Validating flood damage estimations

A framework using plausibility and empirical data

Jason Julius Wever

TU Delft

Validating flood damage estimations

A framework using plausibility and empirical data

by

Jason Julius Wever

Submitted in partial fulfilment of the requirements for the
double degree of Master of Science in

Civil Engineering
at the department of Hydraulic Engineering

and

Construction Management and Engineering

at

Delft University of Technology

To be publicly defended on Wednesday the 13th of July, 2022 15:00

An electronic version of this thesis is available at https://repository.tudelft.nl/

Student number:	4565142	
Project duration:	September 6th, 2021 – July, 2022	
Graduation committee:	Prof. Dr. ir. Matthijs Kok	TU Delft - Flood Risk, Chair
	Dr. ir. Maria Nogal Macho	TU Delft - Engineering & Systems
	Dr. ir. Martine Rutten	TU Delft - Water Management
	Ir. Laurens Bart	RHDHV
	Dr. Lars de Ruig	RHDHV

TU Delft

Abstract

Flood risk models are frequently used to analyse the climate- and socio-economic-driven impact of flooding hazards. However, model validation is rarely done adequately due to the rare occurrence of floods and even less frequent reporting of corresponding damages. In this research, validation is defined as the process of ensuring that a model performs within a range of accuracy and precision, satisfactory for its intended use. To guide experts in their validation efforts, a four-phased framework is developed to validate flood-event damage estimations, created with hazard x exposure x vulnerability models.

The framework was applied to two damage estimates created by the Global Flood Risk Tool (GFRT). 1) For damage caused by the Limburg 2021 river flood (The Netherlands - Europe) and 2), for damage caused by a 2019 hurricane-induced coastal flood in Beira (Mozambique - Africa). For the Limburg case, total direct damage was determined at 349,4 million euro. An initial model overestimation of 34% was caused primarily by a large exaggeration of exposed agricultural surface area, and significant modelling errors of linear infrastructure. Furthermore, an uncertainty range was quantified between 271,8 (-23%) and 388,2 million euro (+11%) due to uncertainty in residential assets (across all three model parameters) and an uncertain exposure parameter of agricultural assets. To create additional damage estimates for verification, a Structured Expert Judgement (SEJ) experiment was executed with ten flood-damage experts. Due to the high experiment cost and low expert-informativeness, the method is currently not advised as a validation approach. In situations with limited data, experts may still be a relevant information source.

For Beira, damage was determined at 8,1 million US dollar. The model underestimated damage by 82% due to errors in infrastructure, industrial, and commercial assets. Besides, overestimations were found for informal residential- and agricultural assets. The estimate ranges between 5,2 (-36%) and 13,2 million US dollar (+62%). This range excludes uncertainties at port and industrial assets, as insufficient information was available. Contrary to the Limburg case study, insights from the plausibility assessment were too uncertain for quantification, thus the validated estimate is based on damage- and construction cost data. Novel techniques were used to disaggregate the compound damage data, such as comparing wind and flood vulnerability curves and applying employee-based estimations.

The significantly altered damage estimate for both case studies demonstrates the usefulness of the framework. However two main limitations remain: first, lacking information on direct damage to critical infrastructure hinders validation. Second, additional detail in data is required to allow parameter calibration that increases accuracy across multiple flooding scenarios. Therefore, the main recommendation for future research is to increase the detail in damage data reporting so that parameter calibration is supported. This may be done by increasing the spatial resolution of reported damages or adding additional variables such as inundation depth in reports.

Preface

This project was the end of a six-year study period at Delft University of Technology. Through many opportunities and projects available at the campus, I formed a broad interest that drove me to combine the two MSc. programmes Hydraulic Engineering and Construction Management Engineering. This project is truly representative of this period, focusing both on engineering analyses as well as the critical thinking and structured approach that was taught so well here in Delft.

I'm thankful for the opportunities the campus created, and the people and organisations it allowed me to meet over the past time. I met many friends here that supported me throughout the project and with whom I created many great memories. I want to thank all my friends who supported me throughout the project and made my study time an unforgettable experience. Especially Shakthi, Ligaya and Roberto, who have been there from the start. Furthermore, I want to thank my parents who provided an ideal base to explore from. Both mentally, to develop myself in Delft and always supported my projects, as well as at home whenever I needed a rest from my student life or thesis. Finally, I want to thank my girlfriend Kavya, for all our relaxing ice cream walks.

This project would have been of much lower quality, and much less enjoyable without all of my project supervisors.Laurens and Lars, thank you for your approachability and all the interesting discussions we had at the RHDHV office, in which you never let it feel that an extension was problematic. I'd also like to thank all the other people at RHDHV for their practical insights, relaxing coffee conversations and many fun events. You provided a great working environment for this long project.

Thank you Matthijs and Maria, for all the critical thinking, easy communication and thoughtful guidance. After every meeting, I felt more energised and motivated to analyse the new questions and directions, which for me is the perfect sign of how I enjoyed our collaboration. Finally thank you Martine, for connecting me with other projects and stakeholders through the Delta Futures Lab. It greatly helped me to have this broad overview at the start of the project. I thank you as a committee member for your valuable and many comments in the last months of the project.

Jason Julius Wever
Delft, July 2022

Contents

Nomenclature

List of Abbrevations

DEM Digital Elevation Model

DM Decision Maker

DR Damage Ratio

EAD Expected Annual Damage

$EMDAT$ Emergency Events Database

$GFRT$ Global Flood Risk Tool

$HIS - SSM$ Hoogwater Informatie Systeem module Schade en Slachtoffers

IDF Intensity - Duration - Frequency

ND No Data

NWB Nationaal Wegenbestand - National Road Repository

PC Postal Code area

$PDNA$ Post Disaster Needs Assessment

$RHDHV$ Royal HaskoningDHV

SEJ Structured Expert Judgement

WTS Wet Tegemoetkoming Schade - Law Damage Compensation

List of Symbols

Cov Coverage Factor

Dam_{Achmea} Achmea reported residential damage

$DR_{compound}$ Damage Ratio caused by compound flooding and wind

DR_{flood} Damage Ratio caused by flooding

DR_{wind} Damage Ratio caused by wind forces

$F_{Ins,Pen}$ Insurance Penetration factor

$F_{Marketshare}$ Achmea residential housing insurance market share factor

$F_{Minordamage}$ Minor unreported damage factor

M Million

P P Significance value

$Q - 50$ 50th Quantile

List of Figures

List of Tables

I

Background & Approach

1 Introduction

Floods are one of the most impactful and frequent natural hazards globally. In the past century, more than 1000 floods occurred in Europe (Paprotny, Morales-Nápoles, et al., 2018). In the last decade, floods resulted in up to 200 billion dollars in the US (Munich Re., 2021). More recently in the summer of 2021, a river flood in North-Western Europe caused damage worth tens of billions of euro. Next to the economical damage, many lives, cultural heritage and nature are lost each year due to floods.

In the future, the impact of floods is only expected to increase due to climate change and socio-economic developments. An increase is expected from 5,3 to 20-40 Billion euro Expected Annual Damage (EAD) in Europe by 2050 (Alfieri et al., 2015). This effect is seen on a global scale, with Arnell and Gosling (2014) estimating even worse adverse effects occurring in Asia.

1.1. Background

The impact of flooding hazards is large and expected to increase. To mitigate this, significant investments in flood protection measures are required (UNDRR, 2019; World Bank, 2021). To economically optimize flood-mitigation investments, it is essential that the mitigated flood risk is estimated accurately. Generally, this is done with models that calculate flood risk as a function of hazard scenarios with their probability p, assets that are exposed to the flood scenario, and their vulnerability as in Equation 1.1, defined by Kron (2005).

$$Risk = \int_{p=0}^{p=1} Hazard(p) * Exposure(p) * Vulnerability(p)dp \tag{1.1}$$

Unfortunately, these flood models are generally not validated properly, with large uncertainties remaining in the hazard and vulnerability component (Arrighi et al., 2021; Merz et al., 2008; Molinari et al., 2019). *Validation* is defined as the process of ensuring that a model performs within a satisfactory range of accuracy and precision[1] for its intended use. *Verification* is a main validation tool and is defined as the comparing of observations with simulated values. Subsequently after verification, the validity of a model can be improved through *calibration*. *Calibration* is defined as the altering of a model's parameters to fit more accurately to the observations.

Inadequate flood model validation is mainly caused by a lack of observational damage data. Generally, flood damage is characterised across four axes as shown in Table 1.1. Direct damages occur immediately at the flooded location, whereas indirect effects can occur outside of the area or at a later point in time[2]. On the other axis, tangible damages can be priced, whereas intangible damages are by definition difficult to quantify (Jonkman et al., 2008).

[1] *Accuracy* describes how close the average of all estimates are to the real value. *Precision* describes the variability of the estimates, and how close the estimates are to each other.

[2] Indirect damage are for instance long-term macro-economical effects on regional development.

For this research, the Global Flood Risk Tool (GFRT) is applied (Bos et al., 2020). This model incorporates Equation 1.1 to estimate direct tangible flood damage as shown in bold in Table 1.1

Table 1.1: Classification of damage types. Bold types are integrated in the Global Flood Risk Tool. Altered from Jonkman et al. (2008)

	Direct	**Indirect**
Tangible	**Property damage or destruction**	Production loss outside of flooded area
	Loss of inventory	Temporary housing
	Infrastructure damage	Cost of traffic disruption
	Vehicles	
	Agricultural losses	
	Business disruption	
	Flood defence repairs	
	Clean up costs	
Intangible	Loss of life / injuries	Societal disruption
	Environmental damage	Reduced government trust
	Cultural loss	Psychological Trauma

1.2. Problem statement

All computational models should undergo a strenuous validation process to ensure that their results can be reliably used (Balci, 1994; Carley, 2017; Klügl, 2008). Validation of flood risk models is of particular importance, as their outcomes support decisions on large public investments, public policy, and risk premium decisions in the (re)insurance industry (Kron, 2005). However, the use of flood risk models remains limited by the inadequate validation of their vulnerability analysis and damage estimations (de Moel et al., 2015; Jongman et al., 2012; Merz et al., 2010; Molinari et al., 2019; Wing et al., 2020).

The main reason for the inadequate validation is the limited amount of empirical direct flood damage data. Despite many calls from the academic world to increase and standardize flood damage documentation, too few events are recorded (McBean et al., 1986; Merz et al., 2010; Wagenaar, 2020). Direct flood damage is frequently documented by (Re)insurance companies as damage claims, but not made publicly available for academia. When flood damage is publicly documented, it is generally done only for high return period events due to the large societal impact (Downton & Pielke, 2005). As direct flood damage data is highly specific for its geographic region, and flood characteristics data application to other flood events introduces additional transfer uncertainty (Smith, 1994). Data availability also decreases further with a smaller spatial modelling scale (de Moel et al., 2015). Damage observations may be available in aggregated form for a flood event, but observations are less available on a smaller scale (e.g. object-based observations).

Furthermore, it can be questioned how relevant historical direct flood damage observations are for future flood events, due to three reasons: First, the initial damage estimations can differ significantly from actual replacement costs. To give an example: For a 2002 German flood, estimations were altered from €22 billion to € 9 billion in the course of a few months, with actual costs amounting to €11,6 billion (Merz et al., 2010). Second, the accuracy of flood damage data highly depends on the spatial scale, with small local areas being extremely inaccurate (Downton & Pielke, 2005). Third, it has to be mentioned that even when historical data is well documented, it remains questionable how representative the data is for the current situation. Under the 'Building back better' principle, an area usually gets upgraded after a flood occurs (Mannakkara & Wilkinson, 2014; Municipality of Beira, 2019). Furthermore, asset values and characterizations, as well as flood defences change significantly over time (Merz et al., 2008).

Observations for flood model validation are either available for a single hazard event with questionable quality, or not available at all. This poses two problems: First, regarding validation of a probabilistic model with only single event observations, and second, validation when no data is available at all.

Probabilistic flood models attempt to represent all possible events that can occur with a coinciding probability. When only a single event is known, it is difficult to assess the accuracy of the used distribution. This is mainly related to the hazard analysis, where limited documentation of the flood extents and depth usually occurs for a single flood event. This single event documentation is then compared to a probabilistic estimate of a flood extent, that indicates the most likely flood extent for a certain return period. Limited research on validation with this disparity is done (Molinari et al., 2019).

When no local data is available at all, engineers need to look at other data sources that can be used to validate the model. Different options such as damage assessments of similar regions, insurance data, plausibility checks, and (Structured) expert judgement are known (Cooke & Goossens, 2008). However, no systematic approach exists as of yet.

The lack of sufficient empirical flood damage data complicates the verification and calibration procedures of flood-risk experts. Difficult verification results in low credibility of flood damage models, and poor calibration means more accurate damage predictions could be achieved. The credibility and accuracy of flood damage models is directly transferred to the credibility and accuracy of flood risk predictions, which need to be dependable because they support decisions on major industry- and public investments (de Moel & Aerts, 2010; Merz et al., 2004).

1.3. Research objective

The economics of flood mitigation measures are highly affected by the estimated flood risk and risk reduction. As the major part of flood risk uncertainty lies in damage predictions, a need for more reliable and accurate damage predictions is present. It is expected that improved validation procedures can significantly help this process. Therefore, this study aims to develop a guiding framework that can help engineers validate their flood damage models.

1.4. Research question

The defined problem statement and research objective lead to the following research (sub)questions:

How can direct economic damage predictions of a probabilistic flood risk model be validated, using plausibility and empirical data?

1. *What are the root causes for uncertainty and inaccuracy in flood risk model predictions, and what methods are available in literature and industry to validate those predictions?*
    - For probabilistic flood hazard models?
    - For flood vulnerability models?
    - For exposure data?
    - For aggregated direct economic damage data?
2. *What framework can be developed to validate direct economic damage estimates for a single flood event?*
    - How can data in economic damage reports be disaggregated for the verification of damage estimates?
    - How can Structured Expert Judgement be used as a validation tool?
    - How can private and governmental claim data be used to indicate model validity?
3. *Can the GFRT be calibrated further after applying the framework, to improve accuracy and precision?*
4. *What are the remaining found uncertainties in the damage estimate that the framework does not reduce?*

1.5. Approach

A conceptual drawing of the simulated flood damage verification is shown in Figure 1.1, starting with the described problem at the top. Probabilistic flood risk models can generally not be verified directly. Either no historical observations are available at all, or observations of a single event are available. In the second case, it may be that the probabilistic hazard map differs significantly from the unique flood event, as the probabilistic map is a combination of many possible flood events. Therefore, direct verification of a probabilistic hazard map may be difficult.

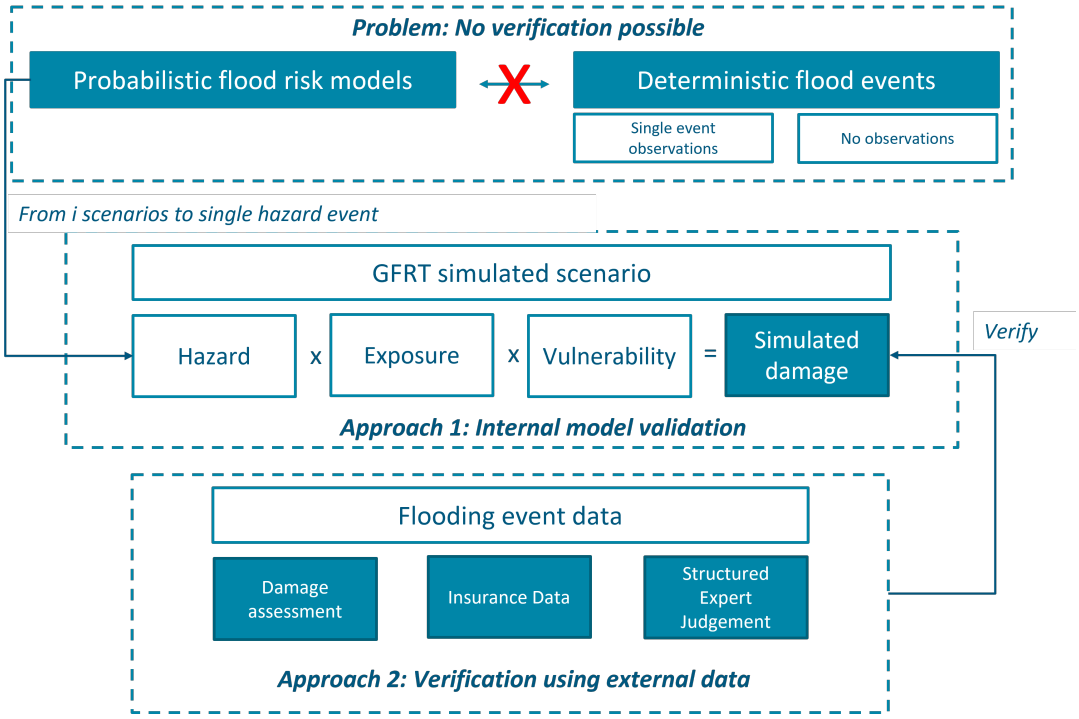


Figure 1.1: Conceptual model of the problem statement and proposed framework contents and their relationships.

Due to this difficulty, it is hypothesized that validation will need to be build upon the simulation of a single flood event. Subsequently, validation can be executed using two approaches. Approach 1 will assess internal model assumptions and results through plausibility. If the model has valid results for a single event, validity for a probabilistic flood assessment can also be argued. Subsequently, approach 2 will aim at damage estimate verification by using empirical data. In the case that no extra data is available, engineers will be guided on using expert judgement to generate new data.

Using this approach, it is expected that the validity of the GFRT can be showcased for the flood events. However, as additional uncertainty remains with probabilistic assessments of unobserved events, it can be doubted how ‘valid’ the final probabilistic model is.

1.6. Report structure

The study is approached using event-based validation of two case studies: The 2021 river floods in Limburg (The Netherlands) and the 2019 hurricane-induced coastal flood in Beira (Mozambique). For the Limburg case study, the focus is placed on internal assumptions and available data such as insurance- and governmental claims. Furthermore, an expert judgement experiment will be designed to create additional empirical data. For the Beira case, data from the widely spread Post Disaster Needs Assessment (PDNA) Methodology is used.

The report is structured in four sections:

- **Part 1: Background & approach** summarizes relevant literature and describes the applied methodology and case studies.
- **Part 2: Results - Limburg** describes the results of the first case study for Limburg. This consists of internal model plausibility, verification using empirical data, and a Structured Expert Judgement research.
- **Part 3: Results - Beira** describes the results of the second case study in Beira Mozambique.
- **Part 4: Discussion & conclusion** gives an academic reflection and highlights this paper's contribution to literature.

2

Literature review

This literature review gives a summarized description of previously executed research related to the validation of flood risk models. First, an overview of different types of floods is given in Section 2.1, followed by a description of how these hazards are modelled. Then, related uncertainty and validation procedures of flood risk models are discussed in Section 2.3 and Section 2.4. Finally in Section 2.5, Structured Expert Judgement (SEJ) background is described.

2.1. Types of flooding

Floods affect millions of people around the globe every year. However, not all floods are produced in the same way. In this section three significantly different types of floods are described: fluvial or river floods, coastal floods, and pluvial floods.

