. Two approaches were considered: (i) the vulnerability indicator method, which is used for identifying regional damage drivers, and (ii) the damage grades approach, used for classifying commonly observed damage patterns.

The review provides a state of the art in physical vulnerability assessment, particularly in expert-based methods, and can serve as a useful information source for further studies with only a limited quantity of empirical data available. We identified the following specific challenges for a further development of these approaches in data-scarce regions: applying vulnerability indicators and indices needs (i) proper data management according to the chosen indicator framework, (ii) a general improvement of sensitivity analysis with respect to vulnerability indices to foster insights for further indicator development, and (iii) an application of performance tests of vulnerability indices and the linkage between the index and the underlying building damage. The application of flood damage models is strongly limited by the lack of data for the

validation of stage–damage curves and thus reducing the high uncertainty of this approach in general. Furthermore, basic information on building types or building quality and damage due to floods is not systematically documented and therefore limits the development of flood damage models. Based on the identified challenges and limitations of different vulnerability methods, the proposed conceptual framework further suggests linking the vulnerability indicator method to the damage grades method using an expert-based approach. Combining these methods has been identified as a useful way to enhance the utility and robustness of individual physical vulnerability assessment methods while limiting their weaknesses. The combination of different vulnerability assessment methods was suggested by different scholars (e.g. Papathoma-Köhle et al. 2017) but extended application is still outstanding.

The proposed framework focuses on enhancing regional adaptability of physical vulnerability assessment methods and fostering model transfer between different data-scarce regions. Three phases are required to operationalize the framework, (i) developing a vulnerability index, (ii) identifying predominant damage grades or patterns, and (iii) carrying out a what-if analysis to link identified flood characteristics to damage grades for each category of building resistance.

In developing the vulnerability index, we considered hazard parameters (BII) and variables relating to the characteristics of a building and its surroundings (BRI). The BRI aggregates information on exposure, susceptibility, and local protection of a building, and hence connects the resistance of a building relative to other buildings assuming the same hazard magnitude. The proposed classification of the BRI is not based on building types (e.g. Maiwald and Schwarz, 2015) but is instead a classification based on aggregated information on exposure, susceptibility, and local protection, such as property-level adaptation measures (Attems et al., 2020b). We recommend such a generic classification of building types (e.g. low, moderate, high resistance) especially in regions with high variation in building quality (Englhardt et al., 2019). Systematic documentation of regional building damage patterns is required for the framework so that frequently observed damage patterns (e.g. moisture defects, cracks on supporting elements, partial collapse, complete collapse) can be integrated into a damage grade classification. As the framework is not case study sensitive, damage categories from other studies can provide a useful basis for categorizing damage grades. Furthermore, expert-based what-if analysis is used to assign identified damage grades to each interval of the BII. Where empirical data are available, even in limited quantity, they should be used to support the what-if analysis. In particular, the potential of citizen-based data sources, such as information taken from interviews or social media, offers a good opportunity for damage data collection (e.g. using Twitter Cervone et al., 2016 or Facebook Sy et al., 2020). The framework is fully expert-based and flexible, allowing vulnerability indicators and damage grades to be updated when new post-flood data become available. Consequently, curves generated between BII, BRI, and damage grades can be continuously updated over time. In this way, the new framework allows temporal changes in damage drivers to be integrated. We further recommend the application of the new framework to evaluate and compare model performance with a data-driven multivariate model. Such an analysis will communicate the success of the framework and also allow for further

improvement. Based on the modular structure of the framework, it has the potential to be adapted for different environments, hazard types, and vulnerability types.

Appendix 1

Table A1: Overview of common elements for framing the vulnerability indicator approach for flood hazards, indicating the hazard type and vulnerability dimensions, the implementation in the risk cycle, the scale, the index output, and the data source (PTVA is the Papathoma-Koehle vulnerability assessment model).

Author	Hazard type	Region of application	Aim of the assessment	Vulnerability dimension	Implementation in risk cycle	Scale	Index output	Data source
Papathoma et al. (2003)	Tsunami	Gulf of Corinth, Greece	Assessing the vulnerability of coastal areas to tsunami	Physical, economic and social	Preparedness	Micro-scale (individual buildings)	Building and human vulnerability index	Field survey
Dominey-Howes and Papathoma (2007)	Tsunami	Maldives, India	Checking the performance of PTVA	Physical, economic, social and environmental	Preparedness	Micro-scale (individual buildings)	Building and human vulnerability index	Field survey
Balica et al. (2009)	River flood	Timisoara, Romania; Mannheim, Germany; Phnom Penh, Cambodia	Assessing the conditions influencing flood damage at various spatial scales	Physical, economic, social and environmental	Preparedness	Meso-scale (regional)	Flood vulnerability index	-
Kienberger et al. (2009)	River flood	Salzach catchment, Austria	Identification of hotspots	Physical, economic and social	Mitigation and preparedness	Meso-scale (regional)	Vulnerability index	Government agency
Dall'Osso et al. (2009)	Tsunami	Sydney, Australia	Assessing the vulnerability of buildings to tsunami and evaluating the use of the PTVA	Physical	Mitigation and preparedness	Micro-scale (individual buildings)	Relative vulnerability index	Field survey

Müller et al. (2011)	(Urban) flood	Peñalolèn and La Reina Municipalities, Santiago de Chile	Empirical investigation of vulnerability towards flood	Physical and social	Mitigation	Micro-scale (entire building blocks)	Vulnerability index (adapted after Haki et al., 2004)	Census data, field survey and satellite data
Kappes et al. (2012)	River flood, flash flood (among others)	Faucon municipality, Barcelonnette basin, France	Assessing the hazard-specific physical vulnerability of buildings towards multi-hazard	Physical, social and environmental	Mitigation and preparedness	Micro-scale (individual building)	Relative vulnerability index	Research agency and aerial-photo-interpretation
Balica et al. (2012)	Coastal flood	Buenos Aires, Argentina; Calcutta, India; Casablanca, Morocco; Dhaka, Bangladesh; Manila, Philippines; Marseille, France; Osaka, Japan; Shanghai, China; Rotterdam, The Netherlands	Developing a coastal city flood vulnerability index	Physical, social, economic and administrative	Preparedness	Meso-scale (regional)	Coastal city flood vulnerability index	Government agencies and data available online
Blanco-Vogt and Schanze (2014)	River flood	Magangué, Columbia	Assessing physical flood susceptibility on a large scale	Physical	Recovery, mitigation, and preparedness	Micro-scale (individual buildings)	Function relating susceptible material volume and water depth	Very high resolution spectral and elevation data and field survey
Thouret et al. (2014)	Flash flood	Arequipa, Peru	Assessing vulnerability	Physical and environmental	Mitigation	Micro-scale (entire building blocks)	Vulnerability index	Field survey
Bagdanavičiute et al. (2015)	Coastal flood	Coast of Lithuania	Assessing coastal vulnerability	Physical	Mitigation	Meso-scale (regional)	Coastal vulnerability index	Field survey

Behanzin et al. (2015)	River flood	Niger River Valley, Bénin	Assess vulnerability and risk	Physical, economic, social and environmental	Mitigation and preparedness	Meso-scale (community)	Vulnerability and risk index	Field survey, other agencies
Godfrey et al. (2015)	River and flash flood, slow-moving landslide, debris flow	Nehoiu City, Buzău County, Romania	Assessing the physical vulnerability of buildings to hydro-meteorological hazards in data-scarce regions	Physical	Mitigation and preparedness	Micro-scale (individual buildings)	Vulnerability index	Field survey and orthophoto interpretation
Akukwe and Ogbodo (2015)	River and coastal flood	Port Harcourt, Nigeria	Showing spatial variations in vulnerability	Physical, economic and social	Mitigation and preparedness	Meso-scale (regional)	Vulnerability Index (adapted after Deressa et al., 2008)	Field survey, survey and map measurements
Fernandez et al. (2016)	River flood	Vila Nova de Gaia, Northern Portugal	Providing an automated framework for classifying vulnerability of neighborhoods	Physical, economic, social and environmental	Preparedness	Micro-scale (neighborhood)	Flood Vulnerability Index and	Government agency
Ntajal et al. (2016)	River flood	Mono River Basin, Togo	Assessing and mapping vulnerable communities	Physical, economic, social and environmental	Mitigation and preparedness	Meso-scale (community)	Index for exposure, susceptibility, capacity, and vulnerability	Field survey, other agencies
Krellenberg and Welz (2017)	Flood (urban)	Metropolitan area of Santiago de Chile	Assessing urban vulnerability	Physical, economic, social and environmental	Mitigation	Micro-scale (building block)	Vulnerability Index	Field survey, government agency, and satellite imagery
Sadeghi-Pouya et al. (2017)	River flood	Mazandaran, Iran	Assessing vulnerability	Physical, economic, social and environmental	Mitigation and preparedness	Micro-scale (building block)	Relative vulnerability index	Field survey and government agency

Carlier et al. (2018)	River flood	Upper Guil catchment, southern French Alps	Assessing the physical and socio-economic consequence of hazards on elements at risk	Physical and social	Mitigation	Micro-scale (individual buildings)	Potential damage index, potential consequence index	Government agency, field survey, and aerial imagery
Yankson et al, (2017)	Coastal flood	Accra, Ghana	Understanding flood risk in coastal communities	Physical and social	Mitigation	Meso-scale (community)	Impact index vulnerability index	Field survey
Percival et al. (2018)	Coastal flood	Portsmouth, United Kingdom	Assessing risk from diurnal floods	Physical, environmental, social, economic	Mitigation	Micro-scale (neighborhood)	Coastal flood vulnerability Index, Coastal flood hazard Index, Coastal flood risk index	Census data
Papathoma-Köhle et al. (2019)	Tsunami	Apulia, Italy	Assessing vulnerability from tsunami hazards to the built environment	Physical vulnerability	Mitigation and preparedness	Micro-scale (neighborhood)	Building vulnerability index	Field survey

Acknowledgements

This study is carried out within the framework of a PhD scholarship funded by the Swiss Government Excellence Scholarships for Foreign Scholars (ESKAS). Thanks to Candace Chow for useful inputs in formulating Fig. 2. Thanks to two anonymous referees for their insightful comments on an earlier version of the paper.

Financial support

This research has been supported by the Swiss government excellence scholarship (grant no. 2017.1027).

References

Adelekan, I., Johnson, C., Manda, M., Matyas, D., Mberu, B., Parnell, S., Pelling, M., Satterthwaite, D., and Vivekananda, J.: Disaster risk and its reduction: an agenda for urban Africa, Int. Dev. Plan. Rev., 37, 33–43, https://doi.org/10.3828/idpr.2015.4, 2015.

Adeleye, B., Popoola, A., Sanni, L., Zitta, N., and Ayangbile, O.: Poor development control as flood vulnerability factor in Suleja, Nigeria, T. Reg. Plan., 74, 23–35, https://doi.org/10.18820/2415-0495/trp74i1.3, 2019.

Akukwe, T. I. and Ogbodo, C.: Spatial analysis of vulnerability to flooding in Port Harcourt metropolis, Nigeria, SAGE Open, 5, https://doi.org/10.1177/2158244015575558, 2015.

Asadzadeh, A., Kötter, T., Salehi, P., and Birkmann, J.: Operationalizing a concept: The systematic review of composite indicator building for measuring community disaster resilience, Int. J. Disast. Risk Re., 25, 147–162, https://doi.org/10.1016/j.ijdrr.2017.09.015, 2017.

Attems, M.-S., Schlögl, M., Thaler, T., Rauter, M., and Fuchs, S.: Risk communication and adaptive behaviour in flood-prone areas of Austria: A Q-methodology study on opinions of affected homeowners, PLoS one, 15, e0233551, https://doi.org/10.1371/journal.pone.0233551, 2020a.

Attems, M.-S., Thaler, T., Genovese, E., and Fuchs, S.: Implementation of property level flood risk adaptation (PLFRA) measures: choices and decisions, WIREs Water, 7, e1404, https://doi.org/10.1002/wat2.1404, 2020b.

Aubrecht, C., Fuchs, S., and Neuhold, C.: Spatio-temporal aspects and dimensions in integrated disaster risk management, Nat. Hazards, 68, 1205–1216, 2013.

Bagdanavičiute, I., Kelpšaite, L., and Soomere, T.: Multi-criteria evaluation approach to coastal vulnerability index development in micro-tidal low-lying areas, Ocean Coast. Manage., 104, 124–135, https://doi.org/10.1016/j.ocecoaman.2014.12.011, 2015.

Balica, S. F., Douben, N., and Wright, N. G.: Flood vulnerability indices at varying spatial scales, Water Sci. Technol., 60, 2571–2580, https://doi.org/10.2166/wst.2009.183, 2009.

Balica, S. F., Wright, N. G., and van der Meulen, F.: A flood vulnerability index for coastal cities and its use in assessing climate change impacts, Nat. Hazards, 64, 73–105, https://doi.org/10.1007/s11069-012-0234-1, 2012.

Barnett, J., Lambert, S., and Fry, I.: The hazards of indicators: insights from the environmental vulnerability index, Ann. Assoc. Am. Geogr., 98, 102–119, https://doi.org/10.1080/00045600701734315, 2008.

Barroca, B., Bernardara, P., Mouchel, J. M., and Hubert, G.: Indicators for identification of urban flooding vulnerability, Nat. Hazards Earth Syst. Sci., 6, 553–561, https://doi.org/10.5194/nhess-6-553-2006, 2006.

Behanzin, I. D., Thiel, M., Szarzynski, J., and Boko, M.: GIS-based mapping of flood vulnerability and risk in the Bénin Niger River Valley, Int. J. Geomatics Geosci., 6, 1653–1669, 2015.

Birkmann, J.: Measuring vulnerability to promote disaster-resilient societies: Conceptual frameworks and definitions, edited by Birkmann, J., Measuring vulnerability to natural Hazards, United Nationas University Press, Tokyo, 9–54, 2006.

Birkmann, J.: Risk and vulnerability indicators at different scales:. Applicability, usefulness and policy implications, Environ. Hazards, 7, 20–31, https://doi.org/10.1016/j.envhaz.2007.04.002, 2007.

Birkmann, J., Cardona, O. D., Carreño, M. L., Barbat, A. H., Pelling, M., Schneiderbauer, S., Kienberger, S., Keiler, M., Alexander, D., and Zeil, P.: Framing vulnerability, risk and societal responses: the MOVE framework, Nat. Hazards, 67, 193–211, https://doi.org/10.1007/s11069-013-0558-5, 2013.

Blanco-Vogt, A. and Schanze, J.: Assessment of the physical flood susceptibility of buildings on a large scale – conceptual and methodological frameworks, Nat. Hazards Earth Syst. Sci., 14, 2105–2117, https://doi.org/10.5194/nhess-14-2105-2014, 2014.

Blong, R.: A new damage index, Nat. Hazards, 30, 1–23, https://doi.org/10.1023/A:1025018822429, 2003a.

Blong, R.: A review of damage intensity scales, Nat. Hazards, 29, 57–76, https://doi.org/10.1023/A:1022960414329, 2003b.

Carlier, B., Puissant, A., Dujarric, C., and Arnaud-Fassetta, G.: Upgrading of an index-oriented methodology for consequence analysis of natural hazards: application to the Upper Guil catchment (southern French Alps), Nat. Hazards Earth Syst. Sci., 18, 2221–2239, https://doi.org/10.5194/nhess-18-2221-2018, 2018.

Cervone, G., Sava, E., Huang, Q., Schnebele, E., Harrison, J., Cervone, G., Sava, E., Huang, Q., Schnebele, E., and Harrison, J.: Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study, Int. J. Remote Sens., 37, 100–124, https://doi.org/10.1080/01431161.2015.1117684, 2016.

Chen, N., Hu, C., Chen, Y., Wang, C., and Gong, J.: Using SensorML to construct a geoprocessing e-Science workflow model under a sensor web environment, Comput. Geosci., 47, 119–129, https://doi.org/10.1016/j.cageo.2011.11.027, 2012.

Chow, C., Andrášik, R., Fischer, B., and Keiler, M.: Application of statistical techniques to proportional loss data: Evaluating the predictive accuracy of physical vulnerability to hazardous hydro- meteorological events, J. Environ. Manage., 246, 85–100, https://doi.org/10.1016/j.jenvman.2019.05.084, 2019.

Cutter, S. L. and Finch, C.: Temporal and spatial changes in social vulnerability to natural hazards, P. Natl. Acad. Sci. USA, 105, 2301–2306, https://doi.org/10.1073/pnas.0710375105, 2008.

Dall'Osso, F. and Dominey-Howes, D.: Coastal vulnerability to multiple inundation sources: COVERMAR project, Literature review report, Sydney, University of New-South Wales and Sydney Coastal Councils Group, 2013.

Dall'Osso, F., Gonella, M., Gabbianelli, G., Withycombe, G., and Dominey-Howes, D.: A revised (PTVA) model for assessing the vulnerability of buildings to tsunami damage, Nat. Hazards Earth Syst. Sci., 9, 1557–1565, https://doi.org/10.5194/nhess-9-1557-2009, 2009.

de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E., and Ward, P. J.: Flood risk assessments at different spatial scales, Mitig. Adapt. Strat. Gl., 20, 865–890, https://doi.org/10.1007/s1102, 2015.

de Ruiter, M. C., Ward, P. J., Daniell, J. E., and Aerts, J. C. J. H.: Review Article: A comparison of flood and earthquake vulnerability assessment indicators, Nat. Hazards Earth Syst. Sci., 17, 1231–1251, https://doi.org/10.5194/nhess-17-1231-2017, 2017.

Deressa, T. T., Hassan, R. M., Ringler, C., Alemu, T., and Yesuf, M.: Analysis of the determinants of farmers' choice of adapta- tion methods and perceptions of climate change in the Nile Basin of Ethiopia [in Amharic] (No. 15 (9) AMH), International Food Policy Research Institute (IFPRI), Dell'Acqua, 2008.

Dominey-Howes, D. and Papathoma, M.: Validating a tsunami vulnerability assessment model (the PTVA Model) using field data from the 2004 Indian Ocean tsunami, Nat. Hazards, 40, 113–136, https://doi.org/10.1007/s11069-006-0007-9, 2007.

Eddy, D. M., Hollingworth, W., Caro, J. J., Tsevat, J., McDonald, K. M., and Wong, J. B.: Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7, Med. Decis. Mak., 32, 733–743, https://doi.org/10.1177/0272989X12454579, 2012.

Englhardt, J., de Moel, H., Huyck, C. K., de Ruiter, M. C., Aerts, J. C. J. H., and Ward, P. J.: Enhancement of large-scale flood risk assessments using building-material-based vulnerability curves for an object-based approach in urban and rural areas, Nat. Hazards Earth Syst. Sci., 19, 1703–1722, https://doi.org/10.5194/nhess-19-1703-2019, 2019.

Eriksen, S. H. and Kelly, P. M.: Developing credible vulnerability indicators for climate adaptation policy assessment, Mitig. Adapt. Strat. Gl., 12, 495–524, https://doi.org/10.1007/s11027-006-3460-6, 2007.

Ettinger, S., Mounaud, L., Magill, C., Yao-Lafourcade, A. F., Thouret, J. C., Manville, V., Negulescu, C., Zuccaro, G., De Gregorio, D., Nardone, S., Uchuchoque, J. A. L., Arguedas, A., Macedo, L., and Manrique Llerena, N.: Building vulnerability to hydro-geomorphic hazards: Estimating damage probability from qualitative vulnerability assessment using logistic regression, J. Hydrol., 541, 563–581, https://doi.org/10.1016/j.jhydrol.2015.04.017, 2016.

Fernandez, P., Mourato, S., Moreira, M., and Pereira, L.: A new approach for computing a flood vulnerability index using cluster analysis, Phys. Chem. Earth, 94, 47–55, https://doi.org/10.1016/j.pce.2016.04.003, 2016.

FGN (Federal Government of Nigeria): Nigeria: Post-disaster needs assessment – 2012 floods, available at: https://www.gfdrr.org/sites/gfdrr/files/NIGERIA_PDNA_PRINT_05_29_2013_WEB.pdf (last access: 1 January 2019), 2013.

Frazier, T. G., Thompson, C. M., and Dezzani, R. J.: A framework for the development of the SERV model: A Spatially Explicit Resilience-Vulnerability model, Appl. Geogr., 51, 158–172, https://doi.org/10.1016/j.apgeog.2014.04.004, 2014.

Fuchs, S.: Susceptibility versus resilience to mountain hazards in Austria – paradigms of vulnerability revisited, Nat. Hazards Earth Syst. Sci., 9, 337–352, https://doi.org/10.5194/nhess-9-337-2009, 2009.

Fuchs, S., Keiler, M., Sokratov, S., and Shnyparkov, A.: Spatiotemporal dynamics: the need for an innovative approach in mountain hazard risk management, Nat. Hazards, 68, 1217–1241, https://doi.org/10.1007/s11069-012-0508-7, 2013.

Fuchs, S., Keiler, M., Ortlepp, R., Schinke, R., and Papathoma-Köhle, M.: Recent advances in vulnerability assessment for the built environment exposed to torrential hazards: challenges and the way forward, J. Hydrol., 575, 587–595, https://doi.org/10.1016/j.jhydrol.2019.05.067, 2019a.

Fuchs, S., Heiser, M., Schlögl, M., Zischg, A., Papathoma-Köhle, M., and Keiler, M.: Short communication: A model to predict flood loss in mountain areas, Environ. Modell. Softw., 117, 176–180, https://doi.org/10.1016/j.envsoft.2019.03.026, 2019b.

Gasparini, P.: Analysis and monitoring of environmental risk: CLUVA Final Report, available at: http://cordis.europa.eu/docs/results/265137/final1-cluva-final-publishable-summary-report.pdf (last access: 1 April 2020), 2013.

Gerl, T., Kreibich, H., Franco, G., Marechal, D., and Schröter, K.: A review of flood loss models as basis for harmonization and benchmarking, PLoS One, 11, 1–22, https://doi.org/10.1371/journal.pone.0159791, 2016.

Godfrey, A., Ciurean, R. L., van Westen, C. J., Kingma, N. C., and Glade, T.: Assessing vulnerability of buildings to hydro-meteorological hazards using an expert based approach – An application in Nehoiu Valley, Romania, Int. J. Disaster Risk Re., 13, 229–241, https://doi.org/10.1016/j.ijdrr.2015.06.001, 2015.

Golz, S.: Resilience in the built environment: How to evaluate the impacts of flood resilient building technologies?, E3S Web Conf., 7, 13001, https://doi.org/10.1051/e3sconf/20160713001, 2016.

Grünthal, G.: European Macroseismic Scale 1992 (up-dated MSK-scale), edited by Grünthal G, Cahiers du Centre Europèen de Gèodynamique et de Seismologie, Conseil de l'Europe, Conseil de l'Europe, 1993.

Grünthal, G.: European Macroseismic Scale 1998, edited by Grünthal G, Cahiers du Centre Europèen de Gèodynamique et de Seismologie, Conseil de l'Europe, Conseil del'Europe, 1998.

Günther, D.: Indicator sets for assessments, available at: http://www.ivm.vu.nl/en/Images/AT10_tcm234-161582.pdf (last access: 1 January 2019), 2006.

Haki, Z., Akyuerek, Z., and Duezguen, S.: Assessment of social vulnerability using Geographic Information Systems: Pendik, Istanbul case study, in: 7th AGILE conference on geographic information science, Heraklion, 2004.

Hammond, M. J. and Chen, A. S.: Urban flood impact assessment: A state-of-the-art review, Urban Water J., 12, 14–29, https://doi.org/10.1080/1573062X.2013.857421, 2015.

Heink, U. and Kowarik, I.: What are indicators? On the definition of indicators in ecology and environmental planning, Ecol. Indic., 10, 584–593, https://doi.org/10.1016/j.ecolind.2009.09.009, 2010.

Hinkel, J.: "Indicators of vulnerability and adaptive capacity": towards a clarification of the science-policy interface, Global Environ. Chang., 21, 198–208, https://doi.org/10.1016/j.gloenvcha.2010.08.002, 2011.

Holub, M. and Fuchs, S.: Benefits of local structural protection to mitigate torrent-related hazards, WIT Trans. Inf. Commun. Technol., 39, 401–411, https://doi.org/10.2495/RISK080391, 2008.

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., Gericke, A., Neal, J., Aerts, J. C. J. H., and Ward, P. J.: Comparative flood damage model assessment: towards a European approach, Nat. Hazards Earth Syst. Sci., 12, 3733–3752, https://doi.org/10.5194/nhess-12-3733-2012, 2012.

JRC and OECD (Joint Research Centre and Organisation for Economic Co-operation and Development): Handbook on constructing composite indicators: Methodology and user guide, Paris, OECD, 2008.

Kappes, M. S., Papathoma-Köhle, M., and Keiler, M.: Assessing physical vulnerability for multi-hazards using an indicator-based methodology, Appl. Geogr., 32, 577–590, https://doi.org/10.1016/j.apgeog.2011.07.002, 2012.

Keiler, M., Sailer, R., Jörg, P., Weber, C., Fuchs, S., Zischg, A., and Sauermoser, S.: Avalanche risk assessment – a multi-temporal approach, results from Galtür, Austria, Nat. Hazards Earth Syst. Sci., 6, 637–651, https://doi.org/10.5194/nhess-6-637-2006, 2006.

Kienberger, S., Lang, S., and Zeil, P.: Spatial vulnerability units – expert-based spatial modelling of socio-economic vulnerability in the Salzach catchment, Austria, Nat. Hazards Earth Syst. Sci., 9, 767–778, https://doi.org/10.5194/nhess-9-767-2009, 2009.

Klein, J. A., Tucker, C. M., Nolin, A. W., Hopping, K. A., Reid, R. S., Steger, C., Grêt-Regamey, A., Lavorel, S., Müller, B., Yeh, E. T., Boone, R. B., Bourgeron, P., Butsic, V., Castellanos, E., Chen, X., Dong, S. K., Greenwood, G., Keiler, M., Marchant, R., Seidl, R., Spies, T., Thorn, J., Yager, K., and the Mountain Sentinels Network: Catalyzing transformations to sustainability in the world's mountains, Earth's Futur., 7, 547–557, https://doi.org/10.1029/2018ef001024, 2019.

Komolafe, A. A., Adegboyega, S. A. A., and Akinluyi, F. O.: A review of flood risk analysis in Nigeria, Am. J. Environ. Sci., 11, 157–166, https://doi.org/10.3844/ajessp.2015.157.166, 2015.

Krellenberg, K. and Welz, J.: Assessing urban vulnerability in the context of flood and heat hazard: Pathways and challenges for indicator-based analysis, Soc. Indic. Res., 132, 709–731, https://doi.org/10.1007/s11205-016-1324-3, 2017.

Kundzewicz, Z. W., Su, B., Wang, Y., Wang, G., Wang, G., Huang, J., and Jiang, T.: Flood risk in a range of spatial perspectives – from global to local scales, Nat. Hazards Earth Syst. Sci., 19, 1319–1328, https://doi.org/10.5194/nhess-19-1319-2019, 2019.

Maiwald, H. and Schwarz, J.: Damage and loss prognosis tools correlating flood action and building's resistance-type parameters, Int. J. Saf. Secur. Eng., 5, 222–250, https://doi.org/10.2495/SAFE-V5-N3-222-250, 2015.

Maiwald, H. and Schwarz, J.: Vereinheitlichte Schadensbeschreibung und Risikobewertung von Bauwerken unter extremen Naturgefahren, Mauerwerk, 23, 95–111, https://doi.org/10.1002/dama.201910014, 2019.

Malgwi, M. B., Ramirez, J. A., Zischg, A., Zimmermann, M., Schürmann, S., and Keiler, M.: Flood reconstruction using field interview data and hydrodynamic modelling: A method for data scarce regions, Int. J. Disaster Risk Sci., submitted, 2020.

Mazzorana, B., Levaggi, L., Keiler, M., and Fuchs, S.: Towards dynamics in flood risk assessment, Nat. Hazards Earth Syst. Sci., 12, 3571–3587, https://doi.org/10.5194/nhess-12-3571-2012, 2012.

Mazzorana, B., Simoni, S., Scherer, C., Gems, B., Fuchs, S., and Keiler, M.: A physical approach on flood risk vulnerability of buildings, Hydrol. Earth Syst. Sci., 18, 3817–3836, https://doi.org/10.5194/hess-18-3817-2014, 2014.

Merz, B., Kreibich, H., Thieken, A., and Schmidtke, R.: Estimation uncertainty of direct monetary flood damage to buildings, Nat. Hazards Earth Syst. Sci., 4, 153–163, https://doi.org/10.5194/nhess-4-153-2004, 2004.

Merz, B., Kreibich, H., Schwarze, R., and Thieken, A.: Review article "Assessment of economic flood damage", Nat. Hazards Earth Syst. Sci., 10, 1697–1724, https://doi.org/10.5194/nhess-10-1697-2010, 2010.

Merz, B., Kreibich, H., and Lall, U.: Multi-variate flood damage assessment: a tree-based data-mining approach, Nat. Hazards Earth Syst. Sci., 13, 53–64, https://doi.org/10.5194/nhess-13-53-2013, 2013.

Meyer, V., Kuhlicke, C., Luther, J., Fuchs, S., Priest, S., Dorner, W., Serrhini, K., Pardoe, J., McCarthy, S., Seidel, J., Palka, G., Unnerstall, H., Viavattene, C., and Scheuer, S.: Recommendations for the user-specific enhancement of flood maps, Nat. Hazards Earth Syst. Sci., 12, 1701–1716, https://doi.org/10.5194/nhess-12-1701-2012, 2012.

Milanesi, L., Pilotti, M., Belleri, A., Marini, A., and Fuchs, S.: Vulnerability to flash floods: a simplified structural model for masonry buildings, Water Resour. Res., 54, 7177–7197, https://doi.org/10.1029/2018WR022577, 2018.

Mirza, M. M. Q.: Climate change and extreme weather events: can developing countries adapt?, Clim. Policy, 3, 233–248, https://doi.org/10.1016/S1469-3062(03)00052-4, 2003.

Mosimann, M., Frossard, L., Keiler, M., Weingartner, R., and Zischg, A.: A robust and transferable model for the prediction of flood losses on household contents, Water, 10, 1596, https://doi.org/10.3390/w10111596, 2018.

Müller, A., Reiter, J., and Weiland, U.: Assessment of urban vulnerability towards floods using an indicator-based approach – a case study for Santiago de Chile, Nat. Hazards Earth Syst. Sci., 11, 2107–2123, https://doi.org/10.5194/nhess-11-2107-2011, 2011.

Naumann, T., Nikolowski, J., and Sebastian, G.: Synthetic depth-damage functions – A detailed tool for analysing flood resilience of building types, edited by: Pasche, E., Evelpidou, N., Zevenbergen, C., Ashley, R., and Garvin, S., Road map towards a flood resilient urban environment, Institut für Wasserbau der TU Hamburg-Harburg, Hamburg, 2009.

Neubert, M., Naumann, T., and Deilmann, C.: Synthetic water level building damage relationships for GIS-supported flood vulnerability modeling of residential properties, edited by Samuels, P., Huntington, S., Allsop, W., and Harrop, J., Flood risk management. Research and practice, London, Taylor & Francis, 1717–1724, https://doi.org/10.1201/9780203883020.ch203, 2008.

Niang, I., Ruppel, O. C., Abdrabo, M. A., Essel, A., Lennard, C., Padgham, J., and Urquhart, P.: Africa, in Climate Change 2014: Impacts, Adaptation and Vulnerability: Part B: Regional Aspects: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Barros, V. R., Field, C. B., Dokken, D. J., Mastrandrea, M. D., and Mach, K. J., Cambridge University Press, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1199–1266, 2015.

NRE (Department of Natural Resources and Environment): Rapid appraisal method (RAM) for floodplain management, Department of Natural Resources and Environment, Victoria, 2000.

Ntajal, J., Lamptey, B. L., and MianikpoSogbedji, J.: Flood vulnerability mapping in the lower mono river basin in Togo, West Africa, Int. J. Sci. Eng. Res., 7, 1553–1562, 2016.

Papathoma-Köhle, M., Kappes, M., Keiler, M., and Glade, T.: Physical vulnerability assessment for alpine hazards: State of the art and future needs, Nat. Hazards, 58, 645–680, https://doi.org/10.1007/s11069-010-9632-4, 2011.

Papathoma-Köhle, M., Gems, B., Sturm, M., and Fuchs, S.: Matrices, curves and indicators: A review of approaches to assess physical vulnerability to debris flows, Earth-Sci. Rev., 171, 272–288, https://doi.org/10.1016/j.earscirev.2017.06.007, 2017.

Papathoma-Köhle, M., Cristofari, G., Wenk, M., and Fuchs, S.: The importance of indicator weights for vulnerability indices and implications for decision making in disaster management, Int. J. Disaster Risk Re., 36, 101103, https://doi.org/10.1016/j.ijdrr.2019.101103, 2019.

Papathoma, M., Dominey-Howes, D., Zong, Y., and Smith, D.: Assessing tsunami vulnerability, an example from Herakleio, Crete, Nat. Hazards Earth Syst. Sci., 3, 377–389, https://doi.org/10.5194/nhess-3-377-2003, 2003.

Penning-Rowsell, E., Johnson, C., Tunstall, S., Tapsell, S., Morris, J., Chatterton, J., and Green, C.: The benefits of flood and coastal risk management: a manual of assessment techniques, Middlesex University Press, Middlesex, 2005.

Percival, S., Gaterell, M., and Teeuw, R.: Urban neighbourhood flood vulnerability and risk assessments at different diurnal levels, J. Flood Risk Manage., 12, 1–14, https://doi.org/10.1111/jfr3.12466, 2018.

Pistrika, A., Tsakiris, G., and Nalbantis, I.: Flood depth-damage functions for built environment, Environ. Process., 1, 553–572, https://doi.org/10.1007/s40710-014-0038-2, 2014.

Quevauviller, P.: Science and Policy Interfacing. In Hydrometeorological Hazards, edited by: Quevauviller, P., https://doi.org/10.1002/9781118629567.ch1d, 2014.

Romali, N. S., Sulaiman, M. A. K., Yusop, Z., and Ismail, Z.: Flood damage assessment: A review of flood stage–damage function curve, edited by: Abu Bakar, S., Tahir, W., Wahid, M., Mohd Nasir, S., and Hassan, R., ISFRAM 2014, Singapore, Springer, 147–159, 2015.

Saaty, T. L.: The Analytical Hierarchy Process, McGraw-Hill, New York, 1980.

Sadeghi-Pouya, A., Nouri, J., Mansouri, N., and Kia-Lashaki, A.: An indexing approach to assess flood vulnerability in the western coastal cities of Mazandaran, Iran, Int. J. Disaster Risk Re., 22, 304–316, https://doi.org/10.1016/j.ijdrr.2017.02.013, 2017.

Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F., and Merz, B.: How useful are complex flood damage models?, Water Resour. Res., 50, 3378–3395, https://doi.org/10.1002/2013WR014396, 2014.

Schwarz, J. and Maiwald, H.: Prognose der Bauwerksschädigung unter Hochwassereinwirkung, Bautechnik, 84, 450–464, https://doi.org/10.1002/bate.200710039, 2007.

Schwarz, J. and Maiwald, H.: Damage and loss prediction model based on the vulnerability of building types, 4th International Symposium on Flood Defence: Managing Flood Risk, Reliability and Vulnerability, Toronto, Ontario, Canada, 6–8 May 2008, 74-1–74-9, https://doi.org/10.13140/2.1.1358.3043, 2008.

Spekkers, M. H., Kok, M., Clemens, F. H. L. R., and ten Veldhuis, J. A. E.: Decision-tree analysis of factors influencing rainfall-related building structure and content damage, Nat. Hazards Earth Syst. Sci., 14, 2531–2547, https://doi.org/10.5194/nhess-14-2531-2014, 2014.

Sturm, M., Gems, B., Keller, F., Mazzorana, B., Fuchs, S., Papathoma-Köhle, M., and Aufleger, M.: Experimental analyses of impact forces on buildings exposed to fluvial hazards, J. Hydrol., 565, 1–13, https://doi.org/10.1016/j.jhydrol.2018.07.070, 2018a.

Sturm, M., Gems, B., Keller, F., Mazzorana, B., Fuchs, S., Papathoma-Köhle, M., and Aufleger, M.: Understanding impact dynamics on buildings caused by fluviatile sediment transport, Geomorphology, 321, 45–59, https://doi.org/10.1016/j.geomorph.2018.08.016, 2018b.

Sy, B., Frischknecht, C., Dao, H., Consuegra, D., and Giuliani, G.: Reconstituting past flood events: the contribution of citizen science, Hydrol. Earth Syst. Sci., 24, 61–74, https://doi.org/10.5194/hess-24-61-2020, 2020.

Tarbotton, C., Dominey-Howes, D., Goff, J. R., Papathoma-Köhle, M., Dall'Osso, F., and Turner, I. L.: GIS-based techniques for assessing the vulnerability of buildings to tsunami: current approaches and future steps, Geol. Soc. London, Spec. Publ., 361, 115–125, https://doi.org/10.1144/SP361.10, 2012.

Tate, E.: Social vulnerability indices: A comparative assessment using uncertainty and sensitivity analysis, Nat. Hazards, 63, 325–347, https://doi.org/10.1007/s11069-012-0152-2, 2012.

Thieken, A. H., Müller, M., Kreibich, H., and Merz, B.: Flood damage and influencing factors: New insights from the August 2002 flood in Germany, Water Resour. Res., 41, 1–16, https://doi.org/10.1029/2005WR004177, 2005.

Thieken, A. H., Olschewski, A., Kreibich, H., Kobsch, S., and Merz, B.: Development and evaluation of FLEMOps - A new Flood Loss Estimation MOdel for the private sector, WIT Trans. Ecol. Envir., 118, 315–324, https://doi.org/10.2495/FRIAR080301, 2008.

Thouret, J. C., Ettinger, S., Guitton, M., Santoni, O., Magill, C., Martelli, K., Zuccaro, G., Revilla, V., Charca, J. A., and Arguedas, A.: Assessing physical vulnerability in large cities exposed to flash floods and debris flows: the case of Arequipa (Peru), Nat. Hazards, 73, 1771–1815, https://doi.org/10.1007/s11069-014-1172-x, 2014.

Totschnig, R. and Fuchs, S.: Mountain torrents: quantifying vulnerability and assessing uncertainties, Eng. Geol., 155, 31–44, https://doi.org/10.1016/j.enggeo.2012.12.019, 2013.

Totschnig, R., Sedlacek, W., and Fuchs, S.: A quantitative vulnerability function for fluvial sediment transport, Nat. Hazards, 58, 681–703, https://doi.org/10.1007/s11069-010-9623-5, 2011.

UNDP (United Nations Development Programme): Human Development Report, reviewed by: Todaro, M. P, in: Population and Development Review, Popul. Counc., 18, 359–363, https://doi.org/10.2307/1973685, 1992.

UNDRR (United Nations Office for Disaster Risk Reduction): Global Assessment Report on Disaster Risk Reduction 2019, Geneva, Switzerland, available at: https://gar.undrr.org/ (last access: 29 March 2020), 2019.

UNISDR (United Nations International Strategy for Disaster Reduction): Terminology on disaster risk reduction, United Nations International Strategy for Disaster Reduction Geneva, Geneva, Switzerland, available at: https://www.undrr.org/publication/2009-unisdr-terminology-disaster-risk-reduction (last access: 15 March 2020), 2009.

UNISDR (United Nations International Strategy for Disaster Reduction): Sendai framework for disaster risk reduction 2015–2030, available at: https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030 (last access: 26 February 2020), 2015.

Vogel, K., Riggelsen, C., Merz, B., Kreibich, H., and Scherbaum, F.: Flood damage and influencing factors: a Bayesian network perspective, in: Proceedings of the 6th European workshop on Probabilistic Graphical Models (PGM 2012), Granada, Spain, edited by: Cano, A., Gómez-Olmedo, M. G., and Nielsen, T. D., 19–21 September 2012, 314–354, 2012.

Wagenaar, D., de Jong, J., and Bouwer, L. M.: Multi-variable flood damage modelling with limited data using supervised learning approaches, Nat. Hazards Earth Syst. Sci., 17, 1683–1696, https://doi.org/10.5194/nhess-17-1683-2017, 2017.

Walliman, N., Ogden, R., Baiche, B., Tagg, A., and Escarameia, M.: Development of a tool to estimate individual building vulnerability to floods, WIT Trans. Ecol. Envir., 155, 1005–1016, https://doi.org/10.2495/SC120842, 2011.

WHO (World Health Orgaonization): WHO Guidelines for indoor air quality: dampness and mould, Copenhagen, Denmark, available at: https://www.who.int/airpollution/guidelines/dampness-mould/en/ (last access: 2 February 2020), 2009.

Yankson, P. W. K., Owusu, A. B., Owusu, G., Boakye-Danquah, J., and Tetteh, J. D.: Assessment of coastal communities' vulnerability to floods using indicator-based approach: a case study of Greater Accra Metropolitan Area, Ghana, Nat. Hazards, 89, 661–689, https://doi.org/10.1007/s11069-017-3006-0, 2017.

Zimmermann, M. and Keiler, M.: International frameworks for disaster risk reduction: Useful guidance for sustainable mountain development, Mt. Res. Dev., 35, 195–202, https://doi.org/10.1659/MRD-JOURNAL-D-15-00006.1, 2015

Chapter 3: Flood scenario reconstruction using field interview data and hydrodynamic modelling: A method for data-scarce regions

Mark Bawa Malgwi[1,2*], Jorge Alberto Ramirez[1,2], Andreas Zischg[1,2,3], Markus Zimmermann[1,2], Stefan Schürmann [1,2,3], Margreth Keiler[1,2,3]

[1] University of Bern, Institute of Geography, Hallerstrasse 12, 3012 Bern, Switzerland

[2] University of Bern, Oeschger Centre for Climate Change Research, Hochschulstrasse 4, 3012 Bern, Switzerland

[3] University of Bern, Mobiliar Lab for Natural Risks, Hallerstrasse 12, 3012 Bern, Switzerland

Author contributions: MBM: Conceptualizing, Methodology, Formal Analysis, Investigation, Writing – Original draft, Visualization, Writing – Review and editing. JAR: Conceptualizing, Methodology, Formal Analysis, Investigation, Visualization, Writing – Review and editing. AZ: Conceptualizing, Methodology, Writing – Review and editing. MZ: Validation, Writing – Review and editing. SS: Validation, Visualization, Writing – Review and editing. MK: Supervision, Funding acquisition, Writing – Review and editing.

Under review in *International Journal of Disaster Risk Science*

Submitted on: 03 July, 2020.

Abstract

The scarcity of input and calibration data have limited efforts in reconstructing scenarios of past floods in data-scarce regions. Although recently, the number of studies that use distributed flood observation data collected throughout flood-affected communities (e.g., photos and texts from social media or face-to-face interviews) are increasing, a systematic method that applies such data for hydrodynamic modelling of past floods is lacking. In this study, we developed a method for reconstructing plausible scenarios of past flood events in data-scarce regions by applying flood observation data collected through field interviews to a hydrodynamic model CAESAR-Lisflood. The method implements a stepwise iteration that aims to systematically minimize the root mean square error (RMSE) between modelled and observed flood depth and duration at building locations. We tested the method using 300 spatially distributed flood depths and duration data collected using questionnaires on five river reaches after the 2017 flood event in Suleja/Tafa region, Nigeria. Results from the reconstructed flood depth scenario produced a RMSE of 0.61 m for all observed and modelled locations and lies in the range of error produced by similar studies using comparable hydrodynamic models. The study demonstrates the potential of utilizing interview data for hydrodynamic modelling applications in data-scarce regions. Extrapolated data from the reconstructed flood event, i.e., flow depths and durations at houses without observations, will provide useful input data for physical vulnerability assessment (e.g., developing flood damage models) to complement disaster risk reduction efforts.

Keywords: hydrodynamic modelling, interviews, data-scarce, floods, peak discharge, flood duration

1 Introduction

Floods are likely the most recurrent and destructive disaster worldwide (Teng et al., 2017). Consequently, it is becoming more important to develop techniques to help strategize against floods and to reduce associated impacts (Sarhadi et al. 2012; Smith et al. 2019; Wang et al. 2018). For example, the recent occurrence of two extreme flood events produced by cyclones Idai and Kenneth (USAID, 2019) with their resulting high human and economic losses in Southern Africa, have reemphasized the need to further investigate risks related to floods especially in regions with limited coping capacity, which are mostly data-scarce. Generally, a first step to investigate past flood events is to simulate flow characteristics (e.g., water depth, duration, velocity) using hydrodynamic models. Hydrodynamic models have enabled a better evaluation of the effects and dynamics of past floods, hence providing inputs for assessing the physical vulnerability of the built environment (e.g., developing flood damage models) (Fuchs et al., 2019). In particular, flood characteristics (water depth, velocity, duration) derived from hydrodynamic models provide reliable and sizeable input data at building locations for developing robust flood damage models (Chow et al. 2019; Wagenaar et al. 2017).

The use of hydrodynamic models to reconstruct past flood events typically require i) input data such as high-resolution Digital Elevation Models (DEM) to represent floodplain topography and river channel bathymetry (river width, depth and bed slope), and ii) calibration data such as river discharge. While input data are generally required for model set up, calibration data are used to tune model parameters for a specific study location and period. Although these input and calibration data are mostly available in data-rich regions (Schumann et al., 2015), they are rarely available in data-scarce regions and this has limited progress in simulating floods in many parts of the world (Komi et al., 2017; Sanyal et al., 2013; Schumann et al., 2015; Yan et al., 2014). Focusing on input data, several studies have developed methods for using globally available data as input and calibration data for hydrodynamic models. In particular, the studies explore the use of low-resolution DEMs (e.g., Jung et al. 2010; Rolim et al. 2011; Neal et al. 2012; Yan et al. 2014), while others have a focus on improving poorly represented channel bathymetry (e.g., Gichamo et al. 2012; Neal et al. 2015). These studies have provided very useful methods to overcome some of the limitations stemming from data-scarcity and have consequently improved the accuracy of hydrodynamic models in data-scarce areas. However, despite current efforts, difficulties still exist in the application of these methods in small to medium scale (50 – 100 km^2) data-scarce catchments. For example, while developed methods that use globally available DEMs require input data (e.g., precipitation or discharge data) or are developed for large-scale (more than 100 km^2 catchment) applications, studies aimed at characterizing river channel bathymetry require upstream or downstream flow or stage measurements (Gichamo et al., 2012) - which are often unavailable in many data-scarce regions. Consequently, a need to refine current methods for better representation of channel and flood plain persist.

Hunter et al. (2005) highlighted the potential of utilizing distributed post-event observations for calibrating hydrodynamic models. Observation data can be water depths or extents measured after a flood and can be

retrieved from gauging stations or observed as marks on buildings or trees. For example, in a study by Neal et al. (2009), distributed water and wrack marks (263 points) collected immediately after the 2005 Carlisle flood (UK) were used to reconstruct the flood event using a hydrodynamic model. Similarly, other studies that use post-event observations include Bronstert et al. (2018) and Borga et al. (2019) using distributed water level data, Zischg et al. (2018c) and Bernet et al. (2019) using geolocalized insurance claims. Recently, Wang et al. (2018) demonstrated an extended application of post-event data for model calibration by using documented reports and photos shared by the flooded community through social media feeds. Such an approach, using knowledge from affected communities, are particularly encouraged in data-scarce regions given their potential to provide low cost and sizeable data with good spatial and temporal coverage (Sy et al., 2019, 2020; Assumpção et al., 2018). An earlier study by Tran et al. (2009) also recommended the use of data retrieved from flooded communities given they are either victims or eye-witnesses of the events and have a good understanding of their environment.

Recent reviews by Assumpção et al. (2018) and Sy et al. (2019) showed that studies that utilize data, collected from affected communities, to reconstruct past flood scenario are increasing. For example, several studies have simulated past floods using either water depths and velocity retrieved from texts, pictures and videos uploaded to social media platforms (e.g., Smith et al. 2017) or flood duration derived using interviews (e.g., Sy et al. 2016). Since data retrieved from communities may require extensive pre-processing (e.g., geo-referencing) or susceptible to errors since they are reported by non-experts, some studies (McDougall and Temple-Watts 2012; McDougall 2011) have suggested field visits to validate reported observations. Focusing on field visits, several researchers have integrated face-to-face interviews for the collection of post-event observations (e.g., Poser and Dransch 2010; Singh 2014; Sy et al. 2016, 2020). One of the most recent advancement in mapping past-floods in typical data-scarce regions using interviews was carried out by Sy et al. (2020) in Yeumbeul North, Senegal. In the study, flood depths and extents were retrieved for multiple flood events (between 2005, 2009, 2012) based on personal recollections using two (independent) sets of community representatives. The use of two sets of community representatives allowed a check on the consistency of the information provided and showed that i) distributed water depths from 64 sites were similar (with a maximum difference of 0.23 m), and ii) a good spatial agreement for flooded areas was observed (with a maximum difference within 1.8 %). In addition, Sy et al. (2020) observed a good agreement after comparing flood spatial extents derived with a community based approach and remotely sensed data. Given that current methods that use interview data are limited to GIS-based mapping of floods (e.g., Musungu et al. 2012; Singh 2014; Sy et al. 2016; Canevari-Luzardo et al. 2017, Sy et al. 2020), Assumpção et al. (2018) and Sy et al. (2020) have recommended exploring new alternatives in integrating such data into hydrodynamic models for flood hazard assessment especially in data-scarce regions. Moreover, Assumpção et al. (2018) noted that current methods require modification to be tailored for specific user needs, for example, data collected from interviews can be applied for flood damage estimation as demonstrated by Poser and Dransch (2010).

Herein we present a method that leverages the use of post-event observations collected through field interviews, and a hydrodynamic model to reconstruct a plausible scenario of a past flood event in a typical data-scarce location. The model output is intended to enrich data sets to develop flood damage models. For example, the method can extrapolate from a sample of buildings for which flood depths are known to the remaining buildings affected by the flood event. Given that i) reconstructing past flood scenarios depends on the optimal agreement between observed and modelled characteristics, and ii) flood damage models require reliable estimates of flood characteristics at building locations, the modelling approach developed uses several iterative rounds of simulations to minimize the error (root mean squared error (RMSE)) between modelled and observed flood depths and durations at building locations. The use of a hydrodynamic model allows the characterization of channel and floodplain features (e.g., Manning coefficient) and interactions between multiple reaches, which cannot be implemented using current GIS approaches. The study provides a framework that integrates a globally available DEM, reconstructed river channel, and synthetic hydrographs as model initial conditions and input data. As such, our study represents one of the first applications using house-to-house field data collected using questionnaires to develop a hydrodynamic model. Herein we explain and test the developed method in a typical data-scarce small to medium scale ungauged catchment with five river reaches.

2 Study area and data

2.1 Study region

The study area is located in Suleja and Tafa local government areas in Niger State, Nigeria (Figure 1). The region is situated in the central part of the country; an area typically characterized by several inland rivers. There are five major reaches in the study area (Figure 1). Reaches 1, 4 and 5 are located in Suleja while reaches 2 and 3 are located in Tafa. The river channels are semi-natural with no flood protection work and partly characterized by grassland vegetation or sediments. All river channels are mostly shallow (≤ 2 m) and have a frequently varying cross-section. The main river traversing the study area is the River Iku and has a reservoir upstream as shown in Figure 1. A summary of reach characteristics based on field observation and the World Imagery (ESRI et al., 2020) on the ArcGIS platform is presented in Table 1. Qualitative characterization of the channel and floodplain as shown in Table 1 were partly based on methods in Phillips and Tadayon (2006).

A 30 m resolution map of the built-up areas in the location can be found in Figure 1. The built-up areas were compiled under the Global Human Settlement (GHS) project of the European Commission (Pesaresi et al., 2013). It includes multi-temporal built-up layers derived from Landsat image collections from Global Land Survey

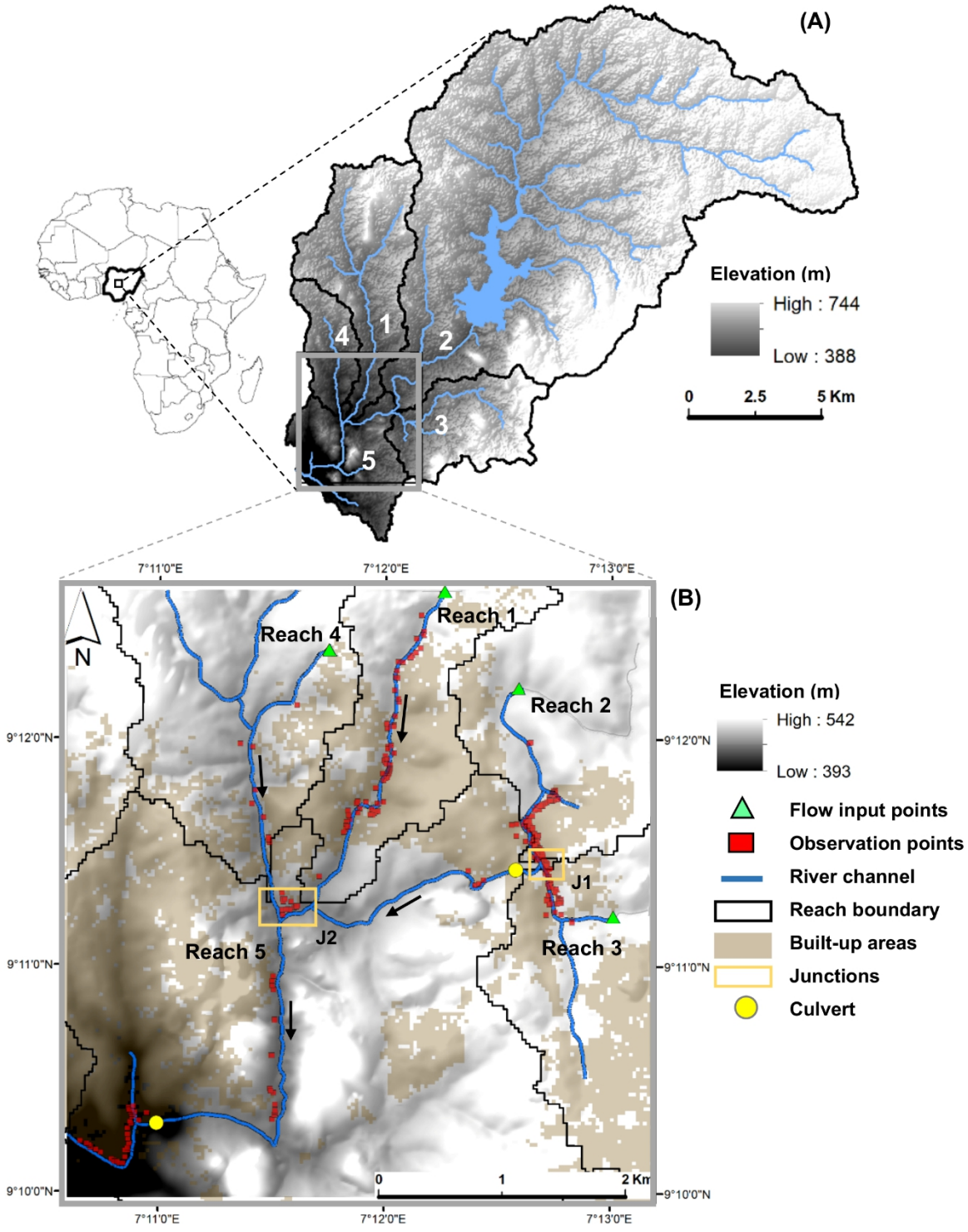


Figure 1: WorldDEM DTM of study region showing (A) catchments 1 - 5 (B) inspected buildings and model flow input points, junctions, built-up areas with arrows indicating river flow direction. Inset map indicate location of Suleja and Tafa area, Nigeria.

Tables 1: Reach characteristics in the study region

Location	Reach	Catchment area (Km^2)	Reach characteristics	
			Channel	**Flood plain**
Upstream	Reach 1	25	Moderate vegetation and sediments	U-shaped valley and high number of closely spaced buildings
	Reach 2	160 (12 downstream reservoir)	Minor vegetation and sediments	Relatively flat and high number of buildings on one side of the channel
	Reach 3	17	Moderate vegetation and sediments	Relatively flat and moderate number of buildings
	Reach 4	8	Mostly vegetated	Relatively flat mostly vegetated. Sparse buildings widely spaced
Downstream	Reach 5		Moderate vegetation, sediments and boulders	U-shaped valley. Moderate to high number of buildings especially around channel outlet

(GLS) between 1975 and 2000, and ad-hoc Landsat 8 collection 2013/2014 (for more information, see Florczyk et al. (2019)). A 100 m spatial resolution land cover map from 2018 (Buchhorn et al., 2020) indicates the catchment encompassing the study area is comprised of 44% agriculture, 34% forest, 10% urban, 9% shrubs, and 3% other. Even though some studies have identified this region as a low flood risk area (e.g., Mayomi et al. 2014; Ndanusa et al. 2018), the magnitude of a recent flood event has indicated otherwise.

2.2 Flood event and data availability

In 2017, from 8th - 9th of July, prolonged heavy rainfall caused severe flooding in Suleja/Tafa region. According to local reports, the floods lasted for about 12 hours and led to the death of over fifteen people (OCHA, 2019). Additionally, hundreds of residential houses and infrastructural facilities were damaged by the floods (Adeleye et al., 2019). Reports from locals in the community indicated that the entire flood was a one-day event (from early morning on 9th July to late evening) and the flood approached at high velocities and resulted in multiple wave-like surges. There were also reports of multiple backwater effects at two river junctions marked junction 1 (J1) and junction 2 (J2) in Figure 1. Additionally, local reports indicated that the

blockage of a bridge culvert by transported materials (tree trunks and cars) close to J1, caused a damming of the floodwater at surrounding areas and resulted in ponding that increased water depths and durations.

For our study, we considered replicating the Suleja/Tafa flood event using hourly precipitation to drive a catchment scale hydrological model that produces hourly discharge for input into a detailed reach scale model of the flooded urban area. However, hourly observed precipitation before and during the flood (July 7-9, 2017) is not available because no weather station exists in the Suleja/Tafa region. Recent advances in reanalysis (Hersbach et al., 2018), which combines model data with observations, were also considered as a source for precipitation data. Specifically, we acquired precipitation from the ERA5-Land reanalysis dataset (Muñoz-Sabater et al., 2018) with a relatively fine spatial (9 km) and temporal (1 hr) resolution. However, the ERA5-Land precipitation did not reproduce the storm event that caused flooding in the Suleja/Tafa region. Instead, the modelled precipitation prior to and during (July 7-9, 2017) the flood is low in intensity (< 12 mm d^{-1}). In addition to the unavailability of local precipitation data, no discharge records exist for the rivers in the study area. As such, it would prove difficult to calibrate rainfall-runoff processes in a catchment scale hydrological model.

2.3 Topographical data

A WorldDEM DTM tile from the TanDEM-X database was acquired from the European Space Agency. The WorldDEM DTM is a bare earth surface without vegetation and man-made objects and has a spatial resolution of 12 m, an absolute vertical accuracy ≤ 10 m, and an absolute horizontal accuracy ≤ 6 m (Archer, 2018). Figure 1 shows the WorldDEM DTM of the study region from which high and low elevation areas within the study region can be visualized. The WorldDEM DTM data set can be acquired free of charge for scientific purpose.

3 Methods

3.1 Data collection

A field investigation was carried out eight months after the flood event. An initial mapping of the affected areas included the analysis of media reports, videos, and photos from different sources. The questionnaire for interviews was developed covering information relating to flood characteristics, depth, duration, qualitative description of velocity, and GPS coordinates of the affected buildings. An extensive house-to-house survey was conducted from March to May 2018, within the initially mapped area. If watermarks were still existing on building walls or block fences, the water depth was measured in situ with a measuring tape. If watermarks were no longer visible, building occupants were asked based on their recollection. Observations were collected using the developed questionnaire (attached in supplementary material). The main criteria for selecting a building for an interview was depending on if it was flooded. The selection was further restricted by the availability of residents or a community representative to provide the required flood information. Many

house inhabitants, especially along reaches 2, 3 and 5, reported having experienced two flood waves within 24 hrs. The second wave is suspected to have been influenced by the reservoir upstream reach 2. From an interview with the reservoir personnel, it was gathered that: i) the spillways were closed during the entire flood event, and ii) discharge from the reservoir, which was reported to arrive many hours after the first flood, was mainly due to the overtopping of the spillway. Since no further information was available from the reservoir management regarding the commencement, duration and estimated discharge of the overtopping, the contribution of the second flood wave was not considered during the modelling.

3.2 Data pre-processing

Before using the WorldDEM DTM, several preprocessing steps were carried out. Firstly, in order to represent sections of the river with small widths (less than the DEM resolution of 12 m), the DEM was resampled to 4 m x 4 m grid cells. We generated a flow accumulation layer using the DEM, however, the generated river network from flow accumulation did not properly align with the actual river channel digitized from a World Imagery (ESRI et al., 2020) satellite image available on the ArcGIS platform. For a selected area in the study region, we show a typical deviation of the channel generated from flow accumulation with the actual channel from the satellite image (Figure 2A). A simple procedure was implemented to reduce the channel deviation by reconstructing a new river channel. First, cross-sectional lines with a downstream distance spacing between 5 m (where channel width is frequently irregular) to 30 m (where channel width is averagely consistent) were manually drawn across the digitized channel centerline extracted from the World Imagery satellite image (see Figure 2A). A total of 362 cross-sections were produced for extracting channel width in order to systematically capture the channel planform: this has been shown by Cook and Merwade (2009) to improve modelling results. Two data sources were used to extract information on channel widths that were assigned to each cross-section. Firstly, we use documented photographs from fieldwork. Secondly, we use a remotely sensed image within the World Imagery archive that was collected by a DigitalGlobe satellite on the 10th January 2017 with a resolution of 0.5 m and a horizontal positional accuracy of 10.2 m. The extracted widths were rounded up to the nearest 4 m for all digitized cross-sections to maintain channel widths that are multiples of the cell size of the resampled DEM. Thereafter, we use the WorldDEM DTM to generate a contour map of the study region and manually digitize the centerline of the river with the aid of the World Imagery satellite image (Figure 2B). This centerline was preferred over the flow accumulation centerline because it better represented the actual channel location in the satellite imagery, whilst a flow accumulation centerline was disregarded because it contained significant channel deviations stemming from the required DEM reconditioning (e.g., DEM pit filling). At this stage, the digitized centerline was segmented using the 5 – 30 m distance spacing and each segment was assigned a width

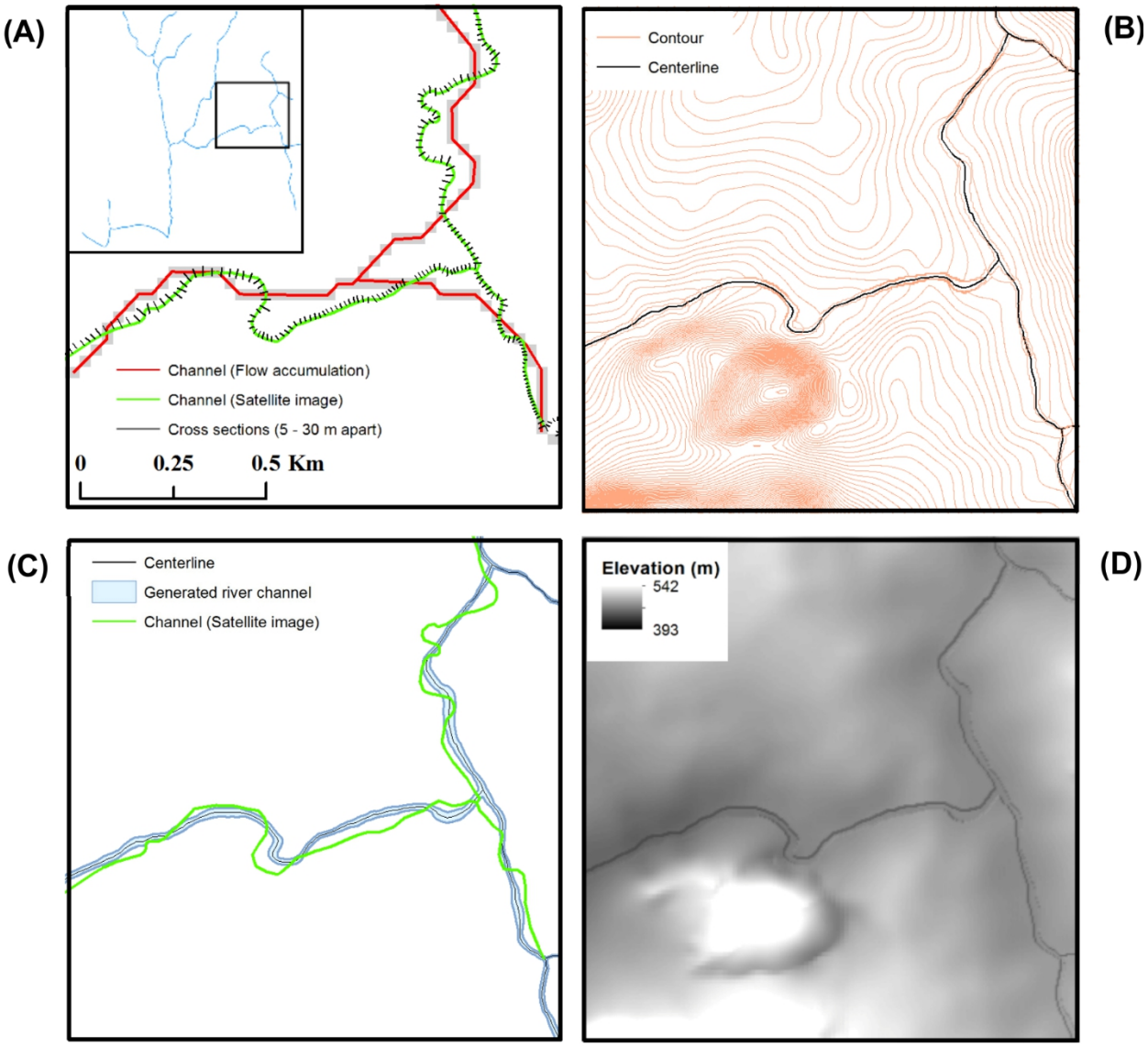


Figure 2: Procedure for reconstructing river channel; (A) deviation of channel derived from flow accumulation of WorldDEM DTM with digitized channel from high-resolution Esri satellite imagery. Cross sections of 5 – 30 m drawn for width extraction (B) digitized channel center using generated contour map from WorldDEM DTM (we use 1 m contour interval for visualization purpose) (C) generated (reconstructed) river channel (shows improvement in channel deviation compared to Fig. 2a) (D) 1m drop of generated river channel.

value that corresponded to the nearest cross-section. Afterwards, a new channel outline was produced by buffering the centerline with a distance equal to the width value of each segment (Figure 2C). The resulting channel outline is a compromise between the positional and dimensional representation of the river. Moreover, this method reduces the spatial deviation between the channel generated from DEM flow accumulation

(Figure 2A) and the channel from the contour map (Figure 2C). In the next step, DEM locations that spatially coincided with the newly developed channel outline were reduced 1 m in elevation. This produced a simple rectangular cross-sectional profile (Figure 2D). In a final DEM pre-processing step, channel blockages existing in the DEM (e.g., bridges and culverts) were identified, using satellite imagery, as anomalies in the channel downstream elevation profiles. Blockages in the DEM were removed using the open source RasterEdit tool (https://sourceforge.net/projects/caesar-lisflood/files/) to ensure unobstructed routing of water downstream. Observation points collected from the field survey were georeferenced and converted to a 4 m x 4 m raster to correspond with the spatial resolution of the resampled DEM. Each georeferenced observation point contained information on observed flood depth and duration. Buildings were not extruded on the DEM considering that most houses were built close to each other (< 1 m) and this would require a DEM with a finer spatial resolution (< 1 m) to represent spacing that permits water flow between houses. Thus the representation of buildings would drastically increase the number of raster cells in our study area and significantly increase computational overhead.

3.3 Model selection and build

The proposed procedure for reconstructing the flood event of 8th - 9th of July, 2017 in the Suleja and Tafa is principally based on iteratively changing the boundary conditions (i.e. inflow hydrographs) and model parameters (i.e. roughness) of a hydrodynamic model until modelled and observed flow depths are similar at inundated building locations. In addition to flow depth, our method considers temporal aspects of the flood because the river network consists of five different river reaches and thus the superimposition of flood waves from the upstream tributaries ultimately determines the flood in the downstream river reach. Consequently, we consider a method postulated by Pattison et al. (2014) which demonstrates the effect of upstream channel space and time shifts on downstream catchments.

For this study, we used the CAESAR-Lisflood model (Coulthard et al. 2013). CAESAR-Lisflood is the output of an innovative integration of a non-steady hydrodynamic model LISFLOOD-FP 2D (Bates and De Roo, 2000) and the landscape evolution model CAESAR (Coulthard et al. 2002). CAESAR-Lisflood is a full 2D model that routes water and sediment over a regular grid of cells representing channel and floodplain topography. In CAESAR-Lisflood, sediment transport can be optionally turned off, as in our study, and the model essentially operates like LISFLOOD-FP by simulating two-dimensional flow with a simplified form of the shallow water equation (for more information on governing equations see Coulthard et al. 2013). CAESAR-Lisflood operates with minimum parameterization (only Manning's coefficient) and is open source (https://sourceforge.net/projects/caesar-lisflood/). In our study we used CAESAR-Lisflood version 1.9g. Input data for CAESAR-Lisflood are a DEM, discharge, and friction values for channel and floodplain locations. Discharge is added to the DEM at specified locations (see Figure 1 on reaches 1 – 4). Smaller tributaries flow into the channels on reaches 2, 3 and 4 (see Figure 1), however, these additional tributaries were not included in the modelling due to uncertainty associated with rationing input discharge.

The modelling was carried out in four main steps; i) model sensitivity analysis of roughness parameters and determination of optimum peak discharge, ii) investigating optimum duration, iii) investigating the optimum combination of extracted upstream hydrographs (reaches 1 - 4) to match optimum discharge downstream (reach 5) and partly account for the spatiotemporal heterogeneity of the rainfall within the catchment, and iv) simulating the constructed flood event over the whole river network in the model domain. A systematic description of the modelling steps are shown in a flow chart in Figure 3. The first two steps were carried out for each reach separately; consequently, observations at the river junctions (J1 and J2 - see Figure 1) were not included. Running separate simulations for each reach also reduced computational demand since we use a 4 m resampled DEM with about 1.46 million grid cells. The processes of evaporation or infiltration were not directly accounted for in the model due to a lack of data for calibration. Additionally, leveraging the availability of CAESAR-Lisflood source code, modifications were made to the model inputs and outputs. Changes to the model included the uploading of surveyed building locations and outputting to file, for every 10 minutes of a model run, the corresponding time step, input and output discharge on each reach, as well as the simulated water depths at each building location. To allow model errors to penalize for underprediction of flood extent, all observation points, flooded or not, were included in calculating model errors.

3.3.1 Simulation round 1: Model calibration and extraction of optimum peak discharge

The first round of simulation aimed to test the sensitivity of the hydraulic model to the Manning's coefficient (n). The n value accounts for the roughness of either the channel (n_c) or floodplain (n_{fp}) and partly serves to control how water is routed from one grid cell to another in the model. The n to be selected for use in the model is not completely physically based, rather it allows the model to account for physical processes not fully replicated by the model (e.g., infiltration and backwater effects) (Coulthard et al. 2013). To account for a wide range of channel bed characteristics and floodplain land cover types, we select n values of 0.01 (cemented surface), 0.07 (medium shrubs and weeds or trees) 0.14 (dense shrubs and weeds or trees) (see Chow 1959). A similar approach has been used in several studies (e.g., Fewtrell et al. 2011; Neal et al. 2009; Wood et al. 2016) whereby a wide range of n values was selected for both channel and floodplain to ensure that the optimal value is bracketed. These three selected n values (0.01, 0.07, 0.14) will give nine combinations for different n_{fp} and n_c (Table 2). Hence, for five river reaches we have a total of 45 model runs for simulation round 1. Model calibration was carried out using distributed observed flood depths. The simulations were performed by applying a linear hydrograph (Figure 4A) at indicated flow input points (see Figure 1). A linearly increasing hydrograph with a duration of 24 hrs and a maximum discharge of 250 m^3s^{-1} was used to capture the maximum

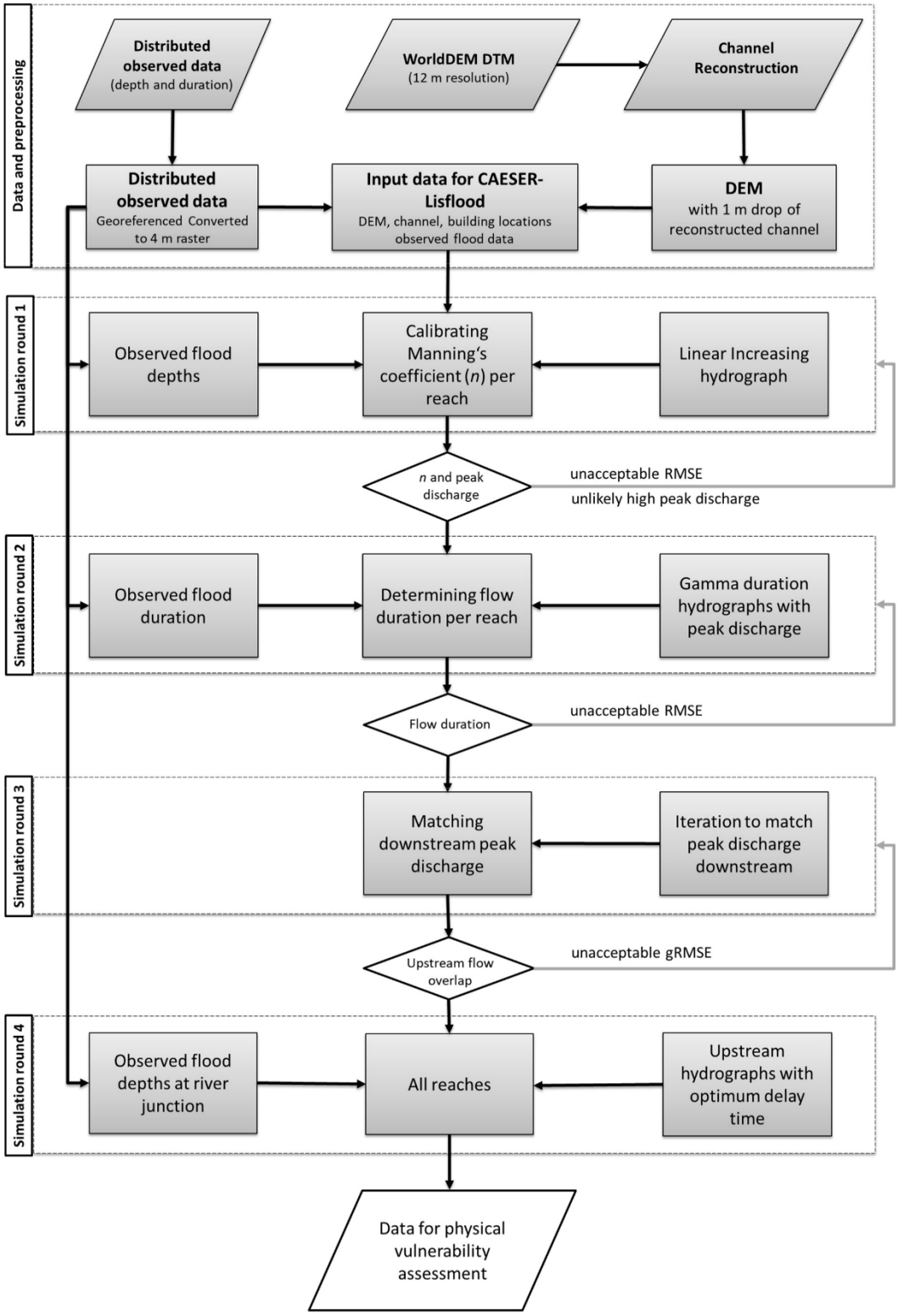


Figure 3: Flowchart of model development.

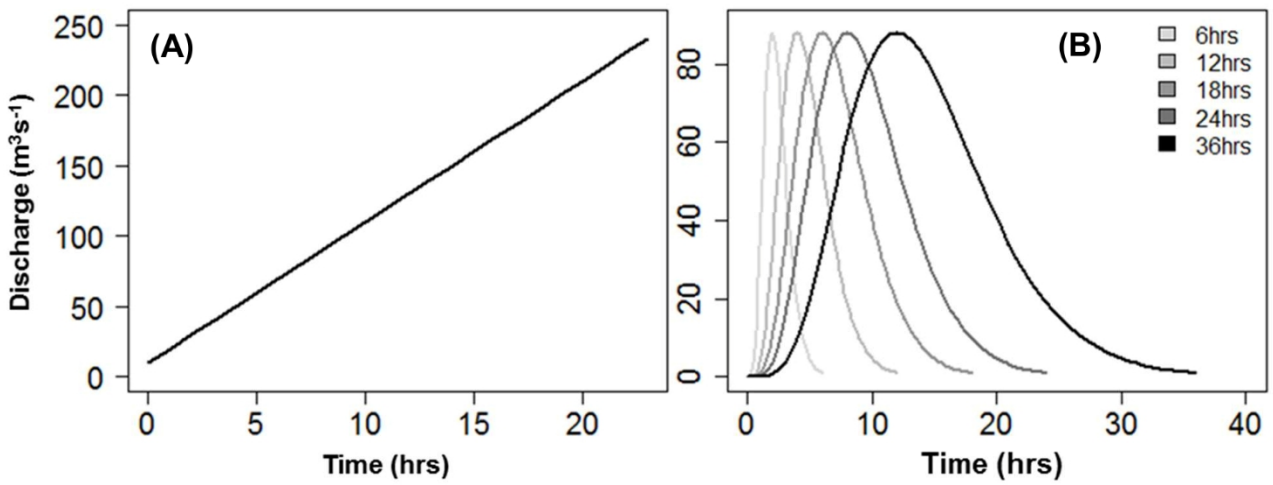


Figure 4: Input hydrographs to determine flood (A) peak and (B) duration. Hydrograph examples are provided for reach 1.

Table 2: Selected combinations of Manning's coefficient (*n*)

		River channel (n_c)		
		0.01	0.07	0.14
Flood plain	0.01	0.01, 0.01	0.01, 0.07	0.01, 0.14
(n_{fp})	0.07	0.07, 0.01	0.07, 0.07	0.07, 0.14
	0.14	0.14, 0.01	0.14, 0.07	0.14, 0.14

discharge at which simulated depths closely matched observed depths on each reach. For each simulation, model flood depths at each building location are output to file every 10 minutes. Since buildings in this region are mostly over 80 m^2 (determined from sampling 648 buildings in the area using remote sensing), recorded modelled flood depths are extracted from the exact building location (4 m x 4 m) and cells in the immediate neighbourhood of the building (12 m x 12 m). From these locations, the simulated water depth with the least difference to the observed building water depth was selected. This selection allows a spatial tolerance for model output since the exact orientation of the building footprint is not known. We compute a Root Mean Square Error (RMSE) between observed and modelled flood depth every 10 minutes of a model run. For each simulation, the input discharge was plotted against the RMSE to understand model error response to combinations of input discharge and roughness. For each river reach, the peak discharge value, corresponding to a minimum RMSE for a specific combination of *n*, is selected as the optimal discharge.

3.3.2 Simulation round 2: Investigating optimum duration

In this step, observed flood duration data is used to estimate the optimum duration of the flow on upstream reaches (1 - 4). Similar to a recent study by Zischg et al. (2018b), we adopt a two parametric gamma function to develop synthetic hydrographs with durations of 6, 12, 18, 24 and 36 hrs. The choice of utilizing a gamma function was because the catchments are: i) relatively small to medium size in area and ii) largely unforested (66%) with low amounts of rainfall interception and these characteristics more than likely produce a flashier hydrological response that is represented by the selected distribution (USDA, 2007; Welsh et al., 2009). For each upstream reach, the synthetic hydrographs are scaled to the peak discharge value determined from simulation round 1. For example, Figure 4B shows gamma duration hydrographs with the peak discharge of reach 1 (88 m^3 s^{-1}). Model runs were carried out for each upstream reach (reaches 1 – 4) using all five synthetic hydrographs; hence, 20 model runs were performed in total for simulation round 2. All simulations were carried out using a single n_{fp} and n_c combination deduced from simulation round 1. Model depths were written to a file after every 10 minutes of each model run. To compute simulated duration at each observation point, we sum up the total time water depth was ≥ 0.20 m at the grid cell. The use of 0.20 m is to allow a minimum building floor elevation (typical in this region) before water enters a building. Pappenberger et al. (2006) used a similar minimum depth, which was related to the observations of occupants when water was entering a building. An RMSE is computed to check the difference between the modelled and observed flood duration at the building locations. The synthetic hydrograph, that provides a minimum RMSE value, is selected as the optimum duration hydrograph.

3.3.3 Simulation round 3: Determining optimal downstream hydrograph

Reports from the field survey indicated that the upstream reaches (reach 1 - 4) did not flood downstream at the same time. This suggests that the downstream hydrograph (reach 5) is not a simple summation of upstream hydrographs across time. Moreover, the optimal peak discharge of reach 5 from simulation round 1 was higher than individual peak discharge on reaches 1 – 4; this further supports the idea of temporal overlap between upstream hydrographs. Without such temporal overlaps, the expected peak discharge on reach 5 will be the same with the upstream reach that has the maximum peak discharge. Neal et al. (2013) and Zischg et al. (2018a) recently emphasized the importance of considering the effects of the superimpositions from sub-catchments for flood risk analysis. To investigate these superimpositions, we carried out three steps to explore all possible combinations of the upstream hydrographs (reach 1 - 4, determined from simulation round 1 and 2) to match the optimum downstream peak discharge (reach 5, determined from simulation round 1). The steps carried out are as follows;

I. We constructed all possible permutation sets between the hydrographs of reaches 1 – 4 using a 24 hrs upper bound as the maximum duration because local reports suggest it was a one-day event. Maintaining an upper bound of 24 hrs was achieved by setting up a permutation algorithm such that temporal shifts for any individual reach hydrograph were possible between 1 to 6 hrs.

II. We established a condition to only select permutated sets where at least one hydrograph is not shifted in time. This condition was to constrain the temporal shifts to a specific event start time.

III. Thereafter, hydrographs from selected permutations sets were summed up across time. Results of the permutation set with the closest match to optimal peak discharge for reach 5 was selected as the optimal downstream hydrograph.

We ran two simulations; one using the selected optimal hydrograph that matches the peak discharge for reach 5, and for comparison, another assuming all upstream reaches flowed downstream at the same time (referred to as no shift in time). The simulation for no shift in time means all the upstream reaches commence flowing downstream at the same time: a direct summation of the hydrographs in time (simplified assumption). The aim of running the two simulations was to assess the plausibility of the spatiotemporal iterations of the upstream reaches compared to a simplified assumption.

3.3.4 Simulation round 4: Running the entire river network

Lastly, the simulation for the entire river network was ran using the temporally shifted reach hydrographs that resulted in the optimum downstream hydrograph with regards to peak discharge (simulation round 3). All 300 sampled observations were used to compute a global RMSE (gRMSE) to compare observed and simulated flood depths across the entire river network. Secondly, we compute a separate RMSE using only observations at J1, J2 and 4 observations located about 500 m downstream J1 (see Figure 1): this is to assess the performance of the superimposition technique and to check the dependence of the computed gRMSE on calibration data.

4 Results

Out of the all 300 interviews, about 100 buildings still had flood marks and the corresponding flood depths could be measured directly. Generally, collected observation data on flood depths were more than observations on flood duration. This was partly because it was easier for residents to remember maximum water depth than the time between which the flood reached the building and when it receded. Summary statistics of flood depth and duration data collected from the field and used for the modelling are presented in Table 3.