Fluvial flooding

Fluvial flooding is the overflowing of a river's banks that results in inundation of its surroundings, caused by upstream extreme precipitation, or snowfall. The flood progresses downstream as a wave, changing shape as it progresses.

The impact of river floods depends largely on the location and propagation of a flood wave. Low damages can occur when unprotected assets next to small rivers are exposed, as well as extreme damages when levee failure results in the inundation of large urban areas. The inundation depth is directly related to the terrain. In shallow areas, floods will propagate and rise slowly, whereas hilly terrain can result in quick flooding of canyons with high depths. When polders are flooded after a levee failure, they can be filled up entirely, leading to large inundation depth as the water can not flow out naturally. Damage from river floods is affected not only by inundation depth, but also by flow velocity, large debris that is transported through the water and pollution in the form of mud or chemicals (a major factor in the Elbe 2002 floods (Thieken et al., 2005)) Especially in steep areas this can be major contributors with knock-on effects as blockages and dynamic forces of debris. Next to economic damage, large amounts of river bank erosion and damage to nature can be caused by river floods.

High water levels from river floods can last multiple weeks, but are luckily also predictable with flood forecasting. As it takes time for a flood wave to transverse downstream, areas can be protected with temporary flood protection measures such as sandbags or flood walls, or evacuated to reduce damage and loss. Flood experience and forecasting can have major reductions of flood damage, as the Meuse floods of 1993 and 1995 showed with a reduction of 80% (Wind et al., 1999).

Coastal flooding

Coastal floods are caused by advancing seawater onto the coast. Depending on the global location, these floods can occur due to wind from seasonal storms like hurricanes, or due to earthquake-caused tsunamis. Storm surges are the most common, with the increase in seawater level highly dependent on the size and duration of the storm, as well as coastal geometry. Floods caused by tropical storms receive much public attention due to their large impact, with frequent occurrence located in the North-Western U.S.A., the East coast of Africa and South-East Asia. Climate change induced sea level rise is expected to increase the risk of coastal flooding significantly in the coming century (Parodi et al., 2020).

Coastal floods are infrequent events as mitigation structures like dunes, dykes and beach sand suppletions protect from low-impact floods. When they occur, massive amounts of damage are inflicted due to the size of the event and the high density of exposed asset value (Hinkel et al., 2014). Because storm surges are caused by strong winds, damage is inflicted both by inundation and strong winds, which is often difficult to differentiate in high-level damage reports. Furthermore, saline seawater can cause stronger damages than freshwater flooding (Wing et al., 2020).

Pluvial flooding

Pluvial flooding concerns floods caused directly by precipitation. This can occur in urban environments when the wastewater system is not able to discharge all surface water. This will result in inundation occurring on the streets that can subsequently affect nearby houses. Next to this, also flash floods can occur when extreme precipitation falls in a short durationIn mountainous areas, the steep gradient and amount of water can result in large flow velocities and debris, rapidly causing large damages in a matter of hours. Pluvial flooding generally lasts shorter than other types of flooding, especially if it concerns small, local evens where the water can disperse easily. As climate change is expected to increase the size and frequency of extreme precipitation events, the risk of pluvial flooding will increase jointly.

Pluvial flooding differs significantly from river and coastal flooding because it is characterised by more frequent events that are spatially distributed, and have lower total damage. An event can for instance cause damage by flooding cellars in a single neighbourhood, which is much lower in comparison to a dike breach caused by a river flood. Due to these large changes, this research will not focus on pluvial flooding. However, some contributions from pluvial flooding during river/coastal flooding events cannot be ruled out in compound events.

Compound events

The simultaneous occurrence of multiple flooding types is called a compound flooding event. Storms causing coastal floods can simultaneously cause extreme precipitation as well as wind damage. Similarly, also fluvial floods can coincide with a storm surge, leading to additional river flooding. The raised downstream water level will raise river water level through a backwater effect, thereby increasing flood intensity and duration.

Compound events are characterised by many difficulties. First, disaggregating damages of compound hazard events is obscure, especially if both hazards affect the same asset component within the same time frame. Second, additional damage caused by secondary effects such as the backwater effect increases complexity. Third, the return period for compound events is complicated due to a possible correlation.

2.2. Flood risk models

Flood damage is generally calculated as a function of a hazard scenario with a coinciding probability of occurrence, exposed asset value and vulnerability, as defined in Equation 2.1 by Kron (2005). As damage processes can vary significantly between locations, many different flood risk estimation tools have been developed and calibrated for a specific purpose. Examples for the Netherlands are developed by Jonkman et al. (2008), The Hoogwater Informatie Systeem-Schade en Slachtoffermodule (HIS-SSM) by Kok et al. (2005) and the Waterschadeschatter by STOWA (2013). More recent examples are FLORES by van Berchum et al. (2020), and a global scale model by Ward et al. (2020).

$$Risk = \int_{p=0}^{p=11} Hazard(p) * Exposure(p) * Vulnerability(p)dp \tag{2.1}$$

The Global Flood Risk Tool

The Global Flood Risk Tool, developed by the engineering firm Royal HaskoningDHV (RHDHV), is used for this study. This application can estimate the flood risk of a study area and subsequently compare flood protection measures underwritten by a business case. The GFRT calculates the flood risk as shown in Figure 2.1. The model takes as input a hazard analysis in the form of an inundation depth raster and return period[1], land-use map with exposure data, and vulnerability data in the form of depth-damage functions[2]. With this information, damage maps can be created showing the estimated damage to occur from hazard scenarios with different return periods. Aggregating the results of the different scenarios in a risk map of the area. This can show the spatial distribution of yearly expected damage in monetary units. Summing the risk spatially gives the Expected Annual Damage (EAD) for the region (Bos et al., 2020).

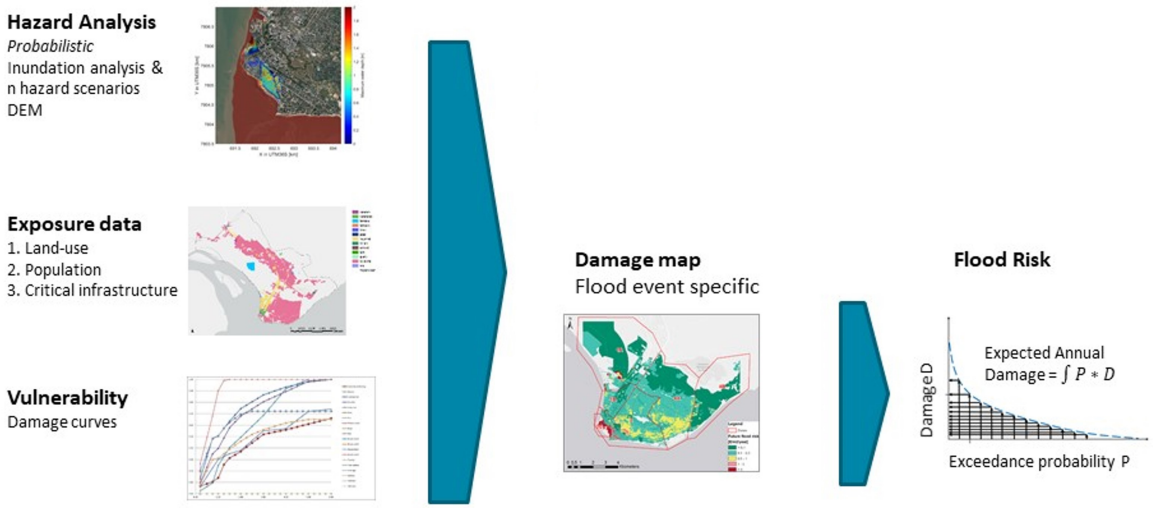


Figure 2.1: Workflow of the Global Flood Risk Tool (GFRT).

[1] This hazard analysis can be made externally using different approaches

[2] So only damage is only related to inundation depth in this model. Although other variables such as flow velocity, pollution and warning time are shown to be relevant, they are not used due to data limitations (Arrighi et al., 2021)

Input variables

Models structured as in Equation 2.1 have three main input parameters: the hazard analysis, exposed asset values and related vulnerability of those assets.

The **Hazard** analysis is used to illustrate different flooding scenarios in the case area. Different types of floods (pluvial, coastal, riverine) with different return periods can be simulated, resulting in a specific inundation reach over the region. The inundation raster can be made externally using different approaches, but is generally created using further input data such as the Digital Elevation Model (DEM), discharge measurements and roughness parameters. Also, model decisions such as the spatial and temporal scale and hydraulic type of the analysis may influence the used hazard analysis.

Exposure is defined as the assets that are affected by an occurring hazard. The exposure can be shown through land-use maps with different asset types and their corresponding value. Asset type classification can differ per region and goal, but generally include:

- *Residential*: Different classifications can be made to distinguish house types and asset values.
- *Industrial*: Assets are highly variable and dependent on asset function and therefore generally more difficult to transfer than damage models for other asset classes (Booysen et al., 1999).
- *Infrastructure*: These mostly include roads, railways and critical infrastructure facilities such as power plants and utilities.
- *Agriculture*: The loss of harvest can result in high damages which is generally crop-specific.

The total exposed value is asset-specific and can be defined in multiple ways. For direct damages, Merz et al. (2010) concludes preference towards using depreciated values over replacement values, following general economical theory. However, in practice it is also often found that replacement values are used instead (Arrighi et al., 2021; Jonkman et al., 2008; Middelmann-Fernandes, 2010). Finally, the exposure itself can be modelled on an object basis or by land use. Generally, data is not detailed enough to give information on the exact building footprint, which is why land use is preferred. Clarity on this decision is important, as the choice strongly affects the size of the maximum damage through the assumed density[3] (Huizinga et al., 2017).

Vulnerability is expressed in the damage ratio that an exposed asset will suffer from a hazard, which is generally explained in depth-damage functions. Depth-damage functions, shown in theFigure 2.1 bottom left[4], are a method to show the relation between inundation depth and asset damage.

Two main types of functions exist. First, *empirical depth-damage functions* are based on past flood events where the damage is assessed after a specific inundation depth was reached. Second, *synthetic depth-damage functions* are created by researchers based purely on theory and assumptions about a specific asset structure (e.g. from housing surveys) and its economic value. The functions usually project the normalized damage rate between 0 to 1, where 1 means all exposed economic value of the asset is lost. Total damage functions in monetary value also exist, but are less flexible (McBean et al., 1986).

As it is plausible that other factors than merely the inundation depth affect the damage to an asset, new types of damage models with a larger amount of variables are proposed. Amongst others, warning time, flow velocity, pollution and flood duration (especially for crops) were found to be relevant (Middelmann-Fernandes, 2010; A. K. Pistrika & Jonkman, 2010). In practice, the cost of finding the additional data is too high, so the industry remains using the classical depth-damage curves. Instead, these variables are implicitly present in the functions and therefore flood-, asset- and region-specific (Davis & Skaggs, 1992; Huizinga et al., 2017). The implicit uncertainty can be found by comparing multiple depth-damage curves for similar assets (but in different regions created for different floods). Research by Moel et al. (2014) and Jongman et al. (2012) found factors 4-8 and 4-11 between different functions.

[3] If the residential density of houses is assumed at 25%, the Landuse maximum damage needs to be reduced to 25% of the object-based maximum damage.

[4] Larger examples of depth-damage functions can be found later in the report, for instance Chapter 5 Figure 5.3

Spatial scale

Flood risk models are designed for either *Macro*, *Meso* or *Micro* spatial scale[5], which can significantly impact its data requirements and validation options. Hazard analyses are generally easily executable on a Macro scale (e.g. global), but can become very computationally intensive on a micro scale due to the increased influence of single objects. Exposure data is usually only available at meso/regional scale, which is rougher than most vulnerability data (Wünsch et al., 2009). Finally, empirical vulnerability data is generally always object-based and therefore suited for Micro scale. Upscaling to Meso- or Macro scale has proven difficult due to uncertainty in the number of objects present per m^2 (de Moel et al., 2015).

Model validation is also scale-dependent. At Macro scale, few datasets on complete flood extents are present, whereas damage estimates are usually done for an entire region or country. On the other hand at Micro scale, complete flood extents are well documented, whereas individual damage reports are less common and highly variable (de Moel et al., 2015).

2.3. Root causes of uncertainty in flood risk models

An essential difference between different research papers is that some research quantifies the uncertainty of *flood risk*, so including the probability of events (Apel et al., 2008; Parodi et al., 2020; Wagenaar et al., 2016). Other papers instead describe the uncertainty in *flood damage*, focusing on the parameters within a specific flood event irrespective of the extreme value statistics that is used for that flood hazard (Jongman et al., 2012; Merz et al., 2004). This is an important difference due to the large amount of uncertainty found within the extreme value statistics.

Total flood damage uncertainty

When looking at the overall impact of the three model components for flood damage uncertainty, many researchers find the vulnerability and hazard components to be the largest contributors, whereas the exposure parameter is found to be a subsidiary contributor (Merz et al., 2008; Moel et al., 2014; Parodi et al., 2020; Winter et al., 2017).

For vulnerability, de Moel et al. (2015) found asset value and the shape of the depth-damage curve to each be responsible for a factor of two, for likely fluctuations. Contrarily, the hazard component required a difference of more than 1m inundation depth to have the same effect. Also, Parodi et al. (2020) found a vulnerability contribution with a bandwidth of 0,25 - 4 times the mean damage, by using different possible damage curves for a coastal flooding. Similarly, Wagenaar et al. (2016) found uncertainties in the vulnerability of a factor 2 - 5 with increasing uncertainty for larger events. The different resolutions and types of exposure datasets were found to have a much smaller impact, giving bandwidths of around 10% around the mean estimated damage.

These findings give the model component ranking on uncertainty contribution as presented in Table 2.1. In the following sections, more in-depth descriptions of the causes are presented.

[5] Macro refers to country-wide studies, Meso scale assesses flood risk at provincial scale whereas micro concerns assessments on city-scale (de Moel et al., 2015)

Table 2.1: Identified sources of uncertainty in flood damage and flood risk assessment. The three main model components are ranked 1-3 from most contributing to least, with their major uncertainty contributors in bold. H, M, L & ND denotes High, Medium, Low and No Data respectively.

Parameter	Size	Comment	Reference
Vulnerability	1		
• **Curve shape**	High	Depends on expert assumptions, large natural variability	Moel et al. (2014)
• Size of case area	Med	Estimates for small areas are highly uncertain	de Moel et al. (2015)
• Not included variables	Med		Winter et al. (2017)
Hazard	2		
• **Digital elevation model**	High	Main factor in hazard analysis	Apel et al. (2008) and Parodi et al. (2020)
• Hydraulic model and parameters	Low	Uncertainty increases with smaller spatial scale	Romanowicz and Beven (2003)
• Spatial resolution	Low		Winter et al. (2017)
Exposure	3		
• **Value of asset**	High	Highly dependent on assuming replacement/ market value and what objects to include	Messner et al. (2007)
• **Object vs land-use based**	High		de Moel et al. (2015)
• Amount of classifications	ND		
• Age of data	ND		
• Spatial resolution	Low		(de Moel et al., 2015)
Risk			
• **Extreme value statistics**	High	Uncertainty increases with extremity of events	Merz et al. (2004)
• Failure of flood defences	ND	Large damage uncertainty around failure probability of flood defence depending on failure	Wagenaar et al. (2016)
• Continuous loss-probability function	ND	Uncertain effect of interpolation	Winter et al. (2017)

Uncertainty in hazard analysis

The flood extent and -depth were found to be the second-largest contributorsSection 2.3. Two main uncertainty sources were found: input data and the hydraulic model.

Regarding the input data, the impact of uncertainty in the DEM is found to be large in multiple papers (Apel et al., 2008). Parodi et al. (2020) described it as the second-largest impact after depth-damage functions, giving a bandwidth of 0,25 - 2,9 of the mean damage with a DEM bias of 3-7m. This is mostly attributed to the vertical accuracy of the DEM, which directly impacts the flood depth, but also the flood extent and the hydraulic model itself. The impact on the flood extent is dependent on the size of the vertical error and the present slope. For instance, if a slope of 0.5m vertical to 1m horizontal is assumed, a vertical error of 5m can result in a 10m horizontal error of the shoreline (Bales & Wagner, 2009). The uncertainty of this effect is dependent on the linearity of the slope, as well as water storage effects and the duration of the flood. That this effect is larger in low slope areas might be compensated by findings from Ludwig and Schneider (2006) that indicate that vertical errors tend to be smaller in those low sloping areas. Furthermore, large errors in the DEM could result in wrong flood dispersion patterns. A falsely elevated grid cell can for instance work as a dike and protect parts of the floodplain, or contrarily, an elevation that is uncaptured in the DEM can result in an overestimation of the flood extent. The effect of probabilities corresponding to the hydrological data is discussed later in Table 2.3.

Apel et al. (2008) the roughness parameters of the hydraulic model as another main error source. These parameters can have a significant impact on the flood dispersion, and are difficult to get 'right'. For instance, Romanowicz and Beven (2003) showed that multiple values of roughness parameters can result in coinciding simulated and observed flood extents, and indicated that floods of different sizes are likely to have different parameters. Especially at the small spatial scale of the GFRT, these epistemic uncertainties in parameters become significant. At this small scale, more complicated hydraulic modelling processes such as 2D floodplain modelling coupled with a 1D river channel model can be required to take into account water storage in the flood plain or polder.

Uncertainty in exposure data

Exposure datasets contain two types of information: Asset value and Asset location. Asset value is distinguished as an important contributor to uncertainty, together with damage models, whereas asset location is mostly indicated as a minor contributor.

For asset values, assumptions on which damage types to include and how to estimate these can give large differences between methods and reports. Different definitions of direct costs, such as the inclusion of dike repair and cleanup costs or loss of harvest can result in large differences. Finally, assuming market values (depreciated using economic theory), full replacement values or insurance payouts can give significant differences[6]. These assumptions should always be researched thoroughly and align with the goal of the model study. To indicate the size of the uncertainty: de Moel et al. (2015) found a factor of two in estimated damage by varying the asset value, which was the most sensitive parameter together with the depth-damage curves. Also Wagenaar et al. (2016) identified asset value as one of the main uncertainty parameters. Parodi et al. (2020) incorporated asset value uncertainty in the depth-damage curves, and found both to be the major contributor, varying the damage by a factor of four.

Asset location is found to be a minor contributor to overall flood damage uncertainty. Total uncertainty depends on asset presentation through points, building footprint or land-use maps. The simplest point representation was reported to underestimate the total damage by 30% (Winter et al., 2017). It can be imagined that in other cases an equally sized overestimation is also possible. When using land-use maps, no difference is made between buildings and other areas (e.g. a garden). This can bring further uncertainties by disaggregating the total value over the house and the garden, and spatial uncertainties related to the location of the flood depth and extents. These uncertainties are expected to even out when a large enough research area is taken (assuming the law of large numbers), but can be significant in small case areas. Finally, de Moel et al. (2015) found the effect of spatial resolution to be subsidiary, but referred to some cases where large differences did occur. Therefore this depends on the case event.