4.1 Optimized discharge

Results for all 9 combinations of n_c and n_{fp} for all five reaches are shown in Figure 5. The results show how the RMSE between observed and modelled flood depths increases or decreases as the discharge linearly increases (Figure 5). Simulations with n values between 0.01 and 0.07 for floodplain and channel produced flood depths that closely match observed depths at unusually high discharge. This discharge is regarded as unusually high given the size of the catchment (see section 2.2 and Figure 1) and based on the rainfall duration,

which was reported by the locals to be about 12 hours. For all reaches, simulations with n combinations of n_{fp} = 0.14 and n_c = 0.01, n_{fp} = 0.14 and n_c = 0.07, and n_{fp} = 0.14 and n_c = 0.14, consistently maintained the lowest RMSE values compared to other n combinations. Further investigation of these three optimally performing n combinations showed that whilst simulations with n_{fp} = 0.14 and n_c = 0.01 resulted in fast routing of the water causing some observation points not to be flooded, simulations with n_{fp} = 0.14 and n_c = 0.14 achieved an overestimation of

Table 3: Summary statistic of flood observations reporting minimum (min), interquartile range (IQR), mean and maximum (max) flood depth and duration on each reach.

	Location	**Buildings**	**Depth (m)**				**Duration (hrs)**			
			min	IQR	mean	max	min	IQR	mean	max
Upstream	Reach 1	67	0.20	1.03	1.23	2.80	1	8	10.40	48
	Reach 2	102	0.10	1.46	1.48	3.03	1	12	29.70	168
	Reach 3	37	0.20	1.48	1.44	2.64	1	8	18.40	72
	Reach 4	10	0.20	0.95	0.93	1.53	4	7	8.70	23
Downstream	Reach 5	55	0.30	1.10	1.07	2.20	7	48	57.80	178
Others	Junction 1 (J1)	11	0.41	1.40	1.40	2.43	5	42	53.20	178
	Junction 2 (J2)	14								
	Downstream J1	4								
Entire river network		300								

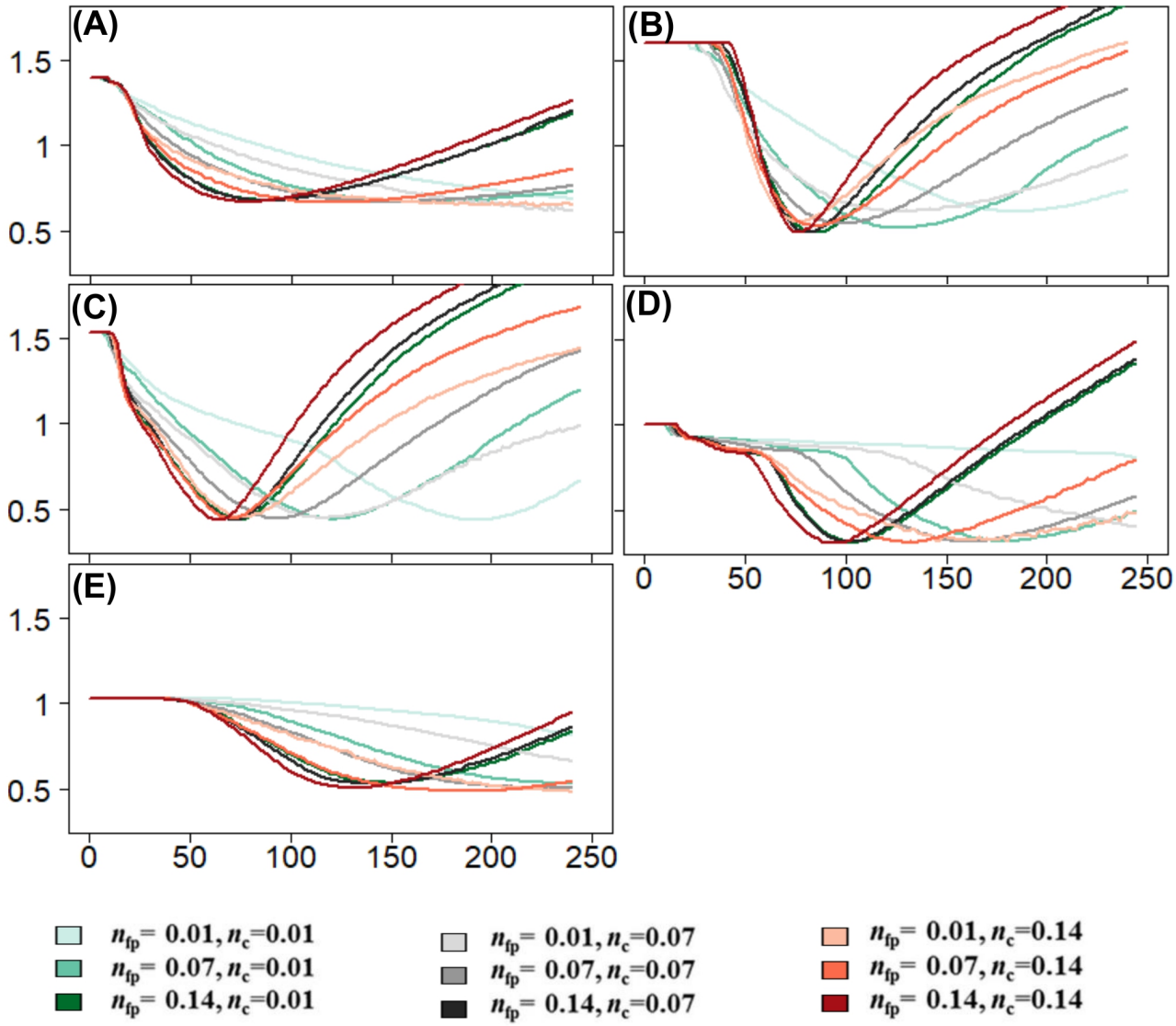


Figure 5: RMSE between observed and simulated flood depths from models driven with linearly increasing discharge and Manning's roughness (fp = floodplain, c = river channel) combinations for (A) Reach 1, (B) Reach 2, (C) Reach 3, (D) Reach 4, and (E) Reach 5. Y-axis limits are shown for the range 0.3 m to 1.75 m.

the flood levels at the observation points. Hence, we selected n_{fp} = 0.14 and n_c = 0.07 as optimum n combination for the study region and the combination was used for other simulation rounds. Table 4 shows the values of optimal peak discharge and the minimum RMSE between simulated and observed water depths for each reach. Optimal peak discharge for reach 5 was 147 m^3 s^{-1} and the summation of peak discharges from upstream reaches (1 - 4) was 340 m^3 s^{-1}: hence, the lower optimal peak discharge for reach 5 indicates that the upstream reaches did not synchronize peak discharges in time. Minimum RMSE was generally comparable across all reaches except for reach 1 with RMSE of 0.67 m (see Table 4).

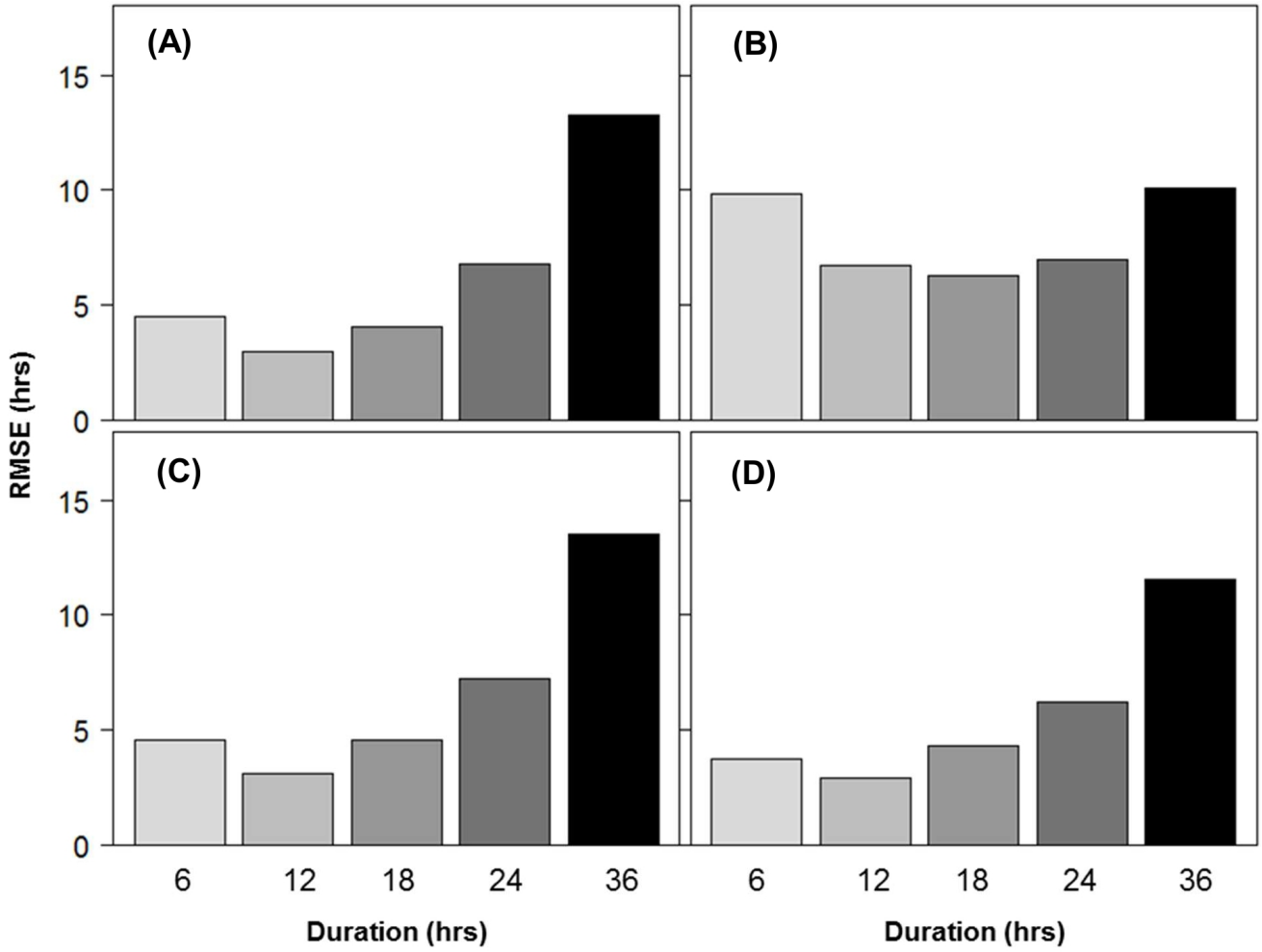


Figure 6: RMSE between observed and simulated flood duration for (A) reach 1(B) reach 2 (C) reach 3 (D) reach 4.

4.2 Optimized duration

Calculated RMSE between observed and modelled durations per reach are shown in Figure 6. On reaches 1, 3 and 4 (Figure 6 A, C, D), the 6 and 12 hrs hydrograph resulted in a low RMSE (≤ 5 hrs), hence suggesting that the flood duration at the building locations was likely within that range. Simulations with longer durations (18 and 36 hrs) resulted in high water volume and overestimated observed flood depths and durations. Consequently, for reaches 1, 3 and 4 the 12 hrs hydrograph duration was selected to be optimal since it results in the minimal RMSE. Generally, RMSE for reach 2 (Figure 6B) was comparably higher with both the lowest (6 hrs) and highest (36 hrs) duration hydrographs showing high errors. Results of the inundation extents showed

Table 4: Results summary showing optimal peak discharge and duration with RMSE's for corresponding reaches

	Location	Buildings	Optimal peak Discharge (m^3s^{-1})	RMSE (m)	Optimum duration (hrs)	RMSE (hrs)
Upstream	Reach 1	67	88	0.67	12	3.5
	Reach 2	102	80	0.46	18	9.5
	Reach 3	37	70	0.41	12	4.3
	Reach 4	10	102	0.33	12	3.7
Downstream	Reach 5	55	147	0.52	22	
Others	Junction 1 (J1) Junction 2 (J2) Downstream J1	29		0.67		
Entire river network		300		0.61		

that in reach 2, the 6 and 12 hrs simulations did not flood many observations points, hence resulting in a high RMSE. Conversely, high duration of 24 and 36 hrs resulted in comparatively high water depths at observation points and prolonged water duration. The 18 hrs duration hydrograph was selected for reach 2 since it had the lowest RMSE.

4.3 Combining hydrographs

Following the implemented spatiotemporal iterations for upstream reaches, a total of 1296 sets of superimpositions were determined. Figure 7A shows the plots of multiple superimpositions resulting in a peak discharge greater than 300 m^3 s^{-1} and peak discharge between 140 m^3 s^{-1} and 160 m^3 s^{-1}. Out of 1296 hydrograph combinations, one combination exactly matched the peak discharge of 147 m^3 s^{-1} (Figure 7A - blue curve) previously determined for reach 5 in isolation. In addition, the time lag for the combination was such that floods from smaller catchments were the first to flow downstream. RMSE between observed and modelled depths on reach 5 for the two simulations using the optimal shift and the simplified assumption of a no shift (Figure 7A – blue and red curve) are shown in Figure 7B.

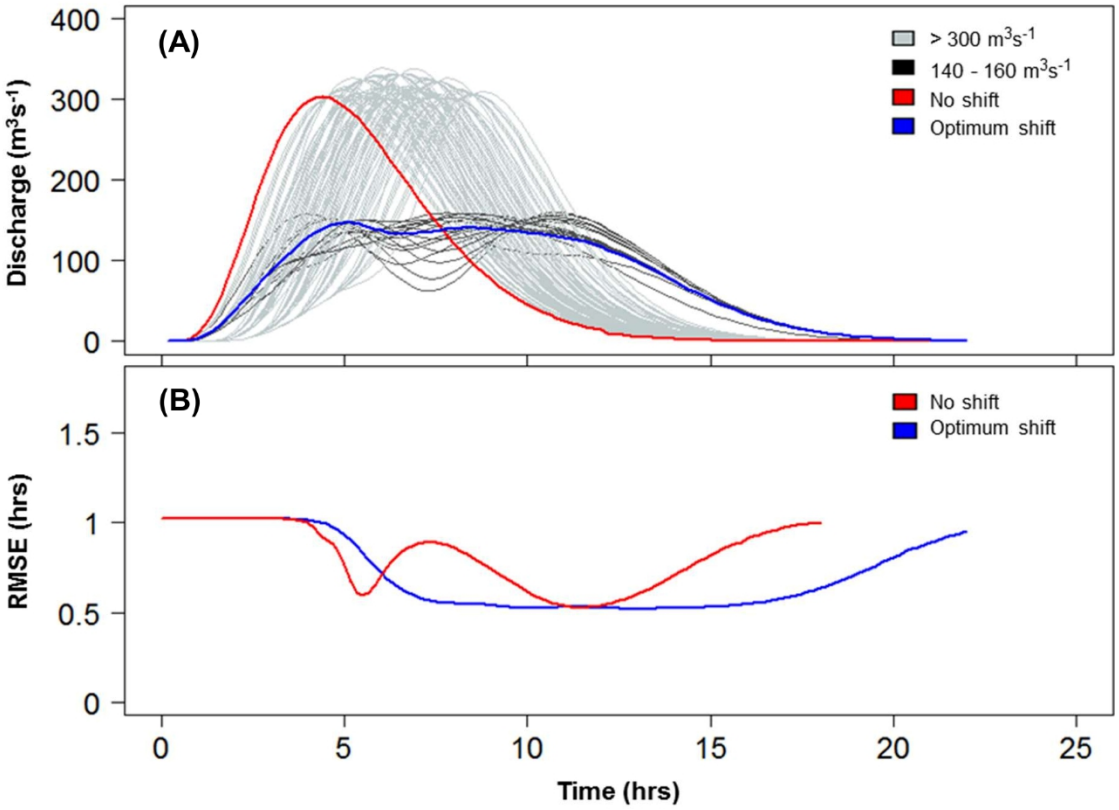


Figure 7: (A) Iteration of hydrographs with temporal shift for reaches 1 - 4. (B) RMSE between observed and simulated water depths.

4.4 Entire river network

Upstream input hydrographs, with optimal spatiotemporal shifts, used to simulate the entire river network are shown in Figure 8A. The global RMSE (gRMSE) between observed and modelled flood depths for the entire river network is shown in Figure 8B. Out of 300 observations, 12 were not flooded, consequently affecting the computed gRMSE. However, their inclusion in calculating gRMSE allows model errors to be sensitive to flood extents. Most of the observations points that were not flooded were located either on reach 1 or 2. The minimum gRMSE for all observations on the entire river network is 0.61 m and this corresponds to a simulated time of about 11 hrs. Figure 9 shows a pictorial representation of the model time step with minimum gRMSE and model errors corresponding to the difference between observed and modelled depths. For better visualization and discussion, selected parts of the river network on reach 1, 2, 3 and 5 are presented as Figures 9 a, b, c and d respectively. Negative errors indicate observation points where flood depths were overestimated and positive values indicate observation points where flood depths were underestimated

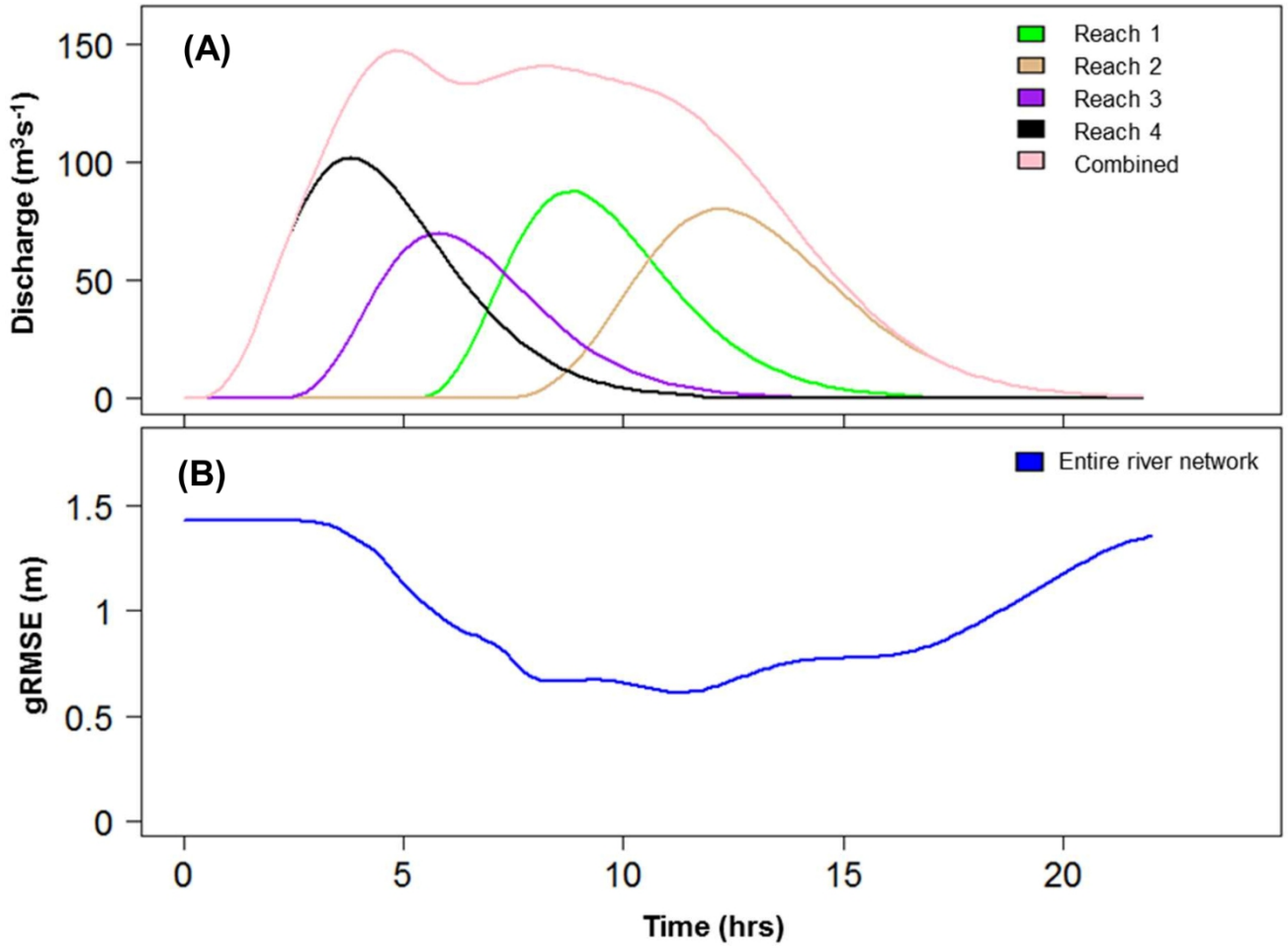


Figure 8: (A) Optimal upstream hydrographs for simulating the entire catchment (B) RMSE between observed and modelled depths.

To check the performance of the superimposition method and the dependence of computed RMSE on calibration data, we compute a separate RMSE only for observations associated with river junctions, which were initially excluded from the model calibration (see sec 3.3.4). For all 29 observations associated with river junctions (see Figure 1 and Table 3), RMSE between observed and modelled depth was 0.67 m.

5 Discussion

5.1 Model input and output evaluation

House-to-house data collected from field interviews provided adequate input to reconstruct the scenarios of the 2017 Suleja/Tafa flood event. Several simulation rounds were needed to closely match spatially distributed flood depths and duration for 300 buildings using the RMSE metric. The hydrodynamic model used, CAESAR-Lisflood, allowed the characterization of physical features of the flood plain and channel (Manning's roughness) and flow dynamics associated with multiple river networks..

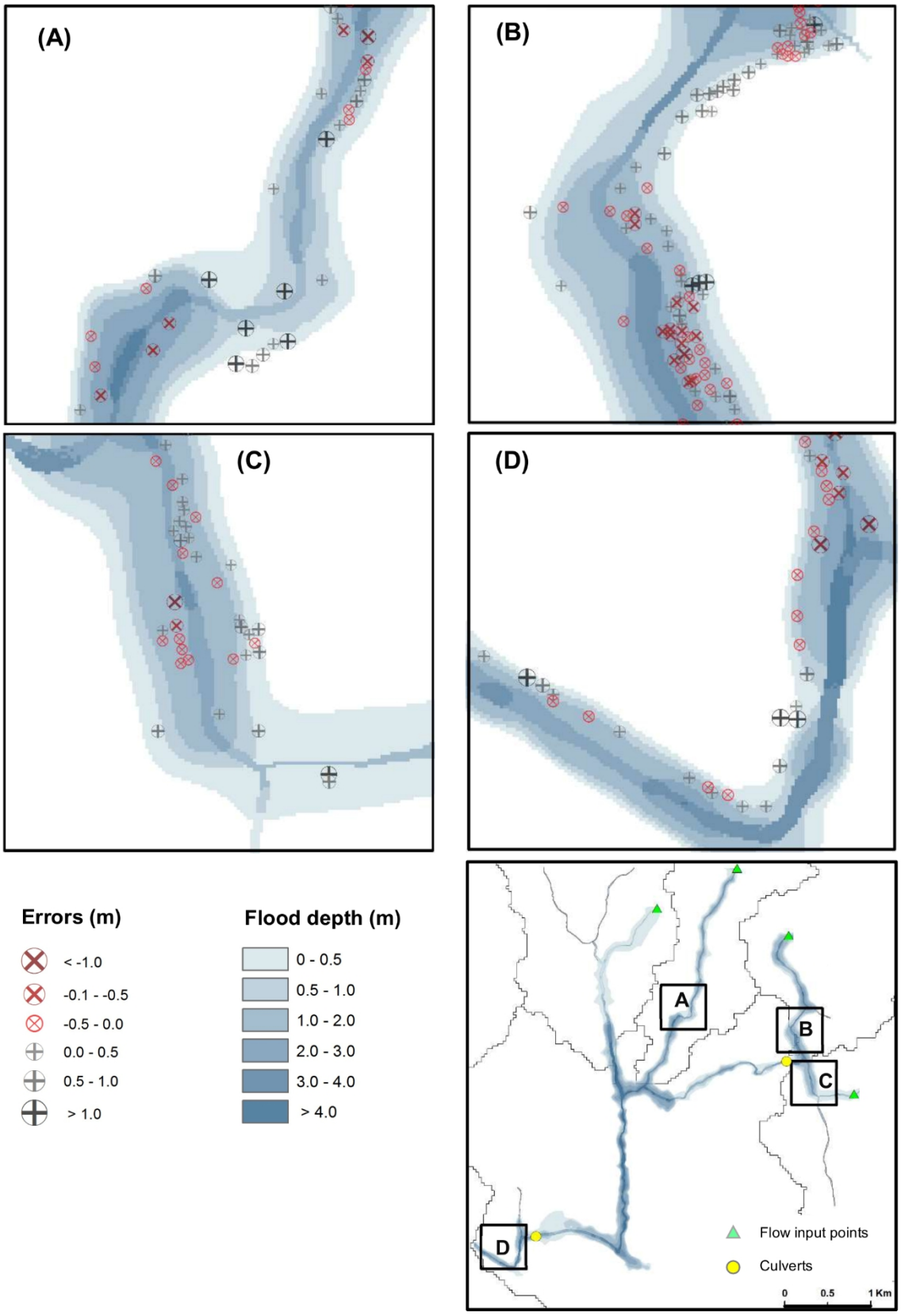


Figure 9: Flood model for the entire catchment with errors showing the difference between observed flood depths and maximum simulated depths for selected areas (A), (B), (C), and (D).

The range of n selected (0.01, 0.07, 0.14), allowed a model evaluation of the sensitivity of CAESAR-Lisflood to different combinations of n_c and n_{fp}. It demonstrates the dependence of peak discharge on n and reach characteristics. Combinations of n_c and n_{fp} with lower values (i.e. low roughness) allow water to be quickly routed towards the outlet, hence it takes a longer time and higher discharge for the water to propagate laterally and flood the building locations (see Figure 5). Curves relating discharge and RMSE (Figure 5) show similar patterns between i) reaches 1 and 5, and ii) reaches 2, 3 and 4. Similar patterns for reaches 1 and 5 is likely due to a characteristically U-shaped valley in both reaches (Table 1) meaning that lower values of n will require more time and discharge to route water to the locations of the observations points that are further away from the channel. On the other hand, reaches 2, 3 and 4 have floodplains with relatively flat terrains, hence a small increase in discharge is more likely to propagate faster and laterally on the flood plain. In general, these observed patterns suggest the influence of the shape of flood plain on peak discharge. Higher RMSE observed for reach 1 (simulation round 1) (Table 4) compared to other reaches is likely due to the deviation of the reconstructed channel: this deviation was highest in reach 1 (~50 m) at a specific location with a sharp channel bend (see Figure 1). The deviation of the channel resulted in having observed water depths locations either being closer or further away from the channel, consequently resulting in higher RMSE. For example, RMSE on reach 1 reduces by 0.1 m if observation points where the channel deviates were removed.

The second round of simulation approximated the flow propagation time on upstream reaches using flood durations at building locations. The highest difference between modelled and observed duration occurred on reach 2. However, this may be attributed to a culvert blockage under a road bridge at J1 (see Figure 1) and backwater effects, causing water to be retained in low terrain areas close to the outlet along reach 2. Although replication of flow at partially and fully blocked culverts (or bridges) is possible in CAESAR-Lisflood by increasing cell n value at blocked locations, adding this level of complexity to the model introduces a number of unknowns that need to be calibrated. First, we would need to determine the n value required to represent the blockage and this roughness will vary over the course of the flood event because the blockage more than likely increases and decrease over time. In addition, we would also have to estimate the time of blockage onset and duration. As such, with insufficient information on the evolution of the culvert blockage, the replication of such effects becomes difficult, and could potentially produce higher RMSEs on reach 2.

The spatial and temporal superimposition performed to characterized downstream flow pattern by upstream reaches resulted in the selection of an optimal flood wave with a peak discharge of 147 m^3 s^{-1} and a total duration of 21 hours. In general, although results for both simulations with optimal shift and the no shift (Figure 7A and B) showed comparable minimum RMSE of 0.52 m, RMSE for the no shift case highly overestimated the peak discharge (303 m^3s^{-1}). The sinuous shape of the RMSE in the no shift simulation (Figure 7B) resulted because the model flood depths closely matched the observed depths at both the rising and falling limb of the hydrograph. Similar results were found in Neal et al. (2013) when peak flows were simply allowed to coincide in time (no shift) which resulted in an overestimation of observed depths. The

optimal shift hydrograph showed a steady, but stable, minimum RMSE meaning that all upstream channels flowed downstream in such a way as to consistently maintain the closest match to the 55 observed water depths downstream. To further evaluate the value of the superimpositions and the choice of the 147 m^3 s^{-1} as the optimal peak discharge, we investigated the simulated depths at all the building locations on reach 5. Simulated water depths for both simulations with a peak discharge of 147 m^3 s^{-1} and 303 m^3 s^{-1} were jointly plotted against simulation time (see S2-Figure 2A). In addition, we show the range of the observed flood depths at all building locations on reach 5 with a boxplot (S2-Figure 2B). While the flood depths for the simulation with the selected optimal peak discharge of 147 m^3s^{-1} are mostly within the range of observed depths, the simulation with a peak discharge of 303 m^3s^{-1} contains many instances of simulated flood depths that lie outside the range of observed depths (see S2-Figure 2). Although both hydrographs resulted in a similar minimum RMSE (Figure 7B), the amount of water in the observed locations differ to a large extent. The optimal combination of upstream hydrographs selected based on the implemented superimposition (with a peak discharge of 147 m^3s^{-1}) more suitably captures the flood depths observed at the building locations compared to a simplified assumption of a direct overlap in time.

Results of the simulation for the entire river network produce interesting patterns. For example, in Figure 9A, high errors were most likely due to the deviation (about 50 m) of the reconstructed channel highlighted earlier. The deviation was caused by a sharp bend of the river channel, which was not well represented by the 12 m resolution WorldDEM DTM. Due to the deviation, most of the buildings were further away from the channel, hence model errors were predominantly underestimations of observed flood depths. Few of the observation points in Figure 9A were not flooded, further increasing errors and computed gRMSE. In Figure 9B, differences in observed and modelled flood depths were most likely due to reported backwater effects and culvert blockages close to J1. Errors in Figure 9C were also likely related to channel deviation resulting in both overestimation and underestimation of observed depths. Additionally, the contribution of an additional tributary (see Figure 9C) must have further influenced flow characteristics at the junction due to channel interactions. Since additional tributaries were not included in the modelling, their influence on flow characteristics at the junction may not be well captured. Similar interaction effects from an additional tributary might likely have influenced model errors on observations in Figure 9D.

The RMSE between observed and modelled flood depths at the river junctions (see Figure 1) was 0.67 m. Due to associated complexities at such locations and exclusion of the data in model calibration, an RMSE (0.67 m) higher than the gRMSE (0.61 m) was expected. However, such a small difference in RMSE of 0.06 m indicated that the superimposition method closely captured the flow processes at the river junctions. This indicates the applicability of this method in a complex multichannel river network. Additionally, it also showed that model results were not dependent on the calibration data used, which shows the potential for the transferability of the method. The gRMSE of 0.61 m is within ±0.10 m of other studies that use similar hydrodynamic models and had comparatively higher data availability and quality (e.g., Mignot et al. 2006;

Fewtrell et al. 2011; Neal et al. 2011; Yan et al. 2015; Ramirez et al. 2016; Altenau et al. 2017). Schumann et al. (2015) noted that generally, water level accuracies vary from few centimetres to 1 – 2 m during calibration, and our results are within this range. Hence, given limitations due to DEM resolution, the geometric representation of the river channel, and unavailability of data (rainfall or discharge) to set up boundary conditions, a gRMSE of 0.61 m for 300 observations is relatively good and demonstrates the applicability of interview data for hydrodynamic modelling in data-scarce regions.

The results obtained from this study may be replicated using different input discharge, shape of duration hydrographs and (or) different spatiotemporal combinations of the upstream peak discharge. As such various model setups can be equifinal (Beven and Freer, 2001) but this is less important in our study because the aim was not to reconstruct the exact hydrographs. Our aim was to arrive at a plausible scenario that minimizes the RMSE on each reach and later the gRMSE on the entire river network. At such minimum gRMSE, flood depths can be reliably extrapolated to other buildings to enrich a data set for developing flood damage models. Information we gained from the field guided the selection of model parameters such that the model output is physically plausible. The use of CAESAR-Lisflood allows the consideration of process characteristics (e.g., Manning's coefficient) and spatiotemporal dynamics of the upstream reaches and in doing this extrapolation of flood depths from a sample of building to all remaining buildings is possible using a physically informed method, which is usually not guaranteed by geo-statistical approaches.

Our study extends current methods in the application of house-to-house interview data collected using questionnaires. In particular, it extends the application of interview data from methods using GIS to map past flood events (e.g., Poser and Dransch 2010; Singh 2014; Sy et al. 2016, 2020) to a method using a hydrodynamic model. In such a way, the physical and dynamic characteristic of past flood events can also be represented. Focusing on hydrodynamic approaches, our study extends the method applied by Borga et al. (2019) and Bronstert et al. (2018) in utilizing post-event data to estimate peak discharge for a single channel; here we applied a similar method for a complex multichannel network with further consideration of spatiotemporal dynamics between upstream and downstream catchments. The key for extending the single-channel approach to a multichannel approach was to model upstream and downstream river reaches separately with synthetic hydrographs as proposed by Zischg (2018b) and finally by analyzing the role of relative tributary timing of upstream peak discharges. This study agrees with Neal et al. (2013), Pattison et al. (2014) and Zischg et al. (2018a), on the importance of considering spatiotemporal dependence of multiple upstream channels on downstream catchment for flood analysis. A simplified direct summation of upstream hydrographs would have led to a large overestimation of observed flood depths in the downstream reach. Furthermore, we demonstrate the use of distributed post-event data in a small to medium-scale data-scarce catchment by utilizing a globally available DEM and completely reconstructing the river channel. In addition to developing a method for reconstructing plausible scenario of past flood events in data-scarce regions, this

study contributes to knowledge for evaluating the sensitivity and in testing the performance of CAESAR-Lisflood in different regions as well as its usability with field data as recommended by Coulthard et al. (2013).

5.2 Model uncertainties

Several studies (e.g., Saksena and Merwade 2015; Ramirez et al. 2016) have shown that model errors are directly related to the accuracy of the elevation model used since the DEM directly affects where water is routed. Consequently, the coarser a DEM is, the more likely it will not correctly represent floodplain and channel characteristics. Saksena and Merwade (2015) suggested that errors from coarse resolution DEMs are specifically related to reach length and width, valley shape and land-use. In this study, we use a 12 m resolution WorldDEM DTM tile. Although the WorldDEM DTM has a better resolution compared to the globally available 30 m SRTM, this study has shown that the WorldDEM DTM still requires pre-processing to represent river channels. The location, width and depth of the river channel, presence of artefacts like bridges and culverts, and shape of channel valley contribute to model uncertainty. In this study, to reduce these uncertainties, we reconstructed the channel location and width with the support of field photographs and satellite imagery. However, considering that the reconstructed channel deviated in multiple areas, which have likely resulted in local errors (see section 4.4), it is apparent that data limitations directly affect model uncertainty. Another source of uncertainty is the characterization of the channel using a simplified assumption of a rectangular shape. Neal et al. (2015) showed that errors introduced due to the simplification of channel bed shape have comparatively the same accuracy in water depth and extent compared with models using more complex channel bed representation. Neal et al. (2015) also indicated that a major difference in simplified representations of channel bed is the increase in flood wave propagation, and this is only more pronounced for large catchments.

A uniform drop of the entire channel by 1 m has been chosen, a value that is close to what we observed in the field. Without defining a channel in the DEM, the floods in some locations commenced nearly at the level of the floodplain and this would underestimate optimal discharge values. In contrast, excessively dropping the channel in elevation would unrealistically confine the river and severely overestimate optimal discharge. To test the sensitivity of the selected drop in channel elevation, we ran additional simulations with different channel elevations. Channel elevation drops generated for the simulations were 0.5 m, 0.75 m, 1.25 m and 1.5 m. All simulations were ran using the same n values for flood plain (0.14) and channel (0.07), which is consistent with the selected optimal n values. In total, we ran 20 simulations, which corresponds to 4 simulations (0.5, 0.75, 1.25 and 1.5 m bed elevation) on reaches 1 – 5. Results of the simulations generally show low sensitivity of the peak discharge and RMSE on changing channel elevations (see S2, Figure 1, Table 1).

Furthermore, uncertainties relating to interview data exist. The fieldwork was carried out eight months after the flood event, and the accuracy of observed flood depths and duration was partly dependent on the personal

reflection of people living in the affected houses. Generally, while peoples recollection of past events can become vague after some time (Lacy and Stark, 2013), floods are traumatic events and people are very likely to correctly remember details over long periods (Sotgiu and Galati, 2007). In our data, more than 100 observation points had flood marks and it was possible to measure the flood depths directly. Such field visits provide an opportunity to reduce uncertainties related to inaccurately reported flood depths by community residents (McDougall and Temple-Watts 2012; McDougall 2011). Conversely, unlike flood depths where it is possible to use flood marks, the case is slightly different for flood duration. Flood durations were completely based on personal reflections, which makes the data susceptible to uncertainties. Also, some studies (e.g. Mbow et al. 2008; Sy et al. 2016) have observed that local factors such as lack of means for evacuation and topographic depressions can influence flood duration data at building locations, and consequently affect model performance. For example, in Table 3 such high variations in flood duration can be observed on different reaches with the maximum range at Junction 1 having a minimum of 1 hour and a maximum of 178 hrs. In general, recently demonstrated approaches in citizen science projects can be used to reduce the uncertainty in interview data collection. For example, two separate individuals (within a household) can be used to check the consistency of information provided: a similar approach was demonstrated by Sy et al. (2020) who interviewed two sets of community representatives to reconstitute three past flood events in Senegal. Sy et al. (2020) reported that this approach improved the validity and reliability of their method since they can check the consistency of information using both sets of participants.

Flood extent data was unavailable to support model calibration or to validate model results for the entire catchment. Due to the use of a DEM without buildings, determined peak discharge on each reach (Table 3) are likely to be slightly overestimated. This is because of limited confinement and friction by such a DEM, meaning that more water volume and higher discharge will be required to closely match observed depths. Bermúdez and Zischg et al. (2018c) and Neal et al. (2009) made similar conclusions where building representation was found to affect model depths. In addition, a common practice in the Suleja/Tafa region is the construction of fencing walls around buildings which serve as a flood protection. The fencing walls are especially predominant on reaches 1 and 2 and are very close to the channels. Generally, such fencing walls alter the amount of floodwater that enters the building except if they have been damaged by the flood. Our model uses a 4 m x 4 m resampled DEM and cannot accurately represent fine-scale features like fencing walls and adding these features to the DEM, at the current resolution, would substantially confine water to the channel and prevent it from propagating to the flood plain. Such limitation in representing the fencing walls will generally lead to overestimated flood depths at building locations.

6 Conclusion

In this study, we developed a systematic method to replicate a plausible scenario of past flood events in a typical data-scarce small to medium scale catchment using field interview data and hydraulic modelling. Flood depth and duration data, collected from interviewing 300 households, were used to reconstruct a flood

event from 2017 in Suleja/Tafa, Nigeria, through a four-step simulation procedure using CAESAR-Lisflood. At first, distributed observed flood depths on each reach were used to calibrate the Manning's coefficient and extract optimal peak discharge that gives the least difference between observed and modelled flood depths. Synthetic hydrographs having 6, 12, 18, 24, 36 hrs durations were used to simulate each reach using determined peak discharge to capture the model duration that closely matches observed duration at observed locations. Combined effects of upstream catchments (reaches 1 - 4) were investigated to explore all possible flood peak synchronization possibilities that will match simulated downstream (reach 5) peak discharge. Model output for simulating the entire river network resulted in a gRMSE of 0.61 m for all 300 observations. Major uncertainties that contributed to model errors relate to observation data, DEM quality, and quality of reconstructed river location, shape and depth.

Our study demonstrates the potential in the application of flood modelling techniques for small to medium-scale, single- or multichannel data-scarce regions, using spatially distributed post-event data acquired through house-to-house interviews. These data are mostly easy to acquire using questionnaires and can serve as a viable option for providing sizeable calibration data in areas with no gauging stations or satellite data. The method developed in this study allows the determination of flood depths, durations, and possibly, flood velocities at building locations that were not surveyed (e.g., see urbanized areas in Figure 1) to enrich data sets for developing flood damage models. Application of the method opens a pathway for the reconstruction of flood events in data-scarce regions and investigating future flood scenarios so that relevant information can be provided to help build capacities and reduce flood risk in the face of recently observed flood extremes.

Acknowledgments

We thank the residents of Suleja and Tafa community for their support during the field studies. Thanks to the University of Bern and ESKAS for covering the data collection cost. Also, thanks to the European Space Agency for granting us permission to use the WorldDEM DTM free of charge. This study is carried out within a Ph.D. funded by the Swiss Government Excellence Scholarship (ESKAS).

Financial support

This research has been supported by the Swiss government excellence scholarship (grant no. 2017.1027).

References

Adeleye, B., Popoola, A., Sanni, L., Zitta, N. and Ayangbile, O.: Poor development control as flood vulnerability factor in Suleja , Nigeria, T. Reg. Plan., (74), 23–35, doi:10.18820/2415-0495/trp74i1.3, 2019.

Altenau, E. H., Pavelsky, T. M., Bates, P. D. and Neal, J. C.: The effects of spatial resolution and dimensionality on modeling regional-scale hydraulics in a multichannel river, Water Resour. Res., 53(2), 1683–1701, 2017.

Archer, L.: Comparing TanDEM-X data with frequently used DEMs for flood inundation modeling, Water Resour. Res., 54(10), 205–222, doi:10.1029/2018WR023688, 2018.

Assumpção, T. H., Popescu, I., Jonoski, A. and Solomatine, D. P.: Citizen observations contributing to flood modelling: opportunities and challenges, Hydrol. Earth Syst. Sci., 22(2), 1473–1489, doi:10.5194/hess-22-1473-2018, 2018.

Bates, P. D. and De Roo, A. P. J.: A simple raster-based model for flood inundation simulation, J. Hydrol., 236(1–2), 54–77, doi:10.1016/S0022-1694(00)00278-X, 2000.

Bermúdez, M. and Zischg, A. P.: Sensitivity of flood loss estimates to building representation and flow depth attribution methods in micro-scale flood modelling, Nat. Hazards, 92(3), 1633–1648, doi:10.1007/s11069-018-3270-7, 2018.

Bernet, D. B., Trefalt, S., Martius, O., Weingartner, R., Mosimann, M., Röthlisberger, V. and Zischg, A. P.: Characterizing precipitation events leading to surface water flood damage over large regions of complex terrain, Environ. Res. Lett., 14(6), doi:10.1088/1748-9326/ab127c, 2019.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249(1), 11–29, doi:https://doi.org/10.1016/S0022-1694(01)00421-8, 2001.

Borga, M., Comiti, F., Ruin, I. and Marra, F.: Forensic analysis of flash flood response, WIREs Water, 6(e1338), doi:10.1002/wat2.1338, 2019.

Bronstert, A., Agarwal, A., Boessenkool, B., Crisologo, I., Fischer, M., Heistermann, M., Köhn-reich, L., López-tarazón, J. A., Moran, T., Ozturk, U., Reinhardt-Imjela, C. and Wendi, D.: Forensic hydro-meteorological analysis of an extreme flash flood: The 2016-05-29 event in Braunsbach, SW Germany, Sci. Total Environ., 630, 977–991, doi:10.1016/j.scitotenv.2018.02.241, 2018.

Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L. and Smets, B.: Copernicus Global Land Cover Layers—Collection 2, Remote Sens., 12(6), 1044, 2020.

Canevari-Luzardo, L., Bastide, J., Choutet, I. and Liverman, D.: Using partial participatory GIS in vulnerability and disaster risk reduction in Grenada, Clim. Dev., 9(2), 95–109, doi:10.1080/17565529.2015.1067593, 2017.

Chow, C., Andrášik, R., Fischer, B. and Keiler, M.: Application of statistical techniques to proportional loss data: Evaluating the predictive accuracy of physical vulnerability to hazardous hydro- meteorological events, J. Environ. Manage., 246, 85–100, doi:10.1016/j.jenvman.2019.05.084, 2019.

Chow, V. T.: 1959, Open channel hydraulics, McGraw-Hill, New York, 1959.

Cook, A. and Merwade, V.: Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping, J. Hydrol., 377, 131–142, doi:10.1016/j.jhydrol.2009.08.015, 2009.

Coulthard, T. J., Macklin, M. G. and Kirkby, M. J.: A cellular model of Holocene upland river basin and alluvial fan evolution, Earth Surf. Process. Landforms, 27(3), 269–288, doi:10.1002/esp.318, 2002.

Coulthard, T. J., Neal, J. C., Bates, P. D., Ramirez, J., de Almeida, G. A. M. and Hancock, G. R.: Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: Implications for modelling landscape evolution, Earth Surf. Process. Landforms, 38(15), 1897–1906, doi:10.1002/esp.3478, 2013.

Fewtrell, T. J., Neal, J. C., Bates, P. D. and Harrison, P. J.: Geometric and structural river channel complexity and the prediction of urban inundation, Hydrol. Process, 25, 3173–3186, doi:10.1002/hyp.8035, 2011.

Florczyk, A. J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., Sabo, F. and Zanchetta, L.: GHSL Data Package 2019 - Technical report by the Joint Research Centre (JRC), European Union, Luxembourg., 2019.

Fuchs, S., Keiler, M., Ortlepp, R., Schinke, R. and Papathoma-Köhle, M.: Recent advances in vulnerability assessment for the built environment exposed to torrential hazards: Challenges and the way forward, J. Hydrol., 575(May), 587–595, doi:10.1016/j.jhydrol.2019.05.067, 2019.

Gichamo, T. Z., Popescu, I., Jonoski, A. and Solomatine, D.: River cross-section extraction from the ASTER global DEM for flood modeling, Environ. Model. Softw., 31, 37–46, doi:10.1016/j.envsoft.2011.12.003, 2012.

Hersbach, H., de Rosnay, P., Bell, B., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Alonso-Balmaseda, M., Balsamo, G. and Bechtold, P.: Operational global reanalysis: progress, future directions and synergies with NWP, ERA Report Series, Eur. Cent. Mediu. Range Weather Forecast. Shinf. Park, doi:10.21957/tkic6g3wm, 2018.

Hunter, N. M., Horritt, M. S., Bates, P. D., Wilson, M. D. and Werner, M. G. F.: An adaptive time step solution for raster-based storage cell modelling of floodplain inundation, Adv. Water Resour., 28, 975–991, doi:10.1016/j.advwatres.2005.03.007, 2005.

Jung, H. C., Hamski, J., Durand, M., Alsdorf, D., Hossain, F., Lee, H., Azad Hossain, A. K. M., Hasan, K., Khan, A. S. and Zeaul Hoque, A. K. M.: Characterization of complex fluvial systems using remote sensing of spatial and temporal water level variations in the Amazon, Congo, and Brahmaputra rivers, Earth Surf. Process. Landforms, 35(3), 294–304, doi:10.1002/esp.1914, 2010.

Komi, K., Neal, J., Trigg, M. A. and Diekkrüger, B.: Modelling of flood hazard extent in data sparse areas: a case study of the Oti River basin, West Africa, J. Hydrol. Reg. Stud., 10, 122–132, doi:10.1016/j.ejrh.2017.03.001, 2017.

Lacy, J. W. and Stark, C. E. L.: The neuroscience of memory: implications for the courtroom, Nat. Rev. Neurosci., 14(9), 649–658, doi:10.1038/nrn3563, 2013.

Mayomi, I., Kolawole, M. S. and Martins, A. K.: Terrain Analysis for Flood Disaster Vulnerability Assessment : A Case Study of Niger State, Nigeria, Am. J. Geogr. Inf. Syst., 3(3), 122–134, doi:10.5923/j.ajgis.20140303.02, 2014.

Mbow, C., Diop, A., Diaw, A. T. and Niang, C. I.: Urban sprawl development and flooding at Yeumbeul suburb (Dakar-Senegal), African J. Environ. Sci. Technol., 2, 68–74, 2008.

McDougall, K.: Using volunteered information to map the Queensland floods, in Proceedings of the Surveying & Spatial Sciences Biennial Conference 21-25 November, Wellington, New Zealand., 2011.

McDougall, K. and Temple-Watts, P.: The Use of LIDAR and Volunteered Geographic Information to Map

Flood Extents and Inundation, ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., I4, 251–256, doi:10.5194/isprsannals-I-4-251-2012, 2012.

Mignot, E., Paquier, A. and Haider, S.: Modeling floods in a dense urban area using 2D shallow water equations, J. Hydrol., 327, 186–199, doi:10.1016/j.jhydrol.2005.11.026, 2006.

Muñoz-Sabater, J., Dutra, E., Balsamo, G., Boussetta, S., Zsoter, E., Albergel, C. and Agusti-Panareda, A.: ERA5-Land: an improved version of the ERA5 reanalysis land component, in Joint ISWG and LSA-SAF Workshop IPMA, Lisbon, pp. 26–28., 2018.

Musungu, K., Motala, S. and Smit, J.: Using Multi-criteria Evaluation and GIS for Flood Risk Analysis in Informal Settlements of Cape Town : The Case of Graveyard Pond, , 1(1), 77–91, 2012.

Ndanusa, A. B., Dahalin, Z. and Ta, A.: Topographic-based framework for flood vulnerability classification: a case of Niger state, Nigeria, J. Inf. Syst. Technol. Manag., 3(9), 27–38, 2018.

Neal, J., Schumann, G., Fewtrell, T., Budimir, M., Bates, P. and Mason, D.: Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK, J. Flood Risk Manag., 4(2), 88–95, doi:10.1111/j.1753-318X.2011.01093.x, 2011.

Neal, J., Schumann, G. and Bates, P.: A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas, Water Resour. Res., 48(11), doi:10.1029/2012WR012514, 2012.

Neal, J., Keef, C., Bates, P., Beven, K. and Leedal, D.: Probabilistic flood risk mapping including spatial dependence, Hydrol. Process, 27(9), 1349–1363, doi:10.1002/hyp.9572, 2013.

Neal, J. C., Bates, P. D., Fewtrell, T. J., Hunter, N. M., Wilson, M. D. and Horritt, M. S.: Distributed whole city water level measurements from the Carlisle 2005 urban flood event and comparison with hydraulic model simulations, J. Hydrol., 368, 42–55, doi:10.1016/j.jhydrol.2009.01.026, 2009.

Neal, J. C., Odoni, N. A., Trigg, M. A., Freer, J. E., Garcia-Pintado, J., Mason, D. C., Wood, M. and Bates, P. D.: Efficient incorporation of channel cross-section geometry uncertainty into regional and global scale flood inundation models, J. Hydrol., 529, 169–183, doi:10.1016/j.jhydrol.2015.07.026, 2015.

OCHA, (United Nations Office for the Coordination of Humanitarian Affairs): NIGERIA : Floods in Borno , Adamawa and Yobe, Situation Report No. 2. [online] Available from: https://reliefweb.int/sites/reliefweb.int/files/resources/ocha_nga_floodoverview_13092019.pdf, 2019.

Pappenberger, F., Matgen, P., Beven, K. J., Henry, J., Pfister, L. and De, P. F.: Influence of uncertain boundary conditions and model structure on flood inundation predictions, Adv. Water Resour., 29(10), 1430–1449, doi:10.1016/j.advwatres.2005.11.012, 2006.

Pattison, I., Lane, S. N., Hardy, R. J. and Reaney, S. M.: The role of tributary relative timing and sequencing in controlling large floods, Water Resour. Res., 50(7), 5444–5458, doi:10.1002/2013WR014067, 2014.

Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kauffmann, M., Kemper, T. and Lu, L.: A global human settlement layer from optical HR/VHR RS data: concept and first results, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 6(5), 2102–2131, 2013.

Phillips, B. J. V and Tadayon, S.: Selection of Manning's roughness coefficient for natural and constructed vegetated and non-vegetated channels, and vegetation maintenance Plan guidelines for vegetated channels in central Arizona: Scientific investigations report 2006 – 5108, Arizona., 2006.

Poser, K. and Dransch, D.: Volunteered geographic information for disaster management with application to

rapid flood damage estimation, Geomatica, 64(1), 89–98, 2010.

Ramirez, J. A., Rajasekar, U., Patel, D. P., Coulthard, T. J. and Keiler, M.: Flood modeling can make a difference: Disaster risk-reduction and resilience-building in urban areas, Hydrol. Earth Syst. Sci. Discuss., 1–21, doi:10.5194/hess-2016-544, 2016.

Rolim, A., Collischonn, W., Tucci, C. E. M. and Padovani, C. R.: Large-scale modelling of channel flow and floodplain inundation dynamics and its application to the Pantanal (Brazil), Hydrol. Process, 25(9), 1498–1516, doi:10.1002/hyp.7926, 2011.

Saksena, S. and Merwade, V.: Incorporating the effect of DEM resolution and accuracy for improved flood inundation mapping, J. Hydrol., 530, 180–194, doi:10.1016/j.jhydrol.2015.09.069, 2015.

Sanyal, J., Carbonneau, P. and Densmore, A. L.: Hydraulic routing of extreme floods in a large ungauged river and the estimation of associated uncertainties: a case study of the Damodar River, India, Nat. Hazards, 66, 1153–1177, doi:10.1007/s11069-012-0540-7, 2013.

Sarhadi, A., Soltani, S. and Modarres, R.: Probabilistic flood inundation mapping of ungauged rivers: Linking GIS techniques and frequency analysis, J. Hydrol., 458–459, 68–86, doi:10.1016/j.jhydrol.2012.06.039, 2012.

Schumann, G. J. P., Bates, P. D., Neal, J. C. and Andreadis, K. M.: Measuring and Mapping Flood Processes, in Hydro-Meteorological Hazards, Risks, and Disasters, edited by F. S. John, D. B. Giuliano, and P. Paolo, pp. 35–64, Elsevier Inc., 2015.

Singh, B. K.: Flood Hazard Mapping with Participatory GIS: The Case of Gorakhpur, Environ. Urban. ASIA, 5(1), 161–173, doi:10.1177/0975425314521546, 2014.

Smith, A., Bates, P. D., Wing, O., Sampson, C., Quinn, N. and Neal, J.: New estimates of flood exposure in developing countries using high-resolution population data, Nat. Commun., 1–7, doi:10.1038/s41467-019-09282-y, 2019.

Smith, L., Liang, Q., James, P. and Lin, W.: Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework, J. Flood Risk Manag., 10(3), 370–380, doi:10.1111/jfr3.12154, 2017.

Sotgiu, I. and Galati, D.: Long-Term Memory for Traumatic Events: Experiences and Emotional Reactions During the 2000 Flood in Italy, J. Psychol., 141(1), 91–108, doi:10.3200/JRLP.141.1.91-108, 2007.

Sy, B., Frischknecht, C., Dao, H., Giuliani, G. and Consuegra, D.: Participatory approach for flood risk assessment : the case of Yeumbeul Nord (YN), Dakar , Senegal, in Proceedings of the 5 International Conference th on Flood Risk Management and Response (FRIAR 2016), vol. 165., 2016.

Sy, B., Frischknecht, C., Dao, H., Consuegra, D. and Giuliani, G.: Flood hazard assessment and the role of citizen science, J. Flood Risk Manag., 12(S2), e12519, doi:10.1111/jfr3.12519, 2019.

Sy, B., Frischknecht, C., Dao, H., Consuegra, D. and Giuliani, G.: Reconstituting past flood events: The contribution of citizen science, Hydrol. Earth Syst. Sci., 24(1), 61–74, doi:10.5194/hess-24-61-2020, 2020.

Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F. W., Dutta, D. and Kim, S.: Flood inundation modelling: A review of methods, recent advances and uncertainty analysis, Environ. Model. Softw., 90, 201–216, doi:10.1016/j.envsoft.2017.01.006, 2017.

Tran, P., Shaw, R., Chantry, G. and Norton, J.: GIS and local knowledge in disaster management: a case study of flood risk mapping in Viet Nam, Disasters, 33(1), 152–169, doi:10.1111/j.1467-7717.2008.01067.x, 2009.

USAID: Southern Africa Tropical Cyclones Fact Sheet #13 - 05-31-2019., 2019.

USDA: National Engineering Handbook, Hydrology, p. 23, United States Department of Agriculture, Soil Conservation Service., 2007.

Wagenaar, D., De Jong, J. and Bouwer, L. M.: Multi-variable flood damage modelling with limited data using supervised learning approaches, Nat. Hazards Earth Syst. Sci., 17(9), 1683–1696, doi:10.5194/nhess-17-1683-2017, 2017.

Wang, Y., Chen, A. S., Fu, G., Djordjevi, S., Zhang, C. and Savi, D. A.: An integrated framework for high-resolution urban flood modelling considering multiple information sources and urban features, Environ. Model. Softw., 107(7), 85–95, doi:10.1016/j.envsoft.2018.06.010, 2018.

Welsh, K. E., Dearing, J. A., Chiverrell, R. C. and Coulthard, T. J.: Testing a cellular modelling approach to simulating late-Holocene sediment and water transfer from catchment to lake in the French Alps since 1826, The Holocene, 19(5), 785–798, 2009.

Wood, M., Hostache, R., Neal, J., Wagener, T., Giustarini, L., Chini, M., Corato, G., Matgen, P. and Bates, P.: Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium-resolution SAR data and identifiability techniques, , 4983–4997, doi:10.5194/hess-20-4983-2016, 2016.

Yan, K., Pappenberger, F. and Solomatine, D. P.: Regional versus physically-based methods for flood inundation modelling in data scarce areas: an application to the Blue Nile, Int. Conf. Hydroinformatics 8-1-2014, 2014.

Yan, K., Neal, J. C., Solomatine, D. P. and Baldassarre, G. Di: Hydro-Meteorological Hazards, Risks, and Disasters: Chapter four-Global and Low-Cost Topographic Data to Support Flood Studies, edited by J. F. Shroder, G. Di Baldassarre, and P. Paron, Elsevier Inc., 2015.

Zischg, A. P., Felder, G., Weingartner, R., Quinn, N., Coxon, G., Neal, J. and Freer, J.: Effects of variability in probable maximum precipitation patterns on flood losses, Hydrol. Earth Syst. Sci., 22(5), 2759–2773, doi:10.5194/hess-22-2759-2018, 2018a.

Zischg, A. P., Felder, G., Mosimann, M., Röthlisberger, V. and Weingartner, R.: Extending coupled hydrological-hydraulic model chains with a surrogate model for the estimation of flood losses, Environ. Model. Softw., 108, 174–185, doi:10.1016/j.envsoft.2018.08.009, 2018b.

Zischg, A. P., Mosimann, M., Bernet, D. B. and Röthlisberger, V.: Validation of 2D flood models with insurance claims, J. Hydrol., 557, 350–361, doi:10.1016/j.jhydrol.2017.12.042, 2018c.

Chapter 4: Expert-based versus data-driven flood damage models: a comparative evaluation for data-scarce regions

Mark Bawa Malgwi[1,2*], Matthias Schlögl[3], Margreth Keiler[1,2,4]

[1] University of Bern, Institute of Geography, Hallerstrasse 12, 3012 Bern, Switzerland

[2] University of Bern, Oeschger Centre for Climate Change Research, Hochschulstrasse 6, 3012 Bern, Switzerland

[3] University of Natural Resources and Life Sciences, Institute of Mountain Risk Engineering, Peter-Jordan-Str. 82, 1190 Vienna, Austria

[4] University of Bern, Mobiliar Lab for Natural Risks, Hallerstrasse 12, 3012 Bern, Switzerland

Author contributions

MBM designed the study with the support of MK. MBM was responsible for the collection and pre-processing of data from the two study regions and implementation of the expert-based approach. MS provided for statistical support and implemented the data-driven approach. MBM and MS were responsible for data visualization. All authors were jointly involved in manuscript preparation and editing.

Under review in *International Journal of Disaster Risk Reduction*
Submitted on: 21 July, 2020.

Abstract

The knowledge about potential flood damage is a key issue for disaster risk reduction. However, the scarcity of empirical data has limited flood damage modelling in several regions. As a result, current approaches in data-scarce regions have so far been restricted to either exposure assessment or identification of vulnerability indicators. As expert-based approaches do not require empirical damage data, they have a high potential for flood damage modelling in data-scarce regions. In this study, we carried out a comparative assessment between an expert-based and a data-driven approach. The expert-based approach systematically combines the vulnerability indicator method and synthetic what-if analysis based on the knowledge of regional experts. The data-driven approach integrates empirical flood damage data in the analysis applying a multivariate random forest model. Flood damage data, collected through interviews after two flood events in 2017 and 2019 at separate locations in Nigeria, were used to evaluate the performance of both methods based on developed damage grades. Results from both methods showed i) a mean predictive accuracy of 30 % and 38 % for the expert-based and data-driven approaches respectively, ii) that distance to channel, wall material, building condition, and building quality are significant regional damage drivers, and iii) comparable model performance can be achieved even with a reduced number of variables. Furthermore, the study demonstrated how experts are likely to underestimate damage at low water depths and how a difference in conformity to building standards can add to challenges in flood damage prediction.

Keywords: flood, expert-based, data-scarce, data-driven, damage prediction, damage grades,

1 Introduction

Globally, flood disasters continue to give cause for growing concern. The increase in frequencies and severity of flood events likely driven by climate change (Hoegh-Guldberg et al., 2018) and changing exposure (Fuchs et al., 2015; Röthlisberger et al., 2017) has resulted in considerable human and economic losses (CRED, 2020). The occurrence of flood events in vulnerable communities is even more critical due to low coping and adaptive capacities. Vulnerability relates to conditions that make communities prone or susceptible to disasters (UNISDR, 2009). Large scale building damage caused by floods from Cyclones Idai and Kenneth in several regions in the southern part of Africa (CRED, 2019) further underlines the need to facilitate efforts for physical vulnerability assessment in vulnerable regions, which are mostly data-scarce (Malgwi et al., 2020). Physical vulnerability assessment methods, such as flood damage models (stage-damage curves, multivariate models) and vulnerability indicators, explore the relationship between flood damage and corresponding damage influencing variables (Papathoma-Köhle et al., 2017). As a result, physical vulnerability assessment provides an important basis for physical resilience assessment and mitigation planning (Papathoma-Köhle et al., 2011), evaluating economic losses (Blanco-Vogt and Schanze, 2014), and cost-benefit analysis, which supports resource allocation for hazard protection (Fuchs, 2009; UNISDR, 2015). In general, such efforts tailored to understanding and reducing vulnerability are considered important steps for disaster risk reduction (UNISDR, 2015).

Flood damage models either show the relationship between flood damage (or monetary loss) and water depth (referred to as stage-damage curves) or include other additional variables (referred to as multivariate models). While stage-damage curves use a continuous curve to relate water depth and damage state (Fuchs et al., 2019a, 2019b), multivariate models use different statistical approaches such as Bayesian network (Vogel et al., 2012), regression and ensembles of bagged decision trees (Merz et al., 2013) or logistic regression (Ettinger et al., 2016). Studies have shown that multivariate flood damage models have better prediction accuracy both spatially and temporarily (Schröter et al., 2014), and better explain the variance in damage data (Merz et al., 2004) compared to stage-damage curves that use only water depths. While most flood damage models are developed using empirical damage or monetary loss data (see Gerl et al., 2016; Hammond and Chen, 2015; Jongman et al., 2012; Merz et al., 2010), scarcity of such data, especially in developing countries, has hindered the application of multivariate models, consequently limiting efforts for disaster risk reduction (Englhardt et al., 2019; Niang et al., 2015). Although, synthetic stage-damage curves, developed using expert what-if analysis (for example Naumann et al., 2009; Neubert et al., 2008; Penning-Rowsell et al., 2005) provide a provisional alternative to empirical data, uncertainties related to damage prediction persist (Fuchs et al., 2019a; Merz et al., 2004, 2013; Pistrika et al., 2014; Schröter et al., 2014).

A combination of physical vulnerability methods was recommended by several studies (Fuchs et al., 2019a; Papathoma-Köhle et al., 2017; 2019) to systematically balance data-scarcity and model uncertainty. Recently, a method tailored to flood damage prediction in data-scarce regions was proposed (Malgwi et al., 2020). This

method is fully expert-based and relies on the deduction that buildings within the same resistance (vulnerability) class will incur comparable damage when impacted by the same level of hazard (Maiwald and Schwarz, 2015, 2019; Schwarz and Maiwald, 2008). The method classifies buildings using an index generated from the vulnerability indicator approach (see (Dall'Osso et al., 2009a)) and then implements a what-if analysis adapted for specific building vulnerability classes. The integration of the vulnerability index allows the consideration of multiple damage influencing variables similar to the multivariate approach. Damage is assessed using damage grades that represent repeatedly observed damage patterns within a region (Maiwald and Schwarz, 2015) since they: i) allow temporal and spatial comparison of impacts between different regions more easily (Blong, 2003), ii) improve transferability of flood damage models (Wagenaar et al., 2017), and iii) are comparable for similar building types (Maiwald and Schwarz, 2015).

In the last decade, the observed increase in frequency and intensity of climate extremes in Africa (CRED, 2019; Hoegh-Guldberg et al., 2018) has consequently resulted in an increased number of fatalities and affected people (CRED, 2019): a trend that is expected to continue given the impacts of climate change and the socio-economic development (Hoegh-Guldberg et al., 2018). Flood hazards, in particular, accounted for over 64 % of all 1164 events recorded between 2000 – 2019 in Africa (CRED, 2019). Despite huge losses to floods, studies on flood vulnerability for common building types typical in many African countries remain under-investigated. For example, while sandcrete block and clay buildings make up a high percentage of buildings in Africa (Gasparini, 2013), studies on their vulnerability to floods and damage patterns remain largely unknown (Komolafe et al., 2015). Generally, studies within these regions have been limited to exposure assessment (e.g. (Adeleye et al., 2019; Ndanusa et al., 2018; Nwilo et al., 2012; Zumo, 2014)) or identification of vulnerability indicators within selected study regions (e.g. (Akukwe and Ogbodo, 2015; Okoye et al., 2015)) without further relating indicators with expected building damage. More recently, Englhardt et al. (Englhardt et al., 2019) presented one of the first studies that relate building vulnerability and absolute damage values for different building types within an urban and rural setting in Ethiopia. The study, which aims at large scale building damage assessment, uses a 15" X 15" grid resolution database of building inventory from imageCAT (http://www.imagecatinc.com/) to reclassify buildings into vulnerability classes before applying stage-damage curves to estimate absolute damage using hazard maps with different return periods. The study by Englhardt et al. (2019) represents an important step for understanding building vulnerability and damage assessment for selected building types in Africa. However, a limitation of the study by Englhardt et al. (2019) is that out of 23 studies considered as a basis for generating the stage-damage curve, only two were from regions with comparable building types similar to those in many African countries. Such regional differences in building characteristics have been pointed to limit model transferability (Maiwald and Schwarz, 2015; Papathoma-Köhle et al., 2017).

In Nigeria, floods have become a yearly event with huge human and economic losses (Komolafe et al., 2015). In 2012, Nigeria experienced one of the worst flooding on record with 28 out of the 36 states being affected (FGN, 2013). The 2012 floods resulted in 364 fatalities, and over 3.8 million displaced people across the

country (FGN, 2013). An assessment of the 12 most-affected states from the 2012 flood showed that over 1.3 million houses were fully or partially destroyed with an estimated monetary value exceeding 6.8 billion USD (FGN, 2013). The 2012 Nigerian flooding further exposed the considerable physical vulnerability of buildings and a stronger need for increased efforts in risk reduction targeted on the built environment in Nigeria. Such disaster risk reduction efforts are required even more urgently given observed trends in extreme events resulting from climate change and the current increase in population in exposed regions within the country (CRED, 2019).

In this paper, we implement and evaluate the performance of i) an expert-based flood damage model (Malgwi et al., 2020), and ii) a multivariate data-driven model using random forests. This paper aims to: i) assess the prediction accuracy of a fully expert-based model, ii) comparatively assess the expert-based and data-driven method, and iii) gain insights on main regional damage drivers typical for Nigeria. Both the expert-based and data-driven methods are evaluated using flood damage data collected from two separate study regions and flood events in Nigeria. This study provides one of the first attempts at using flood damage data for typical building types in Nigeria. The study demonstrates the potential of using expert-based methods for flood damage prediction in data-scarce areas and provides recommendations for improved performance.

2 Study regions and flood events

Two study regions are used for this study. Both are located in the north-central (study region 1) and north-eastern (study region 2) parts of Nigeria (Figure 1). Although the regions are about 500 km apart, they share similar climatic characteristics typical of a guinea savannah with distinct dry and wet season (Ayanlade, 2009; Eludoyin and Adelekan, 2013). Both regions are predominated by low-lying areas and small river channels. A light gray canvas base map provided by Esri et al. (2020) is used to show the location of inspected buildings and river channels in both study regions.

Study region 1 is located between Suleja and Tafa in Niger State, Nigeria (Figure 1 A). Located about 30 km north-west of the capital Abuja, the population in Suleja and Tafa is around 215,000 and 83,000 respectively (NBS, 2012). The area has several river reaches passing through the settlement. Although some studies identified the region as non-vulnerable to floods based on terrain analysis (see (Mayomi et al., 2014; Ndanusa et al., 2018)), recently, high-magnitude events occurred. Heavy rainfall was reported on 8 and 9 June 2017 resulting in severe flooding of settlements located in the area. The floodplains of all five river reaches were flooded (see Figure 1). Several hundred people were affected and 18 fatalities claimed as well as substantial damage to hundreds of residential buildings and infrastructural facilities were reported. Losses from the flood event were mostly attributed to buildings constructed very close to the river channels and blockage of drainages as a result of transported material (Adeleye et al., 2019). Google Earth satellite images from June 2016 (prior to the flood event) (Google Earth Pro, n.d.) show that the area is dominated by dense settlements and sparse grassland vegetation.

Study region 2 is located in Wuro-Jebbe in Yola-north, Adamawa State, Nigeria (Figure 1 B). The population of Yola-north is 198,000 (NBS, 2012). Wuro-Jebbe is about 4 km from River Benue, which is the second major

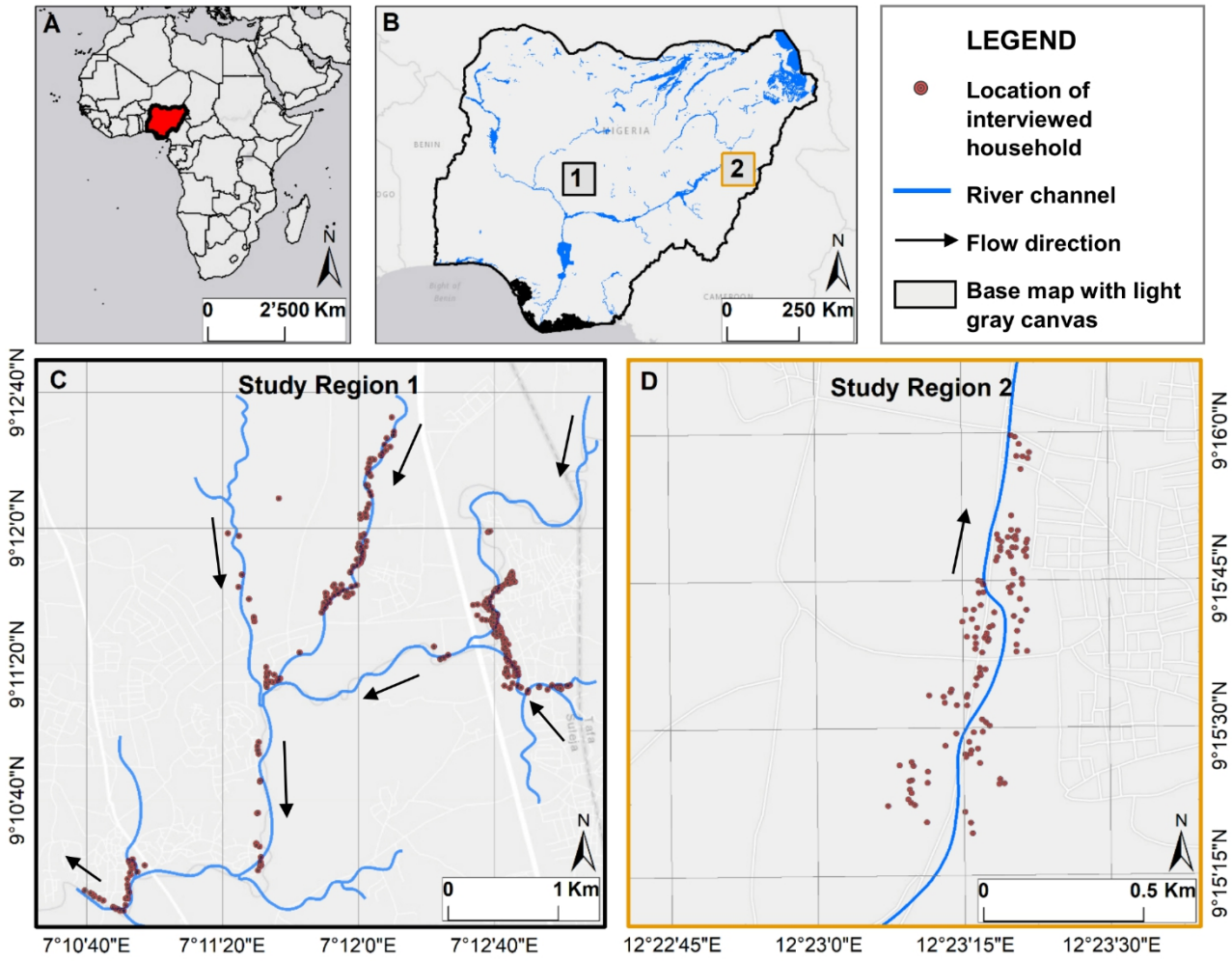


Figure 1: (A, B) Inset maps showing the location of Nigeria, (C) Study region 1, and (D) study region 2. A global light gray canvas satellite image provided by ESRI (Esri et al., 2020) is used to show the location of inspected buildings and river channels.

rivers in Nigeria (see inset map, Figure 1). The region is characterized by scattered grassland vegetation and a single river channel that passes through the settlements and flows downstream into river Benue (Figure 1 B). High-intensity rainfall on 1st August 2019 resulted in the flooding of many regions in Yola-north (OCHA, 2019). Available reports compiled for all affected areas showed that 15 persons were reported to have died from the flood and 5000 persons were displaced (OCHA, 2019). Within the selected study region (Wurro-Jebbe), hundreds of houses were reported to have been damaged as a result of the flood. Locals in Wurro-Jebbe reported that the rainfall lasted for about 5 hours (between 12:00 and 17:00) and resulted in high flow velocities. Google Earth satellite image (Google Earth Pro, n.d.), taken in April 2019, shows that buildings are relatively sparse and very likely a more recent settlement compared to study region 1.

3 Methods

This study uses a combination of flood damage data from field interviews, expert interviews, and statistical analysis to gain insights into regional drivers and how they contribute to model performance. A general categorization of damage drivers into i) impact (or action), and ii) resistance variables, as proposed by several studies (see Kreibich et al., 2010; Schwarz and Maiwald, 2007b; Thieken et al., 2005) was adopted as shown in Table 1. The definition of all variables in Table 1 is presented in the Appendix. Action variables are related to flood hazard characteristics such as flood depth. Resistance variables relate to building and exposure characteristics that influence the degree of hazard impact on a building. In this study, we further categorized resistance variables into susceptibility, local protection, and exposure variables (Malgwi et al., 2020) (Table 1). Susceptibility relates to the inherent structural characteristics of a building without considering the measure for flood protection. Local protection variables refer to features of a building that directly or indirectly served to reduce the impacts of the flood. Local protection includes additional building features that are not necessarily required for the functionality (or stability) of a building but helps to reduce flood impact (Holub and Fuchs, 2008). Exposure variables relate to the characteristics of the natural or built environment that can reduce or exacerbate the impact of floods on a building.

3.1 Data collection

Field data collection consisted of a house-to-house interview using structured questionnaires. Before the data collection, preliminary assessments of the study regions and flood events using media reports, photos, and videos from different sources was carried out. Information gathered from preliminary assessments were used to locate and map affected areas using Google Earth satellite images. A sampling of buildings for interviews was limited to i) houses affected by the floods, and ii) availability of household owners (or a community representative) to provide the required information. Questionnaires were developed such that interview questions covered i) building and exposure characteristics, ii) flood characteristics, and iii) damage sustained and/or repairs done after the flood (see supplementary 1 (S1), questionnaire 1). Flood depths were either measured directly (where flood marks exist) or based on personal recollections from residents. A Garmin etrex 2000 handheld GPS device was used to track spatial location (longitude, latitude, and elevation above sea level) to allow for georeferencing of each inspected building (Figure 1). Exposure variables such as distance to the river channel and sheltering were extracted (or supported) by information from Google Earth satellite maps. In study region 1, data collection was carried out between 1 March – 31 May 2018 (three months), i.e eight months after the flood event. In study region 2, the data collection started three months after the flood event and was conducted between 7 December – 15 January 2019 (approx. 1 month).

Observations from the field show that both study regions share a comparable building, floodplain, and channel characteristics. Buildings in both regions are predominantly one-story constructed from either sandcrete block, clay, or a mixture of both sandcrete and clay material. The river channels in both regions are semi-natural and

Table 1: Summary of damage influencing parameters

Category	Sub-category	Variable	Sub-variable	Abbreviation
Action variables		Flood depth		wat_hei
Resistance variables	Susceptibility variables (s)	Construction material	Wall material	wall_mat
			Wall thickness	wall_thic
		Building condition		buil_con
		Building quality		buil_qua
		Height of opening		hei_open
		Building footprint		buil_footprint
	Local protection variables (l)	Wall Finishes/cover	Wall plaster material	wallplas_mat
			Extent wall plaster	wallplas_ext
		Fencing	Fencing material	fenc_mat
			Extent of fencing	fenc_ext
		Ground floor elevation		grofloor_ele
	Exposure Variables (e)	Distance to channel		dis_chan
		Functional drainage		drainage
		Natural barriers		nat_barr
		Sheltering		shelt

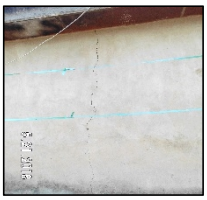
shallow with channel bed elevation mostly between 0.5 m to 1.5 m. In both regions, the cross-sectional width of the river channels highly varies between 4 m to 18 m. Typical for guinea savannah, the rivers are dry in the dry season and surface run-off begins at the start of the rainy season (around April).

Damage classes were compiled based on repeatedly observed damage patterns within the region as proposed by Schwarz and Maiwald (2008). Before the field data collection, a combination of literature review and damage reports was used to compile a list of damage patterns (c.f. S1, questionnaire 1). While on the field, damage data was documented based on i) visual observations of damage (where the repair was not yet carried out), or ii) deduced from visual observations of repairs that were already carried out. Where new patterns were observed, which were not included in the preliminary damage list, an update is made. For example, ground floor settlement was included based on observations after the data collection in study region 1. Documented damage was specifically related to the building structure itself and not for external components such as fencing walls.

3.2 Damage grades

All observed damage patterns were systematically classified using an ordinal interval into a six-class damage grade based on an increasing level of severity (Table 2). The damage grades were adapted after Schwarz and Maiwald (Schwarz and Maiwald, 2007a) with some modifications tailored for regional building characteristics. A pictorial representation of each damage pattern based on observations compiled from the field is also

Table 2: Damage grades for buildings in Nigeria (modified after Schwarz and Maiwald (Schwarz and Maiwald, 2007a))

Class	Description	Example
6	**Collapse** Complete collapse of entire building or more than 2/3 of the building	 
5	**Very Heavy** Partial collapse of building element	 
4	**Heavy** Heavy structural cracks on building elements (Walls, floors, beams, Columns) Settlement of ground floor material	 
3	**Moderate** Slight to moderate cracks on building elements (Walls, floors, beams, Columns)	 
2	**Slight** De-bonding and (or) peeling-off of building finishes	 
1	**Negligible** Water contact Moisture defects Surface cracks on floor or wall finishes	 

presented in Table 2. Damage grade class 1 represents moisture defects or non-structural (only on finishes or plaster) cracks resulting from short or moderate contact with water. Moisture defects are associated with dampness and a weakening of surface material (finishes) (Table 2). Damage grade class 2 indicates de-bonding or peeling-off (falling away) of building finishes (Table 2) resulting from an extensive weakening of the surface material which can be caused by prolonged inundation. While damage grade class 3 represents

light structural cracks, damage grade class 4 represents severe cracks on the building (Table 2). Both damage grade class 3 and 4 are structural cracks, hence, extend beyond the finishes and occur on the main building material. Damage grade classes 3 and 4 can occur vertically (resulting from movement of the soil), or horizontally, resulting from differential water pressure. In addition to severe structural cracks, damage grade class 4 includes damage to the ground flood material, which can occur due to compaction problems from the foundation or soil layer underneath (Table 2). Damage grade class 5 represents a partial collapse of the building usually resulting from overstress or increased weakness of the building. Usually, damage classified as 5 includes a collapse of less than or about one-third of the entire building (Table 2). Damage grade class 6 represents the collapse of the entire building or more than two-thirds of the structure.

3.3 Data pre-processing and analysis

The overall data set consists of observations from study regions 1 and 2. Prior to using the data for analysis, few preprocessing steps were taken. Firstly, to maintain a consistent range across all variables (recommended in (JRC, 2008)), we implement a variable level (scaling) as shown in Table 3. A variable level is a form of internal weighting for sub-variables whereby low or high scores, with a range between 1 to 6, are assigned based on deduced influence on building vulnerability to floods (Dall'Osso et al., 2009b; Ettinger et al., 2016; JRC, 2008). The variable level is implemented in such a way that sub-variables, which are indicative of high vulnerability were given lower scores and sub-variables indicative of lower vulnerability are given higher scores (see Table 3). Secondly, a data reorganization was carried out which involves merging damage grade classes 1 and 2 (Class 1+2), as well as classes 5 and 6 (Class 5+6), while damage grade classes 3 and 4 are maintained as observed on the field. Merging damage grade classes were necessary given the low number of observations in each of the merged classes and low variation in explanatory variables between these neighboring classes. In addition, merging the initial six-factor levels into just four levels facilitated model training, as this resulted in an almost perfectly balanced data set, thereby limiting bias of the data-driven model toward damage grade classes with higher observations. Merging the damage classes was also warranted from a contextual point of view, as they exhibit very similar damage patterns: Damage classes 1 and 2 represent minor damage including moisture defects or damage to wall finishes, and damage classes 5 and 6 represent severe damage including the partial or complete collapse of the building.

Furthermore, we carried out an Exploratory Data Analysis (EDA) to better understand the distribution of the data and empirical relationships between observed variables. Given that the scale of measurement of most variables is categorical, Spearman's ρ rank correlation coefficient was computed for all variables using a variable scale based on Table 3.

For the data analysis, two approaches were used: i) all observations were considered together and model training/validation/testing was performed on the full data set, and ii) the full data set was stratified by study region, using one subset for training/validation and the second subset for testing to check transferability from

Table 3: Building resistance variables with corresponding variable levels as defined for this study

Variable	Variable levels (y)					
	6	**5**	**4**	**3**	**2**	**1**
Wall material (s_1)	Sandcrete		Mixed		Clay	
Wall thickness (s_2)	22.5			15		
Building condition (s_3)	3.5 - 4 (Very good)	2.6 - 3.5 (good)		1.6 - 2.5 (moderate)		1 - 1.5 (poor)
Building quality (s_4)	2.6 - 3 (Very good)	2.1 - 2.5 (good)		1.6 - 2 (moderate)		1 - 1.5 (poor)
Height of opening (cm) (s_5)	above 120		81 - 120			0 - 80
Building footprint (s_6)	rectangle	L-shape		Irregular		
Wall plaster material (l_1)	Cement/sand			Clay		None
Extent wall plaster (l_2)	complete			Partial		None
Fencing material (l_3)	Sandcrete	Clay or mixture of sandcrete and clay			Zinc/thatched	none
Extent of fencing (l_4)	complete			Partial		none
ground floor elevation (cm) (l_5)	81 - 100	61 - 80	41 - 60	21 - 40	5 - 20	0
Distance to channel (m) (e_1)	above 60		41 - 60	21 - 40		0 - 20
Functional channel (e_2)	yes			no		
Natural barriers (e_3)	yes			no		
Sheltering (e_4)	direct			Partial		none

Table 4: Table of influence for indicator weighting, ranging from slight influence of an indicator (1) to extreme influence (9) (modified after Saaty (Saaty, 1980)).