Two other relevant factors are the number of classifications in the exposure dataset and dataset age. In a perfect situation, a specific classification would be available for almost every different house. Unfortunately, usually exposure does not go further than differentiating between a small amount of residential (e.g. low-, mid-, and high-rise) buildings, industry and infrastructure. Subsequently, dataset age may be relevant to ensuring proper representation of reality in the model, especially in fast-developing areas.

Uncertainty in vulnerability analysis

Vulnerability is found to be the largest contributor to the uncertainty of flood damage estimates(de Moel et al., 2015; Parodi et al., 2020; Wagenaar et al., 2016; Winter et al., 2017). In some research, vulnerability analyses incorporate both asset value and depth-damage curve, however in this section we only focus on the relative depth-damage curve .

An often used exercise in literature is the comparison of multiple depth-damage models. The root cause of these differences is that each model is made using different assumptions and datasets, even when they are created for similar building types. After synthetic model creation by experts, a fitting procedure on limited available data can be followed in a validation attempt. Due to the large variability and very

[6] Highly dependent on the situation, but NAIC (2022) indicates for instance a payout of only 80% of the asset value and depreciated values that can be around 1/3th of replacement values.

limited amount of data, large differences can remain in the curve shapes (Winter et al., 2017).

This variability is present on both the flood side as well as in the modelled assets: for the flood side, vulnerability models only relate flood depth to the damage, whereas reality is more complex. Flow velocity and debris can significantly change structural damage ratios, especially in steep areas, which d-D models do not capture explicitly (Parodi et al., 2020). Also factors such as pollution, flood duration (especially with crops) and flood experience/preparation can have large impacts on flood damage. Wind et al. (1999) attributed a 35% decrease in reported flood damage between similar Meuse floods of 1995 and 1993 solely to flood experience and an increase in warning times.

On the asset side, large variability in residential houses regarding geometry, material and spatial location of valuables inside the houses is present. These models do not capture small but highly important differences such as sill height, possibly resulting in no inundation at all, presence of cellars, and the location of valuables on the ground floor, on heightened positions, or the first floor. Other asset types such as industrial buildings can have an even higher amount of variability due to their natural heterogeneity (Wagenaar et al., 2016).

Following the law of large numbers, the effect of variability is expected to decrease when a larger amount of assets is modelled. Downton and Pielke (2005) analysed damage estimates from different institutions and sizes and confirmed this expectation. For very large floods of more than $500 Million independent estimates disagreed by less than 40%. On the other hand, estimates of small events were found to be “extremely inaccurate”. Estimations for events causing less than $50 million in damages were often off by a factor of four, with over half being off by a factor of 1,5. This shows that the size of the area or hazard under study can have a significant impact on the uncertainty bandwidth of the estimation. Whenever possible, a larger study area should be preferred to reduce uncertainty.

Uncertainty in extreme value statistics

After calculating the flood damage corresponding to specific hazards, the flood risk can be defined. This is done in two stages: first, an exceedance probability is assigned to each hazard scenario, and second, a continuous loss-probability function is created. The integral of the loss-probability function gives the total risk or Expected Annual Damage (EAD) for the area.

The uncertainty of the first stage is rooted in the definition of the exceedance probability. Theoretically, exceedance probabilities are based upon observations that are continuously updated. New large flood events can significantly change the observed distribution[7], whereas the ‘real’ distribution can never be observed. Furthermore, the high return period events that are of interest for flood risk analyses, are generally never observed. Their probability of occurrence is defined by the extrapolation of fitted probabilistic functions. As the size of the event fluctuates with different types of functions and fitted parameters, also here uncertainty is present (Sayers et al., 2016).

The definition of an event with a return period is inherently problematic, due to different extrapolation assumptions and definition ambiguity. Many different definitions are available. For pluvial flooding, return period definition generally occurs through Intensity-Duration-Frequency (IDF) curves, which introduces ambiguity on the simulated spatial distribution[8] (Grahn & Nyberg, 2017). For fluvial flooding, two main methods exist based on measurements of peak river discharges: a spatially uniform discharge with a related return period can be set over the entire river section, or synthetic flood events can be simulated. The first approach is most commonly used and works well for small-scale local research. However, on a larger scale the flood extent will be overestimated as it is not realistic to assume that the event will occur in every location of the river. Then, modelling techniques incorporating the spatial dependence between river reaches and reach-specific return periods may be required (de Moel et al., 2015). For coastal flooding, options directly using sea water level statistics or indirect simulations using wind speed-direction statistics can be used.

[7] For instance, a certain peak discharge of x m^3/s that was measured once in the past 10 years at Lobith, can roughly be defined as a 1/10 year event. If in the year after this definition the same discharge occurs again, there are suddenly two observations of this size in the last 11 years, which would significantly change the deemed distribution.

[8] In models it is often assumed that a 1/x year precipitation occurs spatially uniform over the entire case area, which is unlikely in reality

Defining the return period of compound events is even more complex because the joint probability distribution and dependence need to be defined using advanced statistical modelling techniques (Paprotny, Vousdoukas, et al., 2018).

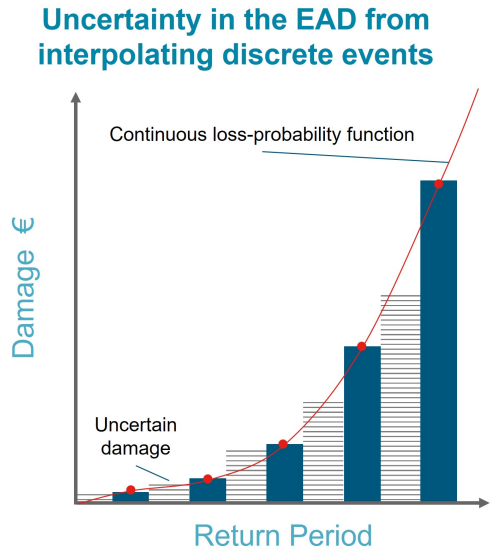


Figure 2.2: A continuous loss-probability function is created by interpolating a discrete set of estimations. The expected annual damage (EAD) is the integral of this function. Due to the unknown damages between the discrete events, extra uncertainty is added to the calculated risk.

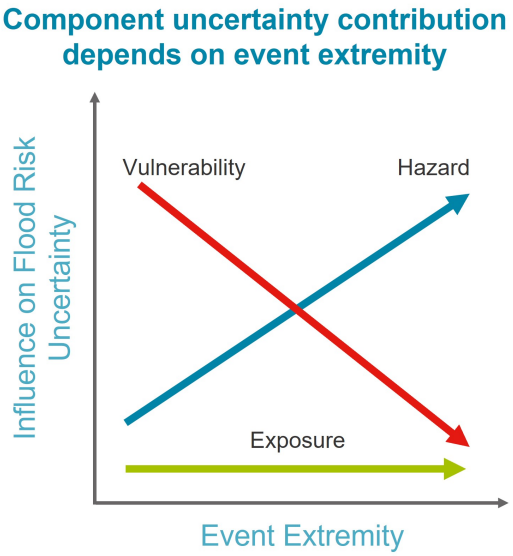


Figure 2.3: The conceptual change of influence that flood risk model aspects have on total flood risk uncertainty. The quantitative ratios between components are case-dependent and generally unknown.

The second stage of creating flood risk includes the creation of a continuous loss-probability function. After a discrete set of probabilistic hazard events and their corresponding probabilities of occurrence and expected damage is defined, interpolation is required to acquire the 'unknown' part of the function, as visualised in Figure 2.2. The exact amount of uncertainty is unknown but likely to be significant (Sayers et al., 2016).

Finally, it should also be noted that all relative uncertainty factors for flood risk are dependent on the extremity of the event. As visualised in Figure 2.3, Merz et al. (2004) found that the uncertainty of vulnerability aspect is generally largest at frequent events, whereas the uncertainty contribution of the hazard component increases with the extremity of the events. This is mostly due to the impact of the extreme value statistics. As an example: it is much more uncertain to indicate if an event has a 1/1000 or 1/5000 return period than it is to indicate if an event has a 1/25 or 1/125 return period, whereas the difference in risk is an equal factor five.

2.4. Flood risk model validation methods

To ensure that flood risk uncertainty is reduced enough for the modelling goal, validation methods are applied. In this section, literature methods to validate flood risk model components are described. As in previous chapters, the aspects of flood damage and flood risk are separated. Flood damage is an 'observable' quantity, and can therefore be verified and validated. 'Risk' is not a directly observable quantity and can not be verified accurately with a small number of observations. Through this definition, we focus on the validation of directly observable quantities concerning flood damage.

Flood damage model validation can be approached in two ways: first, a bottom-up approach can attempt to validate the data and parameters in each of the model's components. By ensuring that all separate model components are valid, it can be inferred that the total estimation is valid. Furthermore, this method gives insight into where main uncertainties lie and which parameters may require further calibration. Second, a top-down approach can be used by validating the final estimations of the model. This directly links the predictions to real-life observations and gives credibility to aggregated estimations (Sayers et al., 2016).

Using either of the approaches, the most relevant validation tools for this research are:

1. **Event based verification** can be used to verify model components and total damage estimations. By comparing model results with event observations, quantitative error measures can be created to clearly show a model's performance (Green & Stephenson, 1986). Two main issues accompany this method: Lack of enough historic event data, and remaining transfer uncertainty to new situations. Many validated parameters can change in unknown ways between events of different sizes meaning extrapolation is often impossible (Romanowicz & Beven, 2003)
2. **Sensitivity analysis** can be used to assess the effects of uncertain model parameters on the model outcomes. This can give insight into which parameters validation efforts should be focused (Moel et al., 2014). Different damage models that may all be considered valid for a case region can be used as different parameter inputs to show epistemic uncertainty (Arrighi et al., 2021). For extreme values in the hazard analysis, specific sensitivity methods should be applied as widely-used Sobol indices are found to be unsuitable (Nogal & Nogal, 2021).
3. **Model-model comparisons** are often executed in academia and provide uncertainty bandwidths (Trigg et al., 2016).
4. **Expert judgement** are specifically useful for situations where models cannot be used, are very expensive, or if no clear observations are available. For instance in a case with complex relations between variables that are difficult to model. Examples are flood extents that can be verified with local residents or experts with knowledge of the river basin, but also damage curves and aggregated total damage estimations can be discussed with experts.

Flood hazard

For a hazard analysis, the following aspects can be validated:

1. **Extreme river water levels** are often a starting point for the hazard analysis, used as input for hydrodynamic simulations that show the propagation of flood water throughout the basin. For some locations, studies predicting water level development under specific discharge time series have been done, which can be compared to model results to discover anomalies[9]. Furthermore, water level measurements can be compared to hind-casted models.
2. **Maximum flood extents** can be compared to theoretical flood maps, aerial observations and local observations. Theoretical hydraulic flood risk maps with return periods may give indications of where the flood extent is likely to be, but can locally be far off from reality. Aerial pictures from a plane or helicopter can give good indications of inundated areas. Satellite images can show similar extents, however these are historically less useful due to their temporal (and spatial) resolution. New industry developments enable the quick setting up of real-time flood monitoring using satellites which may be applicable to future events (ICEYE, 2022; Matgen et al., 2020).

[9] In Dutch results of these studies are called Betrekkingslijn

However, it remains unclear if the picture was taken at the peak flood, or if the floods still extended further. Finally, local residents who experienced can indicate flood extents on a small scale and thereby validate maps.

3. **Maximum inundation depths** are the hardest observable quantity to validate. Again, theoretical maps can be compared, but local observed information about damaged houses and high water marks is more valuable. When high water marks are created it remains important to closely document from where height is measured, as mixed measurements from the road, sidewalk or inside the house could give wrong indications (A. Pistrika et al., 2014)
4. **Digital elevation models** are of high importance for an accurate hazard analysis. Research by Ludwig and Schneider (2006) and Pakoksung and Takagi (2015) describes methods to analyse the validity of DEMs. Generally, DEM products are supplied with information on their accuracy, in means of vertical and horizontal bias and mean error. Although the size of an acceptable error is dependent on the modelling resolution, analysing this remains an easy first step to determine the quality of the DEM. If this information is not present, benchmark comparisons can be made with globally available virtual reference stations.

Exposure data

Exposure data consists of geospatial information regarding the locations of different asset types and their maximum monetary value. To validate this information, the following aspects can be considered:

1. Asset type
2. Asset location
3. Asset value
4. Total exposed asset size/Area

Assets are classified as different damage classes based on land use or building type. As these asset types are connected to depth-damage curves, their classification needs to be based on how they experience damage from inundation depth. No literature is available on a proper classification process. Rather, these decisions are based on the availability of depth-damage curves and data.

Asset location can be indicated by point locations, land-use plans, and building contours. A sensitivity analysis by de Moel et al. (2015) showed medium sensitivity to point locations, but relatively small sensitivity to commonly applied land-use features. Furthermore, the age of the exposure data can be important due to newly built assets in the area.

Asset value inhibits a large amount of variability stemming from assumptions on contents and value types (Section 2.3). Therefore, all assumptions on what is calculated and which aspects are included should be validated against the goal of the analysis. The following aspects need to be confirmed:

1. **Considered exposure type:** Market value versus replacement value versus insurance payout.
2. **Land-use assumptions:** for instance the density of residential houses.
3. **Which building content is included:** for instance house inventory and cars.

Vulnerability analyses

The depth-damage functions of the vulnerability analysis are a vital point of validation as Section 2.3 found it to be the largest contributor to uncertainty. At the core, it should be validated that these functions represent the average damage development in the corresponding land-use class. This consists of validating three main points:

1. **Initial damage** does not have to start at 0 m inundation. First, it can be common to protect houses locally using sills, completely elevated houses, or other measures. This will result in houses not being flooded, even though the model indicates a small inundation surrounding the houses. A second reason can be that buildings are composed in such a way, that no exposed asset value is located on the ground floor (wet-proofing)
2. **Damage development** differs per asset depending on their characteristics.
3. **Maximum damage** differs per asset. Some vulnerability curves do not reach a damage ratio of 1,0 because the assumption is made that not all asset value is damaged. Other curves may reach one, but have a reduced asset exposure. This complicates comparing curves, and points out the importance of aligning assumptions on maximum damage and maximum damage ratio.

There are three main methods to validate these aspects: comparison with historical observations, expert estimations, and model-model comparisons. When historical observations are available on a building level, statistical distributions can be made showing average damage factors at a certain inundation depth. These averages can then be made for building type classifications, and compared with used depth-damage curves. When these measurements are not available, expert housing surveys can be used. Appraisers can execute detailed surveys of housing types, documenting common valuables and their height above floor level. The aggregated averages can then be compared to the damage functions to validate their estimations (Smith, 1994).

In the case that the application of historical data and housing surveys is not possible, model-model comparisons can be used. Differences with other, more validated, damage models can be distinguished and related to regional differences. The main aspects to consider are differences in flood type characteristics and elements at risk.

Flood events have specific characteristics such as flow velocity, max inundation depth, arrival time, and pollution which induce differences in the resulting damage. Especially differences between damage curves for pluvial flood damage with other types can be significant due to different water pathways and flood duration (Grahn & Nyberg, 2017). Differences in the placement of elements at risk can also be significant between locations. First, cultural and temporal changes between validation studies can result in a different placement of inventory inside the houses. Second, differences in construction materials could alter the damage response from inundation depth. The impact of these changes should be reflected in the different depth-damage functions.

Aggregated economic damage estimates

Finally, also the top-down method can be used by validating total aggregated model results. Total direct economic damage from an event can be compared to the following sources:

1. Historical damage reports in the same region
2. Historical damage reports in similar regions
3. Insurance databases

First, damage amounts from old reports can be transformed to current day values and compared. However, differences in the reporting protocol, assets and flood need to be taken into account. Damage reports are created by appraisers who follow a damage report. The type of damage that is reported through this protocol can change over time, which can make direct comparisons impossible (Downton & Pielke, 2005).

Next, a different spatial distribution of exposed assets and their inventory can result in different damage responses for a similar inundation. In a report by Wind et al. (1999), two very similar flood volumes only two years apart resulted in approximately 80% less flood damage. This was probably caused due to flood preparation. This shows that even as floods are roughly the same, there can still be large variability between damage amounts. When flood events differ, even larger differences can occur. Major changed characteristics can be new flood mitigation measures, seasonal changes, pollution and a differently exposed asset distribution(Expertise Netwerk Waterveiligheid, 2021).

Second, when damage reports from similar regions are used, the same uncertainties as before apply, with added uncertainties due to the geographical translation. This is similar to transporting damage functions as described in the Vulnerability analysis section on the previous page.

Third, insurance databases can be a good information source due to the size of their dataset (Botzen & Bergh, 2008; Wing et al., 2020). Though insurers generally have very accurate object-based damage information, they may be hesitant to share them due to privacy and business interests. This restriction may be eased by aggregating the damage and then comparing it to modelled damage, by taking into account relevant factors such as insurance penetration and market share (Patankar & Patwardhan, 2015).

2.5. Structured Expert Judgement for validation

In the case that no observations as empirical data are available for the validation process, it is common to depend on the judgement of experts. From the 1990s onwards, a method to systematically use expert judgement was developed by Cooke and Goossens (2008) as *The Classical Model* at Delft University of Technology. More than a hundred studies have been executed since on a wide variety of topics where observations are limited, such as climate change, natural hazards, aviation and nuclear safety. More recent cross-validation indicates in Colson and Cooke (2017) indicates its usefulness and validity as a method.

The classical model works in the following way. A set of experts is asked to give estimations for the 5th, 50th, and 95th percentile of two types of variables: Calibration variables (Also called Seed variables) and Target variables. Calibration variables are numbers that cannot be known to the experts at the moment of the experiment, but are related to their expertise [10] . The answers should be not publicly known or become available relatively soon, so that experts cannot know the answers beforehand, but will have to estimate them similarly to the target variables. Target variables are the variables that the experiment aims to estimate[11]

Subsequently, the target variable will be estimated through a Decision Maker function (DM), a weighted combination of expert estimations of the target variable. The weight an expert gets assigned depends on the Informativeness and Calibration of his Calibration variable estimations. Informativeness relates to the spread of his estimation, where high informativeness is defined as a narrow estimation (The expert is very sure of his answer). Calibration relates to the accuracy of the expert. Is the realization of the Calibration variable captured within the expert's bandwidth?

Decision-makers can be created in two ways: performance weighting and manual equal weighting. With performance weighting, a function maximizes the sum of Informativeness multiplied by Calibration for all experts, by varying weights to the experts. Performance weighting follows the thought that some experts are better at estimating the variable and should be given more impact. This can even lead to the exclusion of some experts' estimations when their calibration and informativeness are too low.

Equal Weighting can be done, however Colson and Cooke (2017) showed that a performance weighted DM almost always outperforms an Equal Weighted DM. The DM is then used to create a 5th, 50th and 95th percentile estimation of the target variable.

Findings from Cooke and Goossens (2008) should be taken into account when designing a Structured Expert Judgement experiment:

- The validity of the target variable estimation depends highly on how closely related the Calibration variables are.
- The Amount of experts in an expert panel ranges between 4 and 77. As a certain amount of experts generally gets excluded in the DM, it can be assumed beneficial to have a large (at least 20) amount of experts.
- The number of Calibration variables ranged between 5 and 55. The amount can be related to the number of target variables.
- Expert panels have assessed between 7 and 17 target variables.