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Slight influence	Slight to moderate influence	Moderate influence	Moderate to strong influence	Strong influence	Strong to very strong influence	Very strong influence	Very strong to extreme influence	Extreme influence

one case study to another. The combination of data from study regions 1 and 2 was plausible given the highlighted similarity in (i) channel and flood plain geomorphology (sec. 2 and 3.1), and (ii) reported inundation sequence in both regions. Combining data from both study regions is beneficial since it alleviates the problem of overfitting a single event (Wagenaar et al., 2017). Two methods were used to analyze the data

sets: an expert-based and a data-driven approach. A confusion matrix was used to enable a systematic comparison between predicted and observed damage grades using a matrix: it enables a simple visualization for the categorical response variable whereby the diagonal of the matrix represents correct predictions.

3.3.1 Expert-based method

The expert-based method was based on three phases used to develop the model including i) development of vulnerability index, ii) development of damage grade, and iii) synthetic what-if analysis (Malgwi et al., 2020). Experts selected to participate in the study were chosen based on their background in the area of building vulnerability to floods in Nigeria. Selected experts were also from different geographical regions of the country (northeast, north-central, and south-west), and disciplines (such as geography, building or civil engineering, and environmental studies) so that evaluations received are representative from different expert communities within Nigeria.

In phase 1, the selection, weighting, and aggregation of indicators to form an index were carried out. Firstly, a literature review was used to provide a preliminary list of vulnerability indicators for Nigeria. Thereafter, we develop a questionnaire in which experts were requested to comment on the preliminary selected indicators and suggest other indicators that were not included. A second questionnaire was developed using a compiled list of indicators from different experts for indicator weighting (S1, questionnaire 2). Both the indicator selection and weighting was carried out by seven experts. A table of influence was provided (Table 4), containing a scale of influence, to maintain consistent weighting across different experts. Experts were requested to assign a weight for each indicator (variable) and indicator category: the weight is a quantitative value that represents the extent an indicator influences flood damage based on definitions from Table 4 adapted after Saaty (Saaty, 1980). Mean expert weights were calculated and a threshold of weight 3 (moderate influence), representing the first quartile of the influence indicator weighting scale (c.f. Table 4) was chosen to enable a final indicator selection for aggregation. A lower threshold value was not suitable given that it would result in selecting variables that on average were either having only slight (for a threshold of 1) or slight to moderate (for a threshold of 2) influence on vulnerability. Indicator aggregation was carried out using a simple weighted additive method (equation 1) to form a building resistance index (BRI). The BRI sums the product between the weight of each variable and the variable levels (Table 3) for all exposure (e_i, y_i), susceptibility (s_j, y_j), and local protection (l_k, y_k) variables. Variables e_i, s_i, and l_i are shown in Table 3, while y_i, y_j, and y_k are for corresponding variable levels. Each variable category (exposure, susceptibility and local protection) is divided by the number of variables (n, p, k) (equation 1) so that results are not biased towards categories with higher number of variables. Lastly, each variable category is multiplied by the corresponding weights ($Exposure_{weight}$, $Susceptibilty_{weight}$, $Local\ protection_{weight}$) (equation 1). The BRI measures the resistance to flooding that a building can offer given its susceptibility, local protection, and exposure. A maximum-minimum normalization (see (JRC, 2008)) was implemented to confine the upper and lower bounds of the BRI between 0 to 100 (equation 2). The maximum-minimum normalization uses the BRI calculated from equation 1 and a minimum and maximum BRI (BRI_{min}, BRI_{max}) values which are computed using the

minimum variable level (y_{min}) and maximum variable level (y_{max}) for all variables. The formulae for calculating the BRI_{min}, BRI_{max} are shown in the appendix in equation A4 and A5 respectively. Using a quartile classification, the normalized BRI values (BRI_{norm}) are classified into buildings with poor (low), moderate (average) and good (high) resistance to floods.

$$BRI = \frac{Exposure_{weight}}{n}\sum_{i=1}^{n}(e_i y_i) + \frac{Susceptibility_{weight}}{p}\sum_{j=1}^{p}(s_j y_j) + \frac{Local\ protection_{weight}}{r}\sum_{k=1}^{r}(l_k y_k) \qquad equation\ 1$$

$$BRI_{norm} = 100\left(\frac{BRI - BRI_{min}}{BRI_{max} - BRI_{min}}\right) \qquad equation\ 2$$

Phase 2 relies on the classification of commonly observed damage patterns into damage grades. Building damage patterns used in this study (sec. 3.2.1) were developed from a combination of literature review, building damage reports, and evaluation of field data. Additional details on damage grades are given in section 3.2.2 and Table 3.

In phase 3, a synthetic what-if analysis for the three BRI classes was carried out. For selected representative buildings in each BRI class, seven experts were asked to predict expected building damage patterns using synthetic flood depths 1 – 5 m at 1 m intervals (see S1, questionnaire 3). For each flood depth interval, experts predicted three damage states; i) Low Probable Damage (LPD), ii) Most Probable Damage (MPD), and iii) High Probable Damage (HPD). While the LPD and HPD define the lowest and highest likely damage expected, the MPD defines the most likely damage. Predicted damage from different experts was used to compute a single mean damage grade class per flood depth interval for each BRI class. The mean damage grade represents an average expected damage given a vulnerability class (BRI class) and specific water depth interval (Schwarz and Maiwald, 2007a). A Bayesian ordered logistic regression model (Andrew et al., 2020) was used to fit mean damage grades to water depth. The model fit was implemented by leveraging the 'arm' package in R. Specifically, the function 'bayespolr' was used, employing a logistic link function to assign class probabilities for each damage grade through a maximum likelihood approach. Resulting damage grade probabilities for synthetic water depth intervals represent damage curves for the three damage states MPD, LPD, and HPD for all BRI classes.

A model performance check was carried out using the MPD curves given that they define the most likely damage expected. The MPD curves for each BRI class is used to predict damage grade class for data set 1. Damage grade with the highest probability is assigned as the predicted class for the building. Separate confusion matrices were generated the three BRI classes, which were later merged using a matrix addition to generate one confusion matrix.

3.3.2 Data-driven method

The data-driven approach is based on ensembles of random forest models (Breiman, 2001). To obtain robust models and unbiased estimates of model generalization performance, a modelling strategy featuring repeated nested resampling was pursued. Since honest model quality assessment is only possible if all elements of model building are included in the resampling procedure, models that require hyperparameter optimization necessitate two nested resampling loops. The outer resampling, which provides information for model performance assessment, was realized by means of five-fold cross-validation. In the inner resampling, which is targeted at hyperparameter tuning, out-of-bag predictions were used for evaluation. Hyperparameter tuning was carried out by means of model-based optimization (Bischl et al., 2017; Probst et al., 2019). In order to minimize the variance of random partitioning within folds, the whole nested resampling procedure was repeated ten times.

3.4 Further analysis

To additionally evaluate collective outputs of both the expert-based and data-driven method, few additional steps were taken. Firstly, the selected threshold for indicator importance was chosen at weight 3 (Table 4) to ensure that variables with low mean weights were not included in the final indicator selection and index aggregation. To evaluate the uncertainty of the selected threshold was optimal, we carried out a sensitivity analysis for different threshold values and re-evaluate model performance. To do this, we varied the threshold between 4, 5, 6, and 'no threshold', and in each case, we recalculated the BRI value, reclassified each building into a BRI class using the quartile classification and predicted damage grades using the MPD curve. The sensitivity analysis was carried out using both the combined and separate data from the two study regions. Secondly, to further examine model transferability on a 'new' data, we use data from study region 1 to train a random forest model using the same procedure described in section 3.3.2. The prediction accuracy of the model is evaluated thereafter using data from study region 2. Furthermore, we carried out a critical comparison between both methods, focusing on variable importance and model performance.

4 Results and discussion

4.1 Exploratory data analysis

A total of 324 observed buildings are located in study region 1 and 120 buildings in study region 2. The maximum inspected flood depth in study region 1 is 333 cm compared to study region 2 with 147 cm. Bivariate analysis showed that water depth correlates rather poorly with damage grades (Figure 2). For example, in study region 1, over 40% of buildings with a damage grade 6 (complete collapse) were from the lowest category of water depths (0 – 50 cm) (Figure 2A). Also, out of 12 buildings with water depths above 250 cm, none sustained a damage grade 6 (Figure 2A). In study region 2, most of the buildings incurred damage grade 4 especially at

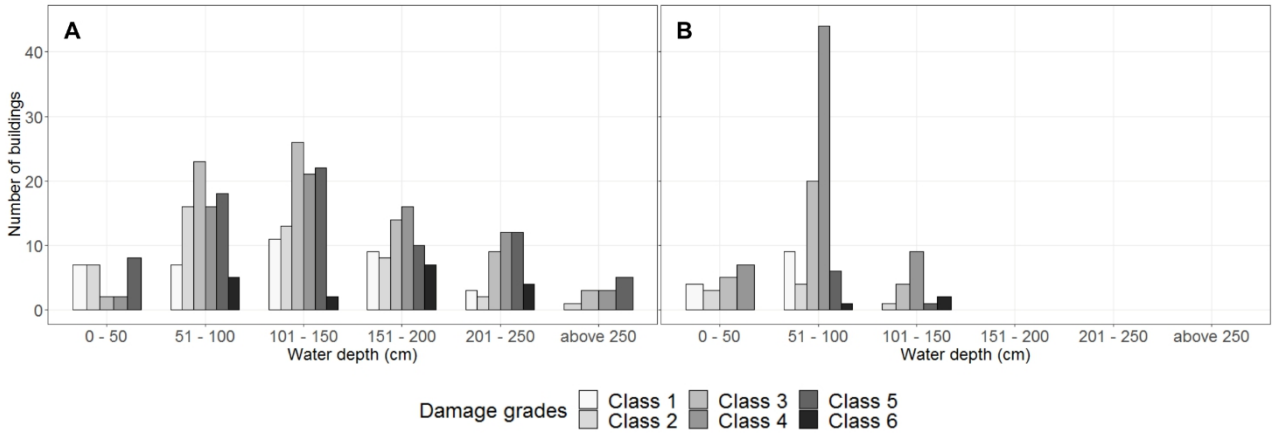


Figure 2: Bivariate plots for water depth and damage grade classes for (A) study region 1, and (B) study region 2

water depths between 51 – 100 cm (Figure 2A). In particular, damage grade 4 for in study region 2 is mostly as a result of damage to the ground floor (Table 2). Spearman correlation coefficient for the combined data showed a high positive correlation between sheltering and distance to channel (Figure 3). The correlation can partly be explained by the fact that the more distanced a building is from a river channel, the more likely it will be surrounded by other buildings. Other high correlations observed were between wall plaster with wall plaster extent and fencing material with fencing plaster extent. However, these are directly related since only buildings with wall plaster (or fencing material) have features for wall plaster extent (or fencing extent). A summary of the bivariate analyses for all resistance variables for both study regions and the combined data is presented in supplementary 2 (Figure S2-1 and S2-2). Results show that sandcrete block buildings are predominant in both regions followed by the mixed building (sandcrete and clay) and clay buildings. Another interesting result of the EDA was that while building condition and building quality were positively correlated in study region 1, they were however negatively correlated in study region 2. The reason for such difference is not completely clear, however, the negative correlation in study region 2 might be related to the fact that it is relatively a new settlement (10 – 20 years) (as visualized from Google Earth satellite image from 2004). As a result, while many buildings were identified to be in a good condition (recently built), their construction quality was not well according to standard specified by NBC (2006). Similar problems relating to building quality were identified in studies by FGN (2013) where over 60 % of inspected buildings were found to have built their houses without using the services of formal institutions.

The distribution of building damage grades in the data is shown in Figure S2-3. Damage grade 6 was generally less represented in both data sets (Figure S2-3A): this was partly because in many completely damaged buildings, residents were no longer available or no community representative could ascertain the flood depth at the building location. Merging damage classes 1 and 2 (class 1+2), and classes 5 and 6 (5+6) (Figure S2-3B) resulted in a distribution that limits bias from over (or under) representation especially for the combined data set.

4.2 Expert-based method

The list of indicators selected by experts is given in Table 1. Mean indicator weights, representing a ranking of variable and variable category importance are presented in Figure 4. The results for mean weights assigned to variables categories (exposure, susceptibility, and local protection) are shown in Figure 4A. The exposure variable was identified to have the highest influence on damage followed by local protection and susceptibility. An additive aggregation of indicators resulted in the calculation of the BRI value for each building (equation 1). For the variable (Figure 4A), distance to channel has been identified by experts to have the highest influence (weight of 7.1) on building damage to floods in Nigeria. Distance to channel is followed by building condition and functional drainage (weight of 6.4) and building quality and ground floor elevation (weight of 6.3). A threshold (cut-off) for final indicator selection (Figure 4B, dotted line) shows that two variables, height of opening and building footprint, were on average considered to have only a slight influence on damage. As a result, they were not included for indicator aggregation. Variables included in the final indicator selection (Figure 4B) represent one of the first attempts at compiling a comprehensive list of vulnerability indicators to floods for Nigeria. The additive aggregation allows compensation such that lower values for one variable can be compensated by another variable with a higher value. Results of the quartile classification for the normalized indices for all data sets are shown in Figure S2-4. The quartile classification categorizes all buildings into four equal classes with the lower and upper quartile reassigned as poor and good BRI classes respectively. The buildings in the interquartile range are reassigned as moderate BRI class. The normalized BRI has a range between 10 and 70 with a mean value of 40.

Synthetic flood damage curves for most probable damage (MPD) are shown in Figure 5 for BRI classes poor, moderate, and good. The curves were generated using probabilities estimated by the Bayesian ordinal logistic regression model. A maximum of 350 cm was used for the curves given that buildings in Nigeria are predominantly one story (NBS, 2012). The damage curve shows that a damage grade of class 3 and above is expected for buildings with poor BRI and water depths above 50 cm (Figure 5A). Buildings identified as BRI moderate and good are expected to have damage grades class 1+2 for water depths between 0 to 100 cm (Figure 5 B, C). While the range for damage grades 3 and 4 are small for the category BRI poor (50 – 150 cm), the range is wider for BRI moderate (100 – 250 cm), and good (100 – 300 cm). Complete or partial collapse (damage grade 5+6) is expected for water depths above 150 cm for building categorized as poor BRI: the same damage grade is expected at water depth above 250 cm (for BRI category moderate) and 320 cm (for BRI category good).

The overall prediction accuracy for the expert-driven method is 0.30. Heatmap and mosaic plots (Figure 6) generated from the confusion matrix showed that in general, the MPD curve performed relatively well for low damage classes (1+2, 3) compared to higher damage classes (4, 5+6) (Figure 6A). Mean correct predictions for damage class 1+2 were relatively high at 48 %, while that of damage class 3 is at 38 %. Prediction accuracy

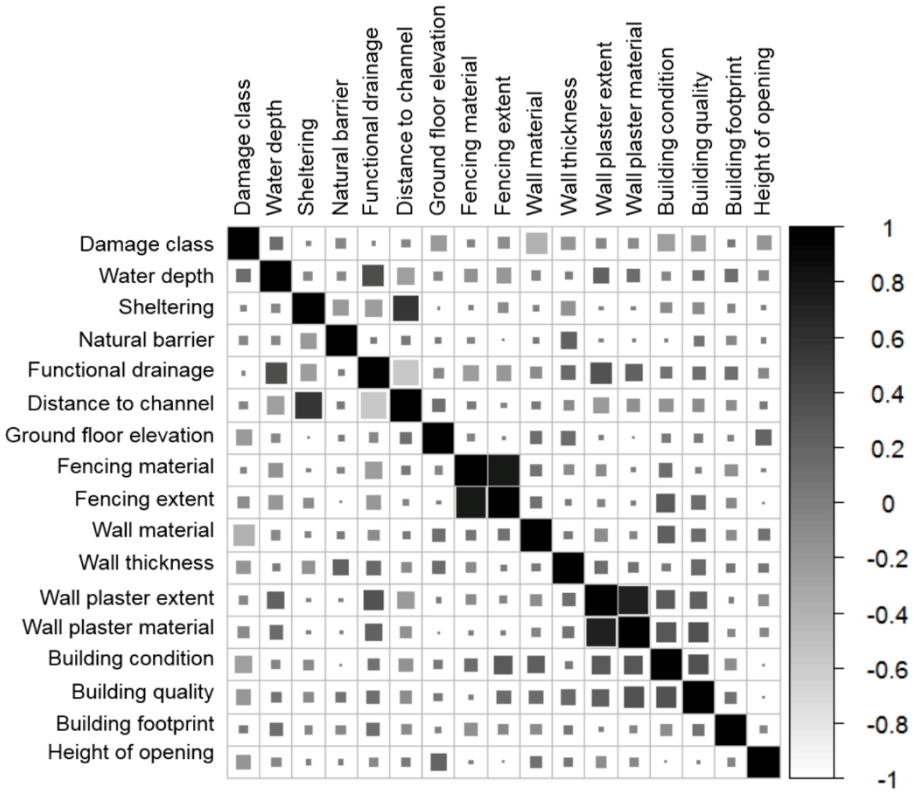


Figure 3: Spearman's correlation for the entire data. Areas of squares represent absolute values of correlation coefficients

dropped to 17 % for damage class 4 and 18 % for damage class 5+6. The MPD curve incorrectly predicted 41 % of damage class 3 as class 1+2 (Figure 6A). In addition, half of the buildings with damage class 4 were incorrectly predicted as damage class 1+2. Further evaluation of the individual confusion matrix generated for BRI classes poor, moderate, and good (Figure S2-5) showed that a substantial part of the misclassifications was resulting from buildings observed as damage classes 3 and 4 being predicted as class 1+2: this is especially high for BRI classes moderate and good with about 50 – 60 % misclassification. In BRI poor, the main misclassification is from observed damage classes 1+2 and 4 being predicted as class 3. Relatively high accuracies in the classification of low damage grades are likely because a high percentage of the observed water depths were less than 150 cm (Figure 2), and within this range, the damage probabilities of the MPD curve is high for low damage (see Figure 6). Conversely, poor accuracy in predicting high damage grades is related to the underestimation of damage grade classes at low water depths. From the MPD curve (Figure 6) higher damage was only assigned for high water depths and low damage for low water depths. However, as seen from the bivariate analysis (Figure 2), higher damage occurs even at lower water depths. The use of a water depth range between 1 to 500 cm for the expert what-if assessment might have influenced the results given that experts become likely to associate higher water depth (400 - 500 cm) with the higher damage grades (class 6).

Results for the mean low probable (LPD) and high probable (HPD) damage are shown in Appendix, Figure A1.

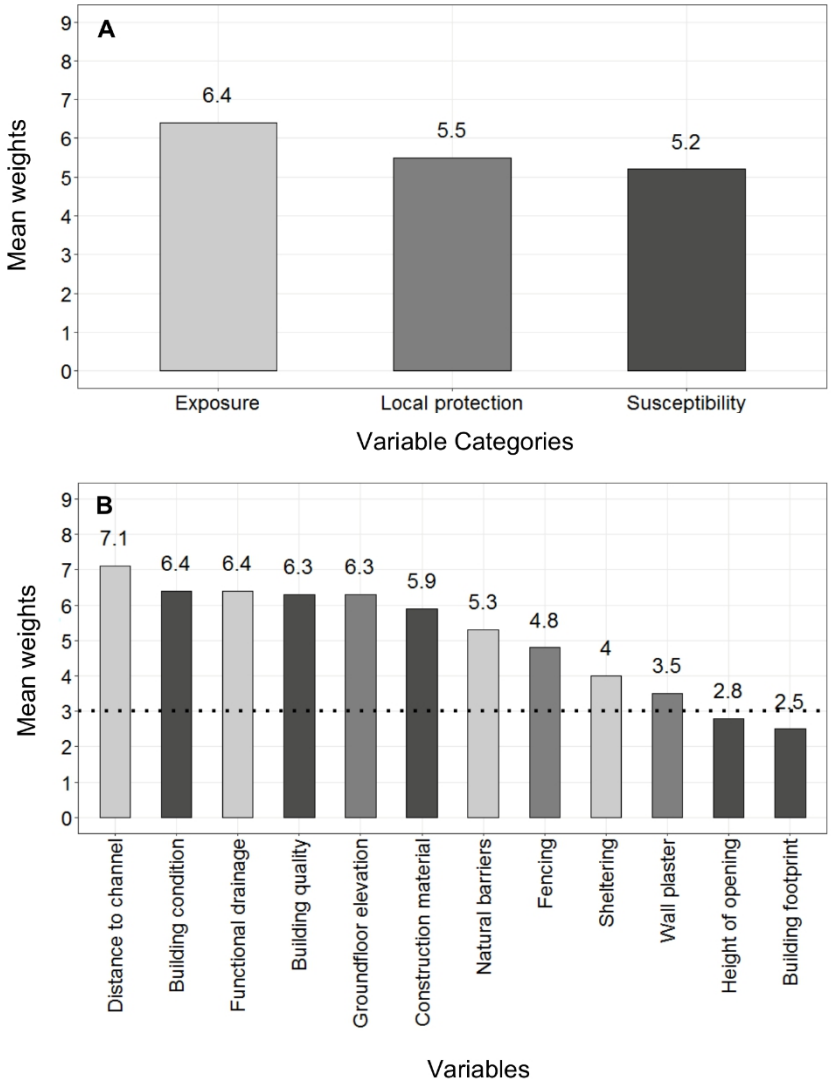


Figure 4: Mean expert weights for (A) variable categories, and (B) variables (or indicators). Dotted line in plot A indicates the selected threshold.

Generally, LPD and HPD aim to accommodate variations in building characteristics within each BRI category. For example, for buildings classified as BRI poor, the lowest damage expected for a 200 cm flood depth is damage class 4, while the highest damage expected is a damage grade 5+6 (Figure A1 A). For the same flood depth (200 cm), we expect the lowest damage of class 4 for BRI moderate or class 3 for BRI good (Figure A1 C, E). The highest probable damage for a 200 cm flood is a class 4 or Class 5+6 (both have equal probabilities) and a class 4 for BRI good (Figure A1 D, F). Between 150 – 250 cm water depth range, the LPD and HPD for BRI moderate overlap, both predicting damage class 4 (Figure A1 B). Generally, for most of the LPD and HPD curves, a change in damage grades class consistently occurs at around 100 cm. The change might be related to the height at which water starts entering the building: in Figure S2-1O, over 60 % of the entire data have a height of openings between 81 – 120 cm.

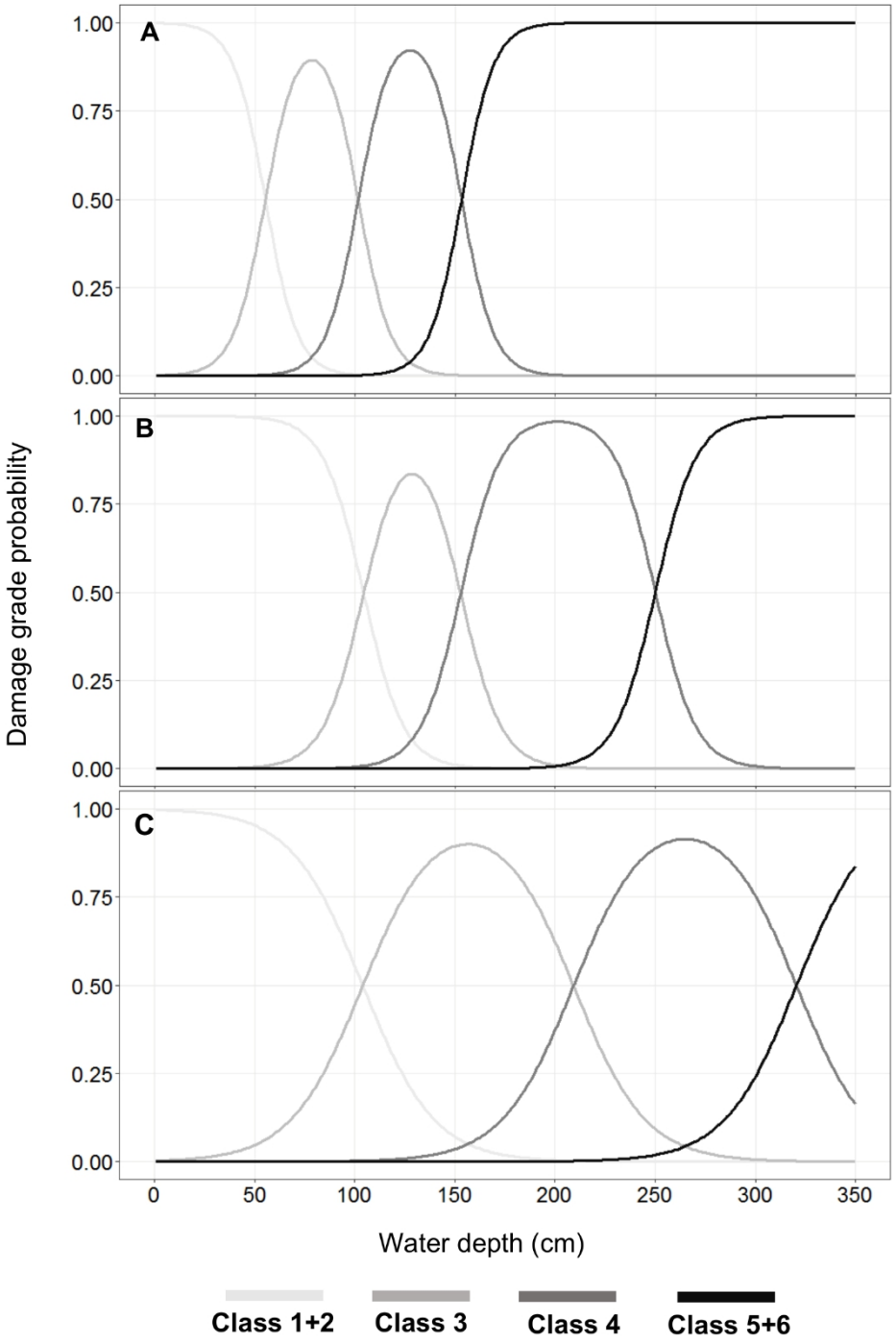


Figure 5: Most Probable Damage (MPD) curves for (A) poor, (B) moderate, and (C) good BRI class

4.3 Data-driven method

The data-driven approach results in an overall multiclass accuracy of 0.38. However, results vary across the different classes as depicted by the heatmap and mosaic plot resulting from the confusion matrix (Figure 7). Generally, the model performed well for severe damage (Classes 4, 5+6,) compared to lower minor or moderate damage (class 1+2, 3). While more than 50% of all instances of classes 4 and 5+6 are predicted

correctly,

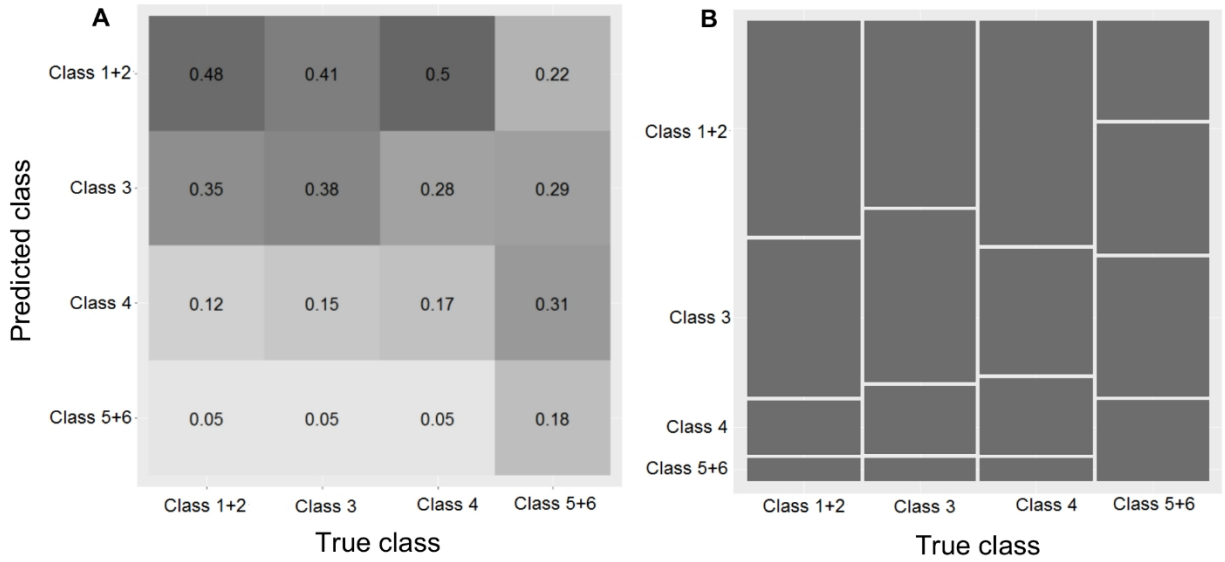


Figure 6: Confusion matrix for observed and predicted classes using the MPD curve. The heatmap (A) indicates the percentage of values in each cell with respect to the total number of instances in the true class (i.e., columns sum up to 1). The mosaic plot (B) is a graphical illustration of the conditional relative frequency for each combination. The area of the tiles is proportional to the number of observations exhibiting the respective combination of factor levels, i.e. the corresponding joint frequency.

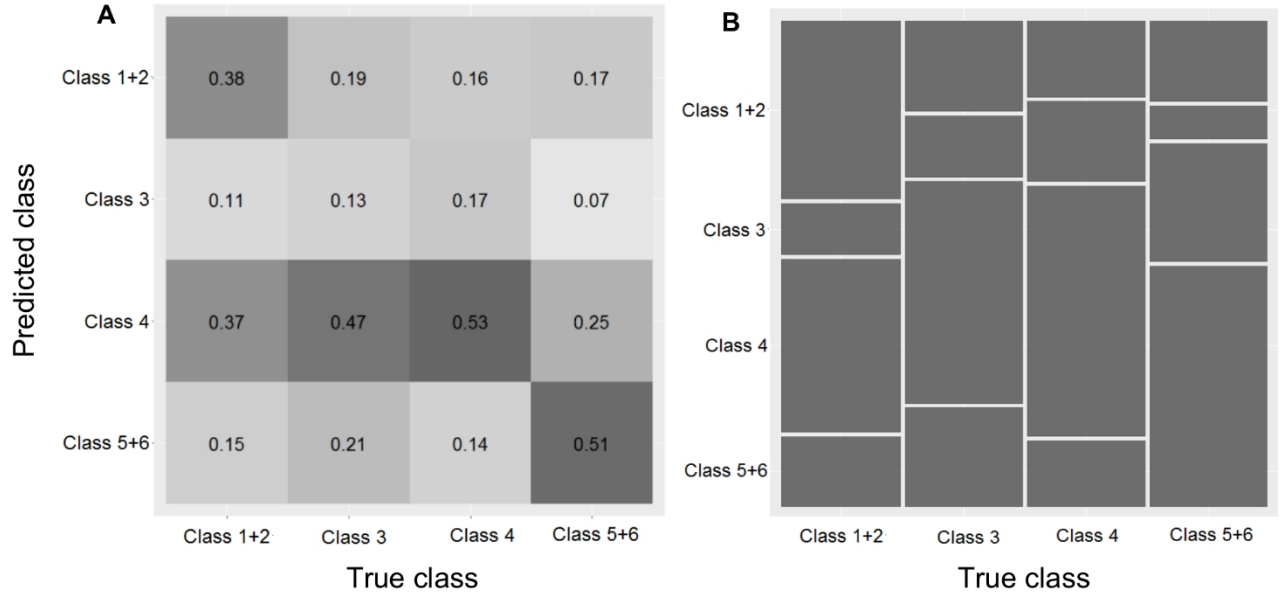


Figure 7: Confusion matrix for observed and predicted classes using random forests. The heatmap (A) indicates the percentage of values in each cell with respect to the total number of instances in the true class (i.e., columns sum up to 1). The mosaic plot (B) is a graphical illustration of the conditional relative frequency

for each combination. The area of the tiles is proportional to the number of observations exhibiting the respective combination of factor levels, i.e. the corresponding joint frequency.

accuracy drops to 38% for class 1+2 and down to only 13% for class 3 (Figure 7). About half of the buildings observed as class 3 (47 %) were predicted as class 4. Also, high number of class 1+2 is predicted as class 4 as well. Density estimates (Figure 8) of predicted class probabilities for all combinations of the confusion matrix show interesting patterns. Some correctly classified instances of class 5+6 values are predicted with relatively high certainty. Predictions for class 4 exhibit a mean clearly above 0.25 (which would correspond to random guessing given a classification problem with four classes) (Figure 7), with predicted class probabilities up to almost 0.7. Class 3 exhibits a very balanced prediction across all four true classes, with distinct peaks around or slightly below 0.25, reflecting the overall low prediction of class 3 instances. Among all class 3 predictions, most belong to observed class 4. Predicted class probabilities for class 1+2 exhibit the best result for true class 1+2 observations, with class probabilities up to 60%, but also show between 15% and 20% false predictions in each of the other classes with single cases exhibiting similar class probability of up to 60%.

In terms of variable importance, wall material clearly prevails as the most important variable (Figure 9). Distance to channel emerges as the second most important variable, exhibiting a large importance gap not only to wall material but also to wall thickness, which is ranked third (Figure 9). The identification of wall material and distance to channel as the two most important damage influencing features was not entirely unexpected, given that i) buildings with sandcrete block have relatively consistent stability after contact with water compared to with clay material (especially unburnt clay), and ii) buildings constructed closer to the channel have higher water depths on average and are less sheltered by other buildings as shown by a high correlation between the two variables (Figure 3). Building condition and quality were also ranked to be relatively important variables. Both wall plaster material and extent were identified to have low variable importance. Given their importance in delaying the intrusion of water into the primary wall material, it remains unclear why they do not significantly contribute to explaining the variance in building damage. The low variability in factor levels of the wall material and extent (Figure S2-1) is a probable suspect to their performance. Functional drainage was identified to be the least important variable. The poor performance of functional drainage was unexpected given its identification as important damage influencing variables in previous studies (e.g. (Adeleye et al., 2019; Musa et al., 2015; Okoye et al., 2015)). However, the distribution of the data for functional drainage (see Figure S2-1D) shows that most observed buildings (about 80 %) do not have drainages or the drainages were blocked (both classified as ‘no functional drainage’), consequently limiting variability in the variable and hence its performance in the model.

Concerning predicted class probabilities, the results of the repeated nested resampling approach present a consistent picture across all ten models (Figure 8). Density plots of predicted probabilities show a favorable distribution for correct class 5+6 predictions, with some instances being predicted with a probability of up to almost 90%. Densities of class 5+6 predictions for all other classes show modes of around 0.1. Predicted

probabilities for observations predicted as class 3 and class 4 are more ambiguous. While densities for class 4

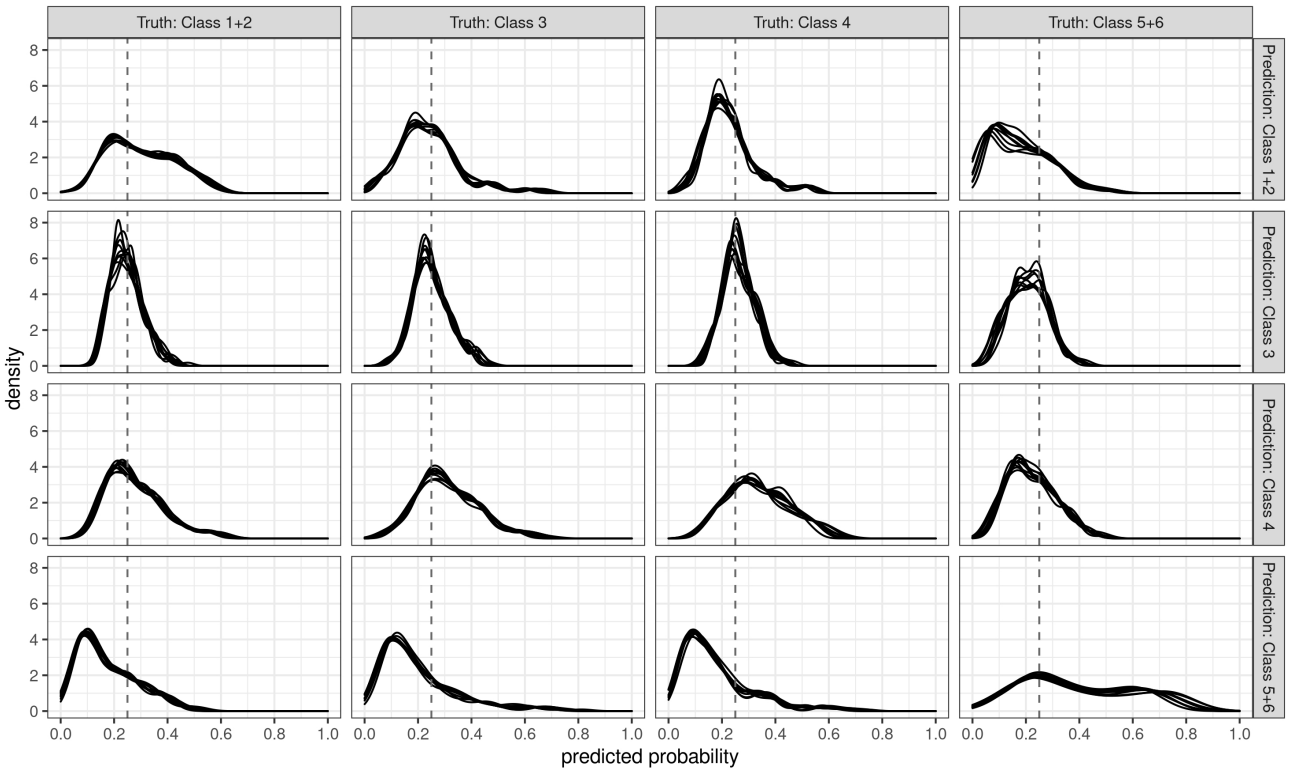


Figure 8: Densities of predicted class probabilities for all class combinations. Results are presented for all ten models obtained via the repeated nesting resampling procedure. The vertical dashed line at 0.25 indicates the threshold for random guessing.

predictions indicate comparatively high predicted probabilities not only for true class 4 instances but across all other classes as well, class 3 exhibits narrower densities with lower-class probabilities. True class 1+2 predictions exhibit a very similar pattern to predicted class 4 probabilities for class 1+2 instances, thereby reflecting the percentage of 38% true class 1+2 predictions and 37% of class 1+2 erroneously predicted as class 4.

4.4 Model comparison

Interesting implications can be drawn from a comparison of the expert-based and data-driven method. Firstly, the variable importance between the two methods reveals both similar and contrasting deductions. In both methods, distance to channel was consistently identified as significant damage influencing variable. The consistency comes from its (i) repeated performance across the ten models of the random forest, and (ii) high mean weights from multiple experts compared to other variables. Wall material, although identified as the most important in the data-driven method, was moderately important in the expert method. Both methods rank building condition as more important than building quality, hence suggesting that a building with a low quality of construction is likely to perform well when it is properly maintained. Functional drainage, ranked

second

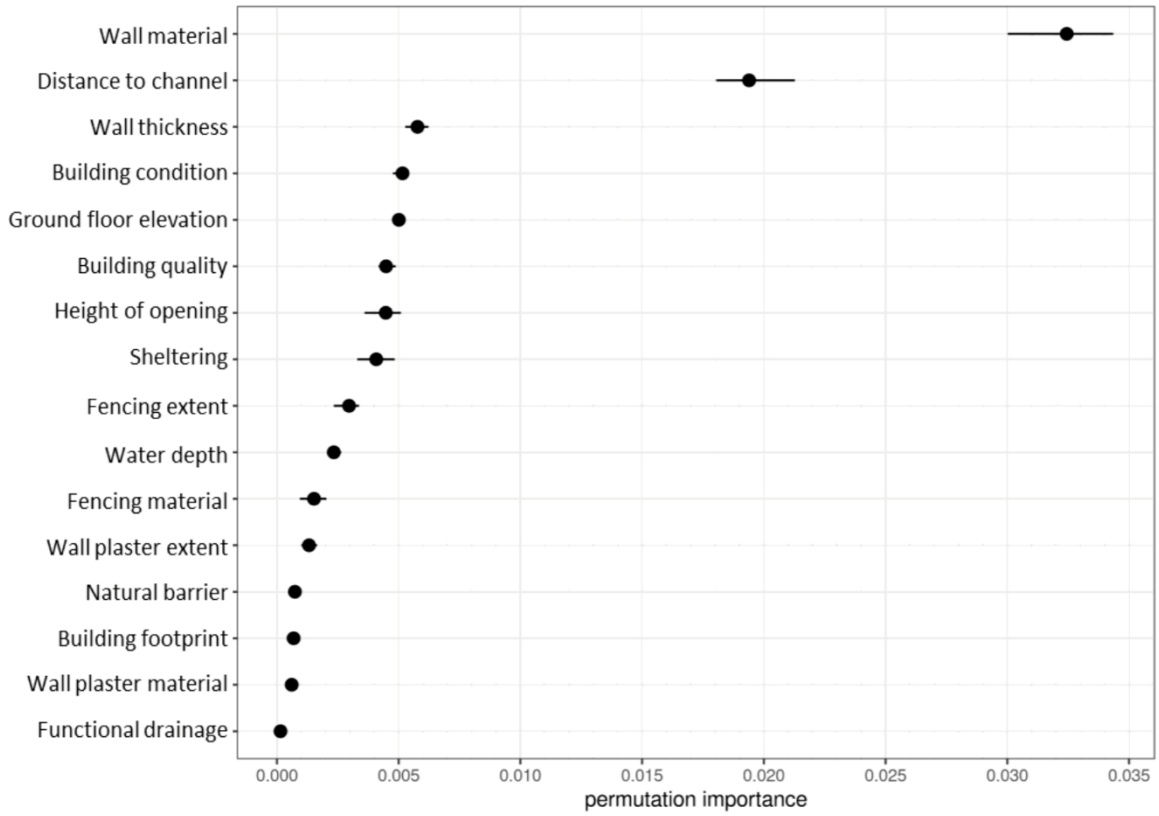


Figure 9: Random forest ensemble variable importance. Both the mean (black dot) and the range across the ten-member ensemble are depicted.

most important variable in the expert-based method, is the least important variable in the data-driven approach. Given the limited variability of functional drainage in our data (Figure S2-1D), this variable may require further evaluation. Furthermore, while the height of opening was identified to be relatively important in the data-driven method, it was averagely weighted as only 'slightly important' by experts. Consequently, it was removed from the analysis based on the selected threshold (Figure 4A).

Model performance was measured using the percentage of correct predictions in each damage class. Overall (multiclass) prediction accuracy for the data-driven method (38 %) was higher than that of the expert-based approach (30 %) by 8 %. Better performance of the data-driven method was not unexpected given that the implemented approach (random forest) is a supervised learning method that uses the observed damage for model training. On the other hand, the expert-based method was 'unsupervised' given that no data on observed damage grade classes were used: it relies on a prior classification of buildings into resistance (vulnerability) classes based on building and exposure characteristics, and a what-if analysis based on expert knowledge. The data-driven method performed well especially for higher damage grades (classes 4, 5+6) with predictions at over 50 % accuracy (Figure 7). The accuracy of the data-driven method dropped for low damage grades, especially for damage class 3 with the least accuracy at 17%. Apparently, class 3 and class 4 are difficult to distinguish based on the available data, implying that the collected variables do not have

enough explanatory power to better separate between these two classes. The low accuracy for class 3 is mainly because observed damage class 3 instances are mostly overestimated and predicted as damage class 4.

Conversely, the expert-based method performed well at low damage grades (class 1+2, 48% and class 3, 38 %) compared to high damage grades (class 4, 17% and class 5+6, 18%) (Figure 6). However, this can be attributed to the fact that the expert-based method is biased towards underestimating the damage class in general. Since low damage classes are predicted much more frequently, the number of correct predictions seems to be high. At the same time, the number of buildings incorrectly predicted to belong to low damage classes is high as well. These results suggest that the data-driven method outperforms the expert-based method, especially for predicting damage classes 4 and 5+6. Results for low damage classes have to be interpreted carefully. While the data-driven method generally overestimates low damage classes, the expert-based method clearly underestimates high damage classes. The highest misclassifications in the data-driven approach were related to 47 % of class 3 being predicted as class 4 (data-driven method). As highlighted earlier, the small difference between moderate cracks (class 3) and heavy crack (class 4), might result in difficulty for the model to distinguish the two classes. In the expert method, 50 % of class 4 and 41 % of class 3 were predicted as class 1+2, further underlining how experts underestimated damage for low water depths. In both methods, while class 1+2 has the highest cumulative prediction accuracy considering both methods (38% in the data-driven and 48% in the expert method) class 3 had the worst performance (13% in the data-driven and 40% in the expert method). In general, given limited research on typical sandcrete, clay, and mixed buildings used in this study, and the unsupervised approach adopted by the expert-based method, a 30 % prediction accuracy is considered satisfactory and demonstrates the potential of the method in data-scarce areas.

4.5 Model transferability

To ensure that the spatial extent of model applicability is not entirely limited to the selected study regions, some steps were taken regarding data used and methods applied. For example, the use of a merged data set (section 3.3) reduces model overfitting to a single event. In the expert-based method, experts were chosen from different locations within Nigeria (section 3.3.1), so that variable (indicator) selection, weighting, and what-if analysis were reflective of different regional damage drivers, physical geomorphology, and hazard characteristics. On the other hand, the data-driven model implemented a nested resampling technique (sec 3.3.2) to allow for unbiased performance estimates on the combined data. Model results for training a random forest model using data from study region 1 and testing on study region 2 show the same multiclass accuracy of 0.38 similar to the model developed from the combined data set. Heatmap and mosaic plot (Figure S2-5) generated from the confusion matrix of true and predicted damage classes show the highest prediction accuracy for class 5+6 at 70%. All the other classes show low accuracies: class 1+2 (29 %), class 3 (24 %), and class 4 (32 %).

Given the highlighted steps and the overall accuracy of the model (0.38) on a different flood event and spatial location, the transferability of both models is highly plausible particularly in regions with similar building

characteristics. In general, we further recommend the application of both methods in regions with comparable building characteristics so that model transferability can be further evaluated. Alternatively, spatial resampling such as spatial cross-validation can be used to alleviate the problem of spatial autocorrelation, which might lead to overoptimistic results (Brenning, 2005; Schratz et al., 2019).

4.6 Model uncertainties and outlook

Several uncertainties regarding data collection and analysis exist. In this section, we discuss these uncertainties and provide recommendations for future studies.

A basic input for the study is the field data collected by interviews. As a result, model output relies on the accuracy of such data, which in turn partly relies on personal recollections by the building residents during field interviews. Such personal reflections present some uncertainties especially if the field surveys were conducted long after the flood event. In study region 1, field data collection was carried out eight months after the flood. While data such as susceptibility, local protection, and exposure could be directly observed on the field and satellite imagery, water depths are not directly deductible. Yet, a large percentage of the water depths used in the study were based on personal reflections by building residents and not on measured watermarks. In study region 2, field surveys were conducted three months after the flood event, hence, residents have a higher tendency to remember flood depths with good accuracy. In general, floods are traumatic events (Fontalba-Navas et al., 2017; Mason et al., 2010; Verger et al., 2003) and people affected by them do not easily forget details about the events. However, where possible, data collections should be carried out as soon as possible after flood occurrence to generally ensure higher accuracy in water depths.

Bivariate analysis showed that the distribution of some variables is skewed and the data-driven model may favor factor levels that are over-represented. For example, variables such as functional drainage, natural barrier, wall thickness are characterized by imbalanced factor levels, with one dominating manifestation of the variable. However, since the random forest algorithm is a non-parametric classifier, there are no prerequisites with respect to the distribution assumptions of input data. The repeated nested resampling strategy was applied to obtain a robust model with honest performance estimates. Models fitted using a simple single train-test split might still be prone to suffer from a slight bias caused by overfitting. For some variables, the disproportionate distribution of the data (e.g. sandcrete block in wall material) is representative of the actual situation (see data from (NBS, 2012)).

Additional uncertainties relating to non-inclusion of other variables that may influence building damage exists (i.e. unobserved heterogeneity). For example, two resistance variables, building age and number of floors, initially selected by experts had to be removed. The building age was highly incomplete because residents could either not remember the year of construction or simply have no knowledge about it since they were not the first to reside in the house. The number of stories was removed due to a lack of variability: out of the entire data comprising 444 buildings, only one building was two-story while all the others were one story (only ground floor). Other hazard variables that could have been interesting for the study were flood duration

and velocity. Many residents qualitatively described how the flow approached at 'high' speed. Other residents, especially those residing in clay buildings, suggested that their houses were damaged due to longer durations of exposure to water. Given high uncertainties in translating qualitative and quantitative reports on flood velocities and duration, both variables were excluded from the data. For example, many residents have stated that during flooding, it is very difficult to keep track of time. Hence, in some cases reported durations for two buildings next to each were more than six hours apart. A method for hydrodynamic modelling for data-scarce regions without hydrological data has been proposed in (Malgwi et al., n.d.). Such methods provide a pathway for flood data extrapolation (flood duration and velocity) to complement current efforts for flood damage modelling in data-scarce regions.

Predicting damage grades presents a rather challenging task which is also evident from the mediocre performance observed in the random forest model. Such difficulties are even more pronounced in regions where policies on building standards are less well implemented, which results in substantial variation in building quality and value (Englhardt et al., 2019; FGN, 2013; Malgwi et al., 2020). The variations in building quality contribute to uncertainty in damage prediction since buildings within the same BRI category (poor, moderate, good) may not incur comparable damage even when impacted by the same flood depth. The LPD and HPD curves (Figure A1) are developed to reduce uncertainties inherent in each BRI class. We recommend a further sub-classification within each BRI class such that MPD is only used for buildings that are more comparable to the representative building. LPD and HPD are then applied to buildings with characteristics (or calculated BRI value) deviating from the representative building. Where high variations in building standards do not exist, the recommended additional sub-setting in each category will not be necessary. In general, we recommend further performance assessment of the MPD, LPD, and HPD curves in different data-scarce regions with comparable building characteristics.

Expert interviews are generally subjective, hence present some uncertainties. As a result, a high number of experts are required so that results are representative. In practice, getting a high number of experts is not always feasible, especially in regions where the required expertise is limited. In Nigeria, the challenge regarding the low number of experts in flood damage and vulnerability assessment has been previously pointed out by Komolafe et al. (2015). Our study included seven experts for indicator selection, weighting, and what-if analysis. Selected experts were chosen from different geographical regions and fields of study, which generally influences how they carry out the assessment. For example, in the what-if analysis, while experts with an engineering background included 'ground floor settlement' in their assessment, it was partly challenging for few others (with other backgrounds) to relate the variable with a water depth range. We generally recommend that for future studies, the what-if analysis should be conducted within the framework of a workshop during which all relevant information regarding damage states, water depth classification, and representative buildings from BRI classes are properly discussed. Such workshops will reduce consequent uncertainties that arise from knowledge gaps based on expert background. As highlighted above, the general poor performance of expert-based assessment for higher damage may be related to using water depth ranges

up to 500 cm. It would be interesting for future studies to re-evaluate the outcome of such what-if analysis using a maximum water depth of 350 cm to evaluate the influence of maximum water depths on damage grade estimates.

The results of the sensitivity analysis are shown in Figure S2-6. Generally, percentage mean correct predictions for varying threshold values were marginally comparable across each data set. The highest variation in prediction accuracy within the same data was within a 5 % difference as observed in data set 2 between a threshold value of 5 and 6. Observed low sensitivity of varying threshold values might be related to two reasons. Firstly, variables with low mean weights do not significantly contribute to the BRI, hence their removal results in only slight (or no) changes on the BRI classification. Secondly, the quartile classification of the BRI was relatively conservative since it maintains the same number of observations between each quartile. Also, for the quartile classification, the classification of BRI moderate is relatively large (interquartile range). As a result, it limits the ease with which small changes in BRI will alter a buildings' class, except for buildings at class boundaries. Generally, the low sensitivity of different threshold values suggests that even with a low number of variables (e.g. only 5 variables at a threshold value of 6), relatively comparable model performance can be achieved. Hence efforts required in undertaking field data collection can be drastically reduced and methods for rapid building vulnerability assessment can be further enhanced. Figure S2-6 indicates that data from region 1 maintained a performance accuracy above 31 % but region 2 showed a consistent low performance with an average of around 20 %. The reason for the low performance in region 2 is most likely attributable to a higher proportion of damage grade 4 at lower water depths (51 – 100 cm) (Figure 2). Field observations showed that a high percentage of the buildings (in region 2) experienced either (i) a ground floor settlement or (ii) disturbance to the compacted soil material directly below the ground floor - in both cases, resulting in damage to the ground floor cover material (class 4). The frequency in damage class 4 for region 2 may be related to the soil properties and will need further investigation. The general low water depths observed in study region 2 contributed to the low performance since the MPD curve underestimates damage for low water depths. The standard deviation for the percentage correct prediction between the BRI classes shows that study region 2 had the highest variation of percentage accuracy between BRI classes compared to study region 1 or the combined data. The high variation in study region 2 shows that poor prediction accuracy is related to one or two BRI classes showing significant low performance.

5 Conclusion

The development of flood damage models represents an important step towards flood risk disaster reduction given its multiple applications in mitigation and emergency planning, economic loss evaluation, the cost-benefit analysis for flood protection measures. However, the unavailability of well documented empirical data has so far limited the application of flood damage models in several data-scarce regions. In addition, current methods have either been limited to exposure assessment, identification of vulnerability indicators, or not well representative of regional building types.

In this study, we carried out a comparative assessment of a data-driven with an expert-based approach that does not require empirical data. The comparative assessment aimed at evaluating the prediction accuracy of the expert-based approach as well as gain understanding into regional damage drivers for typical building types. Flood damage data, collected from two different regions in Nigeria, was used to evaluate model performance. Data from the two study regions were used either combined to reduce model overfitting to a single event or separate to enable an evaluation of model transferability.

Several conclusions can be drawn from the study:

i. Generally, experts underestimated the damage to lower water depths. The MPD curves predict high damage (class 4, 5+6) only for high water depths (above 100 cm). However, the bivariate analysis showed that high damage can occur at lower flood depths, for example, some buildings categorized under BRI low incur a damage grade 6 from a water depth of 50 cm. The underestimation of damage for low water depths consequently resulted in multiclass prediction accuracy of 30 % by the expert-based model. However, given i) limited research on sandcrete and clay building types, ii) observed variation in building standards, a 30 % performance accuracy is considered satisfactory and demonstrates the potential of the method for data-scarce regions. The bivariate analysis supports other studies that demonstrate high uncertainty in predicting building damage using only water depths.

ii. Damage grade prediction presents several challenges especially in regions with relatively high variation in building standards. Consequently, the data-driven method had an average performance accuracy of 38 %.

iii. Both methods suggest that even at a reduced number of variables, comparable model performance can be achieved. Hence, efforts and time spent on field data collection can be reduced or better targeted to the most important variables. Similar conclusions were deduced in a recent study by Papathoma-Köhle et al. (2019).

iv. Buildings within the BRI classes showed considerable differences for similar water depth range. This suggests that achieving better performance with the expert-based method will require i) a re-evaluation of the variables weights or classification scheme used or ii) incorporating additional variables that were not considered in the study.

v. Both the expert-based and data-driven methods suggest that distance to channel, wall material, building condition, and building quality are important variables to be considered for physical vulnerability assessment in typical regions.

vi. Experts assessments, in particular the what-if analysis, might require a formal discussion (e.g. workshop) to bridge knowledge gaps that arise especially when experts are from different fields of study.

vii. The combination of different physical vulnerability assessment methods shows good potential for adapting flood damage models to regional situations typical for data-scarce areas. The inclusion of local experts allows the model to be tailored specifically to regional situations.

Appendix 1

Variable definition and measurement

Flood depth

The height of floodwater, measured from the ground level (not floor level) at a building location. These were measured directly if flood marks were still visible or residents were asked based on personal recollections.

Construction material

The material used for erecting the walls of the building. The construction material was further classified into wall material and thickness.

- Wall material: Three building materials are considered i) sandcrete block, ii) clay (burnt and unburnt), and iii) Mixed (a combination of sandcrete block and clay). Usually, for the mixed class, sandcrete block is used to a height of about 100 cm and the rest are completed using clay. The mixed system is mainly used in regions that are exposed to floods providing a relative balance between safety and construction cost.
- Wall thickness: This refers to the size (width) of the wall unit. Common widths for wall thickness found in the study regions are 15 cm and 24 cm.

Building condition

Assesses the maintenance status or state of the individual components of a building. In our study region, we carried out this assessment using scoring for different building components: i) 4 - very good (as new), ii) 3 - good (light deterioration), iii) 2 - moderate (average or increased deterioration), and iv) 1 - poor (severe deterioration) (Manager, 2017; Straub, 2009). Building components included in this assessment are: i) walls, ii) doors and windows, iii) finishes or plastering, iv) flooring (part of the building the room stands on), and v) roofing. The building condition is computed as an average of the assigned score overall building components as shown in equation A1. The range of the building quality lies between 1 to 4.

$$Building\ Quality = \frac{1}{5}\left(external\ wall_{score} + \ doors\ and\ windows_{score} + flooring\ + \ finishes_{score}\ + roofing_{score}\right) \quad equation\ A1$$

Building quality

Relates to the conformity of the building to country standards. In order to be consistent with established standards outlined in the Nigerian building code (NBC, 2006), building standard is assessed based on three considerations: aesthetics, durability, and functionality. A scoring system is used to qualify conformity to standards using i) 3 - good (high), ii) 2 - moderate (above average), and iii) 1 - poor (below average) conformity. Four building components used for evaluating building standard includes; external walls, doors and windows, roofing, and finishes or plastering. The building quality for each component is computed as an

average of the score for aesthetic, durability, and functionality. An example for external wall is given in equation A1. The building quality score is thereafter computed as the average of the score for all four building components as shown in equation A3. The range of the building quality lies between 1 to 3.

$$External\ wall_{score} = \frac{1}{3}\left(aesthetic_{score} + durability_{score} + functionality_{score}\right) \qquad equation\ A2$$

$$Building\ Quality = \frac{1}{4}\left(external\ walls_{score} + doors\ and\ windows_{score} + finishes_{score} + roofing_{score}\right) \qquad equation\ A3$$

Height of opening

The height of opening is the distance measured from the ground level to the lowest part of the window. If the window heights are different, the lowest window is used. The lower the height of opening, the faster flood water can gain entrance into a building.

Building footprint

Taken as the external form (or outline) of a building. In our study regions, predominant building footprints have been categorized into i) L-shape, ii) Rectangular, iii) irregular.

Ground floor elevation

The elevation of the building ground level (floor) or entrance partly influences the amount and time flood water can gain access inside a building. In certain communities where floods are relatively common, residents either raise the height of the building floor level or erect a small barrier at the doors as a local protection measure.

Wall finishes/plaster

Materials used as cover to the main wall unit are commonly referred to as wall finishes. The use of wall finishes varies from protection and aesthetics. While the recommended practice is the use of plaster on the entire building, (NBC, 2006) this is not necessarily the case in many buildings. To accommodate these differences, we further categorize wall finishes/plaster into i) plaster material, and ii) plaster extent.

- Plaster material: The material used for plaster categorized into i) cement and sand, ii) clay, and iii) none (no plaster).
- Plaster extent: The proportion of wall that is plastered categorized into i) complete, ii) partial, and iii) none (no plaster).

Fencing

Refers to a walling unit erected around a building. In Nigeria, the practice of building a fence is relatively common. Although fences can primarily be for controlling access, they reduce the direct impact of the flood on buildings. We further classify fencing into i) fencing material, and ii) fencing extent.

- Fencing material: The material used for the fencing influences how much pressure it can withstand from the flood. The fencing material is classified into i) sandcrete blocks, ii) clay, iii) mixture of sandcrete blocks and clay, iv) others (Zinc metal and thatched), and iv) none (no fencing erected)
- Fencing extent: Fencing can be erected around the entire building or allowed to cover only part of a building. Three classes of fencing extent used in this study include i) complete (fencing covers the entire building), ii) partial (fencing covers only a part of the building), and iii) none (no fencing erected)

Distance to river channel

The distance between a building and the river channel is expressed using the variable distance to river channel. Distance to channel is usually limited to the flood extent since only buildings affected by the flood are included in this assessment.

Functional drainage

Drainages function to provide channelization of excess runoff. Where such drainages are not available or are non-functional (e.g., blocked by debris), excess runoff can quickly result in floods. In Nigeria, some studies (see (Adeleye et al., 2019; Musa et al., 2015; Okoye et al., 2015)) have shown that the availability (or functionality) of drainages have remarkably contributed to flood occurrence and subsequent building damage. Here, we categorize functional drainage into two i) yes (available), and ii) no (not available or non-functional).

Natural barrier(s)

The presence of vegetation (grasses, trees) around a building is expected to influence the velocity of flood water as it approaches the building. Here, each building is classified into to i) yes (existence of a natural barrier), and ii) no (no natural barrier) around the building.

Sheltering

The availability of a structure in between a building and a flood source or preferable water path (e.g., river channel or roads) influences the impacting force (Maiwald and Schwarz, 2012). Sheltering refers to the relative protection of one building by another. Such that a direct impact is reduced or avoided. In our study, similar to Maiwald and Schwarz (2012), two considerations are used to classify a buildings' sheltering status: spatial location of a building relative to the river channel and other buildings, and direction of river flow. Buildings are classified into i) direct (complete), ii) partial (moderate), and iii) no (none) sheltering.

Equations for maximum-minimum normalization

$$BRI_{min} = \frac{E_{weight}}{n}\sum_{i=1}^{n}(e_i y_{min}) + \frac{S_{weight}}{p}\sum_{j=1}^{p}(s_j y_{min}) + \frac{L_{weight}}{r}\sum_{k=1}^{r}(l_k y_{min})$$

$equation\ A4$

$$BRI_{max} = \frac{E_{weight}}{n}\sum_{i=1}^{n}(e_i y_{max}) + \frac{S_{weight}}{p}\sum_{j=1}^{p}(s_j y_{max}) + \frac{L_{weight}}{r}\sum_{k=1}^{r}(e_k y_{max})$$

$equation\ A5$

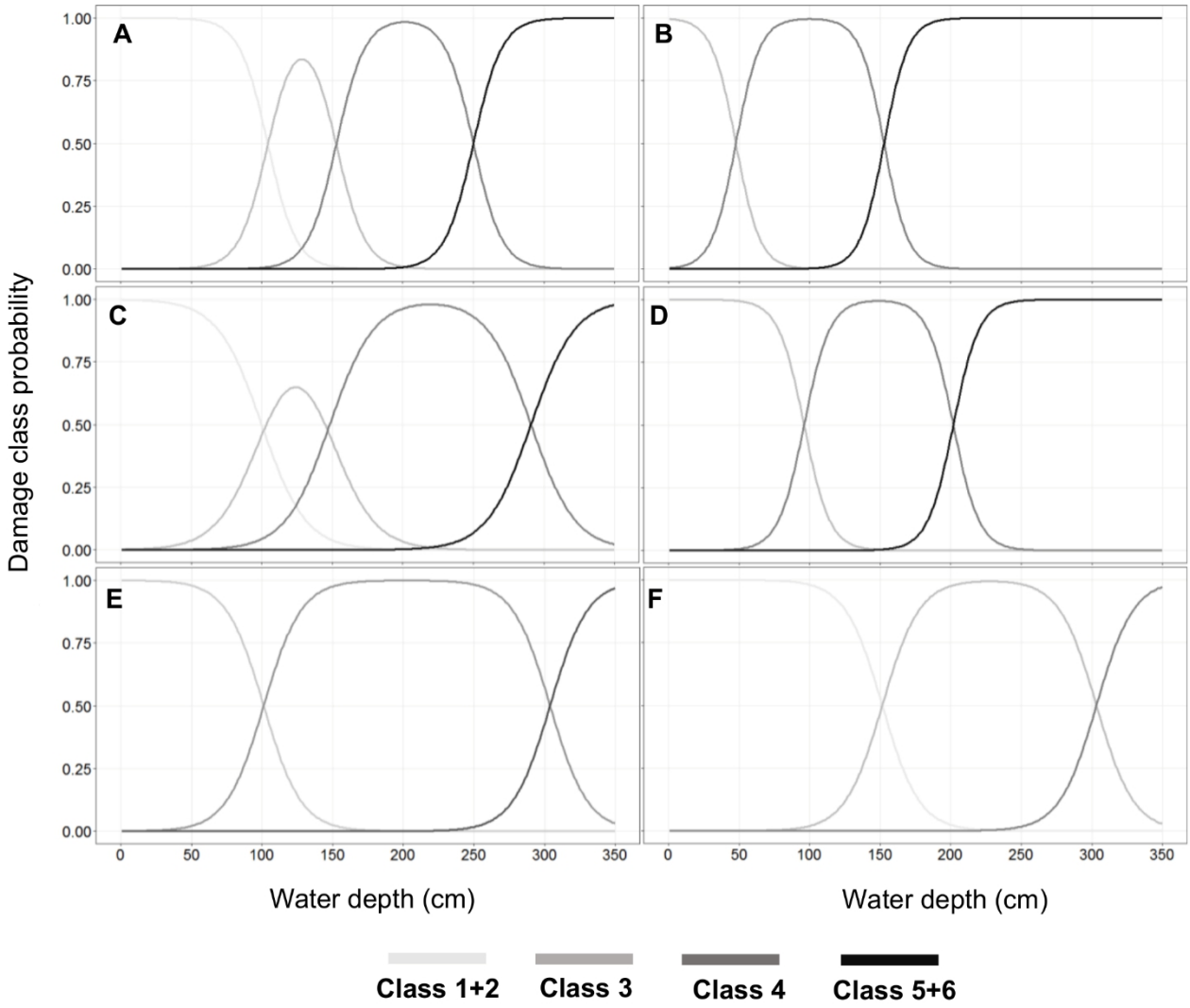


Figure A1: Lower and higher probable damage curves for (A, B) poor, (C, D) moderate, and (E, F) good BRI classes respectively.

Acknowledgment

This study is carried out within the framework of a PhD scholarship funded by the Swiss Government Excellence Scholarships for Foreign Scholars (ESKAS). The authors would like to especially thank all the experts that volunteered to participate in the expert evaluation: Dr. A. A. Komolafe, Dr. M. K. Kawu, Dr. O. S. Rafiu, Dr. A. Richard, Dr. O. Samsideen, Prof. O. D. Jimoh, Prof. I. O. Adelekan, Dr. O. P. Uchenna, Dr. T. Akukwe, Dr. A. B. Ismail. In addition, we would like to thank Bala Inuwa and Ambrose Patrick for guidance and assistance during data collection in study regions 1 and 2 respectively.

Financial support

This research has been supported by the Swiss government excellence scholarship (grant no. 2017.1027).

References

Adeleye, B., Popoola, A., Sanni, L., Zitta, N. and Ayangbile, O.: Poor development control as flood vulnerability factor in Suleja , Nigeria, T. Reg. Plan., (74), 23–35, doi:10.18820/2415-0495/trp74i1.3, 2019.

Akukwe, T. I. and Ogbodo, C.: Spatial analysis of vulnerability to flooding in Port Harcourt metropolis, Nigeria, SAGE Open, 5(1), doi:10.1177/2158244015575558, 2015.

Andrew, G., Yu-Sung, S. and Maria, G. P.: arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.11-1. [online] Available from: https://cran.r-projkect.org/package/arm, 2020.

Ayanlade, A.: Seasonal rainfall variability in Guinea Savanna part of Nigeria: A GIS approach, Int. J. Clim. Chang. Strateg. Manag., 1(3), 282–296, doi:10.1108/17568690910977492, 2009.

Bischl, B., Richter, J., Bossek, J., Horn, D., Thomas, J. and Lang, M.: mlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions, arXiv Mach. Learn., 2017.

Blanco-Vogt, A. and Schanze, J.: Assessment of the physical flood susceptibility of buildings on a large scale - Conceptual and methodological frameworks, Nat. Hazards Earth Syst. Sci., 14(8), 2105–2117, doi:10.5194/nhess-14-2105-2014, 2014.

Blong, R.: A review of damage intensity scales, Nat. Hazards, 29(1), 57–76, doi:10.1023/A:1022960414329, 2003.

Breiman, L.: Random Forests, Mach. Learn., 45(1), 5–32, doi:10.1023/A:1010933404324, 2001.

Brenning, A.: Spatial prediction models for landslide hazards: review, comparison and evaluation, Nat. Hazards Earth Syst. Sci., 5(6), 853–862, doi:10.5194/nhess-5-853-2005, 2005.

CRED, (Centre for Research on the Epidemiology of Disasters): Disasters in Africa: 20 Year Review (2000-2019)., 2019.

CRED, (Centre for Research on the Epidemiology of Disasters): Disaster Year in Review 2019., 2020.

Dall'Osso, F., Gonella, M., Gabbianelli, G., Withycombe, G. and Dominey-Howes, D.: A revised (PTVA) model for assessing the vulnerability of buildings to tsunami damage, Nat. Hazards Earth Syst. Sci., 9(5), 1557–1565, 2009a.

Dall'Osso, F., Gonella, M., Gabbianelli, G., Withycombe, G. and Dominey-Howes, D.: A revised (PTVA) model for assessing the vulnerability of buildings to tsunami damage, Nat. Hazards Earth Syst. Sci., 9(5), 1557–1565, 2009b.

Eludoyin, O. M. and Adelekan, I. O.: The physiologic climate of Nigeria, Int. J. Biometeorol., 57(2), 241–264, doi:10.1007/s00484-012-0549-3, 2013.

Englhardt, J., Moel, H. de, Huyck, C. K., Ruiter, M. C. de, Aerts, J. C. J. H. and Ward, P. J.: Enhancement of large-scale flood risk assessments using building-material-based vulnerability curves for an object-based approach in urban and rural areas, Nat. Hazards Earth Syst. Sci., 19(8), 1703–1722, doi:10.5194/nhess-19-1703-2019, 2019.

Esri, HERE, Garmin, OpenStreetMap and GIS-Community: Light gray canvas base map, [online] Available from: https://services.arcgisonline.com/ArcGIS/rest/services/Canvas/World_Light_Gray_Base/MapServer, 2020.

Ettinger, S., Mounaud, L., Magill, C., Yao-Lafourcade, A. F., Thouret, J. C., Manville, V., Negulescu, C., Zuccaro, G., De Gregorio, D., Nardone, S., Uchuchoque, J. A. L., Arguedas, A., Macedo, L. and Manrique Llerena, N.: Building vulnerability to hydro-geomorphic hazards: Estimating damage probability from qualitative vulnerability assessment using logistic regression, J. Hydrol., 541, 563–581, doi:10.1016/j.jhydrol.2015.04.017, 2016.

FGN, (Federal Government of Nigeria): Nigeria: Post-Disaster needs assessment - 2012 Floods. [online] Available from: https://www.gfdrr.org/sites/gfdrr/files/NIGERIA_PDNA_PRINT_05_29_2013_WEB.pdf (Accessed 1 January 2019), 2013.

Fontalba-Navas, A., Lucas-Borja, M. E., Gil-Aguilar, V., Arrebola, J. P., Pena-Andreu, J. M. and Perez, J.: Incidence and risk factors for post-traumatic stress disorder in a population affected by a severe flood, Public Health, 144, 96–102, doi:https://doi.org/10.1016/j.puhe.2016.12.015, 2017.

Fuchs, S.: Susceptibility versus resilience to mountain hazards in Austria-paradigms of vulnerability revisited., Nat. Hazards Earth Syst. Sci., 9(2), 337–352, doi:10.5194/nhess-9-337-2009, 2009.

Fuchs, S., Keiler, M. and Zischg, A.: A spatiotemporal multi-hazard exposure assessment based on property data, Nat. Hazards Earth Syst. Sci., 15(9), 2127–2142, doi:10.5194/nhess-15-2127-2015, 2015.

Fuchs, S., Keiler, M., Ortlepp, R., Schinke, R. and Papathoma-Köhle, M.: Recent advances in vulnerability assessment for the built environment exposed to torrential hazards: challenges and the way forward, J. Hydrol., 575, 587–595, doi:10.1016/j.jhydrol.2019.05.067, 2019a.

Fuchs, S., Heiser, M., Schlögl, M., Zischg, A., Papathoma-Köhle, M. and Keiler, M.: Short communication: A model to predict flood loss in mountain areas, Environ. Model. Softw., 117, 176–180, doi:https://doi.org/10.1016/j.envsoft.2019.03.026, 2019b.

Gasparini, P.: Analysis and monitoring of environmental risk: CLUVA Final Report, Clim. Chang. Urban Vulnerability Africa, 1–26 [online] Available from: http://cordis.europa.eu/docs/results/265137/final1-cluva-final-publishable-summary-report.pdf (Accessed 1 January 2019), 2013.

Gerl, T., Kreibich, H., Franco, G., Marechal, D. and Schröter, K.: A review of flood loss models as basis for harmonization and benchmarking, PLoS One, 11(7), 1–22, doi:10.1371/journal.pone.0159791, 2016.

Google Earth Pro: Google Earth, [online] Available from: earth.google.com/web/, n.d.

Hammond, M. J. and Chen, A. S.: Urban flood impact assessment : A state-of-the-art review, Urban Water J., 12(1), 14–29, doi:10.1080/1573062X.2013.857421, 2015.

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K. L., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R. and Zhou, G.: Special Report on Global Warming of 1.5 °C - Chapter 3: Impacts of 1.5º C global warming on natural and human systems, Glob. Warm. 1.5°C. An IPCC Spec. Rep. impacts Glob. Warm. 1.5°C above pre-industrial levels Relat. Glob. Greenh. gas Emiss. pathways, Context Strength. Glob. response to Threat Clim. Chang., 175–311, doi:10.1002/ejoc.201200111, 2018.

Holub, M. and Fuchs, S.: Benefits of local structural protection to mitigate torrent-related hazards, WIT Trans. Inf. Commun. Technol., 39, 401–411, 2008.

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., Gericke, A., Neal, J., Aerts, J. C. J. H. and Ward, P. J.: Comparative flood damage model assessment: Towards a European approach, Nat. Hazards Earth Syst. Sci., 12(12), 3733–3752, doi:10.5194/nhess-12-3733-2012, 2012.

JRC, (Joint Research Centre-European Commission and OECD): Handbook on Constructing Composite Indicators: Methodology and User Guide, Joint Research Centre-European Commission and OECD., 2008.

Komolafe, A. A., Adegboyega, S. A. A. and Akinluyi, F. O.: A review of flood risk analysis in Nigeria, Am. J. Environ. Sci., 11(3), 157–166, doi:10.3844/ajessp.2015.157.166, 2015.

Kreibich, H., Seifert, I., Merz, B. and Thieken, A. H.: Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector, Hydrol. Sci. J., 55(8), 1302–1314, doi:10.1080/02626667.2010.529815, 2010.

Maiwald, H. and Schwarz, J.: Damage and loss prediction model considering inundation level, flow velocity and vulnerability of building types, WIT Trans. Ecol. Environ., 159, 53–65, doi:10.2495/FRIAR120051, 2012.

Maiwald, H. and Schwarz, J.: Damage and loss prognosis tools correlating flood action and building's resistance-type parameters, Int. J. Saf. Secur. Eng., 5(3), 222–250, doi:10.2495/SAFE-V5-N3-222-250, 2015.

Maiwald, H. and Schwarz, J.: Unified damage description and risk assessment of buildings under extreme natural hazards, Eur. J. Mason. – Mauerw., 23(2), 95–111, doi:10.1002/dama.201910014, 2019.

Malgwi, M. B., Fuchs, S. and Keiler, M.: A generic physical vulnerability model for floods: review and concept for data-scarce regions, Nat. Hazards Earth Syst. Sci., 20(7), 2067–2090, doi:10.5194/nhess-20-2067-2020, 2020.

Malgwi, M. B., Ramirez, J. A., Zischg, A., Zimmermann, M., Schürmann, S. and Keiler, M.: Flood reconstruction using field interview data and hydrodynamic modelling: A method for data-scarce regions, Int. J. Disaster Risk Sci., Submitted, n.d.

Manager, T.: Maintenance Management Framework Guidline: Building Condition Assessment, Brisbane, Queensland. [online] Available from: https://www.hpw.qld.gov.au/__data/assets/pdf_file/0019/3277/mmfbca.pdf, 2017.

Mason, V., Andrews, H. and Upton, D.: The psychological impact of exposure to floods, Psychol. Health Med., 15(1), 61–73, doi:10.1080/13548500903483478, 2010.

Mayomi, I., Kolawole, M. S. and Martins, A. K.: Terrain Analysis for Flood Disaster Vulnerability Assessment : A Case Study of Niger State, Nigeria, Am. J. Geogr. Inf. Syst., 3(3), 122–134, doi:10.5923/j.ajgis.20140303.02, 2014.

Merz, B., Kreibich, H., Thieken, A. and Schmidtke, R.: Estimation uncertainty of direct monetary flood damage to buildings, Nat. Hazards Earth Syst. Sci., 4(1), 153–163, doi:10.5194/nhess-4-153-2004, 2004.

Merz, B., Kreibich, H., Schwarze, R. and Thieken, A.: Review article "assessment of economic flood damage," Nat. Hazards Earth Syst. Sci., 10(8), 1697–1724, doi:10.5194/nhess-10-1697-2010, 2010.

Merz, B., Kreibich, H. and Lall, U.: Multi-variate flood damage assessment: A tree-based data-mining approach, Nat. Hazards Earth Syst. Sci., 13(1), 53–64, doi:10.5194/nhess-13-53-2013, 2013.

Musa, Nda, H. M., Usman, M. Y., Abdul, H. and Sanni, L. M.: An assessment of flood vulnerability on physical development along drainage channels in Minna, Niger State, Nigeria, African J. Environ. Sci. Technol., 9(1), 38–46, doi:10.5897/AJEST2014.1815, 2015.

Naumann, T., Nikolowski, J. and Sebastian, G.: Synthetic depth-damage functions – A detailed tool for analysing flood resilience of building types, Road Map Towar. a Flood Resilient Urban Environ. Proc. Final Conf. COST Action C, (November), doi:10.1093/CERCOR/8.4.321, 2009.

NBC, (Nigerian Building Code): National Building Code, Federal Republic of Nigeria, 2006.

NBS, (National Bureau of Statistics): Annual Abstract of Statistics , Federal Republic of Nigeria., 2012.

Ndanusa, A. B., Dahalin, Z. and Ta, A.: Topographic-based framework for flood vulnerability classification: a case of Niger state, Nigeria, J. Inf. Syst. Technol. Manag., 3(9), 27–38, 2018.

Neubert, M., Naumann, T. and Deilmann, C.: Synthetic water level building damage relationships for GIS-supported flood vulnerability modeling of residential properties, in Flood Risk Management: Research and Practice. Proceedings of the European Conference on Flood Risk Management Research into Practice, FLOODrisk, p. 294., 2008.

Niang, I., Ruppel, O. C., Abdrabo, M. A., Essel, A., Lennard, C., Padgham, J. and Urquhart, P.: Africa, in Climate Change 2014: Impacts, Adaptation and Vulnerability: Part B: Regional Aspects: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by V. R. Barros, C. B. Field, D. J. Dokken, M. D. Mastrandrea, and K. J. Mach, pp. 1199–1266, Cambridge University Press, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2015.

Nwilo, P. C., Olayinka, D. N. and Adzandeh, A. E.: Flood Modelling and Vulnerability Assessment of Settlements in the Adamawa State Floodplain Using GIS and Cellular Framework Approach, Glob. J. Hum. Soc. Sci., 12(3), 11–20, 2012.

OCHA, (United Nations Office for the Coordination of Humanitarian Affairs): NIGERIA : Floods in Borno , Adamawa and Yobe, Situation Report No. 2. [online] Available from: https://reliefweb.int/sites/reliefweb.int/files/resources/ocha_nga_floodoverview_13092019.pdf, 2019.

Okoye, P. U., Ezeokoli, F. O. and Ugochukwu, J.: Building Development Practice in Flood Prone Area: Case of Ogbaru Council Area of Anambra State Nigeria, Int. J. Eng. Res. Appl., 5(8), 30–40, 2015.

Papathoma-köhle, M., Schlögl, M. and Fuchs, S.: Vulnerability indicators for natural hazards : an innovative selection and weighting approach, Sci. Rep., 9, 1–14, doi:10.1038/s41598-019-50257-2, 2019.

Papathoma-Köhle, M., Kappes, M., Keiler, M. and Glade, T.: Physical vulnerability assessment for alpine hazards: State of the art and future needs, Nat. Hazards, 58(2), 645–680, doi:10.1007/s11069-010-9632-4, 2011.

Papathoma-Köhle, M., Gems, B., Sturm, M. and Fuchs, S.: Matrices, curves and indicators: A review of approaches to assess physical vulnerability to debris flows, Earth-Science Rev., 171, 272–288, doi:10.1016/j.earscirev.2017.06.007, 2017.

Penning-Rowsell, E., Tunstall, S., Tapsell, S., Morris, J., Chatterton, J. and Colin, G.: The benefits of flood and coastal risk management: a manual of assessment techniques, ISBN 19047., 2005.

Pistrika, A., Tsakiris, G. and Nalbantis, I.: Flood Depth-Damage Functions for Built Environment, Environ. Process., 1(4), 553–572, doi:10.1007/s40710-014-0038-2, 2014.

Probst, P., Wright, M. N. and Boulesteix, A.-L.: Hyperparameters and tuning strategies for random forest, WIREs Data Min. Knowl. Discov. , 9(3), e1301, doi:10.1002/widm.1301, 2019.

Röthlisberger, V., Zischg, A. P. and Keiler, M.: Identifying spatial clusters of flood exposure to support decision making in risk management, Sci. Total Environ., 598, 593–603, doi:10.1016/j.scitotenv.2017.03.216, 2017.

Saaty, T. L.: The Analytical Hierarchy Process, Planning, Priority, Resour. Alloc. RWS Publ. USA, 1980.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J. and Brenning, A.: Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data, Ecol. Modell., 406, 109–120, doi:https://doi.org/10.1016/j.ecolmodel.2019.06.002, 2019.

Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F. and Merz, B.: How useful are complex flood damage models?, Water Resour. Res., 50(4), 3378–3395, doi:10.1002/2013WR014396, 2014.

Schwarz, J. and Maiwald, H.: Berücksichtigung struktureller Schäden unter Hochwassereinwirkung, Bautechnik, 84(7), 450–464, 2007a.

Schwarz, J. and Maiwald, H.: Prognose der Bauwerksschdigung unter Hochwassereinwirkung, Bautechnik, 84(7), 450–464, doi:10.1002/bate.200710039, 2007b.

Schwarz, J. and Maiwald, H.: Damage and loss prediction model based on the vulnerability of building types, 4th Int. Symp. Flood Def. Manag. Flood Risk, Reliab. Vulnerability, (May 2008), 9, doi:10.13140/2.1.1358.3043, 2008.

Straub, A.: Dutch standard for condition assessment of buildings, Struct. Surv., 27(1), 23–35, doi:10.1108/02630800910941665, 2009.

Thieken, A. H., Müller, M., Kreibich, H. and Merz, B.: Flood damage and influencing factors: New insights from the August 2002 flood in Germany, Water Resour. Res., 41(12), 1–16, doi:10.1029/2005WR004177, 2005.

UNISDR, (United Nations International Strategy for Disaster Reduction): Terminology on disaster risk reduction, United Nations International Strategy for Disaster Reduction Geneva, Geneva, Switzerland. [online] Available from: https://www.undrr.org/publication/2009-unisdr-terminology-disaster-risk-reduction (Accessed 15 March 2020), 2009.