[10] Examples could be: "What is the 5th, 50th and 95th percentile of next week's daily maximum discharge at the Rhine at Lobith?" "What is the 5th, 50th and 95th percentile of 2021's maximum 24h precipitation sum measured at KNMI stations in the Netherlands.

[11] A Target question variable could be: "What will be the 5th, 50th, and 95th percentile of the mean sea level rise in the Netherlands by 2050?"

3

Methodological framework

The goal of this research is the development of a validation framework that flood-risk experts can follow to validate their flood damage estimations. The proposed framework is described in this section, and explains the steps that were followed to validate two case studies. The framework itself was shaped by the case results and created iteratively, by evaluating which approaches led to insightful results.

3.1. Framework development

Validation is the process of increasing confidence that a model outcome is accurate enough for its goal (Apel et al., 2009; Molinari et al., 2019; Sayers et al., 2016). The proposed validation process is visualised in Figure 3.1 and broadly consists of four phases to build confidence in the model.

The first phase consists of analysing the situation and clarifying the model's goals. This defines what the model should include and at what level of detail. The second phase focuses on the model's internal plausibility, by scrutinizing the input data as well as the output. Subsequently, the third phase verifies the model outcomes using empirical observations. Finally, the fourth phase combines all previous steps in a conclusion resulting in calibration and considerations for probabilistic flood risk studies.

3.2. Framework steps

The phases in Figure 3.1 are broken down into seven corresponding steps and expected results, which are thoroughly explained in this section.

Step 1: Situation assessment

The validation process was initiated with a situation assessment, in which the goal of the model and situational characteristics were clarified. Furthermore, flood damage is highly dependent on local asset- and flooding characteristics (Merz et al., 2010; Wagenaar et al., 2016; Wind et al., 1999). Therefore an overview of key asset- and flood characteristics that may influence flood damage was created.

Step 2: Input data assessment

For phase two, the focus was placed on the model plausibility. In step two, input data were assessed on two aspects: input data quality and data assumptions. By validating the quality of the input data, found root sources of uncertainty helped to identify errors later on in the process. For instance with hazard analyses, an amount of uncertainty frequently remains present in several locations, despite calibration with local high watermarks (Merz et al., 2004; Moel et al., 2014). Assessing the size of the uncertainty contributes to the process of building confidence in the model results.

Data assumptions largely involve the vulnerability functions and exposure data. Examples are assumptions on the house structure, flow velocity, flood duration and water salinity. A table overview of input data assumptions was made to identify errors between area characteristics and input data assumptions in step three.

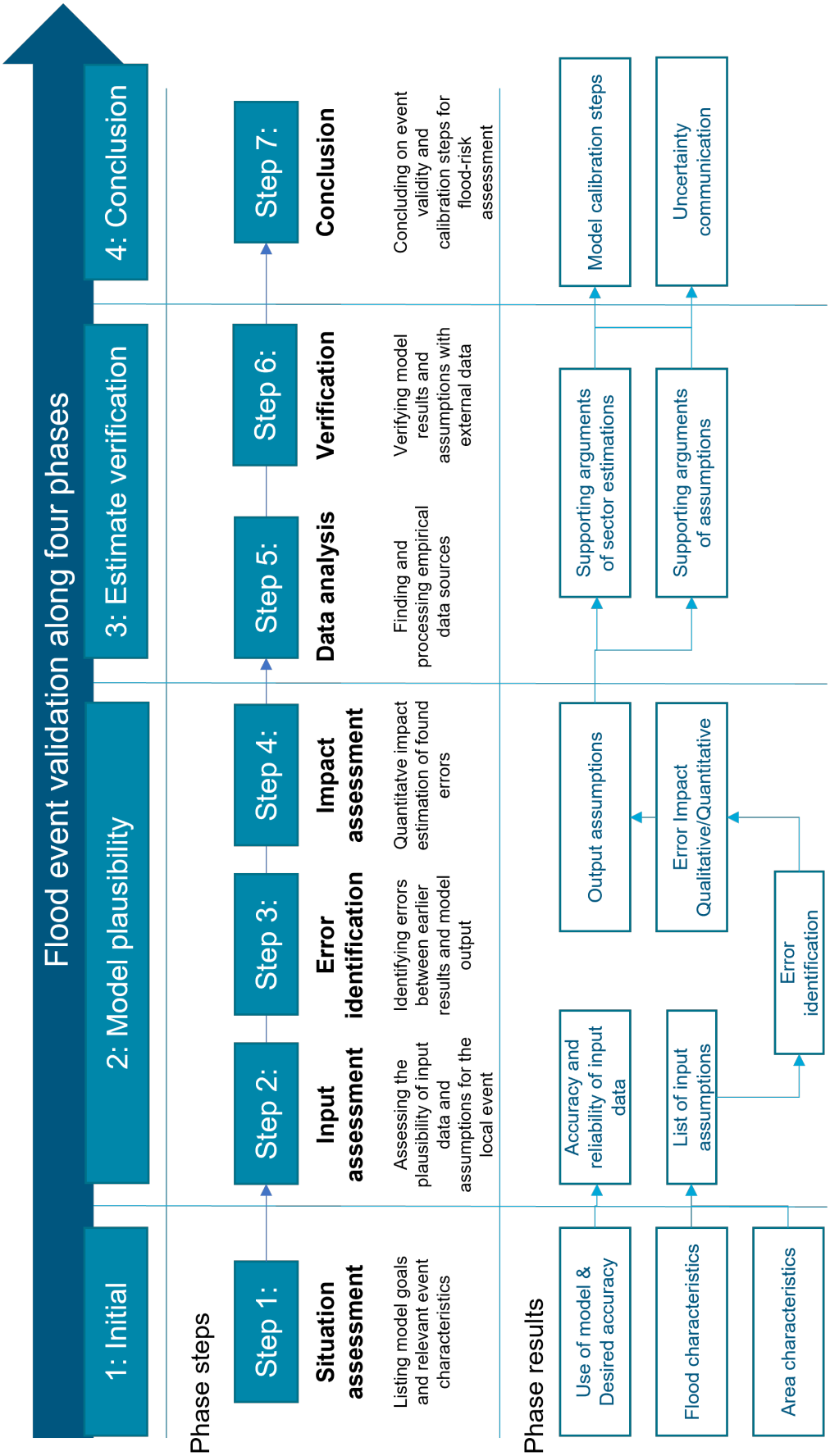


Figure 3.1: Visualisation of the proposed validation framework across four phases. These are further broken down into seven steps, with their results indicated at the bottom of the figure.

Step 3: Error identification

Step three involved identifying errors in the internal workings of the models. These errors could be caused by two sources: Differences in assumptions with reality and errors caused by schematisation of the area. Assumption errors were found by comparing area and flood characteristics from step 1, with model assumptions in step 2. Although it was difficult to quantify the exact error, pointing out the differences was vital to understanding the limitations of the model. Model output errors were found by analysing the output of the GFRT in Python and GIS. In Python, parameters such as the inundation depth distribution were analysed, and a comparison was made with economic data and other damage reports. In GIS, visual inspection of damaged areas was done to check plausibility as well as comparisons with spatial data sets.

Step 4: Error impact estimation

After the errors were identified, an estimation of the impact of the error on the total damage estimation was calculated. The preferred outcome of this step is a quantitative estimation in monetary value, when possible. First, the found errors are classified as one out of four different error types defined below:

1. **Overestimation:** The characteristic resulted in the overestimation of direct economic damage for this class.
2. **Underestimation:** The characteristic resulted in the underestimation of direct economic damage for this class.
3. **Uncertainty:** The characteristic may differ significantly from reality, however it is not possible to conclude if this results in an over- or underestimation. For instance when no information on the inundation depth is available, and a certain depth is assumed.
4. **Not applicable:** The characteristic could result in an over or underestimation, however in this specific situation it does not affect the accuracy. An example of this is to scrutinize the flooding time concerning sowing and harvest time of crops. If the flood occurs after harvest, no crop-loss damage will occur, significantly altering agricultural damage estimates. This category explicitly states the validity of an important variable, and thereby significantly improves confidence in the model.

Subsequently errors were estimated quantitatively. The estimation methods could be clustered in three approaches:

1. **Comprehensive approach:** Aspects that affect the prediction of the entire damage class. For instance, an inconsistency in the maximum damage value, flood preparation or the occurrence of a flood after harvest time. The error is estimated by multiplying the initial damage estimation with a correction factor[1].
2. **Linear approach:** Some features are more suitable to be modelled linearly than in the GFRT's raster approach. For instance roads and railways exhibit this behaviour if they are not properly aligned to the raster orientation and resolution. Here, a sub-model that calculates damage per length unit ($/m$) was used to estimate road damage more accurately.
3. **Individual approach:** Some classes were found to have highly heterogeneous damage patterns, that require individual validation. Examples are elevated highways or unique assets like damaged sluices or power plants. Due to their heterogeneity, the exact damage to these types of assets often remained highly uncertain.

By adding the impact of all over- and underestimations, new damage estimations were made for each sector and the total damage estimate. This resulted in a new 'best estimate' of the modelled damage. Furthermore, an uncertainty bandwidth around this best estimate was created using the summed errors classified as uncertainty. Finally, the main assumptions that substantiated the error quantification and new damage estimate were distinctly presented. These assumptions play a key role in exhibiting the alignment of the damage estimation with the model goal. Furthermore, these assumptions require-substantiating proof from observations, which is found in the next step.

[1] A correction factor for the error. For instance, if the maximum damage is deemed to be 20% less than the input parameter, the estimate can simply be multiplied by a factor of 0,8.

Step 5: Analysis of empirical data

After the model plausibility phase, the next step was to substantiate key assumptions and damage estimates with real-world observations. As these observations are generally lacking in flood damage studies, the first step was to identify any useful data sources. Subsequently, the found data was prepared and transformed to be comparable with the modelled damage. The applied damage sources and transformation processes are described in Table 3.1

Table 3.1: Applied data sources and their corresponding data-transformation processes. Reference to application

Data type	Process	Reference
Expert data	Creation of a Structured Expert Judgement and elicitating data	Chapter 7
Insurance data	Transformation of a single insurer's residential damage to total residential damage.	Chapter 6
Governmental claims data	Regression and pattern analysis of data to analyse the coverage factor and discover outliers	Chapter 6
Event damage data	disaggregation of compound hazard damage and a transformation on spatial scale.	Chapter 11
Global damage databases (EMDAT)	GIS analysis to analyse affected population and GDP with global damage data	Appendix C Appendix D

Step 6: Estimate verification using observations

After the data was prepared and transformed it could be compared to GFRT estimates. As data availability was the main bottleneck, damage verification for each sector differed significantly. Verification was done either on an aggregated level, by comparing total damage estimates for a sector, or at an individual level, by comparing the estimate of average damage for a single asset. For improved validation applying both methods was preferred, however this was not always possible. After exploring differences between GFRT estimates and the observations, a conclusion was presented on possible under- or overestimating factors and the impact this had on the total estimate.

Step 7: Conclusion on model validity

Finally, a total overview of all validated damage quantification, found errors that may be calibrated, and key assumptions is given. Furthermore, this overview indicates the amount of uncertainty expected to remain with the event-based damage estimation.

The conclusion then connects the event validation to the overarching flood-risk study. First, the decision needs to be taken which assets can be calibrated further to improve accuracy, without overfitting the model to the case event. Second, it can be considered if the remaining uncertainty is small enough for the model's purpose, or if further validation should be attempted. Third, major uncertainties and assumptions that are important for decision-makers need to be selected and communicated transparently, so a suitable flood-mitigation strategy can be selected.

Case studies

Models like the Global Flood Risk Tool are applied to various cases, differing in terms of data availability and quality, socio-economic circumstances and flooding types. To mirror this, two case studies were selected to be different in these characteristics. The Limburg river floods of 2021 represent a relatively data-rich environment with European asset characteristics. On the other hand, the Beira case has much less data and is further complicated by compound event damage of strong winds and pluvial flooding. The chapters describing the result of each applied framework step are shown in Table 3.2.

Table 3.2: Overview of applied framework steps to report chapters

Phases		1: Initial	2: Model plausibility	3: Estimate verification	4: Conclusion
Chapter	*Limburg*	Chapter 4	Chapter 5	Chapter 6, Chapter 7	Chapter 8
	Beira	Chapter 9	Chapter 10	Chapter 11, Appendix C	Chapter 12

II

Results - Limburg river flooding

4

Limburg - Case description

The first case study of the report concerns the 2021 river floods in the province of Limburg in the Netherlands. This chapter describes step 1 of the validation framework, by assessing the relevant background of the area and characteristics of the flooding event. Third, applied data are summarized.

Limburg background

Limburg is a province in the southeast of the Netherlands, bordering to Belgium and Germany. The province has a surface area of around 2.200 km^2 with approximately 1,1 million inhabitants.

The main river of interest is the Meuse (Maas), which sources in northern France, runs through Belgium and enters the Netherlands at Eijsden. Subsequently, the Meuse flows to the North of Limburg past Maastricht, Roermond and Venlo and exits the province at the head of Limburg. The upstream part between Eijsden and Maasbracht is often referenced as the border meuse (Grensmaas), whereas downstream of Maasbracht is called the sand-meuse (Zandmaas). The Meuse has multiple tributaries, which may contribute to a more extreme water level. The most notable tributaries in the Netherlands are the Geul and the Roer, which are characterised by a much stronger gradient (Bisschop et al., 2015).

Like most rivers in the Netherlands, the Meuse was heavily trained throughout history. The river has been straightened and has seven weirs to ease shipping (Barneveld et al., 2021). Furthermore, primary dikes surround the Meuse and were recently strengthened due to the introduction of new norms in 2017 (Rijkswaterstaat, 2021a). Contrarily, the smaller tributaries exhibit a more natural flow, and only have secondary or local dike protections. River flooding plays an important role in the history of the region, with large floods occurring along the Meuse in 1993 and 1995, resulting in 201 and 126 million euro respectively (2021 adjusted)[1] (Wind et al., 1999).

Limburg flooding event

The flooding event was initiated on July 12th with extreme precipitation in the drainage area of the Meuse lasting for at least a week. The precipitation event was characterised as extremely rare, especially for a summer event, with return periods of precipitation as well as discharge far exceeding 1:100 years (Expertise Netwerk Waterveiligheid, 2021). The runoff resulted in extreme water levels in the Meuse as well as in tributaries of the Geul, Roer, Gulp and Eyserbeek.

Although all primary dikes along the Meuse resisted the water, thereby minimizing damage, large damages occurred along the tributaries. Especially along the Geul and Roer many settlements were hit, notably Valkenburg, Schin op Geul and Bunde-Geulle. An estimated 2.500 houses and 600 companies were affected directly, with another 50.000 people evacuated preventively. Damage was concluded to be significantly higher than previous floods in 1993 and 1995, totalling approximately 350-600 million euro of direct damages, including business losses (Expertise Netwerk Waterveiligheid, 2021).

[1] Damage numbers in 1993 and 1995 are 254 and 165 million guilders, transformed using an average inflation of 2% and conversion of 1 euro = 2,2 guilder

Limburg main data

Multiple open-source datasets were used to model the fluvial flooding and are shown below in Table 4.1.

Table 4.1: Main Datasets used for the Limburg 2021 case study. The top three datasets were used as input for the GFRT. Other datasets were used either in GIS analysis or as verification data.

Name	Description	Resolution	Author
Inundation raster	The inundation map and flood extents were created to map the event and were provided by Deltares.	5x5m	Expertise Netwerk Waterveiligheid (2021)
Land use map	A combination of multiple open source datasets was used: Bag, Top10NL, CBS soil usage.	Polygon	CBS (2015) and Kadaster (2021a, 2021b)
Vulnerability	Internally created damage curves by RHDHV, similar to HIS-SSM curves	-	De Bruijn et al. (2015)
Fact finding	Report of a task force that mapped the event and its consequences rapidly after the occurrence.	-	Expertise Netwerk Waterveiligheid (2021)
BAG	National data on adres location and footprint.	Polygon	Kadaster (2021a)
NWB	A national repository on the location of publicly maintained roads (Nationaal Wegen Bestand).	Polylines	Rijkswaterstaat (2021b)
WTS data	Spatial data on governmental damage support applications for 6 different damage categories.	Postal Code 6[2]	RVO (2022)

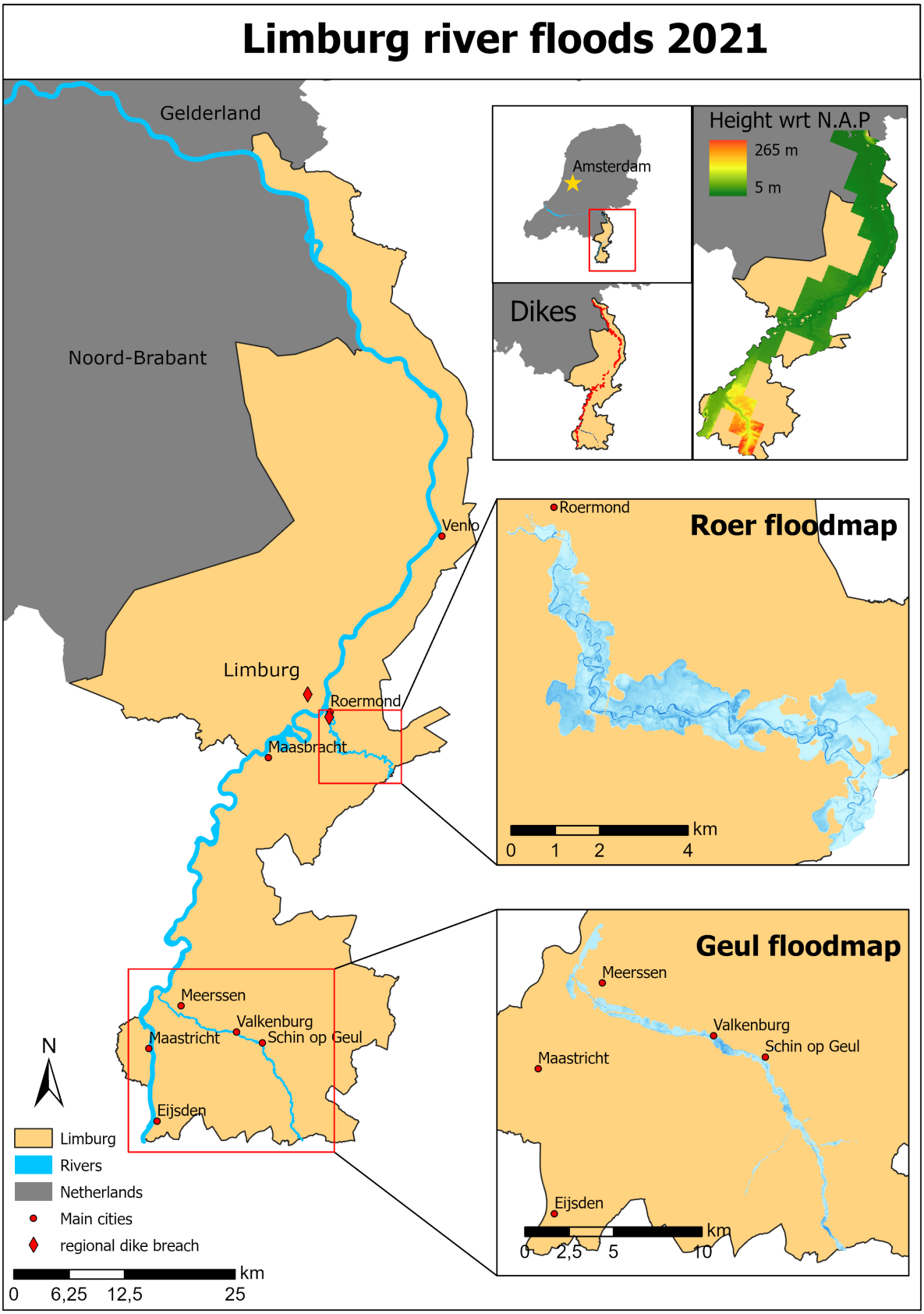


Figure 4.1: Overview map of Limburg flooding situation in July 2021

5

Limburg - Model plausibility

In this chapter, the results of applying framework phase 2 to the Limburg case study are discussed (See Chapter 3). For framework step 2, the results of validating the model input data are discussed in Section 5.1. Step 3 and step 4 are discussed in Section 5.2, where an in-depth analysis of errors in the model output is presented. Finally in Section 5.3, a total conclusion is given.

5.1. Model input validation

Here, input data for the three model components: Hazard, Exposure and Vulnerability was analysed.

Hazard analysis

For the hazard analysis, flood map validation was executed on three aspects: flood extent, inundation depth and water levels. Most of this validation was reported by Slager (2021), when the inundation map of the event was created for the Expertise Netwerk Waterveiligheid (2021).

For the flood extent, the area was validated using aerial pictures taken during the event. It was judged if the pictures were taken at the flood peak or after by analysing surroundings for signs of retreating water[1]. However, it was not possible to judge if a picture was taken before the peak, as there were no signs that could indicate where water would flow next. This may have resulted in an underestimation of the flood extent. Subsequently, interviews with residents were gathered and compared with flood extents to validate if water had reached a location, especially at high-damage locations. Particularly for Valkenburg, an employee had gone around by bike to validate the flood extent.

The depth distribution of the used inundation raster was analysed in ArcGIS, where it was noted that large areas were set to exactly 0,50m (See Figure 5.1). In an interview with Slager (2021) it was mentioned that these are areas where the hydraulic model did not predict a flood depth, but flood extent reports (from residents or aerial pictures) did show inundation. Therefore, it was chosen to manually insert a relatively low inundation depth of 0,50m into the inundation map. The impact of this decision was analysed and is significant for the Geul and the Roer area, where approximately 47% and 17% of the flooded surface area was manually filled. As shown in Figure 5.1, the 0,50m raster alone resulted in 58 million euro of damage for the Geul area, and 6 million euro around the Roer. The residential sector around the Geul accounts for more than half the damage. For the Roer, more than half of the damage was inflicted on agriculture and residential classes combined.

For the water level validation, the inundation map was summed with elevation data to acquire the water level of the model. This was compared with data from Rijkswaterstaat (2021c) and Waterschap Limburg (2021). Water levels for the Maas aligned well with maximum observations (Appendix A Figure A.2). As water level observations for the Geul and Roer were incomplete, no conclusion was drawn[2].

[1] For instance left-behind debris or toppled trashcans

[2] Ongoing research at Delft University of Technology - Delta Futures Lab is reconstructing the water levels.

Table 5.1: Damage caused by the manually insert 0,5m inundation grid cells for the Geul and Roer.

Geul	
Class	**Damage**
Residential house	25,44 M€
Communal building	7,68 M€
Residential area	5,79 M€
Agriculture	3,97 M€
Other	15,38 M€
Total	58,33 M€

Roer	
Class	**Damage**
Agriculture	1,58 M€
Residential building	1,07 M€
Residential area	0,82 M€
Other	2,53 M€
Total	6,00 M€

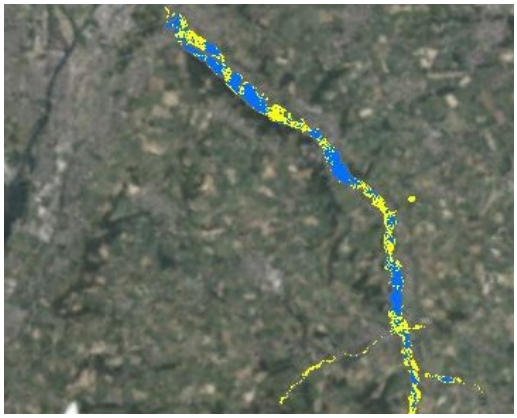


Figure 5.1: Visualisation of the manually inserted 0,5m inundation cells at the middle Geul section. Yellow indicates grid cells with 0,5m whereas blue is the underlain flood extent map.

Exposure dataset

The exposure dataset consists of two aspects: the land use map and the maximum exposed asset value. For the land use map, three aspects were validated: First, the number of land use classes was considered to be detailed enough for the modelling purpose, although this remains a subjective decision. A total of 57 land use classes were used on the map. Second, the geometry and placement of the land use map were compared visually to aerial pictures, and deemed accurate. This step was repeated after uploading in the GFRT, as small displacements could have occurred due to small differences in the coordinate systems. Finally the geometry of the classes was compared with each other, which showed a double counting of overlapping road classes. This was solved by removing one of the road classes.

After the land use map, the exposed asset value was validated. Each class in the land use map was connected to an exposed asset value $/m^2$. Before applying this data to the model it was validated through a model-model comparison with HIS-SSM. This process is visualised in Figure 5.2.

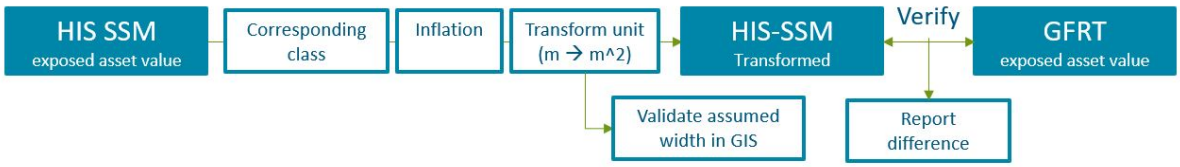


Figure 5.2: Flowchart of validation steps for exposed asset value

As a first result, the differences from the transformed HIS-SSM values with the GFRT exposed asset values were analysed (Shown in Figure A.4). Here, two outliers were found: 'Communal buildings' and 'Hotels' both deviated around 20% from the transformed HIS-SSM class 'Low rise'. Furthermore, the 'Residential Building' maximum damage was lower than HIS-SSM. It assumed only house-structure maximum damage, and excluded the house-inventory aspect.

As a second result, a list of assumed widths used in the unit transformation process of infrastructure was created. These assumed widths should align with the widths of the damaged roads in the GIS model to have an equal exposed asset value. These widths were later compared in Section 5.2

Vulnerability analysis

After the exposure dataset, the used depth-damage curves were validated through a comparison with HIS-SSM vulnerability curves, as can be seen in Figure 5.3. As there were more classes and corresponding vulnerability curves in the GFRT than in HIS-SSM, multiple GFRT curves were compared to a single HIS-SSM curve.

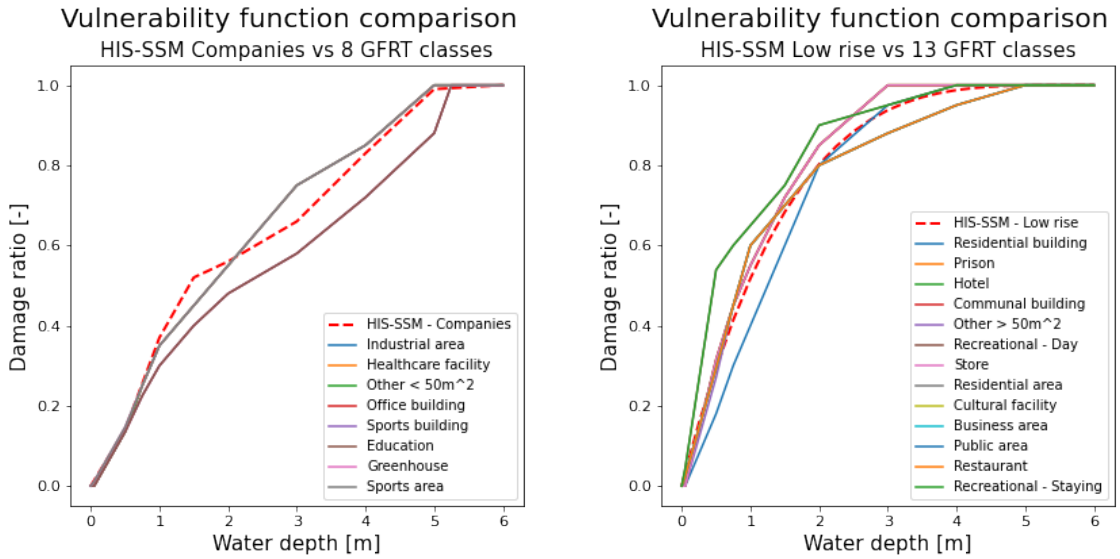


Figure 5.3: Validation method of vulnerability analysis. The red-dashed HIS-SSM line is compared to multiple GFRT vulnerability curves. The full analysis can be found in the Appendix A

13 classes coincided with the residential HIS-SSM class 'low-rise'As can be seen in Figure 5.3 many curves deviated slightly, but overall no significant difference was concluded. For 'companies' eight GFRT classes followed two unique depth-damage functions. One function showed a slightly more vulnerable damage process, and the other a less vulnerable pattern. The final 15 damage classes are similar to HIS-SSM 'agriculture' or the 'transport' vulnerability[3].

Although high similarity with the generally applied vulnerability model in the Netherlands may indicate validity for the area, uncertainty remains on two points: First, HIS-SSM has never been validated for this specific region. Second HIS-SSM was based on data from a large-scale flood with high inundation depths. Therefore, the vulnerability curve may deviate significantly from smaller river floods such as our case study (Slager & Wagenaar, 2017).

Conclusions and limitations for input validation

From the input validation, a better understanding of the model's components was achieved, thereby reducing the 'black-box' application and removing errors. This is a key step in improving the confidence of model results (Sayers et al., 2016). The following conclusions and limitations were found:

1. The inundation depth is mostly unknown and set to 0,5m around the Geul (47% surface area), and to a lesser extent at the Roer (17%). Furthermore, a minimum amount of uncertainty remains in the flood extent as this is validated well with local- and aerial observations.
2. Exposed asset value coincides with HIS-SSM values for most land use classes. Special attention should be given to communal buildings, hotels, and residential houses as these deviate from HIS-SSM. Furthermore the actual width of infrastructure elements can be checked against the width assumed in the maximum damage estimation.
3. The depth-damage curves are mostly similar to vulnerability curves of the generally applied HIS-SSM. However, uncertainty remains as the performance of the vulnerability curves for this specific area (and flood) has never been shown.

[3]The comparison can be seen in Figure A.3

5.2. Model output validation

The total damage for the flood event in the Netherlands was estimated at 526,2 million euro by the GFRT. As shown in Figure 5.4, the main damage driving land use classes are agriculture, residential, minor roads and public area.

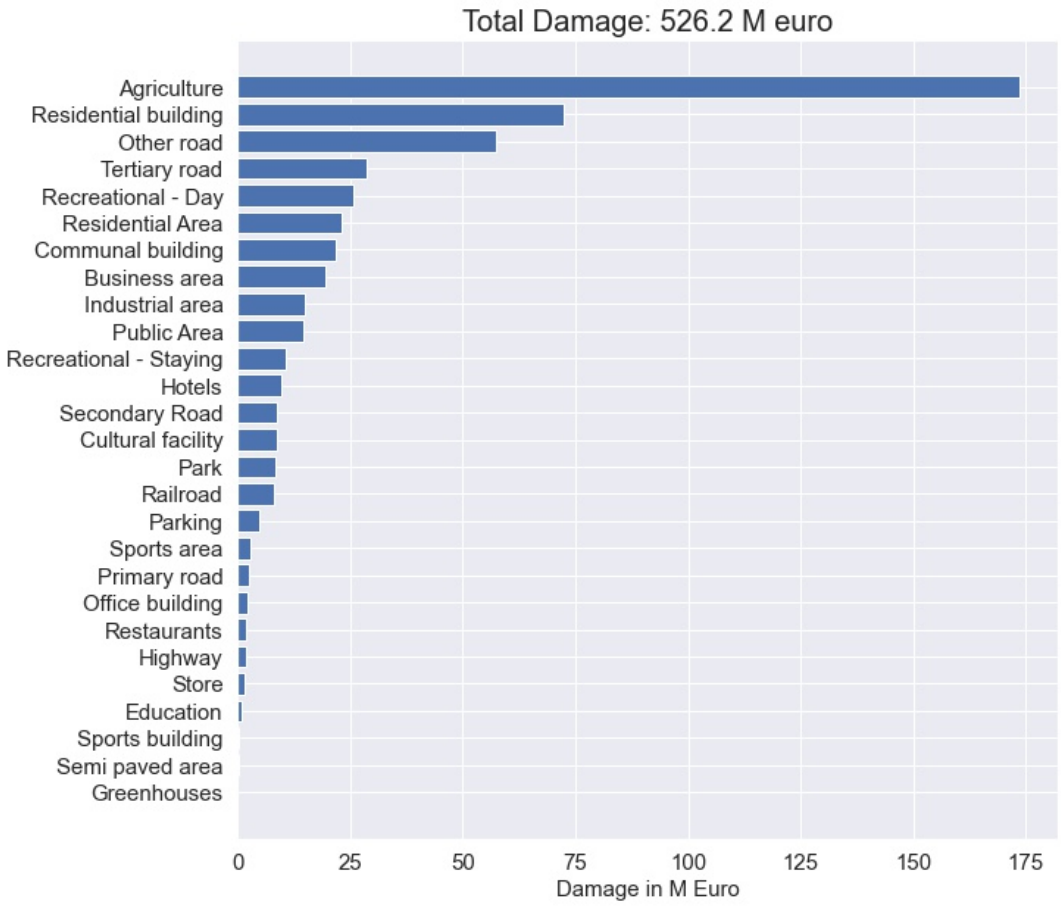


Figure 5.4: Overview of estimated damage in the Limburg case studies. Only classifications where damage is estimated are shown.

From a comparison with a damage estimation of the Expertise Netwerk Waterveiligheid (2021) the following initial conclusions were taken:

1. The GFRT total damage estimation is 96 to 101 M€ higher than the fact finding estimation, especially along the Meuse (Appendix A Figure A.5).
2. The GFRT agricultural damage is more than 135 M€ higher than the fact finding's estimate (Appendix A Figure A.8).
3. The GFRT estimate for road infrastructure is 40 M€ higher along the Meuse, but 20 M€ and 2 M€ lower alongside the Geul and Roer respectively (Appendix A Figure A.9).
4. The GFRT residential damage is higher for all areas. For the Geul and Roer, this may be caused due to the inclusion of residential area (approximately 23 M€ in total), whereas for the Meuse, residential housing is already 7,5 M€ higher than the fact finding (Appendix A Figure A.10).

Following the initial analysis, the estimations for different sectors were scrutinized in order of magnitude. The sectors were analysed in the following steps: First, an overview of the total estimated damage, the observed assets in the land-use class, and the expected damage processes is given. Subsequently, multiple factors leading to an estimation error are described based on their assumptions and effects. Finally, a conclusion on the overall validity of the estimated damage for this sector is given.

Agriculture

Agricultural damage can be classified as damage to crops or damage to fields (Pivot & Martin, 2002). Crop damage is dependent on the type of crops, inundation duration and time[4], whereas field damage depends on erosion, pollution and cleanup costs (Messner et al., 2007; Winteraeken & Spaan, 2010). Furthermore, livestock casualties and evacuation costs can be relevant (Brémond & Grelot, 2013).

The calculated damage for the agriculture classification amounted to 174 million euro. Visual inspection of the classified agricultural surface showed diverse natural areas such as crop fields, flood plains and grass fields. Indicated in Table 5.2, four significant factors were identified.

Table 5.2: Errors in the GFRT estimation of the agriculture land-use class

Overestimating factors	**Impact**
1: 60% of registered affected agriculture contains grass for which damage is unlikely.	90M€ overestimation
2: 14% of damaged classified surface is unregistered in the BRP.	24M€ overestimation
Uncertain factors	**Impact**
1: Maximum damage per m^2 seems overestimated depending on the amount of agricultural pollution.	0 – 45M€ uncertainty
Non-affecting factors	**Impact**
1: Temporal effect of flood occurrence	-

Overestimation 1: Registered cropland

Two reasons led to the hypothesis that GFRT-estimated agricultural damage was overestimated. First, a comparison with flood reports of 1993, 1995 and the Expertise Netwerk Waterveiligheid (2021) indicated agricultural damages in the range of 20 million euros (Wind et al., 1999). Second, a visual inspection in ArcGIS revealed large areas where agricultural damage seemed questionable.

To analyse this hypothesis, a plot-level crop dataset (BRP) was used RVO (2021a). Analysis of the BRP dataset indicated that on average 60% of the damaged area was classified as grassland, with up to 77% for the Geul. As these areas were inundated for less than a week limited damage and loss of production for the grassland is plausible[5].

Table 5.3: Spatial analysis of BRP plots present in the damaged area

Crop type	**Area**	**Damage**
Grass	60%	No
Mais	15%	Yes
Potatoes	5%	Yes
Sugarbeets	5%	Yes
Other	15%	Yes
Total damaged	40%	

$$\begin{aligned} &\text{86\% of agricultural land is BRP registered:} \\ &174M€ * 0.86 = 150M€ \\ &\\ &\text{Of which 60\% is undamaged grassland:} \\ &174M€ * 0.86 * 0.6 = 90M€ \end{aligned} \tag{5.1}$$

Under the assumption that no financial damage is caused by the flooding of grassland, 60% of the estimated damage in the registered BRP parcels is not incurred. 86% of all agricultural land is registered, which leads to a 90 million euro overestimation (Equation 5.1). This considers that no financial loss is incurred from the flooding of these areas due to cleanup of pollution, buying replacement feed for cattle, evacuation or loss of cattle that were grazing in the area at the time of inundation (Hess & Morris, 1988).

[4] concerning harvest and seeding moments

[5] The floods started around the 14th of July and water retreated around the 20th. Most pasture grasses can survive for 1 or 2 weeks (Undersander, 2016)

Overestimation 2: Unregistered cropland

Next to the large amount of grassland classified as damaged, 14% of the inundated Agricultural land was not registered in the crop dataset (RVO, 2021a). These areas have no registered economical contribution to the agricultural sector. Therefore, it is plausible that also for this area no financial damage occurred, resulting in another 14% damage overestimation (Equation 5.2.

$$\text{Unregistered agricultural classified land:} \quad 174M€ * 0.14 \quad = 24M€ \tag{5.2}$$

Underestimation 1: Overestimated maximum damage

Besides the magnified surface area with economic value, it is also argued that the maximum exposed asset value is overestimated. In the GFRT an amount of 2,15 $€/m^2$ was used as the maximum incurred damage, which includes crop loss and pollution. However, a comparison with data on agricultural crop revenue[6] indicates a much lower value of 0,55 $€/m^2$ (CBS, 2021).