UNISDR, (United Nations International Strategy for Disaster Reduction): Sendai framework for disaster risk reduction 2015–2030, in Proceedings of the 3rd United Nations World Conference on DRR, Sendai, Japan, p. 32. [online] Available from: https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030 (Accessed 26 February 2020), 2015.

Verger, P., Rotily, M., Hunault, C., Brenot, J., Baruffol, E. and Bard, D.: Assessment of exposure to a flood disaster in a mental-health study, J. Expo. Sci. Environ. Epidemiol., 13(6), 436–442, doi:10.1038/sj.jea.7500290, 2003.

Vogel, K., Riggelsen, C., Merz, B., Kreibich, H. and Scherbaum, F.: Flood Damage and Influencing Factors: A Bayesian Network Perspective, Proc. 6th Eur. Work. Probabilistic Graph. Model. (PGM 2012), 625, 314–354, 2012.

Wagenaar, D., De Jong, J. and Bouwer, L. M.: Multi-variable flood damage modelling with limited data using supervised learning approaches, Nat. Hazards Earth Syst. Sci., 17(9), 1683–1696, doi:10.5194/nhess-17-1683-2017, 2017.

Zumo, M.: Delineation of built-up areas liable to flood in Yola, Adamawa State, Nigeria using remote sensing and geographic information system technologies, FUTY J. Environ., 8(1), 20–30, 2014.

Chapter 5: Synthesis

5.1 Main findings

In its entirety, this thesis aimed to develop and test a new physical vulnerability model (PVM) to provide insights on floods that occur in data-scarce locations, so that the associated impacts can be better characterized and predicted. To achieve this objective, three research questions were developed and previously outlined in section 1.4 in Chapter 1. This section provides a comprehensive summary of how each of the formulated research questions have been addressed.

Research question 1: *How can existing PVMs be utilized to develop a new model based on i) reduced requirements for empirical data and ii) that considers multiple damage influencing variables?*

In Chapter 2, a new PVM was conceptualized for data-scarce regions. First, an extensive review of state-of-the-art pertinent to how physical vulnerability assessments have been conducted provided the necessary background and motivation that drove the development of the new concept. The review evaluated existing PVMs, focusing on reduced data requirements and the integration of multiple damage influencing variables. Based on the aforementioned criteria, PVMs that were investigated covered a range of vulnerability indicators, damage grades, stage-damage curves and multivariate. The new PVM systematically combined specific aspects, or complete frameworks, past studies and tailored these for data-scarce areas.

In particular, the new PVM was developed on the vulnerability indicator, so that a multi-variable approach is supported. Several damage influencing variables (or indicators) collected through expert interviews are aggregated to characterize building resistance (Building Resistance Index (BRI)), relative to another building considering similar hazard magnitudes. The BRI combined three categories of variables: building characteristics, local protection measures and building exposure. These variables provide an overarching assessment of vulnerability, which influences variables with respect to building resistance. A notable advantage of using regional expert knowledge is the reduced reliance of the new PVM on empirical data. More specifically, the integration of this knowledge supported both the identification of damage drivers through the vulnerability indicator approach and the performance of what-if analysis to relate building damage to damage influencing variables.

The new PVM relies on the working assumption that buildings classified within the same vulnerability category will incur similar damage when exposed to the same hazard level (Schwarz and Maiwald, 2008; Maiwald and Schwarz, 2015). For example, while Schwarz and Maiwald (2008) developed a building-type (e.g., clay, masonry, reinforced concrete) vulnerability classification for buildings in Germany, they also showed that this classification can be extended to buildings with different structural characteristics, from which comparable damage can be expected for a similar range of flood characteristics. In the new PVM, the classification of the BRI into three discrete vulnerability classes (low, moderate and high), before performing the what-if analysis, is a particularly innovative feature. By using these BRI classes, the what-if analysis is better targeted for buildings with similar vulnerability characteristics. For example, in Nigeria, given the

significant difference observed in building standards, a what-if analysis generalized for sandcrete block buildings will be highly uncertain with a large scatter in terms of building damage for the same flood depth. However, the BRI classification improves this uncertainty by categorizing buildings with similar characteristics based on exposure, susceptibility and local protection.

The use of damage grades to communicate hazard consequences provides a good basis for developing PVM in data-scarce areas. This is because damage grades can be compiled and developed from a literature review and the analysis of damage reports or pictures uploaded to social media platforms after the occurrence of a hazard. Apart from providing a usable alternative, especially where monetary loss data is unavailable, damage grades are not influenced by cost variations (or fluctuations in market prices). Hence they can be easily applied to compare damage incurred across hazard events and between different regions. Damage grades are also transferable, especially within areas where comparable building types are found. Damage grades are simplistic and easy to use. Consequently, it can be an effective tool to communicate disaster consequences to stakeholders and to provide support for subsequent decision making. Additionally, damage grades can be used to raise awareness in communities that are exposed to hazards to encourage the adoption of suitable mitigation or local protection measures.

While the new PVM can be fully implemented using expert knowledge, a likely limitation is that such evaluations can be subjective, where the subjectivity stems from differences in knowledge or domain of specific expertise. Methods such as the Analytical Hierarchal Process (AHP) (Saaty, 1980) have been used to assess and report the level of subjectivity using the consistency ratio (e.g. Dall'Osso et al., 2009). However, the AHP requires extensive expert training and efforts to develop (JRC, 2008). In any case, it is considered to be best practice to include several experts to ensure that results are representative and transferrable to more general conditions (i.e have not been overfitted to limited perspectives or interpretations) (JRC, 2008). In the new PVM, three damage states were proposed for the what-if analysis: low probable, most probable and high probable damage. These concept for the damage states, adopted from the vulnerability classification by Schwarz and Maiwald (2008), is used to quantify the range (scatter) within which the expected damage would lie for each BRI class. The damage states show the lowest (low probable), likely (most probable), and highest (high probable) classes of damage grades for each of the synthetic water depths.

A combination of existing physical vulnerability assessment methods have been used to develop a concept for a new PVM for data-scarce regions; in particular, it uses a multi-variable approach and does not require empirical data for implementation. The integration of expert knowledge allows the model to be adapted to reflect specific regional characteristics. The entire workflow for developing the new approach is customizable and modular: the separate components (vulnerability indicators, damage grades, what-if analysis) can be updated individually or in combination, as empirical data becomes available. Alternatively, updates can be performed as new insights from other studies are available.

Research question 2: *How can a hydrodynamic model be utilized to reconstruct plausible scenarios of past flood events in areas without hydrological data, such that modelled results have an acceptable accuracy and can be extrapolated to further develop PVMs?*

In Chapter 3, a method for reconstructing plausible scenario of past-floods without hydrological data was developed and tested using the 2017 flood event in Suleja/Tafa. The selected flood event and case study were particularly challenging to analyze, since it was characterized by interactions from five river reaches. However, the complexity associated with this event was an opportunity to further demonstrate the applicability of the proposed method on multiple river networks. The key feature for the scenario reconstruction is the use of post-event, spatially distributed flood depths and flood duration data, which were collected through interviews. The workflow in the approach utilizes four rounds of simulations that aimed to minimize the root mean square error (RMSE) between collected observed data and simulated data at identified building locations. The iterative steps for minimizing the RMSE were necessary, given that PVMs depend on reliable estimates of flood characteristics at building locations. The global RMSE (gRMSE) after simulating the entire river network, was 0.61 m, a value that falls within the water level accuracy of similar studies using hydrodynamic models. Schumann et al. (2015) highlighted that simulated water level accuracies normally lie between a few centimetres to 1 – 2 m; the study results were within this range. Within this level of accuracy, modelled flood depths could be extrapolated beyond the spatial extent of the collected data as a means to increase the number of observations used to develop PVMs in data-scarce regions.

The use of the hydrodynamic model, CAESAR-Lisflood, in combination with data from interviews offered several interesting applications for data-scarce locations. CAESAR-Lisflood has a reduced data requirement: in particular, only the choice of a representative Manning's coefficient was needed for model calibration. The model can operate on reach mode, whereby discharge is introduced at defined upstream locations. In addition, modelling at reach scale supported modelling to be carried out for a specific location of interest and not necessarily for the whole catchment.

Basic data required for the modelling, flood depths and durations, can be collected from households through interviews. Generally, these interviews should be carried out at the earliest possible chance after the occurrence of floods, so that participants' memories are fresh and they are able to recall details with higher degrees of certainty. Interviews also allow for researchers to interact with residents in affected communities and follow-up questions can be posed to validate reported details. Assumpção et al. (2018) noted that this one-on-one validation is especially important for participatory approaches, given that community residents are usually non-experts. Also, with increasing access to the Internet and technologically equipped devices (e.g. smartphones), participatory approaches involving the use of social media platforms for data collection can be further optimized.

Within a broader context, several interesting deductions, especially for hydrodynamic applications, can be made from the results of this study. Firstly, the study has extended the application of data collected based on a

combination of participatory approaches. In particular, while previous studies that used field interview data were restricted to GIS-based mapping of floods, the new method extends the application to hydrodynamic modelling to provide a more physically based approach. In this way, flood hazard assessment within these regions can include i) floodplain and channel roughness and ii) flow interactions between multiple rivers. Secondly, the study provided a simple pre-processing workflow to support the recreation of river channels that are not well represented in low-resolution digital elevation models (DEMs). The re-created channel, based on contour maps derived from the DEM and satellite images, improved the spatial alignment of the channel, which in turn, contributed to improving overall model accuracy (i.e lower RMSE). Given that in many data-scarce locations, available DEMs used for flood hazard assessment are limited to globally available Shuttle Radar Topography Mission (SRTM) with 30 m or 90 m spatial resolutions, the procedure implemented in this study demonstrated how to improve river channel alignment without the need for additional resources that may otherwise be expensive to acquire. Furthermore, while the current method for estimating peak discharge has only been demonstrated for single river channels (e.g. study by Borga et al., 2019; Bronstert et al., 2018), this study provided a first step to extend the application for multiple river channels. This was achieved by modelling upstream and downstream reaches separately using synthetic hydrographs (as proposed by Zischg et al., 2018) and examining the combined effect of upstream peak discharge on the downstream catchment.

Although the developed method aimed to facilitate the extrapolation of flood characteristics at building locations, only modelled flood depths are currently possible. Although flood velocities can be additionally extrapolated from the CAESER-Lisflood outputs, additional data may be required to calibrate the model to improve the associated accuracy. Collection of flood velocity data in the field is rather difficult since, in many post-event situations, only qualitative descriptions (e.g. slow, moderately fast, very fast) can be provided, without the availability of an extensive network of in situ, ground-based measurement instruments. Such qualitative data are challenging to compare with the quantitative outputs from the hydrodynamic model. On the other hand, where videos of the events are available, the presence of transported objects relative to fixed references per frame can be used to estimate flow velocities. However, this method may require high levels of spatial and temporal referencing to precisely determine at which point of the hydrograph (rising or falling limb) the velocity is defined.

Research question 3: *How does the new PVM perform against a traditional (existing) PVM using damage prediction accuracy and identified damage drivers as a set of evaluation metrics?*

In Chapter 4, the application of the new PVM (here referred to as the expert-based approach) was evaluated against a multivariate (data-driven) approach developed using random forests. The selection of a multivariate method for comparison is based on the recognition that it provides the most reliable results on damage drivers and in terms of damage prediction accuracy to date. The comparison, which was realized with the use of data sets from two study regions in Nigeria, aimed to evaluate the applicability of the new approach in data-scarce locations. Prior to implementing both methods, a set of damage grades was developed. The six-class damage

grades (with classes 1-6) was the first for sandcrete and clay building types. It describes damage patterns that are repeatedly observed in the region. Due to data limitations, however, only four damage grades were tested in this study. Damage grades with similar degree of damage were combined into single classes (i.e class 1+2, class 3, class 4, class 5+6) (supplementary material, Figure S2-3).

For the expert-based method, which was implemented following the steps described in section 5 in Chapter 2, the development of the building resistance index (BRI) provided the basis for a vulnerability classification (poor, moderate and good). For each of the three BRI classes, an ordinal logistic regression model was applied to relate predicted damage grades by experts after the what-if analysis to water depth (Figure 5, Chapter 4). The logistic regression provided a suitable separation for expected damages such that for a given BRI class and water depth, a most probable damage class can be assigned. The data-driven approach addressed the same problem by generating ensembles of random forest models.

The results show that while the expert-based approach obtained relatively higher predictive accuracy of 48% and 38% for lower damage grades (classes 1+2 and 3 respectively), the performance dropped for higher damage grades (classes 4 and 5+6) to 17% and 18% respectively (Chapter 4, Figure 6). Conversely, the data-driven method attained higher predictive accuracy for higher damage grades (53% for Class 4 and 51% for class 5+6), which dropped at lower damage grades (38% for class 1+2 and 13% for class 3) (Chapter 4, Figure 7). The highest misclassification for the data-driven method was misclassifying class 3 as class 4. In the expert-based method, half of class 4 buildings were incorrectly predicted as class 1+2. The comparably higher predictive accuracy of the expert-based approach for low damage grades is likely because a high percentage of the data have water depths less than 150 cm and within this range, the MPD curve (Chapter 4, Figure 5) mostly predicts low damage grades. On the other hand, the low predictive accuracy of the expert method for high damage grades is attributed to a general underestimation of damage at low water depths. The overall multi-class predictive accuracy for the expert-based and data-driven methods was 30 % and 38 % respectively (Chapter 4, Figure 6 and 7).

In addition to the MPD curve, the expert-based method generated high probable damage (HPD) and low probable damage (LPD) curves for all three BRI classes (Chapter 4, Figure A1). Unlike the MDP curves, the LPD and HPD characterize the range for lowest and highest probable damage grade at given flood depths for each BRI class.

In general, the predictive accuracy of both models is likely influenced by a variation in compliance with building standards at both study locations. As reported by FGN (2013), over 60 % of households build their houses without using the services of designated institutions. The implication of this practice is the fact that strict adherence to building codes may not be followed, leading to significant variation in building quality. Owing to this practice, a large amount of scatter is expected in building behaviour (or damage) and likely influences predictive accuracy. In both models, the variable ‘building quality’ was used to characterize building standards using a combination of three factors: functionality, durability and aesthetics (Chapter 4,

Appendix), as defined in the Nigerian building code (NBC, 2006). However, it is unlikely the variable (building quality) adequately captured the difference in building standards and may require further evaluation.

Furthermore, damage influencing variables indicated as high importance were partly similar across both models. A complete list of damage influencing variables used to develop both models is shown in Chapter 4, Table 1. For the expert-based model, distance to channel, building condition, functional drainage, building quality, ground floor elevation and wall material were the most important variables identified by experts, in the order of importance (Chapter 4, Figure 4). Other lower scored variables include natural barriers, fencing, sheltering and wall plaster. In the data-driven method, wall material and distance to channel were identified as the most important variables (Chapter 4, Figure 9). Other successively important variables include wall thickness, building condition, ground floor elevation, building quality, height of opening and sheltering. In comparison, variables identified by both methods as important for PVM include: distance to channel, wall material, building condition and building quality. A disparity in variable importance between both methods was identified with the functional drainage variable. While functional drainage was the third most important variable identified by experts, it was the least important in the data-driven method. However, low variability in the data for functional drainage may have influenced its performance in the random forest model. As a result, the variable is subject to further evaluation.

The application of the expert-based and data-driven models have highlighted several other interesting suggestions that are important to consider to support physical vulnerability assessments. Firstly, experts tend to underestimate damage at low water depths. Experts may linearly relate water depth and damage such that high water depth is associated with high damage and low water depths is associated with low damage. From the bivariate analysis, however (Chapter 4, Figure 2), high damage grades were observed even at low water depths. As a result, we can deduce one of the following possibilities. Firstly, other variables that were not considered strongly contributed to the observed damage. Secondly, given that the proportion of damage at low water depths was predominantly observed in study region 1 than in region 2, damage drivers, even for comparable building types, may differ between the two regions. Furthermore, both methods have shown that flood damage modelling can be carried out even with a reduced (or optimal) number of variables. This was evident from the sensitivity analysis for the expert-method and the results of the data-driven approach; both demonstrated that similar prediction accuracy can be achieved after including main damage variables. This finding was similarly noted by Papathoma-Köhle et al. (2019) and is particularly important, given that it holds the potential for reducing the high variable data requirements of multivariate PVMs. The lower number of variables requirements translate to less time, cost or physical efforts in data collection. Furthermore, since decision makers find it difficult to handle too many indicators or variables (Barroca et al., 2008), an approach that returns usable insights with a lower number of variables offers a viable alternative. Additionally, the new approach has extended the application of vulnerability indicators from their previous use in relative vulnerability assessments to their inclusion to return damage predictions. This was achieved by applying the

developed index for a vulnerability classification (BRI classes) and customizing the what-if analysis for specific vulnerability classes.

The study presents one of the first detailed comparisons of physical vulnerability assessments between expert-based and data-driven methods, using two independent data sets. It provides insights from data-driven methods so that the underlying theory for expert methods can be evaluated, and where possible, improved to support overall disaster risk reduction in data-scarce areas. Using more than one dataset provided a better opportunity for evaluating model transferability more critically than with a single case study alone. The data-driven model achieved the same mean performance accuracy (38%) when it was trained on data from study region 1 and tested on data from study region 2. This suggests that damage influencing variables and building damage mechanisms for both study regions are similar and that the model is transferable between the two cases. On the other hand, the expert-based model showed a 34% and 23% mean performance accuracy for data from study regions 1 and 2, respectively. This difference was partly influenced by high damage grades (class 4, class 5+6) at low water depths. This is especially valid in study region 2.

5.2 Conclusion

As the frequency and severity of floods continue to rise globally, decision makers are faced with the challenge to take effective actions to reduce consequent risk. In many data-scarce regions, flood risk driven by climate change and an increase in urban encroachment on floodplains are already increasing. Physical vulnerability assessments provide a knowledge base for identifying drivers of flood damages and predicting where significant damages are to be expected.

This thesis aimed to develop and test a new PVM to improve the characterization and prediction of flood impacts in data-scarce regions. This overall aim was achieved with three research objectives: i) existing methods were thoroughly reviewed to conceptualize a PVM approach that can be tailored for regional situations in typical data-scarce areas (Chapter 2); ii) a hydrodynamic modelling approach that uses field interview data was developed and evaluated to reconstruct a past flood scenario in a data-scarce location to extrapolate modelled flood characteristics for use in a PVM (Chapter 3); and iii) the applicability and performance of the new PVM approach was tested by comparing the results with existing methods (Chapter 4). Two study regions in Nigeria were selected for as representative data-scarce locations; field data collection campaigns were carried out to evaluate the developed methods.

From this set of objectives, a combination of physical vulnerability assessment methods were considered to develop a new PVM that is tailored for data-scarce regions (Chapter 1). The new PVM has reduced data requirements and can be fully implemented by experts. It systematically combines vulnerability indicators, damage grades and synthetic what-if analysis for modelling flood damage. The method is characterized by a flexible workflow that can be adapted to specific regional situations. Vulnerability indicators, damage grades and results of the what-if analysis can be updated as more information on risk drivers or damage patterns becomes available. The new method is transferable and can be applied to areas without empirical damage data by including knowledge provided by regional experts.

The new PVM requires flood depths at respective building locations to make damage predictions. As a result, an approach was developed for extrapolating higher quantities of point-based flood depths, especially where hydrological data is unavailable (Chapter 2). The method was tested using i) an open-source hydrodynamic model (CAESAR-Lisflood), which requires minimal parameterization (only representative Manning's coefficient) and ii) 300 spatially distributed observation data points with flood depths and flood durations collected through interviews. Furthermore, for further application in typical data-scarce areas, the modelling method uses a globally available DEM and a synthetic hydrograph to set up initial conditions. An overall global RMSE of 0.61 m between the observed and modelled flood depths, demonstrates the applicability of the method for the extrapolation of modelled flood characteristics for further use for the development of PVMs in data-scarce areas. In addition, scenarios with higher flood depths can be investigated with this method, so that high risk buildings (i.e. buildings with high damage grades) can be identified for mitigation and emergency planning. This possibility had not yet been thoroughly explored and tested; the preliminary

findings from this thesis encourage the dedication of further efforts in this direction, including more comprehensive evaluations extended to comparable data-scarce and flood prone locations.

The recommendation is predicated on the successful application of the new PVM for assessing the physical vulnerability of buildings in Nigeria. With an output showing i) moderate predictive accuracy and ii) similarity in important damage drivers, in comparison with a multivariate method, the applicability of the new approach as a first step for physical vulnerability assessment is demonstrated. These outputs are important for disaster management; for instance, policymakers can recommend or pass regulations with respect to the minimum distance buildings should be located from river channels or the type of wall materials that should be allowed within a certain distance to floodwater sources. In addition, the building resistance index (BRI) can quantify buildings that are at higher risk for a given flood scenario. The multi-class predictive accuracy of the new expert-based method is moderate. Given the limited information on sandcrete and clay buildings and the high degree of variability in terms of pre-flood event building standards, the performance of the new PVM is considered to be satisfactory and its application for use in data-scarce locations is well demonstrated. The developed damage grades used in the study are the first for sandcrete and clay building types, which are simple representations of commonly observed damage patterns within a region. Damage grades are a compromise between the need for detailed information and simplicity of use (Blong, 2003). Therefore, they represent a first step towards an alternative to represent flood hazard consequences where detailed monetary loss information cannot be acquired. As data and information on damage drivers and damage patterns become more available, vulnerability indicators, damage grades and results of the what-if analysis can be updated to improve the performance of the new PVM.

Overall, the thesis achieved its central objective, which is to develop and test a new PVM. Through this project, important methods were developed, which have contributed to improving the characterization and prediction of flood impacts. The applicability of the new model is twofold: it can be applied to evaluate the pre- and post-disaster conditions of buildings. These methods bridge the gap between the theory of physical vulnerability and practice. Identified damage drivers (vulnerability indicators) provide decision makers with viable information that can be used to design suitable strategies to address disaster risk. Also, future damage potential in exposed communities can be investigated to support emergency and mitigation planning and economic loss assessments, based on scenario reconstruction and application of the new PVM.

With the expected increase in extreme events and uncertain coping capacity of communities to disasters, the findings of this thesis provide an important first step for disaster management in data-scarce regions. Decision making in Nigeria, as well as in many African countries where sandcrete block and clay buildings are common, can use the findings from this study and adapt the proposed methods to gain a better understanding of physical vulnerability to floods as well as strategize to reduce impacts of hazard occurrence such as the 2012 Nigerian floods. The workflow for the new PVM and hydrodynamic modelling is transferable and can be adapted to meet this urgent need in other flood affected and data-scarce locations.

5.3 Outlook

This thesis describes the development and evaluation of a new physical vulnerability model (PVM) to characterize and to predict flood impacts in data-scarce regions. While the findings are encouraging, especially given the data availability constraint and the high motivation to obtain usable information to characterize flood risk, further research is recommended.

Firstly, the upscaling of the new PVM from a building-scale to a larger (meso-)scale is recommended to facilitate regional physical vulnerability assessments. The availability of large-scale building data sets for several African countries such as the ImageCAT (ImageCat et al., 2017) provides a potential starting point for regional assessments. The ImageCAT database combines data from remote sensing, population census and literature reviews to provide information on the number of buildings, total floor area, and construction types for rural and urban areas. An example of such an upscaling assessment for PVMs was carried out by Kreibich et al. (2016). Such regional assessment is important for disaster management decision-making at the city level by stakeholders.

Secondly, the entire workflow that was developed for reconstructing past flood scenarios should be validated in a typical data-rich region to further evaluate the plausibility of model outputs. In addition, while this study has highlighted several sources of uncertainty inherent in flood scenario reconstruction, these uncertainties have not yet been quantified. For example, in the hydrodynamic modelling component, an important source of input data that was collected is flood depths at building locations. Many of these data points are based on an individual's recollection of the flood events. In reality, the actual flood depth may be within a range of a few centimetres (e.g. +/- 10 cm) of the reported flood depths. Using this range, the peak discharge on each reach and the combined peak discharge of upstream reaches is more realistically represented as a range of values rather than as single values. As a result, future studies should investigate the sensitivity of the developed approach to the variability of reported flood depths. A method shown by Sy et al. (2020), whereby the consistency of information provided by an interviewee is compared between multiple sources can be adopted. For example, two or three people within a household can be independently asked to report the flood depth to establish an approximate upper and lower limit.

A further examination of damage influencing variables for sandcrete and clay buildings is also recommended. This is because buildings within developed BRI classes still show a substantial difference at similar water depths. Although the variation in building standard has already been identified as a likely cause, other variables such as flood velocity and duration may additionally play a role. Future studies should focus on estimating the scatter of structural behaviour within each BRI class.

Furthermore, due to data limitations, it was only possible to evaluate four damage grade classes (class 1+2, class 3, class 4, class 5+6), instead of the original six-class that were defined. Given that this combination of

damage grades may lead to ambiguous classifications, it is recommended that the application of the full six-class damage grades be explored.

Also, another important recommendation for future research is to link the developed damage grade classes to a repair cost so that the application of the new PVM can be extended for monetary loss (absolute or relative) assessments. In general, monetary values of damages are important for i) economic assessment of disaster impacts, ii) cost-benefit assessment of flood protection or local mitigation measures, and iii) estimating compensation cost for victims of flood impacts. Knowledge from civil engineers, quantity surveyors or local technicians can be consulted to provide realistic estimates of repair costs associated with each damage class. Generally, given that repair cost is related to the building type, the susceptibility variables, combining building characteristics such as wall material and thickness and building quality can be integrated as a component to estimate repair costs since it contains information on the structural type. A similar approach was used in Maiwald and Schwarz, (2015) whereby loss is estimated as a function of building type and flood depth.

In general, the thesis explores methods by which limited information in data-scarce areas can be utilized to obtain usable insights to support disaster risk reduction in Nigeria and other typical data-scarce areas. As data on flood characteristics and damage becomes more available, the developed methods should be continually improved and updated to further underpin physical vulnerability assessment in data-scarce areas.

Acknowledgements

Many thanks to Candace Chow for constructive reviews on the document. Also, special thanks to Jorge Ramirez and Sabrina Erlwein for providing insightful feedback.

References

Assumpção, T. H., Popescu, I., Jonoski, A. and Solomatine, D. P.: Citizen observations contributing to flood modelling: opportunities and challenges, Hydrol. Earth Syst. Sci., 22(2), 1473–1489, doi:10.5194/hess-22-1473-2018, 2018.

Barroca, B., Mouchel, J.-M., Bonierbale, T. and Hubert, G.: Flood Vulnerability Assessment Tool (FVAT), DayWater an Adapt. Decis. Support Syst. Urban Stormwater Manag., 123, 2008.

Blong, R.: A review of damage intensity scales, Nat. Hazards, 29(1), 57–76, doi:10.1023/A:1022960414329, 2003.

Borga, M., Comiti, F., Ruin, I. and Marra, F.: Forensic analysis of flash flood response, WIREs Water, 6(e1338), doi:10.1002/wat2.1338, 2019.

Bronstert, A., Agarwal, A., Boessenkool, B., Crisologo, I., Fischer, M., Heistermann, M., Köhn-reich, L., López-tarazón, J. A., Moran, T., Ozturk, U., Reinhardt-Imjela, C. and Wendi, D.: Forensic hydro-meteorological analysis of an extreme flash flood: The 2016-05-29 event in Braunsbach, SW Germany, Sci. Total Environ., 630, 977–991, doi:10.1016/j.scitotenv.2018.02.241, 2018.

Dall'Osso, F., Gonella, M., Gabbianelli, G., Withycombe, G. and Dominey-Howes, D.: A revised (PTVA) model for assessing the vulnerability of buildings to tsunami damage, Nat. Hazards Earth Syst. Sci., 9(5), 1557–1565, 2009.

Englhardt, J., Moel, H. de, Huyck, C. K., Ruiter, M. C. de, Aerts, J. C. J. H. and Ward, P. J.: Enhancement of large-scale flood risk assessments using building-material-based vulnerability curves for an object-based approach in urban and rural areas, Nat. Hazards Earth Syst. Sci., 19(8), 1703–1722, doi:10.5194/nhess-19-1703-2019, 2019.

FGN, (Federal Government of Nigeria): Nigeria: Post-Disaster needs assessment - 2012 Floods. [online] Available from: https://www.gfdrr.org/sites/gfdrr/files/NIGERIA_PDNA_PRINT_05_29_2013_WEB.pdf (Accessed 1 June 2020), 2013.

ImageCat, CIESIN and Porter, K.: Africa Disaster Risk Financing Phase 1 – Result Area 5, Exposure Development for 5 Sub-Saharan African countries, Ethiopia, Kenya, Uganda, Niger, Senegal., 2017.

JRC, (Joint Research Centre-European Commission and OECD): Handbook on Constructing Composite Indicators: Methodology and User Guide, Joint Research Centre-European Commission and OECD., 2008.

Kreibich, H., Schröter, K. and Merz, B.: Up-scaling of multi-variable flood loss models from objects to land use units at the meso-scale, IAHS-AISH Proc. Reports, 373, 179–182, doi:10.5194/piahs-373-179-2016, 2016.

Maiwald, H. and Schwarz, J.: Damage and loss prognosis tools correlating flood action and building's resistance-type parameters, Int. J. Saf. Secur. Eng., 5(3), 222–250, doi:10.2495/SAFE-V5-N3-222-250, 2015.

NBC, (Nigerian Building Code): National Building Code, Federal Republic of Nigeria, 2006.

Papathoma-Köhle, M., Schlögl, M. and Fuchs, S.: Vulnerability indicators for natural hazards : an innovative selection and weighting approach, Sci. Rep., 9, 1–14, doi:10.1038/s41598-019-50257-2, 2019.

Saaty, T. L.: The Analytical Hierarchy Process, Planning, Priority, Resour. Alloc. RWS Publ. USA, 1980.

Schumann, G. J. P., Bates, P. D., Neal, J. C. and Andreadis, K. M.: Measuring and Mapping Flood Processes, in Hydro-Meteorological Hazards, Risks, and Disasters, edited by F. S. John, D. B. Giuliano, and P. Paolo, pp. 35–64, Elsevier Inc., 2015.

Schwarz, J. and Maiwald, H.: Damage and loss prediction model based on the vulnerability of building types, 4th Int. Symp. Flood Def. Manag. Flood Risk, Reliab. Vulnerability, 6–8, doi:10.13140/2.1.1358.3043, 2008.

Sy, B., Frischknecht, C., Dao, H., Consuegra, D. and Giuliani, G.: Reconstituting past flood events: The contribution of citizen science, Hydrol. Earth Syst. Sci., 24(1), 61–74, doi:10.5194/hess-24-61-2020, 2020.

Zischg, A. P., Felder, G., Mosimann, M., Röthlisberger, V. and Weingartner, R.: Extending coupled hydrological-hydraulic model chains with a surrogate model for the estimation of flood losses, Environ. Model. Softw., 108, 174–185, doi:10.1016/j.envsoft.2018.08.009, 2018.

Supplementary material

1 Supplements to Chapter 3

1.1 Supplementary Material 1: Questionnaire for field data collection

1. Was there contact between the building and floodwater? [Yes] [No]

If yes;

- Was it only with the external wall or was there water intrusion? …….
- How long was the contact with the building wall? …… (hours/days/weeks)
- What was the (estimated) height? ……..…… (cm, m)
- Was there a large amount of mud, coarse debris like pebbles or stones (or other transported material) along with the floodwater? [Yes] [No]

If yes, comment on the kind of material and the deposition height on the wall

……..……..

2. Do you have any remaining flood mark on any part of your building? [Yes] [No]

If yes, Take photograph.

3. Were the floodwaters moving fast or slow? How could you simply describe the speed? How would you compare it with a normal human walking pace? On what day was it?

Describe: ……

4. Do you have any video or pictures you took during the flood event? [Yes] [No]

If yes, could I use it for research purpose?

Comment: ……

5. Have you experienced a flood of this magnitude (bigger or smaller) in this area before? [Yes] [No]

If yes, what year? ……what was the flood height?.....

Additional comment about the earlier flood

……

6. Is there any flood mark on any object (building, trees etc.) close by? [Yes] [No]

If yes, describe object……

Coordinates: LONGITUDE ……...LATITUDE……

Flood depth on neighbouring object.....

Additional Comment ……..

1.2 Supplementary Material 2: Figures

Table 1: Sensitivity analysis for various reductions in channel elevation. Shaded entries are values used in the study.

	Channel elevation drop (m)	Peak discharge RMSE (m)	Peak discharge (m^3s^{-1})	Percent change in channel conveyance	Percent change in peak discharge
Reach 1	0.5	0.675	78.3	-50	-11
	0.75	0.674	83.3	-25	-5
	1	0.676	88	—	—
	1.25	0.68	88.3	25	0
	1.5	0.683	98.3	50	12
Reach 2	0.5	0.489	73.3	-50	-8
	0.75	0.492	76.6	-25	-4
	1	0.494	80	—	—
	1.25	0.491	86.6	25	8
	1.5	0.496	91.6	50	15
Reach 3	0.5	0.425	64	-50	-11
	0.75	0.425	69	-25	-5
	1	0.435	72.3	—	—
	1.25	0.432	79	25	9
	1.5	0.44	85.6	50	18
Reach 4	0.5	0.309	90	-50	-12
	0.75	0.312	96	-25	-6
	1	0.313	102	—	—
	1.25	0.323	108	25	6
	1.5	0.331	116	50	14
Reach 5	0.5	0.507	128.3	-50	-13
	0.75	0.512	136.6	-25	-7
	1	0.526	147	—	—
	1.25	0.524	150	25	2
	1.5	0.536	165	50	12

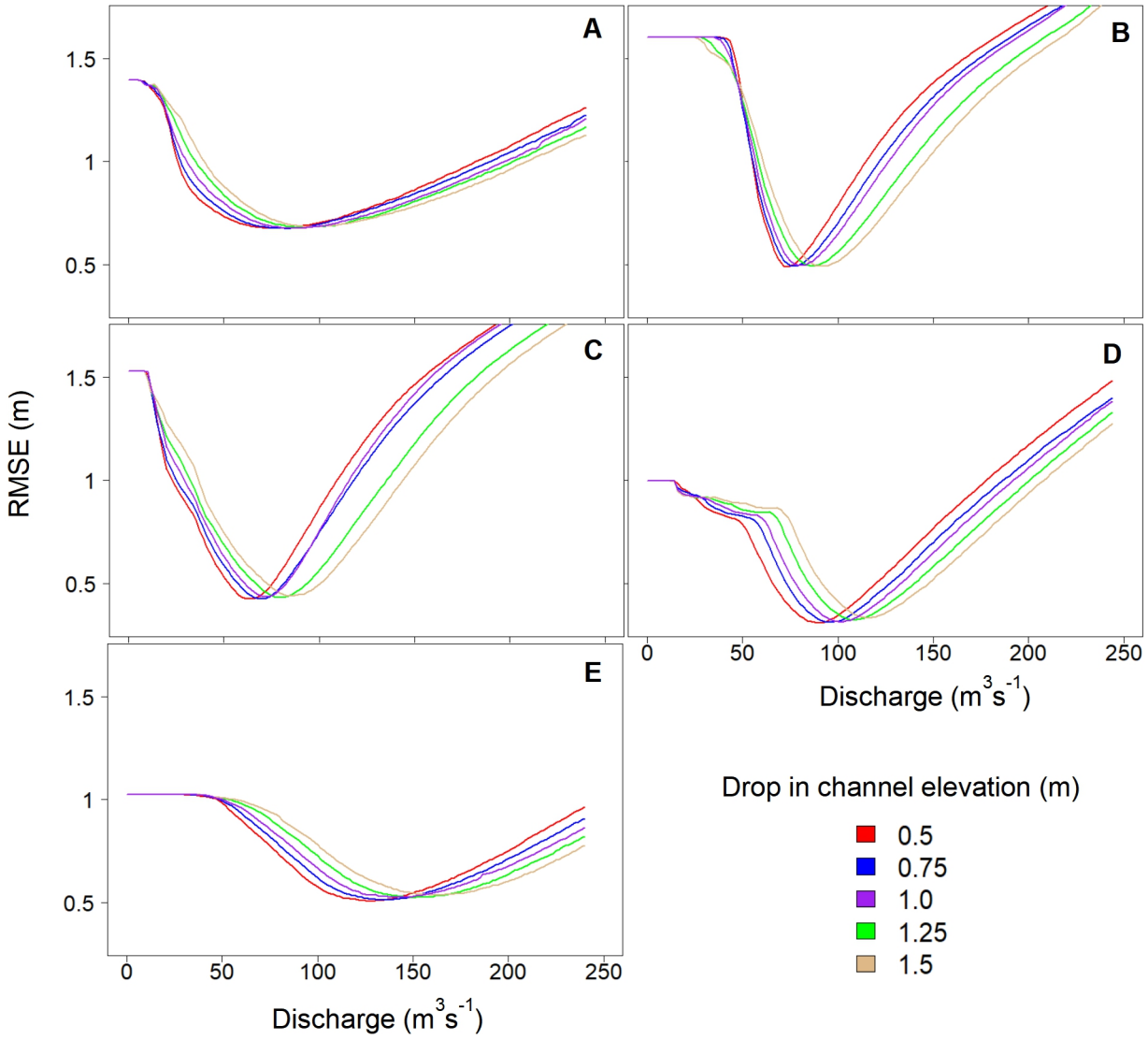


Figure 1: RMSE between observed and simulated flood depths from models driven with linearly increasing discharge and reductions in channel elevations of 0.5 m, 0.75 m, 1 m, 1.25 m and 1.5 m on (a) Reach 1, (b) Reach 2 (c) Reach 3 (d) Reach 4 (e) Reach 5. Y-axis limits are shown for the range 0.3 m to 1.7 m

Simulations representing -/+ 25 % change in channel conveyance (channel lowering or raising by 0.25 m) resulted in peak discharge between -7 % (reach 5) to 9 % (reach 3) relative to a 1 m channel lowering (see S2 Table 1). Whereas, simulations representing -/+ 50 % change in channel conveyance (channel lowering or raising by 0.5 m) resulted in peak discharge between -13 % (reach 5) to 18 % (reach 3) relative to a 1 m channel lowering (see S2 Table 1). While changes in channel conveyance between -/+ 25 % results in a minimal change in peak discharge, changes between -/+ 50 % results in a moderate change in peak discharge. Given our initial choice of a 1 m drop in channel elevation based on field observations, the actual channel elevation is most likely within the -/+ 0.25 m change in channel elevation. However, in a worst case, a -/+ 0.5 m change in channel elevation does not drastically alter the peak discharge.

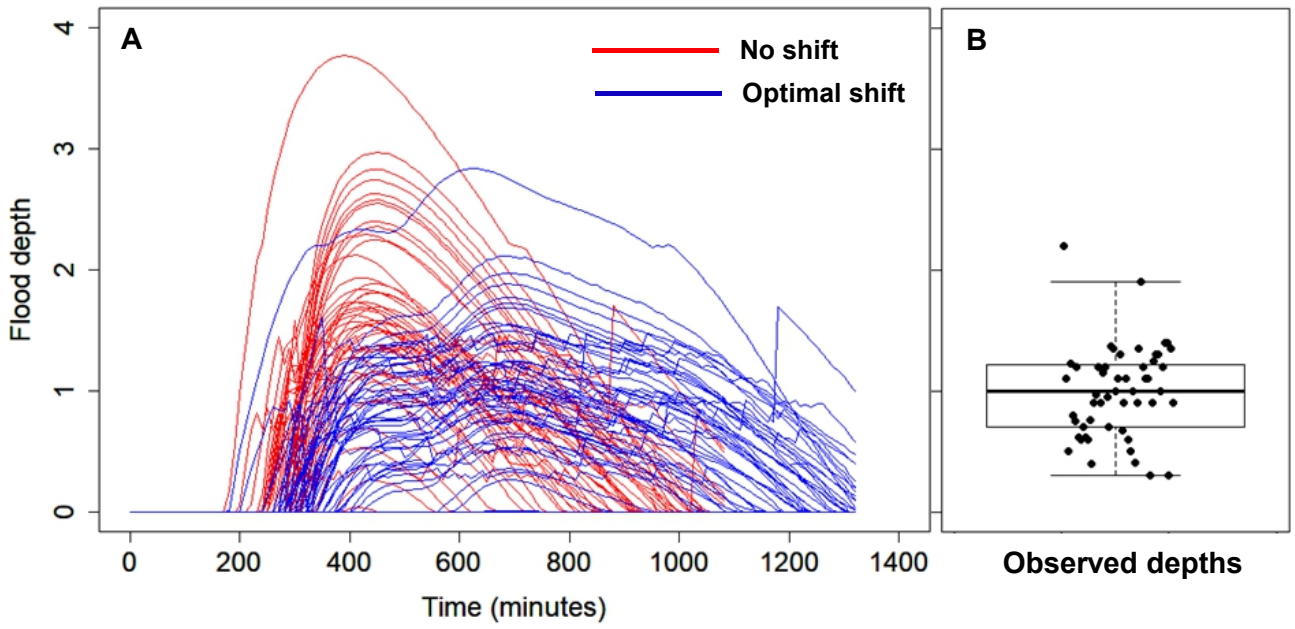


Figure 2: (a) Time-series of simulated flood depths for reach 5 using a hydrograph combination with peak discharge of 147 m^3s^{-1} (optimal shift) and peak discharge of 303 m^3s^{-1} (no shift). (b) Observed flood depths for reach 5.

2 Supplements to Chapter 4

2.1 Supplementary material 1: Questionnaire 1

Building ID:……

General questions (Gq)

Gq.1. Where you residing in the building at the time of the flood? [Yes] [No]

Gq.2. Street name and house number; ……

Gq.3. Use of the building. Residential [ ] Commercial [ ] Religious [ ] Other……

Gq.4. Building coordinates. LONGITUDE……LATITUDE……

SKETCH

PLAN VIEW	SIDE VIEW	FRONT VIEW

Gq.5. Building Photo

Gq.6. Photo Number……

Local protection (Lp) - (NOTE: building measures before flood event)

Lp.1. Was the ground floor elevated above the surrounding ground level? [Yes] [No]

If yes, by how many centimeters? ……

Or estimate (in cm): [0 – 20] [20 - 40] [40 - 60] [60 – 80] [80 – 100] [Above]

Lp.2. Is the building fenced/walled? [Yes] [No]

If yes, by which material? [Thatched] [Clay bricks] [Sandcrete block] Other……

Is the fence plastered? If yes, by which material? [Clay] [Cement and sand] Other?.....