If crop loss is only priced at 0,55 $€/m^2$ the remaining 1,60 $€/m^2$ should be attributed to the cleanup of the fields. However, no significant pollution was reported by the authorities[7], and Hess and Morris (1988) estimated cleanup costs to be significantly less than financial crop loss. Due to both arguments, it can be inferred that limited cleanup costs should be present.

A more conservative assumption was taken that total cleanup costs lie somewhere between 0 and 1,60 $€/m^2$. Relative to the 0,55 $€/m^2$ certain crop loss, the remainder of agricultural damage is overestimated by 0 to 75%.

$$\text{Remaining damage is 0\% -75\% overestimated} \quad 60M€ * [0 - 75\%] \quad = 0M€ - 45M€ \tag{5.3}$$

Non-Affecting 1: Temporal effect of flood occurrence

Finally, the moment of the flood was validated, as crop loss is highly dependent on the flood moment (Brémond & Grelot, 2013; Förster et al., 2008; Hess & Morris, 1988; Pivot & Martin, 2002). When floods occur after the harvest time, only damage to the field can still result in financial damage. However, predicting the total impact of this effect is complex, as the exact harvesting time is difficult to estimate. The moment a farmer will harvest its crop depends on the moment the crop was planted, the type of crop, and the yearly variations of the climate it has been growing in Brémond and Grelot (2013).

With a large temporal bandwidth, it was found that the affected crops were supposed to be harvested between August - November. As the flood occurred in July, it could be safely assumed that in this case damage to crops occurred.

[6] Revenue per m^2 of cropland. Because the total amount of crop loss depends heavily on the type of crop planted, crop types were analysed with the BRP dataset in Table 5.3. The three most common crops in the inundated zone are Mais, Potatoes and Sugarbeets, which is largely consistent with the used CBS data.

[7] The current advice on sludge removal was to wait until it is washed out by rain, for which it has ample time before the major part of next year's crops is planted (RVO, 2021c)

Infrastructure

Infrastructure damage from inundation is complex to model, with literature giving widely varying damage functions and methods (Habermann & Hedel, 2018). Exact damage processes depend on the type of inundation, but are sensitive to high flow velocity and the erosion it causes (Merz et al., 2010).

Calculated damage for the transport sector is the second biggest contributor, amounting to a total of 107 million euro. Two errors resulted in an overestimation of 39,8 million euro. First, false damage calculated at elevated portions of major roads resulted in an 11,3 million euro overestimation. Second, a linear approach to calculating the three most minor road types showed a summed overestimation of 28,5 million euro. Finally, two non-affecting factors were validated, and are described in Appendix A.

Table 5.4: Errors in GFRT estimation of infrastructure

Overestimating factors	**Impact**
1: Calculated damage at elevated, non-inundated primary roads, highways and railroads is damage in only 10% of the cases	12 M€ overestimation
2: Classification errors in 'Other road', 'Tertiary road' and 'Secondary road'	30 M€ & 3,5 M€ overestimation 5 M€ underestimation

Non-affecting factors	**Impact**
1: Surface area error due to gridcell approach [8]	-
2: Sensitivity to specific inundation depth	-

Overestimation 1: Calculated damage at elevated locations in main road-links

Major transport links such as railroads and highways are considered as critical infrastructure links (Habermann & Hedel, 2018). To protect them from hazards and ensure fast and safe crossing with minor roads and waterways, they are often placed at an elevated position using berms and/or bridges.

Through visual inspection two main erroneous situations were discovered that are caused by this elevation: The first situation occurs at bridges. When roads are located on a bridge they are not often inundated, whereas lower-placed floodplains can be. Thus the model flood extent can be drawn completely over the highway, or an overlap could occur on the side due to the rasterization. Both characteristics result in the GFRT modelling (part-of) the highway as damaged, whereas the water does not reach the road.

The second situation occurs at berms, where the flood extent can be drawn up to the side of the berm but does not overtop it. Grid-cells can overlap the berm of the road, and can be identified as inundated due to the steep slope present. This will also result in a damaged road whereas the water does not reach. Visualizations of these situations can be seen in Appendix A Figure A.12.

These situations result in an overestimation that can quickly lead up to multiple millions due to the high economic value of major infrastructure classes. Analysis of 15 zones in the classes 'Primary road', 'Railroad', and 'Highway' revealed only 2 situations where the damage occurred, which was around 10% of the damage. Therefore the conclusion was drawn that for this flood event, the estimated damage for these infrastructure classes was overestimated with 11,3 million euro.

$$\text{90\% of 'Primary road', 'Railroad', and 'Highway':} \quad 0,9 * (8,0 + 2,6 + 2,0)\ M€ = 11,3\ M€ \tag{5.4}$$

[8] Both Non-affecting factor 1 and 2 are further described in Appendix A

Overestimation 2: Classification errors in 'Other road' and 'Tertiary Road'

In the case area, a wide variety of road types ranging from highways to dirt tracks was present, for which three errors were found.

The first error regards road classification and exposure. A total of 6 transport land use classes were used, shown in Figure 5.5. The classes were connected to similar depth-Damage (vulnerability) curves, but used different maximum damage (exposure) values. The main variable driving the maximum damage is the quality level of the road. However, when validating the model, road-quality inconsistencies were found inside the minor road classes. Visually, the 'Other road' class consisted both of asphalted roads as well as dirt tracks and bike lanes. Besides, 'Tertiary road' exhibited a higher consistency but also included a smaller amount of dirt tracks.

Second, the GFRT maximum road damage in $€/m^2$ was transformed from HIS-SSM's road damage in $€/m$, using an assumed road width. However, the 'Other road' polygon width showed a lower mean of 4m width versus the assumed width of 7,5m[9]. For 'Secondary road' this difference was even larger with a polygon average width of 6m versus an assumed width of 12,5m. Theoretically, this would have resulted in a model underestimation, as the maximum damage $/m^2$ would have been higher when the proper width was assumed.

Third, when looking at the representation of roads in the GFRT, it was noticed that the grid-cell-based approach has enormous difficulties representing smaller and thinner roads. Large parts of the road are not classified as damaged roads and simultaneously large berm areas are classified as damaged roads. This misrepresentation creates inaccuracies in the inundation depth, the total surface area of damaged road, and the maximum damage per surface area.

Due to these errors, it was concluded that the grid-cell approach leads to large inaccuracies in both the hazard and exposure component of road damage modelling. Therefore, a linear modelling approach is suggested.

Table 5.5: Land use classes and their maximum damage per m^2 in the transport sector

Classification	Exposure $/m^2$	Exposure $/m$
Other road	42,84 €	353,96 €
Tertiary road	42,84 €	353,96 €
Secundary road	96,39 €	1298,92 €
Primary road	107,10 €	1915,90 €
Highway	107,10 €	1915,90 €
Railroad	584,81 €	5845,13 €

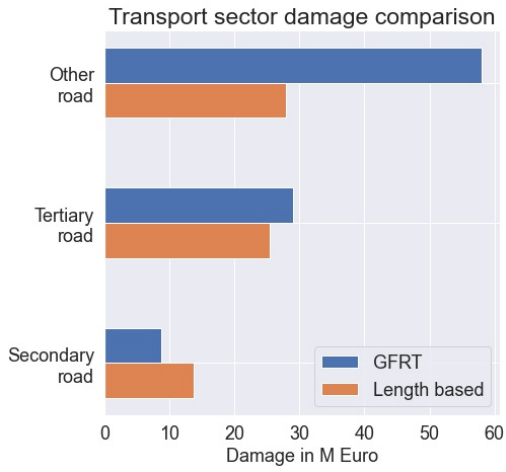


Figure 5.5: Comparison of length-based road damage estimation with GFRT estimated damage.

Linear approach

To eliminate the found issues, a different method to calculate total road damage was used, based on two parts: First, the method calculates the total length of the damaged road instead of the total surface area. Second, a dataset with all government-maintained roads[10] was used, to erase all dirt roads and bike lanesfrom the estimated damage (Rijkswaterstaat, 2021b). The approach is described and visualised in Appendix A Figure A.13.

[9] Found in Section 5.1

[10] The nationaal wegenbestand, NWB. This dataset contains all streets in the Netherlands that have a street name or road number and are maintained by a governmental body (The state, Provinces, Municipalities, Waterboards).

Results of this approach are compared with the GFRT in Figure 5.5. We find that compared to this approach, the GFRT overestimates the 'Other road' class by 30 million euro. As mentioned earlier, this is likely due to the removal of dirt tracks and other paved areas. For 'Tertiary roads', a smaller overestimation of 3,5 million euro was found, which aligns with the fact that fewer dirt roads were present in this class. It may be hypothesized that the underestimation caused by lower road width and the overestimation by the inclusion of (a smaller number than in 'Other road') cancelled each other. For the secondary road we find an opposite result of a 5 million underestimation by the GFRT. This could be explained by the earlier mentioned difference between the GIS road width and assumed road width, which was used in the maximum damage $/m^2$.

Infrastructure conclusion

Combining the differences, a GFRT overestimation of 28,5 million euro was found using the linear approach. Furthermore, an 10,3 million overestimation due to the elevation of major roads leads to a GFRT overestimation of 39,8 million euro. This is based on two main assumptions: First, 90% of damage to primary roads, highways and railroads does not occur. Second, the linear approach using HIS-SSM's maximum damage values $/m$ of road length is valid. This assumption can be further validated using damage data or reconstruction costs.

Residential

The damage process in residential houses is very uncertain, as houses are highly heterogeneous and private response to floods is difficult to predict. The damage can already be significant on low inundation levels and is generally split up into two main parts: house structure and house inventory (McBean et al., 1986; Merz et al., 2010; A. Pistrika et al., 2014; Smith, 1994). Major parameters driving the damage are flood warning time, flood experience and pollution (Merz et al., 2004; Wind et al., 1999).

The calculated damage for the residential sector amounts to 95,5 million euro. This sector consists of two land-use classes: residential houses and residential area. Residential houses represent the exact surface area of residential structures (Kadaster, 2021a). Residential area delineates the surrounding public area, including sidewalks and gardens in the neighbourhood. Shown in Table 5.6, two errors and one main uncertainty were found.

Table 5.6: Errors in GFRT estimation of residential

Overestimating factors	**Impact**
1: Flooded dike berms	6 M€ overestimation
Underestimating factors	**Impact**
1: Absence of house content in exposure	23,7 M€ underestimation
Uncertain factors	**Impact**
1: Manually inserted inundation depth of 0,5m	{-32,6 M€ ; +38,8 M€} uncertainty

Overestimation 1: Flooded dike berms

Alongside the Maas and the Roer, many inundation sites occur in neighbourhoods classified as 'residential area'. Large zones of this inundated area contain gardens, sidewalks and small streets. Here, the damage is assumed to occur due to loss of property and cleanup costs. Around 75% of damage occurs in these large areas, where damage is validated.

However, for the other 25% of the damage, the validity of the estimation is doubted due to two reasons: first, visual inspection shows that many zones are at the edge of the river bed, where the water covers (part of) the berm. As this is a natural area consisting of grass or rocks, financial damage here is unlikely. Second, damage at small inundated zones can also be doubted, because these polygons seem to be located at the edges of the residential area, where no valuable assets are present. Therefore, it is concluded that around 25% of the total damage, adding up to 6 million euro, may not occur.

Underestimation 1: Absence of house content in exposure

Next to the damaged surfaces, also the maximum exposed asset value was scrutinized, using a model-model comparison. When validating the maximum asset damage used for residential houses in Section 5.1, it was noted that only value for the house structure was used. Generally, also inventory is added to the exposed asset value due to its significance (Messner et al., 2007; Molinari et al., 2020; Slager & Wagenaar, 2017).

When calculating damage to house inventory, also an inventory-specific vulnerability curve could be used to link inundation depth to a damage ratio. In this case however, it was assumed that the vulnerability process of house content was equal to that of the house structure. Therefore, a simple scaling factor for the change in maximum exposed asset value could be used.

Slager and Wagenaar (2017) proposed the content value of houses to be at 70.000 euro per object. The additional exposed asset value $/m^2$ by including content can be calculated from this in two steps. First, by using the average surface area of residential houses located within the flood extent, and second by calculating the present value for 2021. Using Equation 5.5, this gave a total of 354,07 €$/m^2$. If it is assumed that house content should be included in the analysis, then the total damage estimate for residential houses should be increased by 32,7%, leading to an additional 23,7 million euro of damages.

$$\begin{aligned} \frac{Content\ value\ [€/object]}{Average\ house\ surface\ [m^2]} * \text{Inflation} &= \frac{70.000}{214} * 1,08 &&= 354,07\ €/m^2 \\ \frac{New\ exposure\ [€/m^2]}{Old\ exposure\ [€/m^2]} * \text{Initial damage} &= \frac{1082,43 + 354,07}{1082,43} * 72,4\ M€ &&= 23,7\ M€ \end{aligned} \tag{5.5}$$

Uncertainty 1: Manually inserted inundation depth of 0,5m

As mentioned in Section 5.1, the inundation depth was unknown for large areas around the Geul, and to a lesser amount around the Roer. In the total area the damage estimates for residential houses and residential area were affected by this uncertainty to an amount of 25,8 million and 6,8 million euro respectively. Looking at the surface area alone, 51% of the damaged residential housing area consists of these uncertain grid cells.

To assess the total sensitivity, a what-if analysis is performed. An upper estimate was created by assessing the total damage if all the grid cells would be affected by 1m inundation depth instead of the set 0,5m, so an increase of +0,5m. Subsequently, a lower bandwidth was created by assessing the damage reduction if the grid cells indeed had no damage at all, as the hydraulic model initially suggested. For the upper bandwidth, the damage ratios change from 0,18 to 0,40 (residential houses) and from 0,29 to 0,60 (residential area), leading to an additional 38,3 M€ damage as in Equation 5.6.

$$\begin{aligned} &\text{Residential house scaling factor:} && \frac{0,40 - 0,18}{0,18} &&= 1,22 \\ &\text{Residential area scaling factor:} && \frac{0,60 - 0,29}{0,29} &&= 1,07 \\ &\text{Damage increase residential house:} && 25,8\ M€ * 1,2 &&= 31,0\ M€ \\ &\text{Damage increase residential area:} && 6,8\ M€ * 1,07 &&= 7,3\ M€ \\ &\text{Total upper bandwidth:} && 31,0 M€ + 7,3\ M€ &&= +38,3\ M€ \end{aligned} \tag{5.6}$$

For the lower bandwidth, a decrease of 0,5m would mean that no inundation depth and damage is present at uncertain locations. The damage reduction would be 32,6 million euro as in Equation 5.7.

$$\text{Lower bandwidth:} \quad -25,8\ M€ - 6,8\ M€ \quad = -32,6\ M€ \tag{5.7}$$

Other

The final category contains three major classes. These classes were characterised by a strong diversity of asset types within their land-use class, requiring an individual validation approach. The land use classes 'Recreational - Day', 'Communal building' and 'Public area' were validated[11].

Table 5.7: 'Other' land use class GFRT damage estimation error sources

Overestimating factors	**Impact**
Recreational - Day: Boat damage in Meuse Marinas	10 M€ overestimation
Public area: No damage at gas power plant and riverside location	7 M€ overestimation
Non-affecting factors	**Impact**
Communal building: Special needs school	

Recreational - Day

The largest damage zones of 'Recreational - Day' consist of 6 Marinas along the Meuse, with a total estimated damage of 17,7 million euro (Table 5.8). This calculation is based on the amount of inundated surface surrounding the marinas that is classified as 'Recreational - Day'. However, as that area consists mostly of grass and trees, it is plausible that most of the financial damage in this area will stem from private boats. Therefore a surface-based damage estimation seems wrong.

Damage is concluded as overestimated due to two reasons. First, the goal of the GFRT is to analyse flood mitigation measures. As it is plausible that boats will still be damaged despite flood mitigation measures such as dikes, this damage may be irrelevant to consider. Second, as no reports of damage to boats were found and no storm with heavy winds was present at the time of the flood, it is assumed that few boats were damaged. Therefore, an overestimation of 10 million euro is concluded.

Table 5.8: Damaged marinas for 'Recreational - Day'.

Location	**Impact M€**
Roermond North	6,6
Roermond South	3,5
Oost maarland	2,7
Zuid maastricht	2,6
Ohé en laak	1,2
Stevensweert	1,1
Total	**17,7 M€**

Table 5.9: Damage areas of 'Public area' land use class.

Location	**Impact M€**	**Damage**
Gas power plant	4,9	No
Riverside	2,2	No
Sluice	2,0	Yes
Water treatment plant	1,9	Yes
Sluice	1,6	Yes
Other	2,1	Yes
Total	**14,7 M€**	

Public area

In 'Public area', four different types of assets were distinguished: A gas power plant (The Clauscentrale), Sluices, water purifying plants, and parking lots. Direct damage to these industrial sites is difficult to predict and cannot be done on a simple land-use basis, as its vulnerability changes spatially. Furthermore, this class contains critical infrastructure, which requires special attention in the damage analysis due to additional indirect damage when it fails (De Bruijn et al., 2016).

An analysis of the damage zones in Table 5.9 showed that the main damage contributor was the Claus centrale, at 4,9 million euro. This was followed by an industrial riverside location estimated at 2,2 million euro, two sluices and a water purifying plant for another combined 7,6 million euro.

Currently, a conclusion was drawn on the gas power plant and the riverside. The gas power plant is one of the largest in the Netherlands and is protected by a dike (Bisschop et al., 2015). As no flood damage news from the gas power plant was found, it was assumed that no damage occurred. Subsequently, it is also plausible that the empty riverside damage of 2,2 million euro is overestimated [12]. Therefore, a total overestimation of 7,1 million euro was concluded.

[11] In dutch: Dagrecreatieve voorziening, Bijeenkomstfunctie, and Openbare voorziening
[12] Most of this area looks like unused river berms next to an industrial area, meaning little damage would be expected.

Communal building

The communal building class includes sports halls, casinos and meeting halls. The exposure is higher than residential housing at 1374,09 €/m^2, and the GFRT total estimated damage equals 20 M€. As it seemed unsuitable to conclude the estimate of such a heterogeneous set, the focus was brought to the largest contributor. A total damage estimate of 7 million euro is attributed to a special education school in the Municipality of Valkenburg aan de Geul. This area was indeed inundated with an average depth of 0.75m, so estimating damage for this location seems valid, however the total amount is difficult to validate due to the uncommon land use.

To grasp the exposed asset value involved in school buildings, governmental data from Municipality Utrecht (2021) was used. It was calculated that a new school building for special education with a similar surface area of 11,550 m^2 would cost around 29 million euro. Therefore the estimated damage of 7 million euro would be approximately 25% of the total damage, which seems to be in the right order of magnitude. For specific conclusions if this is an over or underestimation, the site would need to be inspected with a loss adjuster. Other locations were smaller contributors and consisted of more similar assets like sports halls. Here, it is assumed that on average the damage is valid.

5.3. Conclusion

Summarizing all the found factors in Table 5.10, a total overestimation of 153,1 million euro was found. This led to a GFRT best estimate of 373,1 million euro for the direct flood damage in Limburg, visualised in Figure 5.6. Besides in Figure 5.7, the errors are attributed to their root parameter. The largest error was found in the exposure land use of agriculture exposure. The second-largest overestimation was found in the infrastructure classes, based on the length-based model. Confidence in this approach can still be improved with damage observations. This error was attributed to the vulnerability parameter, but was caused by a rasterization error. Finally, a smaller underestimation in residential and an overestimation in three individual classes were found, with a total sum of approximately zero.

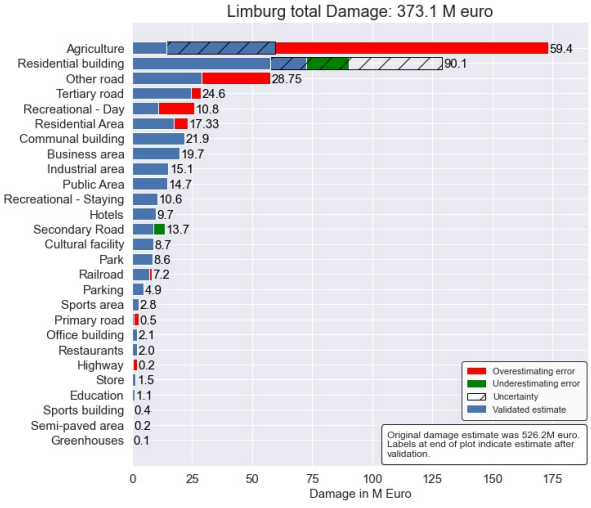


Figure 5.6: GFRT estimated damage after model output validation

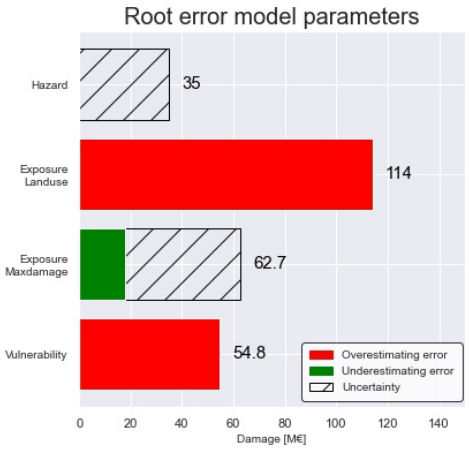


Figure 5.7: GFRT estimated damage after model output validation

Besides the summed error, an uncertainty bandwidth of +77,6 and -38,8 M€ was presented in the hazard parameter for the residential class, and the maximum damage parameter for agriculture. Adding this range to the new best estimate gave a total uncertainty bandwidth of 295,5 M€ to 411,9 M€.

This bandwidth stems from two uncertainties: agricultural pollution and residential inundation depth. Based on governmental reports, it may be that agricultural pollution is low, thereby reducing the best estimate by another 45 million euro (RVO, 2021c). Therefore, it is plausible that the real damage may be low in the bandwidth, between 295,5 and 373,1 million euro. For further reliability, the assumptions in Table 5.10 that underwrite this validation can be further investigated with local authorities.

Table 5.10: Overview of found GFRT errors for the Limburg case study. Error size is indicated in million euro and a positive (+) error indicates an overestimation by the GFRT. At the bottom of the table, the errors are summed and compared to the original GFRT estimate.

Sector GFRT estimate	Error type	Impact M€	Assumptions
Agriculture 174 M€	Overestimation	+ 114,0	• Inundated grassland does not create financial loss.
	Lower uncertainty	45,0	• Uncertainty surrounding amount of pollution requiring clean-up.
Infrastructure 107 M€	Overestimation	+ 39,8	• Only 10% of calculated rail, highway and primary road damage occurs due to their elevation which is modelled incorrectly. • Secondary, Tertiary and Other road damage is highly variable and is instead found using a length-based model.
Residential 95,5 M€	Underestimation	- 17,7	• Small inundated areas in neighbourhoods and river berms do not result in damage (+6 M€). • Inclusion of house content in the maximum damage values (-23,7 M€).
	Lower Uncertainty	32,6	• Uncertainty caused by unknown inundation depth that is set manually to 0,5m.
	Higher uncertainty	38,8	
Other 17,7 + 14,7 + 20,0 = 52,4 M€	Overestimation	+ 17,0	• Most private boats in marinas were moved prior to the highwater and have suffered not much damage. 10M€ out of 17,7M€ damage estimated for Marinas was taken as an overestimation. • Damage at a school is accurate with 7 M€ • The gas power plant 'Clauscentrale' did not suffer 4,9M€ flood damage, and an unused riverside area has a 2,2 M€ overestimation. • 7,7 M€ estimated damage at Communal buildings using manually inserted 0,5m depth is correct.

GFRT original estimate:	**526,2**	**M€**
Total overestimation:	- 170,8	Agriculture, Infrastructure, Other.
Total underestimation:	+ 17,7	Residential.
GFRT adjusted estimate:	**373,1**	**M€**
Total higher uncertainty:	+ 38,8	Residential. Damage could be this amount higher.
Total lower uncertainty:	- 77,6	Agriculture, Residential. Damage could be this amount lower.
GFRT adjusted range:	**295,5 - 411,9 M€**	

6

Limburg - Verification using empirical data

Besides the plausibility assessment, verification was applied with multiple external data sources by applying phase 3, step 5 and step 6 of the framework (Elaborated in Chapter 3).

6.1. Residential insurance data

Insurance databases can be a valuable information source in flood damage modelling (Botzen & Bergh, 2008; Wing et al., 2020). As housing and inventory insurance penetration rates in the Netherlands are high, multiple large insurers were involved and had employees locally present during the flood event.

Damage estimation using insurance data

One insurer, Achmea, communicated their registered total residential damage (Achmea, 2022). This amount was used to estimate the total residential damage using Equation 6.1.

$$\text{Total damage} = \quad Dam_{Achmea} * F_{Marketshare} * F_{Ins.Pen} * F_{Minordamage} \tag{6.1}$$

The formula takes that the total residential damage can be estimated by multiplying the damage reported by Achmea with three factors. First, Achmea is not the only insurer present in this market. Therefore a factor based on Achmea's market share is used to estimate the damage at other insurers. Second, not all houses in the area are insured[1], for which a factor based on the insurance penetration is used. Finally, it was explained that the damage in Achmea's database only concerns large damage events, for which a damage expert was required [2]. To correct for this, a factor $F_{Minordamage}$ is used.

Reference data was found to estimate the factors from the Dutch Central Bank (DNB, 2021) and the Verbond van Verzekeraars (VvV, 2016), presented in Table 6.1. Looking at the assumed factors, the calculation is especially sensitive to insurance penetration and minor damage assumptions. It may be inferred that insurance penetration is likely on the higher end, due to the old reference age, as insurance rates generally increase over time. $F_{Minordamage}$ has no reference data available. Instead, the reference is based upon the fact that when fluvial flood damage occurs, it is generally large. Therefore the not-claimed damage is likely small compared to the events registered in the database. Less sensitivity is concluded for $F_{Marketshare}$. As the case area of Limburg is large, a small regional difference between the region and the entire Netherlands is plausible. Finally, the 10 M€ reported by Achmea is deterministic, as the data quality is assumed high due to its application for underwriting insurance payouts. Because insurance penetration and market share are likely on the higher end, the conclusion is drawn that the real damage is likely between the lower and median estimate.

[1] Uninsured damage is either incurred by residents themselves or claimed at the WTS.

[2] Smaller damages are not present in the database because they were either unreported, or were handled by the front desk without a damage expert.

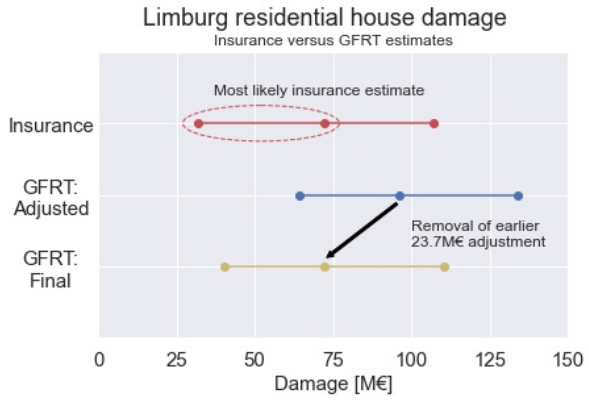
Table 6.1: Assumed values used to quantify formula Equation 6.1. The Low Medium and High factors were estimated using plausible values in the found reference data.

Parameter	Explanation	Data	Reference	Assumed factor		
				Low	Med	High
Dam_{Achmea}	Valued house and inventory damage in database.		Achmea (2022)		10 M€	
$F_{Marketshare}$	Achmea market share.	25 - 29%[3]	DNB (2021) and VvV (2016)	3,03	3,45	4,00
$F_{Ins.Pen}$	Insurance penetration.	56 - 97%[4]	VvV (2016)	1,04	1,67	1,77
$F_{Minordamage}$	Damage not in database.			1,1	1,25	1,5
			Total damage M€:	**31,6**	**71,8**	**107,1**

Verification & conclusion

Residential house damage estimates are visualised in Figure 6.1. The insurance estimation is concluded at 71,8M€. Furthermore, an uncertainty range of 31,6 - 107,1M€ was set, with a high probability that the real damage is between the low and middle estimates. In Chapter 6 GFRT residential house damage[5] was adjusted to 96,1 M€, with a range of 63,5M€ - 134,9M€. Comparing both estimates it is concluded that the bandwidths overlap, but that the insurance estimate may indicate an overestimation, especially because the insurance estimate is expected at the lower end of the bandwidth.

In the previous chapter, the GFRT estimate was increased by 23,7 M€ based on a model-model comparison (see Chapter 5). However, the insurance data does not support this adjustment due to two reasons: First, the most likely insurance estimate is much lower than the adjusted range. Second, the insurance best estimate of 71,8 million euro aligns better with the original model estimate. Therefore, the added 23,7 M€ overestimation is removed, resulting in a final residential house damage estimate of 72.4 M€. Nevertheless, the significant amount of uncertainty should be acknowledged here. As the insurance estimate is so uninformative, a verdict may also be that too little information is available to conclude the validity.



Total residential damage

$$\begin{array}{lllll} \text{Range:} & \{39,8; & 72,4; & 111,2\} & M€ \\ +17,1 & M€ & & & \\ \hline = & \{56,9; & 89,5; & 128,3\} & M€ \end{array}$$

GFRT total damage

$$\begin{array}{lllll} \text{Range:} & \{295,5; & 373,1; & 411,9\} & M€ \\ -23,7 & M€ & & & \\ \hline = & \{271,8; & 349,4; & 388,2\} & M€ \end{array} \quad (6.2)$$

Figure 6.1: Insurance and GFRT residential house damage.

The final estimates are shown in Equation 6.2, with the final house damage estimate shown at the top. Combining this with the remaining 17.1 M€ residential area damage, the total residential damage estimate range is found. Finally, the GFRT total damage estimate and range for the event are adjusted by the same 23,7 M€.

[3] Market share for The Netherlands. The 2021 source quotes a three year stable 29%. The earlier 2016 source quotes 25%
[4] For the whole Netherlands. Inventory insurance rates are much higher (97%) than housing structure rates (56%). Verbal sources quote both rates during the event to be above 95%.
[5] Note that this concerns damage to houses, not including damage to residential area.

6.2. Governmental claim data (WTS)

As the floods in Limburg were classified as a disaster, the national government set up a fund to accommodate (part of) the damage suffered by residents, the WTS. In this section, claim data of this fund is used to spatially verify model predictions.

Data description

Damaged parties[6] and residents were able to claim *uninsurable* damage at the fund. Claim data of valued damage (Thus the actual damage valued by loss-adjusters, not the smaller payout) was shared in a spatially aggregated form on postal code 6 (PC6) level for privacy reasons. Furthermore, the number of applications per area was shared. This provides information on the coverage factor, which is defined as the percentage of the total occurred damage that the WTS covers (RVO, 2021b).

An overview of the valued damage of all 11 categories can be seen in Table 6.2. The largest part was claimed by businesses[7], around 30 million euro, with another 4 million euros being claimed by private residents. The final 3 million euro are public rescue and cleaning costs[8].

Table 6.2: Overview of data in the WTS dataset. The top table rows show the damage and number of applications for 5 damage categories for the business sector. The bottom row indicates this for private residents and general costs made by either residents or organisations. Further explanation is given by RVO (2021b)

	Business costs				
	Crop loss	Fixed assets	Current assets	Startup costs	Company damage
Damage [M€]	10,34	18,61	1,90	5,2e-3	0,051
$N_{applications}$	276	277	113	3	10

	Residential			**General**		
	Structure	Inventory	Cars	Rescue	Evacuation	Cleaning
Damage [M€]	2,93	0,99	0,14	1,03	0,45	1,80
$N_{applications}$	207	190	41	224	133	328

Bias in the data

Although the damage numbers represent valued damage, significant bias was found in the data that indicates they do not resemble the full truth. The following points were found:

1. The damage to many personal cars was valued at exactly 3.000 euros. As the maximum payment is 2.750 euro, loss-adjusters may have simply valued damage at 3.000 euro to indicate that the damage is higher than the maximum payout.
2. PC6 area 6229AN was found to have approximately 150.000 euro of fixed and current asset claims [9], whereas the GFRT calculated very little damage. The area concerned a camping place with a marina, indicating the damage was likely boats or caravans. As boat damage is not calculated by the GFRT it cannot be used for verification. The impact of this effect on fixed and current asset claim data in other areas is unknown.
3. Large differences were found in the areas around Valkenburg. Communication with the data owner indicated that in this area, practically all damage was *insurable*, and therefore could not be covered by the WTS. However, some exceptions were made. Therefore, this area differs from the rest of the dataset, but the exact effect on the coverage ratio is unknown (RVO, 2021b).

[6] Companies, NGOs as sports clubs, religious organisations, municipalities
[7] Businesses had lower insurance coverage and therefore went more to the WTS (VvV, 2016)
[8] The data also contained a minor amount of infrastructure damage, which was not used because the claims were not finished.
[9] Relatively large, and $N_{application}$ was 25, the largest of all PC6 areas

Goal and method

The goal of comparing GFRT predictions to the WTS data is twofold. First, it was attempted to show pattern validation in the region for crop loss, business damage and residential damage. If the aggregated GFRT predictions correlate well with the WTS data, it is plausible that the model performs equally well for the entirety of Limburg. Subsequently, areas that deviate from the regression may be modelled less accurately by the GFRT. This could indicate errors in the vulnerability or the inundation model. As relative damage is compared, exposure values cannot be validated through this approach.

Pearson's correlation was calculated for multiple spatial scales: PC6, PC5 and PC4 level, whereby PC4 has the largest spatial scale[10]. This was done because the data showed significant variability on PC6 level, possibly due to a varying coverage factor. Pearson's correlation was preferred over Spearman's rank correlation, as a linear relationship between the GFRT and the WTS data is expected[11].

As a second goal, it was attempted to approach the absolute amount of residential damage in the area using a coverage ratio correction factor Cov. By estimating the coverage factor for each area, the total damage in that area can be estimated and subsequently used to verify the GFRT's estimation.

Results - Agriculture

In Section 5.2, it was concluded that agriculture was highly overestimated due to the inclusion of undamaged areas such as floodplains and pastures. Hence, these areas were removed from the damage prediction, to allow a more exact comparison. Subsequently, the predicted agriculture damage was aggregated for each PC6 area, to allow the comparison between observations and data.

The comparison for the PC6 area and further aggregated PC5 and PC4 areas are shown in Figure 6.2. On the left two graphs, significant Pearson's correlation coefficients of 0,51 and 0,52 were found. However on the smaller PC6 scale, no significant correlation was found between the WTS observations and the GFRT estimates.

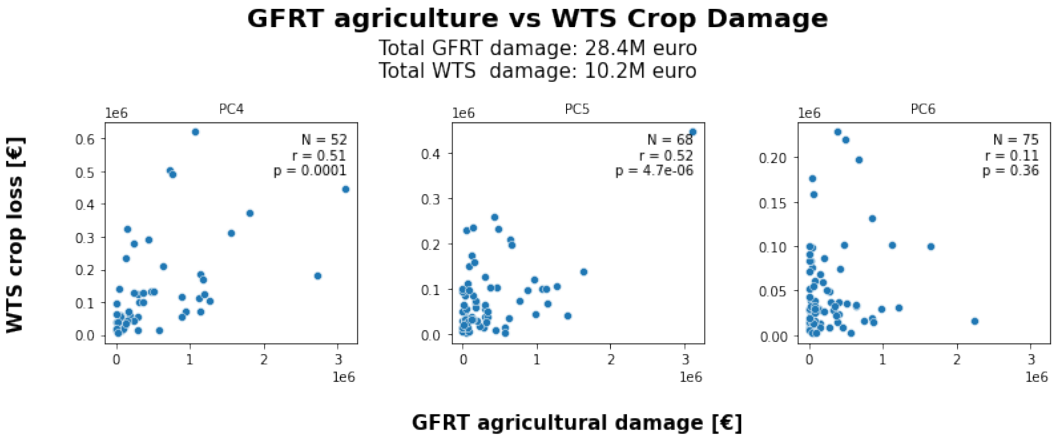


Figure 6.2: Scatterplots of GFRT agricultural damage (x-axes) versus WTS crop loss damage (y-axes). Spatial scale decreases towards the right from PC4 to PC6 level.

This discrepancy can be rooted in the WTS data, or the GFRT. The WTS data may also include pluvial flood damage to crops, leading to a relatively higher WTS damage in some areas. Furthermore, crop loss damage is registered in the PC6 area of the claiming farm, whereas the inundated field could be located in another PC6 area. On the GFRT side, unmodelled variation in crop type value and -vulnerability could explain the large scatter. The GFRT only applies a single vulnerability curve and maximum damage, whereas in reality this is highly crop dependent. Furthermore, the exclusion of other flood parameters such as pollution, flow velocity and inundation time could cause the unexplained variability.

[10] PC6 areas have a mean surface area of 85.000 m^2, PC5 of 300.000 m^2 and PC4 of 750.000 m^2 (Appendix A Figure A.14)

[11] Assuming that the coverage factor is constant (at least at larger spatial scales), it is assumed that the WTS has a fixed proportion of the total damage that the GFRT estimates.

Results - Residential

For residential damage, correlations were hypothesized with both residential, as well as the general categories indicated in Table 6.2. Pearson's correlations are analysed in Figure 6.3. For PC4, it was found that five out of six damage categories correlate strongly and significantly with the GFRT's residential damage [12]. Subsequently, all correlations diminish at PC5 and PC6 scales, with the cars category performing best. Due to the explained major bias in the car data, this category is not further used. Instead, it was attempted to improve the PC6 correlation between the predictions and the WTS residential observations, by correcting the latter for the varying coverage factor.