Lp.3. Are the walls of the building plastered? [Yes] [No]

If yes, by which material? [Clay] [Cement and sand] Other?.....

Lp.4. Are there other local protection measures? [Yes] [No]

If yes, describe; ……

Additional
Comment: ……..

Susceptibility variables (Sq). (NOTE) - Answers should relate to building status before flood event)

Sp.1. What is the building type (wall material)? [Stone Masonry] [Burnt Clay Bricks]

[Rammed Earth Wall] [Sun dried Clay Bricks] [Sandcrete Block] [Reinforced Concrete]

Other? Describe; …….

Sp.2. What is the (mean) external wall thickness? [10 cm] [15 cm] [22.5 cm] Other……

Sp.3. What is the condition of the following building components? Give the condition that best describe the building state before the flood event.

Walls/Columns/Beams (or lintels)	Very Good	Good	Moderate	Poor
Doors/Windows	Very Good	Good	Moderate	Poor
Finishes/Rendering	Very Good	Good	Moderate	Poor
Flooring (ground floor)	Very Good	Good	Moderate	Poor
Roofing	Very Good	Good	Moderate	Poor
Selected (overall) Building condition				

Comment of any structural
deterioration; …….....
..

Sp.4. What is the shape of the building footprint? [L-shape] [Rectangular] [Square] [Round]

if other? ……

Sp.5. What is the material of the openings-doors and windows? [Wood] [Metal] Other ……

Sp.6. What is the height of lowest windows (or other opening) above the floor level? If no precise value, give best range. ……

Sp.7. What is your opinion about the overall quality of the building? [Good] [Moderate] [Poor]

Finishes/Rendering (cover for weather exposed surfaces)	Material: ……	Aesthetics			Durability			Functionality			
		1	2	3	1	2	3	1	2	3	
Doors/Windows	Material: ……	Aesthetics			Durability			Functionality			
		1	2	3	1	2	3	1	2	3	
Roofing	Material: ……	Aesthetics			Durability			Functionality			
		1	2	3	1	2	3	1	2	3	
External wall	Material: ……	Aesthetics			Durability			Functionality			
		1	2	3	1	2	3	1	2	3	
Ceiling height (Standard is 2.4m)	Height of roof eaves ……..				Height of roof ridge ……..						
Access to a public way/yard on not less than one side	Yes	No		One room not less than 12m^2		Yes		No			
Adequate support to overhanging parts		1	2	3							
External wall corner columns		1	2	3							
Properly tied roofing parts and fascia		1	2	3							
Selected (overall) Building standard											

Sp.8. When was the building constructed? ……

If no precise date available, make best estimate. [Before 1995] [1996 - 2000] [2001 - 2005] [2006 - 2010] [2011 - 2015] [After 2015]

Sp.9. What is the renovation status of the building at the time of the flood?

[Renovated] [Not renovated] [Newly constructed] [Newly constructed]

Sp.10. What is the number of storey? [1] [2] [3] [Other…….]

Additional
Comment: …….....
.....

…….....

Exposure variables (Ex)

Ex.1. What is the distance of the building to the closest following water-preferred-parts?

[Road]…… [River/Water Channel]…… [City Drainage]….....

Other ……

Ex.2. What is the mean elevation of the building above sea level? .....

Ex.3. What is the building location relative to other buildings? [Head] [Middle] [Stand Alone]

Ex.4. Are there natural barriers around the building? [Yes] [No]

If yes describe; ……

Ex.5. Are there functional water gutter/drainage around the building? [Yes] [No]

Additional Comment: …….....

.....

…….....

Action variables: see supplements to chapter 2

DAMAGE (NOTE: Answers should relate to the status of situation of the building immediately after the flood)

D.1. Damage on the building

Building Component	Observed Effects		Observ.	(Additional) Comment/Observation
Ground floor Finishes/ Rendering (Plaster, tiles etc.)	Moisture defects			
	Cracks	Minor		
		Moderate		
	De-bonding			
	Peeling off			
External wall Finishes/ Rendering (Plaster, tiles etc.)	Moisture defects (discoloration, algae)			
	Cracks	Minor		
		Moderate		
	De-bonding			
	Peeling off			
Door/Windows	Impressed			
	Distortion of frame			
	Rot (wood)			
Ground floor surface	Cracks	Minor		
		Moderate		
		Heavy		
Walls (External)	Cracks	Minor		
		Moderate		
		Heavy		
	Collapse	Partial < 1/3		
		Partial > 1/3		
		Complete		
Beams (with lintels)	Cracks	Minor		
		Moderate		
		Heavy		
	Stress (Spalling etc.)	Minor		
		Moderate		
		Heavy		
	Collapse	Partial		
		Complete		
Columns	Cracks	Minor		
		Moderate		
		Heavy		
	Stress	Minor		

	(Spalling etc.)	Moderate		
		Heavy		
	Collapse	Partial		
		Complete		
Roof (distortion, leakage)	Sheet			
	Truss/ frame			
	Ceiling			

ASSIGNED DAMAGE GRADE	

D.2. Was there damage to the fencing wall? If yes; Thick appropriate: [Yes] [No]

[Cracks: [Minor] – [Moderate] – [Heavy]] [Collapse: [Partial] – [Complete]]

If other: …….

D.3. Did you observe the building impacted by any debris or other transported material, aiding failure? If yes,

Describe……

D.4. Did you carry out any repair work after the flood event on the building? If yes, briefly state the measure:

Finishes (ground floor)	
Finishes (Walls)	
Doors/Windows	
Walls	
Column/Beams	
Roofing	

Additional

Comment: …….....

.....

…….....

2.2 Supplementary material 1: Questionnaire 2

'LEVEL OF INFLUENCE' TABLE

1	2	3	4	5	6	7	8	9
Slight influence	Slight to moderate influence	Moderate influence	Moderate to strong influence	Strong influence	Strong to Very strong influence	Very strong influence	Very Strong to Extreme influence	Extreme influence

The table above is a 'level of influence' table with scores attached. This will be used to rank the extent to which each parameter can influence vulnerability to buildings to flood in a region leading to damage(s). This damage could be as small-scaled as moisture defects from water intrusion or as large scaled as a complete collapse of the building.
NB: Please note that two parameters can be assigned the same level of influence if from your opinion that is the best evaluation of the level of influence.

1. According to the 'level of influence' table, how would you rank the following SUSCEPTIBILITY Variables based on their level of influence on building vulnerability to damage by flood.

Susceptibility Variables	Level of influence Score
Building type (sandcrete wall, clay wall etc.)	
Building condition	
Shape of the building footprint (square, round, l-shape etc.)	
Height of openings (e.g. windows)	
Building standard (quality of construction)	
Building age	
Number of storey	

2. According to the 'level of influence' table, how would you rank the following EXPOSURE PARAMETERS based on their level of influence on building vulnerability to damage by flood.

Exposure Parameters	**Level of influence Score**
Distance to preferred water path (river channel/road)	
Height above sea level	
Building location (are there buildings beside it or is it alone)	
Presence of natural barriers (trees, bushes etc.)	
Availability of functional drainage channel in the area	

3. According to the 'level of influence' table, how would you rank the following LOCAL PROTECTION/RESILIENCE variables based on their level of influence on building vulnerability to damage by flood.

Local protection variables	**Level of influence Score**
Elevation of building above ground level	
Fencing of the building (protection walls)	
Finishes/rendering (plaster, tiles on building walls)	

4. According to the 'level of influence' table, how would you rank the following ACTION (FLOOD HAZARD) variables based on their level of influence on building vulnerability to damage by flood.

Action (Flood Hazard) variables	**Level of influence Score**
Flood Height (Depth)	
Flood Duration	
Flood Velocity	
Transported Materials	

5. In comparison between SUSCEPTIBILITY, EXPOSURE, LOCAL PROTECTION variables, how would you rank their level of influence on building vulnerability to damage by flood according to the 'level of influence' table.

Building Resistance variables	**Level of Influence Score**
Susceptibility Parameters	
Exposure Parameters	
Resilience Parameters	

<u>LOCAL PROTECTION IMPROVEMENT</u>

7. Do you think the local protection measures (outlined under resilience parameters) presently employed by the communities are effective against building damage?

[Yes] [No]

Additional comment (if any): …….

……

8. Do you have suggestion/recommendation for local protection measures?

[Yes] [No]

If yes, which?……

General Comment (if
any): …….

…….....

2.3 Supplementary material 1: Questionnaire 3

Key issues for expert evaluation

The expert evaluation aims to utilize expert knowledge for predicting the level of damage that a building can incur given a synthetic (scenario-based) flood depth. For this expert evaluation we have prepared following evaluation structure:

We have, as representatives, **three (3) categories of buildings** (low, moderate and high building vulnerability) classified based on building material, condition and quality (Figures 1, 2, 3-see page 2).

For each category of building, we ask that you assign the **expected level of damage** (Tables 1, 2, 3- see expert evaluation sheet on pages 3, 4 and 5) a building can incur given a flood depth scenario between 0 to < 1m, 1 to < 2m, 2 to < 3m, 3 to < 4m, and 4 to < 5 m. Since we are aware that making exact predictions are rather difficult and possibly have high uncertainties, we provide **three probable damage categories** that can be used to capture the range of possible damage;

•**Low Probable Damage (LP):** Least possible damage from a given flood depth for a given building category

•**Most Probable Damage (MP):** Most likely damage from a given flood depth for a given building category

•**High Probable Damage (HP):** Maximum possible damage from a given flood depth for a given building category

Please proceed first with making you familiar with the different categories of buildings on page 2. Afterwards we kindly ask you to evaluate each building categories on the individual expert evaluation sheets. If you have any questions, please contact me. On the last page you have the possibility to add further comments.

Thank you for supporting the research on physical vulnerability of buildings due to floods by sharing your knowledge.

Building categories based on high (Figure 1), moderate (Figure 2), and low (Figure 3) vulnerability


Figure 1: High vulnerability class. Buildings with poor building material, building condition and building quality. NB: Please use these buildings only for ASSESSMENT SHEET 1 - Table 1

Figure 2: Moderate vulnerability class. Buildings with moderate (fair) building material, building condition and building quality. NB: Please use these buildings only for ASSESSMENT SHEET 2 - Table 2

Figure 3: Low vulnerability class. Buildings with good building material, building condition and building quality. NB: Please use these buildings only for ASSESSMENT SHEET 3 - Table 3

TABLE 1 - EXPERT 'WHAT-IF ANALYSIS' ASSESSMENT SHEET 1

Category – HIGH BUILDING VULNERABILITY (see Figure 1)

(Buildings considered to have a generally **low quality of construction**, **poor maintenance condition** and **poor plaster coverage and quality**)

Low Probable (**LP**) damage (lowest possible damage from a given flood depth): **LP** damage should be **lower than** MP and HP	LP
Most probable (**MP**) damage (most possible damage from a given flood depth): **MP** damage should be **between** LP and HP	MP
High probable (**HP**) damage (Highest possible damage from a given flood depth): **HP** damage should be **higher than** LP and MP	HP

BUILDINT ELEMENT	**OBSERVED EFFECTS**		**WATER DEPTH**													
			> 0 to < 1 m			**> 1 m to < 2 m**			**> 2 m to < 3.0**			**> 3 m to < 4 m**			**> 4 m to < 5 m**		
			LP	MP	HP	LP	MP	HP	LP	MP	HP	LP	MP	HP	LP	MP	HP
Wall finishes – plaster or tiles	Only water contact and (or) moisture defects																
	Cracks																
	De-bonding																
	Peeling off																
Floor finishes (plaster/tiles)	Moisture defects																
	Cracks																
	De-bonding																
	Peeling off																
Flooring (ground)	Cracks	Minor															
		Moderate															
		Heavy															
	Settlement (Ground subsidence)																
Walls (Main wall material e.g. Sandcrete block or clay)	Cracks	Minor															
		Moderate															
		Heavy															
	Collapse	Partial															
		Complete															

TABLE 2 - EXPERT 'WHAT-IF ANALYSIS' ASSESSMENT SHEET 2

Category – MODERATE BUILDING VULNERABILITY (see Figure 2)

(Buildings considered to have a generally **moderate quality of construction**, **moderate maintenance condition** and **moderate plaster coverage and quality**)

Low Probable (**LP**) damage (low possible damage from a given flood depth): **LP** damage should be **lower than** MP and HP	LP
Most probable (**MP**) damage (most possible damage from a given flood depth): **MP** damage should be **between** LP and HP	MP
High probable (**HP**) damage (Highest possible damage from a given flood depth): **HP** damage should be **higher than** LP and MP	HP

BUILDINT ELEMENT	**OBSERVED EFFECTS**		**WATER DEPTH**													
			> 0 to < 1 m			**> 1 m to < 2 m**			**> 2 m to < 3.0**			**> 3 m to < 4 m**			**> 4 m to < 5 m**		
			LP	MP	HP	LP	MP	HP	LP	MP	HP	LP	MP	HP	LP	MP	HP
Wall finishes – plaster or tiles	Only water contact and (or) moisture defects																
	Cracks																
	De-bonding																
	Peeling off																
Floor finishes (plaster/tiles)	Moisture defects																
	Cracks																
	De-bonding																
	Peeling off																
Flooring (ground)	Cracks	Minor															
		Moderate															
		Heavy															
	Settlement (Ground subsidence)																
Walls (Main wall material e.g. Sandcrete block or clay)	Cracks	Minor															
		Moderate															
		Heavy															
	Collapse	Partial															
		Complete															

TABLE 3 - EXPERT 'WHAT-IF ANALYSIS' ASSESSMENT SHEET 3

Category – LOW BUILDING VULNERABILITY (see Figure 3)

(Buildings considered to have a generally **high quality of construction**, **good maintenance condition** and **good plaster coverage and quality**)

Low Probable (**LP**) damage (low possible damage from a given flood depth): **LP** damage should be **lower than** MP and HP	LP
Most probable (**MP**) damage (most possible damage from a given flood depth): **MP** damage should be **between** LP and HP	MP
High probable (**HP**) damage (Highest possible damage from a given flood depth) : **HP** damage should be **higher than** LP and MP	HP

BUILDINT ELEMENT	**OBSERVED EFFECTS**		**WATER DEPTH**													
			> 0 to < 1 m			**> 1 m to < 2 m**			**> 2 m to < 3.0**			**> 3 m to < 4 m**			**> 4 m to < 5 m**		
			LP	MP	HP	LP	MP	HP	LP	MP	HP	LP	MP	HP	LP	MP	HP
Wall finishes – plaster or tiles	Only water contact and (or) moisture defects																
	Cracks																
	De-bonding																
	Peeling off																
Floor finishes (plaster/tiles)	Moisture defects																
	Cracks																
	De-bonding																
	Peeling off																
Flooring (ground)	Cracks	Minor															
		Moderate															
		Heavy															
	Settlement (Ground subsidence)																
Walls (Main wall material e.g. Sandcrete block or clay)	Cracks	Minor															
		Moderate															
		Heavy															
	Collapse	Partial															
		Complete															

Additional comments if any:

General comments:

Table 1:

Table 2:

Table 3:

2.4 Supplementary material 2: Figures

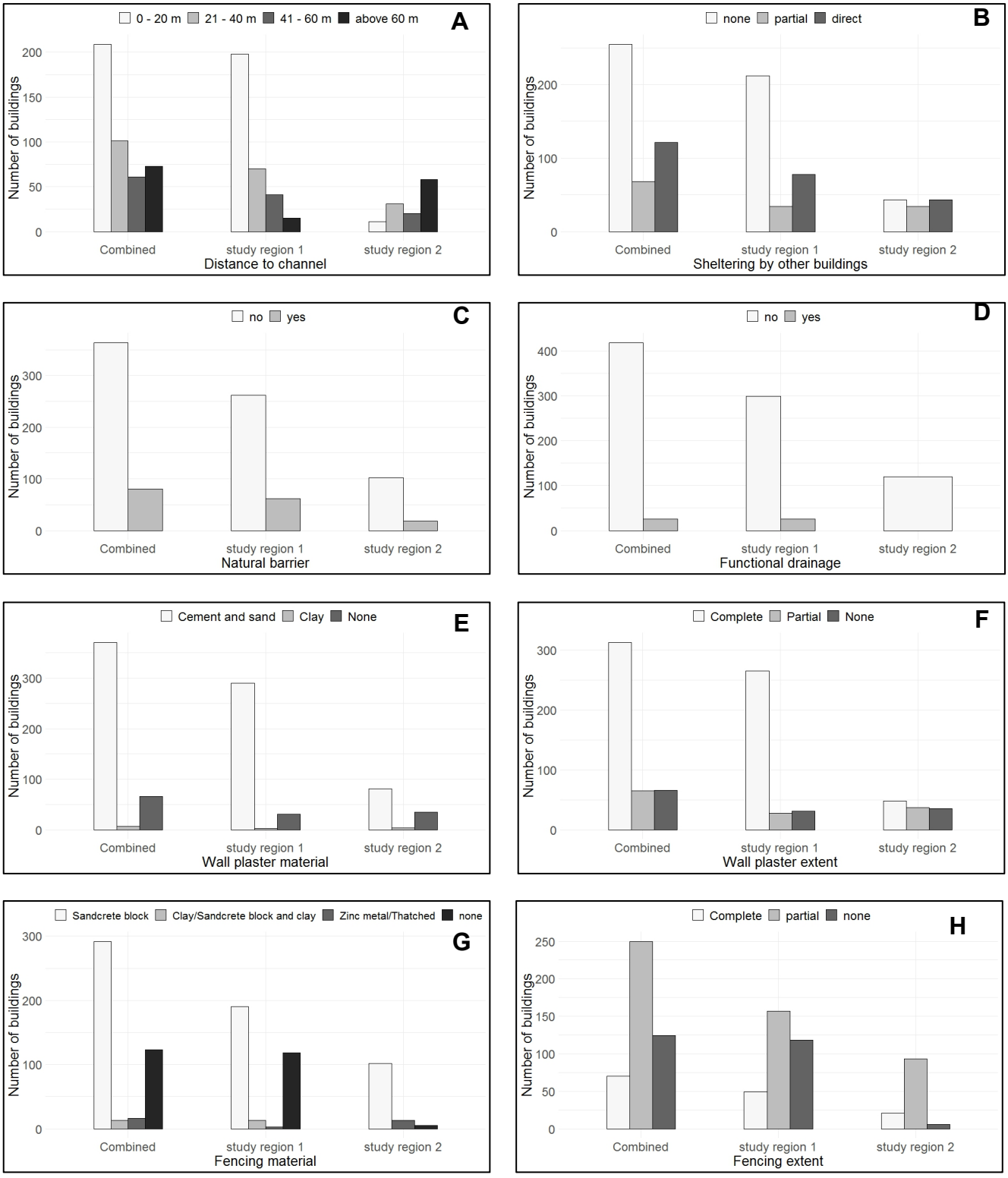
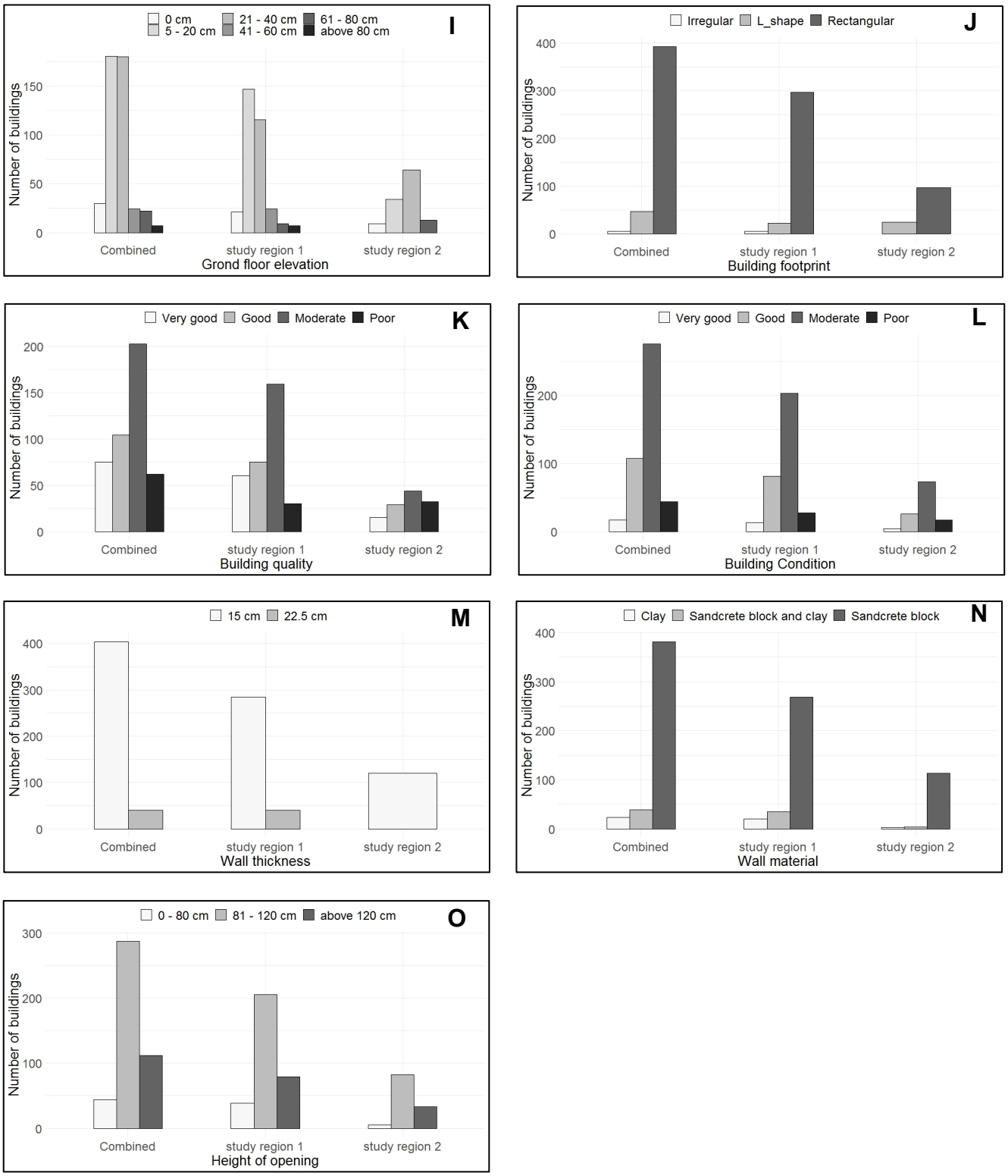


Figure S1: Variable factor distribution for combined data and individual data from study regions and 2



Cont. Figure S1: Variable factor distribution for combined data and individual data from study region

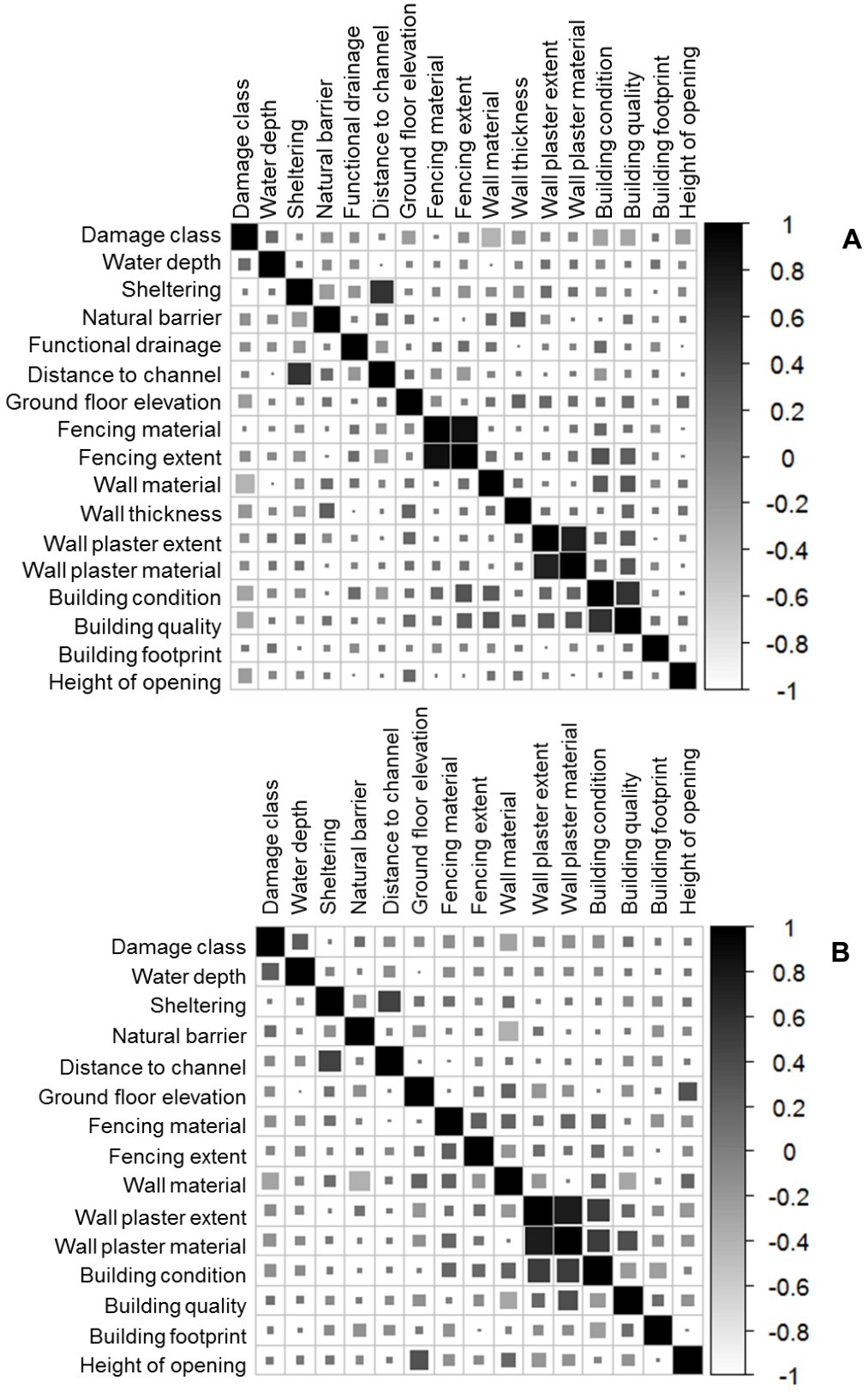


Figure S2: Spearmans' correlation coefficient for (A) study region 1, and (B) Study region 2. Areas of squares represent absolute values of correlation coefficients

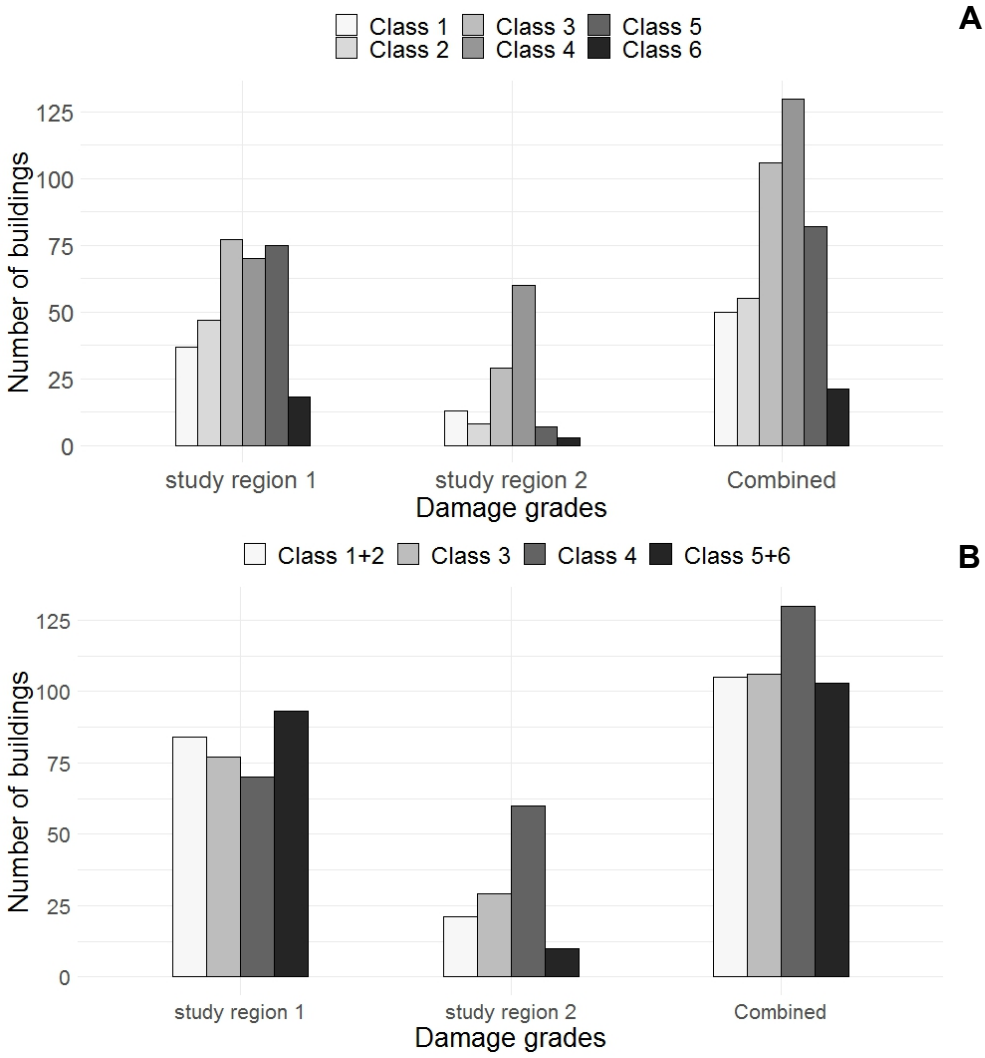


Figure S3: Distribution of damage grade classes with a (A) 6-classes, and (B) 4 classes

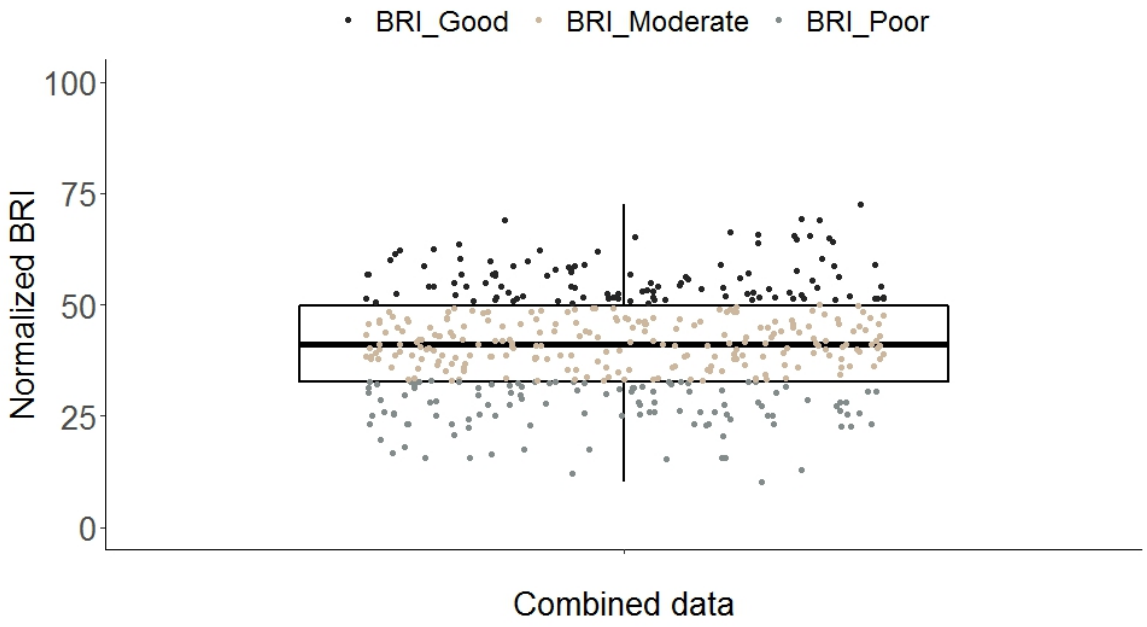


Figure S4: Normalized building resistance index (BRI) indices showing quantile classification used to categorize buildings into BRI classes good (range from upper quartile and the maximum), moderate (interquartile range) and poor (range between the lower quartile and minimum).

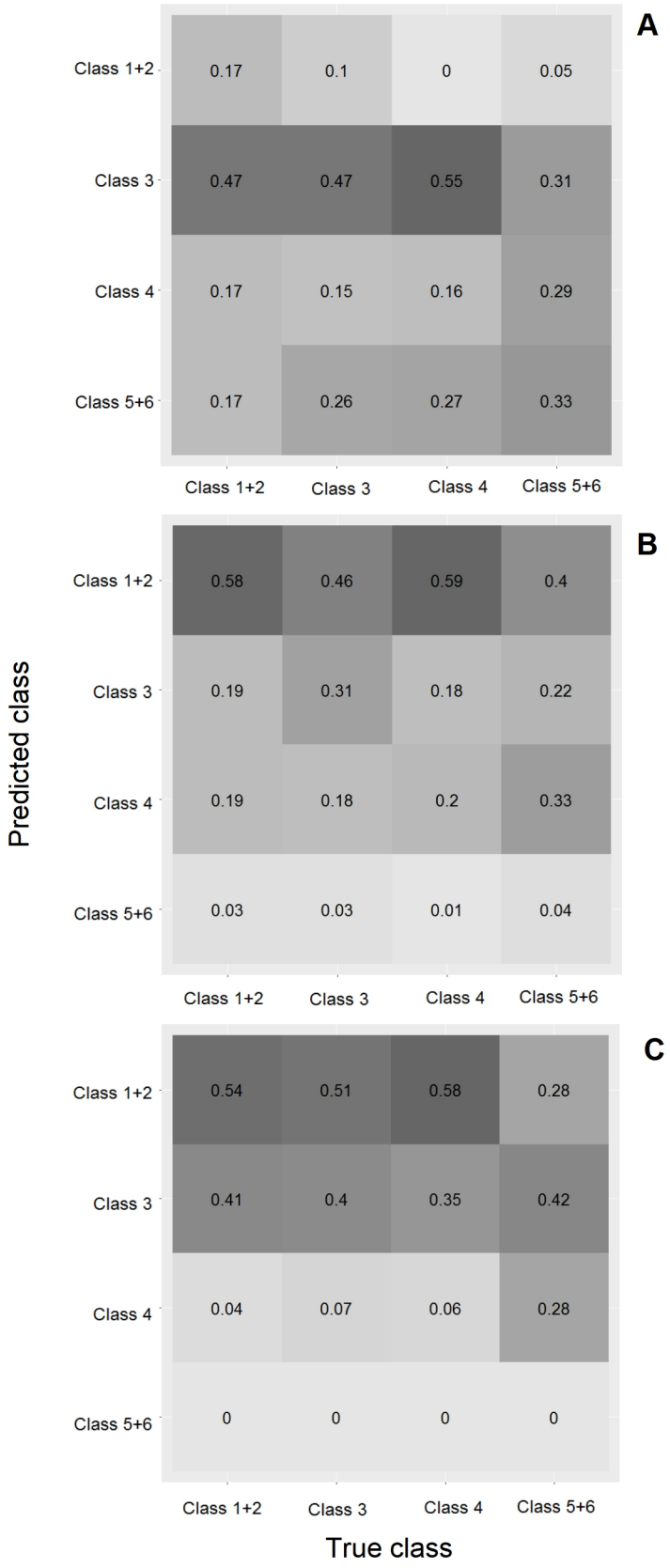


Figure S5: Heatmaps generated from the confusion matrix for BRI classes (A) poor, (B) moderate, and (C) good categories. Results for data-driven method trained on study region 1and tested on study region 2

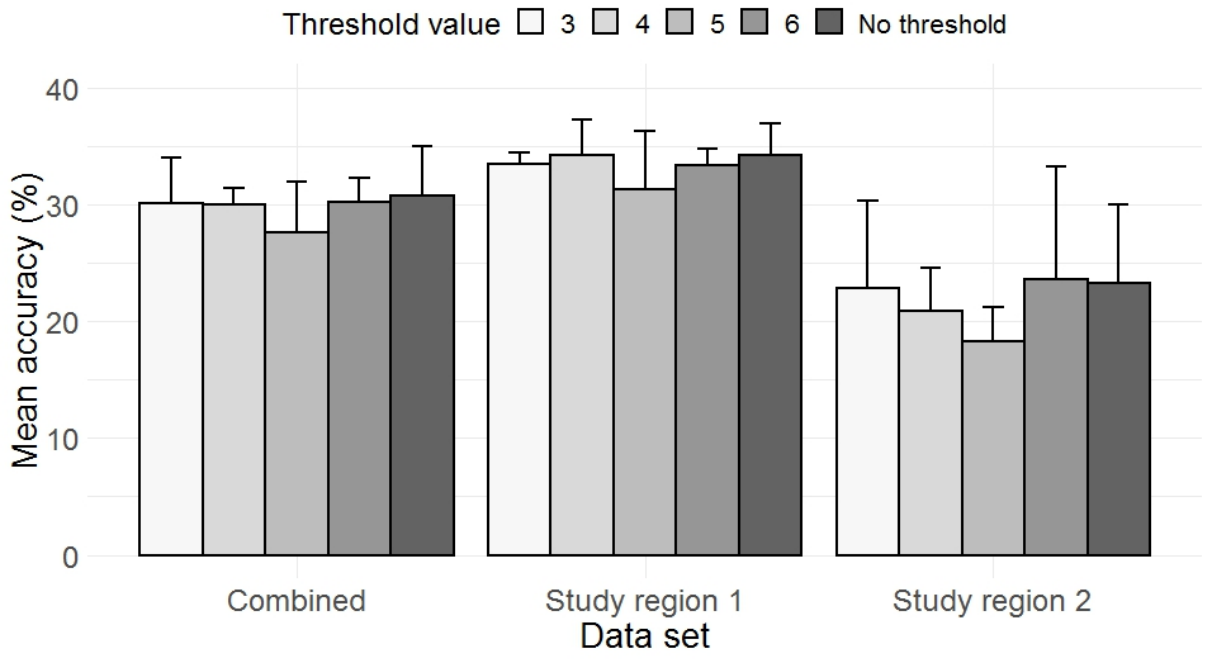


Figure S6: Sensitivity analysis for threshold values showing mean correct prediction (barplots) and standard deviation (errorbars) across different data sets. Mean and standard deviations are calculated from percentage accuracy of the BRI classes poor, moderate and good.

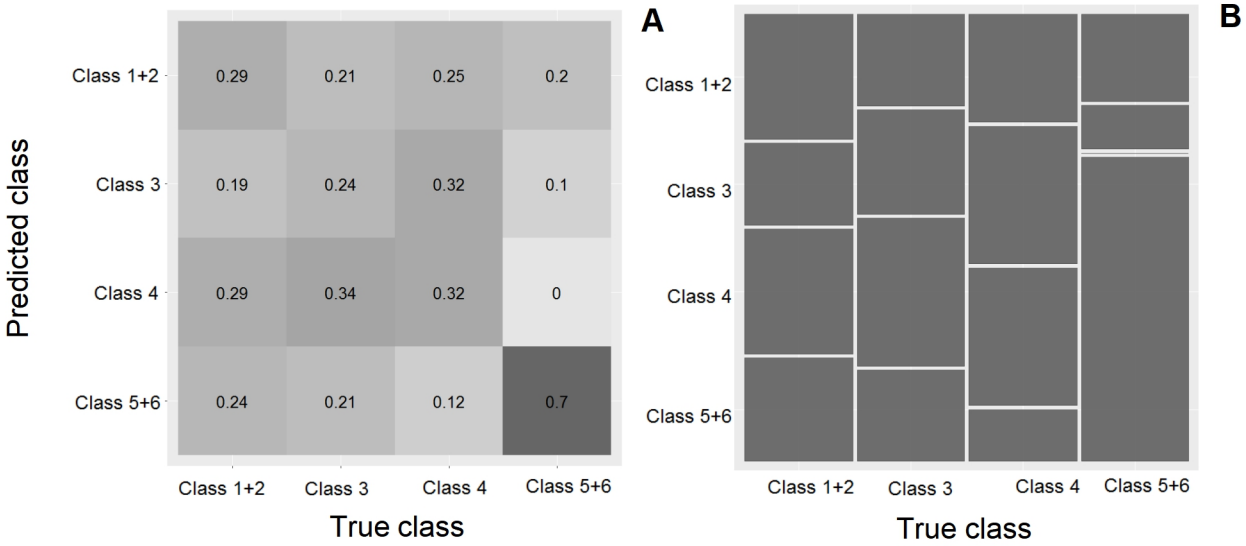


Figure S7: (A) Heatmap and (B) Mosaic plot for evaluating model transferability in study region 2 based on model training on data from study region 1 and testing on study region 2

Declaration of consent

on the basis of Article 18 of the PromR Phil.-nat. 19

Name/First Name: Mark, Bawa Malgwi

Registration Number: 17-133-125

Study program: Geography

Bachelor ☐ Master ☐ Dissertation ☑

Title of the thesis: A physical vulnerability model for floods in data-scarce regions: a case study of Nigeria

Supervisor: Prof. Dr. Margreth Keiler

I declare herewith that this thesis is my own work and that I have not used any sources other than those stated. I have indicated the adoption of quotations as well as thoughts taken from other authors as such in the thesis. I am aware that the Senate pursuant to Article 36 paragraph 1 litera r of the University Act of September 5th, 1996 and Article 69 of the University Statute of June 7th, 2011 is authorized to revoke the doctoral degree awarded on the basis of this thesis.
For the purposes of evaluation and verification of compliance with the declaration of originality and the regulations governing plagiarism, I hereby grant the University of Bern the right to process my personal data and to perform the acts of use this requires, in particular, to reproduce the written thesis and to store it permanently in a database, and to use said database, or to make said database available, to enable comparison with theses submitted by others.

Bern, 30.09.2020

Place/Date

Mark, Bawa Malgwi

Signature