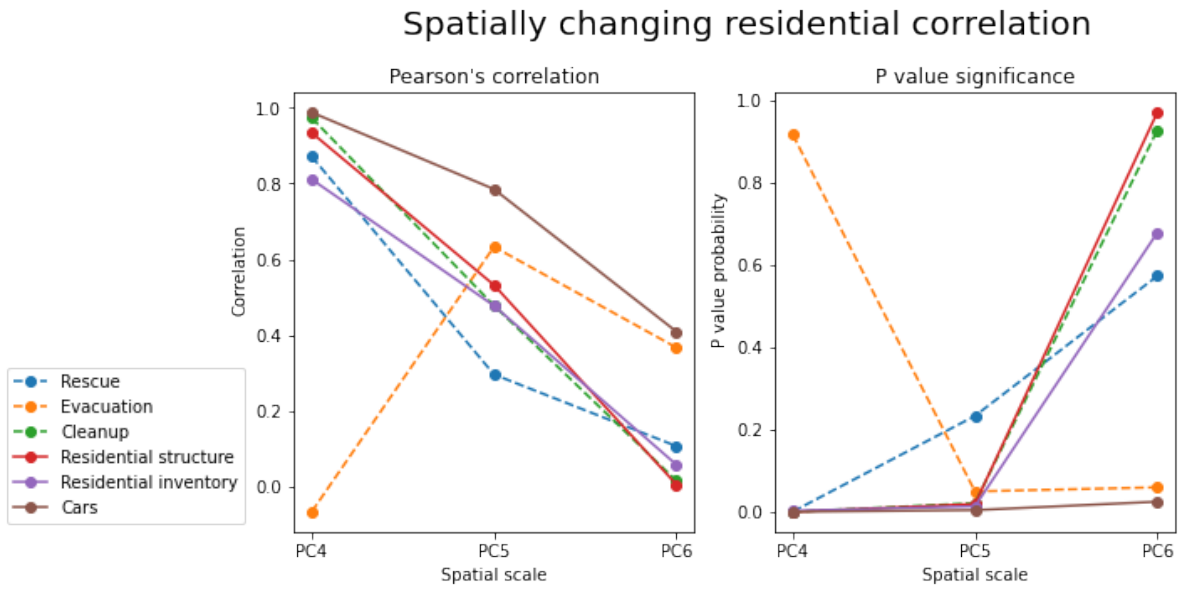


Figure 6.3: The left plot shows the correlation of GFRT residential damage with six WTS damage categories on the y axis, for three different spatial scales on the x-axis. The right plot indicates the corresponding significance probabilities.

Results - Residential coverage factor estimation

The coverage ratio can be estimated through two approaches: an overall order of magnitude estimate and a GIS analysis. First, the order of magnitude estimate. As illustrated in Figure 6.4, the WTS data will only represent part of the total damage, as other damage is claimed at insurance or incurred self. This distribution is sector-specific, as both insurance rates as well as WTS requirements differ per sector. For residential, it is known that insurance rates lie at 90 - 95% for Limburg (VvV, 2016). Assuming that this rate is equal in the affected area, around 5 - 10% of damages should be distributed over the WTS and residents themselves. Assuming an upper limit split of 50-50 between WTS and residents, the WTS coverage ratio may be assumed to be below 5%.

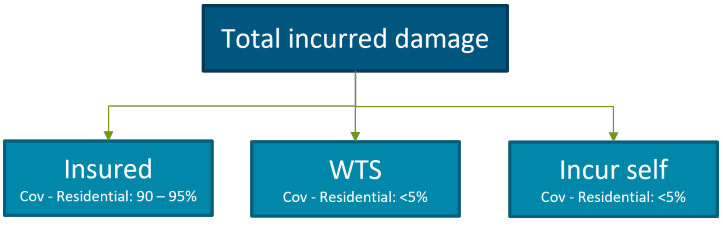


Figure 6.4: Illustration of damage distribution over three areas. Some of the affected people will be insured, bringing costs there. Subsequently if no insurance policy was taken, people may get reimbursed by the WTS. Finally, if the damage is not eligible for the WTS and not insured (or claimed), damage can be incurred by the damaged party self.

$$Cov = \frac{N_{WTS}}{N_{GIS}} \qquad (6.3)$$

$$Dam_{total} = \frac{Dam_{WTS}}{Cov}$$

[12] PC4 Evacuation costs have a low correlation but an insignifcant P-value due to a sample size of five.

The second method of estimating the coverage factor applies $N_{applications}$ and relates this to the number of affected houses, found with GIS analysis and BAG data. This was done in the following method: First, building surfaces were intersected with the flood extent to find all affected buildings. Second, the number of addresses in each affected house was aggregated over the postal code 6 area, to find the total amount of affected houses in the area[13]. Third, the coverage factor was estimated by dividing the amount of WTS applications over the amount of GIS-estimated affected houses as shown in Equation 6.3. However, the GIS-found average coverage factor was 37%[14], which is much higher than the expected 0-5%. This is likely caused due to the following four modelling errors and uncertainties:

1. The WTS coverage factor for private residents and businesses is significantly different, but cannot be disaggregated because $N_{applications}$ indicates no further differentiation. Also the BAG data does not indicate exactly which addresses are private or business-related. Therefore, Cov is likely too high for residential damage, and too low for business damage.
2. The amount of WTS claims does not align with the GIS-found affected addresses. An example of this is the boat example elaborated on in the previous bias section.
3. Currently all addresses are used, whereas especially in Valkenburg, many higher buildings with unaffected apartments are present. Contrarily, these houses may have affected cellars, but the exact distribution is unknown.
4. With the current method, all addresses in affected buildings are counted as affected. However, large buildings on the border of the flood extent will also have unaffected apartments on the not-flooded side of the building.

These reasons indicate why the coverage error will deviate significantly from reality. Unfortunately, the current data capabilities did not allow an increase in accuracy using GIS. Therefore, two further steps were taken in Equation 6.4 to improve the factor.

$$\text{Removal: } Cov = COV < 0,5 \qquad \text{Linear transformation: } Cov = \frac{Cov}{12,97} \tag{6.4}$$

First, all areas where Cov was larger than 0,5 were removed, under the assumption that the indicated errors were too large to allow verification. Subsequently, the assumption was made that the distribution of the coverage factor in the other areas is correct, but that a linear transformation is required to reach the correct order of magnitude (estimated using Figure 6.4). A regression analysis indicated that for residential damage, a further division of 12,97 was required to minimize the RMSE (elaborated below in Figure 6.5. Finally at the bottom of Equation 6.3, the found Cov was used to transform the residential WTS data (structure + inventory) to total residential damage in each PC6 area.

Results - Residential coverage factor correction

Results are shown in four graphs in Figure 6.5, with WTS data on the y-axes and GFRT data on the x-axes. Starting with the top left figure, the entire dataset is shown, without coverage correction and poor correlation (as found earlier). Subsequently moving to the left-below graph, the coverage correction is applied to each data point, with a mean Coverage ratio of 1,77%. In this graph, it was noticed that most outliers (far away from the diagonal) are from the PC4 areas 6301 (brown) and 6243 (green).

[13] It may be plausible that only ground-level apartments were affected, especially in Valkenburg. However, both methods were tested and applying the total amount gave more plausible results.

[14] Histograms are shown in Appendix A Figure A.15

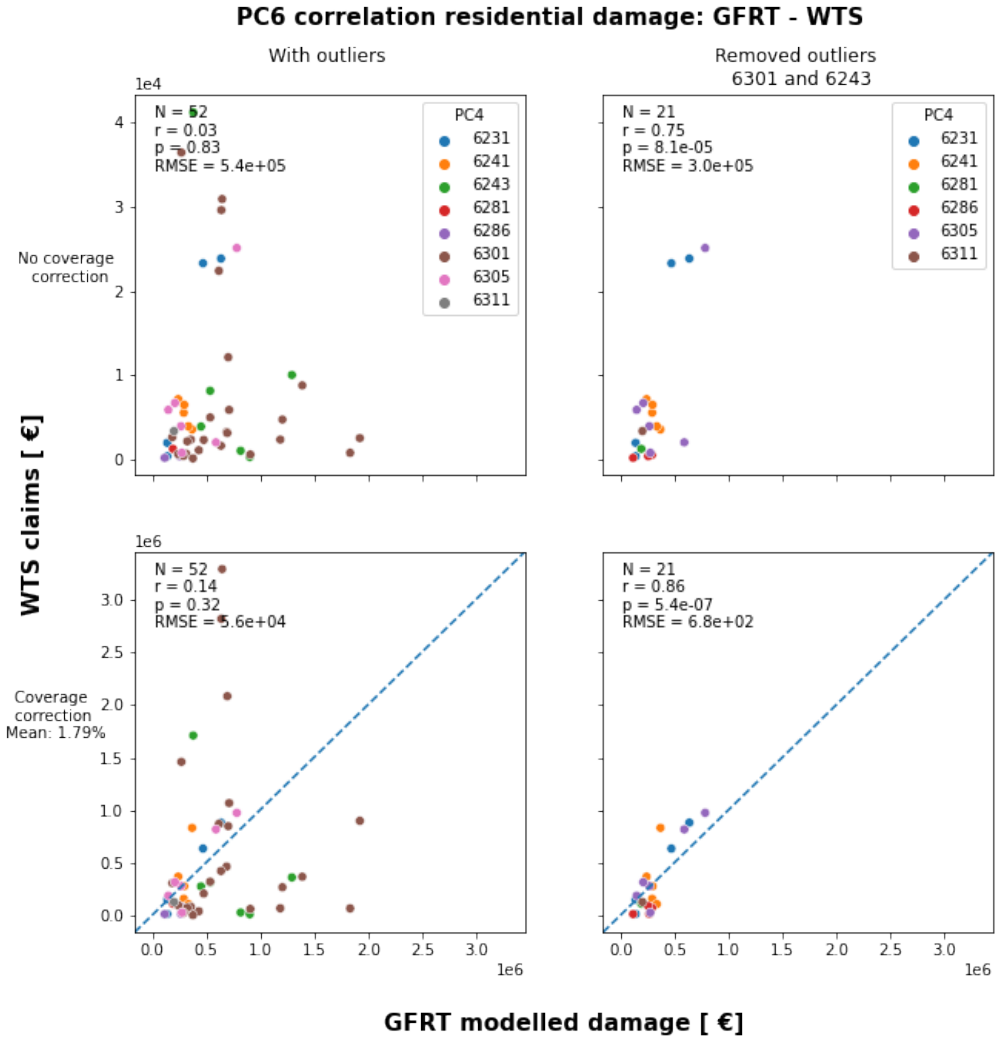


Figure 6.5: All plots show GFRT modelled damage on the x-axis vs WTS structure + inventory damage on the y-axis. The left top indicates the base data. The left bottom figure indicates the base data corrected for each coverage factor. The dashed line indicates the 1-1 line that would indicate r=1 The right top figure indicates the base data without removing outliers. The right bottom figure indicates data without removed outliers that are also corrected for the coverage factor.

When analysing the outlier areas, it was concluded that both WTS data errors and modelling errors could cause the deviations. The 6301 area concerns the Valkenburg area, where it was earlier concluded that highly irregular coverage factors are present. It may be that the applied coverage correction is unsuitable for this area. On the other hand, it may also indicate the larger uncertainty of the model, as large amounts of unknown inundation grid cells with a depth of 0,5m [15] were found here.

The 6243 area concerns the municipality of Brommelen, where the entire flood extent was set to 0,5m. The large discrepancy between the data and the model, may be caused by the occurred spatial variability of the inundation depth, which was not inserted into the model. Three areas with lower WTS damage, and one with much higher WTS damage were found (Appendix A Figure A.17). It is not exactly clear why one area has much higher damage, as this area has a higher average elevation[16] Examination of the area through google street view gave the impression that houses in the less-damaged areas had higher sills than buildings in the more damaged area. This could have limited the water from entering the houses. Further, local research is required to draw any final conclusions.

[15]Elaborated in Section 5.1, 0,5m inundation was set where the inundation depth was unknown
[16]Analysed AHN Digital Terrain Model (DTM) with a resolution of 0,5m.

Because of these irregularities, the outliers were removed from the plots on the right. The bottom right plot has both a coverage correction as well as the removed outliers, which results in a significant Pearson's correlation of 0,86. This indicates that the model shows a similar pattern with the remaining six PC4 locations throughout the area.

The earlier mentioned regression analysis was executed on the data in the right bottom plot, to minimize the RMSE. The resulting mean coverage factor of 1,79% is plausible taking into account the analysis presented in Figure 6.4. By dividing all WTS residential damage over this coverage factor, a total residential damage estimate of 25,6 million euro is found. Nevertheless, large uncertainty remains here, as we do not know the exact value of the coverage factor. A varying coverage factor between 3,59% and 0,90%, which are also plausible, creates estimates of 12,80 and 51,19 million euro respectively (See Appendix A. The uncertainty seems too large to draw a definite conclusion on the quantitative estimation, but as the bandwidth is lower than the model's estimated mean value, it may indicate a model overestimation.

Results - Companies

A similar analysis was executed by verifying fixed, current and total assets[17] with GFRT company estimated damage[18]. Results are shown in Figure 6.6, indicating a similar pattern to residential correlation. On PC6 level, total assets correlate best, with a correlation value of 0,47 and a significance of 0,009. Further variability may be caused by the earlier shown ambiguity in the WTS data or modelling errors.

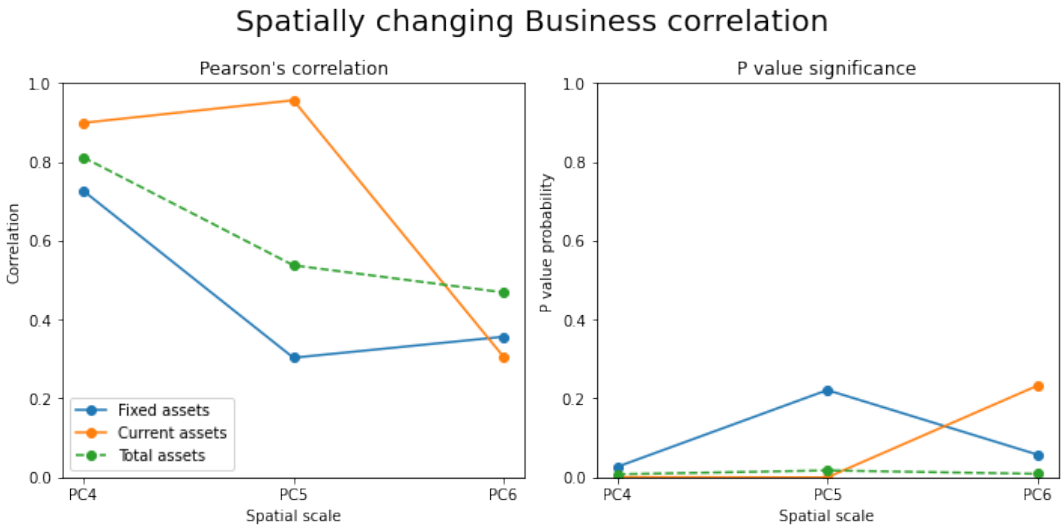


Figure 6.6: The left plot shows the correlation of GFRT - business damage (the classes Industry, Office, Store, Hotel, Company terrain) with WTS damage categories fixed, current and total assets (fixed + current) on the y axis, for three different spatial scales on the x-axis. The right plot indicates the corresponding significance probabilities.

It is counterintuitive that current assets (Machines, stock) exhibits a much higher correlation than fixed assets (buildings) on PC4 and PC5 levels. North of Meerssen, PC6 6241ND was identified as a major outlier when analysing the scatterplots (shown in Appendix A Figure A.16). Here, the GFRT estimated damage 5x as large as the next largest P6 areas, however WTS damage showed no further improvement. GIS analysis indicated that this was again an unknown inundation depth area of 0,5m. Therefore, it is plausible that many buildings here did not flood.

It was not attempted to improve correlations further due to expected difficulties with the data quality and the coverage factor, as shown with residential assets. In any case the comparison shows that large variability can be expected on a smaller spatial scale, but that on a large scale the model captures damage patterns for businesses.

[17] Total assets is calculated as fixed + current assets

[18] There are many business damage classes in the GFRT. For the comparison, an aggregation of Industry, Office, Store, Hotel, and Company terrain was used.

Conclusion

This chapter found strong and significant Pearson's correlations between governmental claim data and GFRT estimations. For business and residential damage, Pearson's correlations decreased from approximately 0,90 on the large spatial scale (PC4) down to approximately 0,40 on the smaller PC6 scale, due to variability in the data and the data coverage factor.

The comparison provided two insights. First, regions where the model deviated significantly from the data were indicated in both residential and business sectors. This is likely caused by a difference between the occurred inundation and the assumed 0,5m inundation map. Furthermore for residential assets, it was partly possible to correct for the varying coverage factor. This allowed a data transformation which indicated pattern validation of the estimations as well as remaining uncertainty. The mean coverage factor found through regression analysis seems plausible, which further supports the accuracy of the residential damage estimate (as concluded in Equation 6.2). Second, for agricultural damage, a comparison on a small spatial scale was hindered by the fact that damages are reported in a different location than where they occurred. Large variability in the data remains unexplained by the model, likely due to varying crop types and flood parameters.

This analysis pointed out that the model can capture part of the variability seen in the data, as well as indicate error locations for this flood event. However, as many errors seem attributable to the applied hazard map, it remains difficult to transfer this knowledge to other flood scenarios. Therefore, the analysis does not provide additional information on the validity of the model for risk studies.

7

Limburg - Structured Expert Judgement

During phase 3 - step 5 of the validation process, it is possible that limited observations are available. Then, it is possible to create new verification data using expert judgement as is done in this chapter for Limburg. Subsequently in step 6, the data is used to verify model components and outcomes.

7.1. Method

Although experts are frequently used in flood risk modelling and validation, applied methodologies vary significantly (Budiyono et al., 2015; Chen et al., 2011; Sayers et al., 2016). The currently applied method is based upon the classical model, developed by Cooke and Goossens (2008) (See Section 2.5).

Goals

Expert elicitation was applied to the Limburg 2021 floods case study with three particular goals:

1. The verification of GFRT predictions
2. The verification of GFRT input data
3. Evaluating the effectiveness of structured expert judgement as a validation methodology.

(1) Verification of GFRT predictions requires damage observations, that are generally not fully available for flood events. It was hypothesised that experts are able to combine the available fragmented information such as partial damage numbers from an insurer, governmental claims and historical damage reports. The expert elicitation method is especially proficient at combining this information with expert experience, to create comprehensive observations usable for verification.

For the first goal, estimations on two different system scales were elicited: Average residential scale and total damage scale. First, experts were asked to estimate average residential damage to houses and inventory in a small urban area. Second, a large scale estimation of total direct flood damage along a river stream was attempted. Experts were asked to consider direct physical damage to houses, agriculture, companies and infrastructure for three locations: the Meuse, the Geul and the Roer. Their total estimate for this damage was compared to the model's estimate. This separation was executed under the hypothesis that some experts may have knowledge especially about residential houses, whereas others could be more adept at estimating damage across different sectors.

(2) Besides verification of model predictions, the second goal was to verify the used input data. A potentially large error was found in the modelled exposure of residential assets (Section 5.1). It was hypothesised that the exposure was underestimated by around 32,7%, possible leading to an additional 23,7 million euro of damage. To test this hypothesis, experts were asked to give their estimate of the maximum damage that could have occurred to an average house in Limburg, due to the 2021 floods.

(3) The third goal was achieved through post-elicitation interviews where experts gave feedback on the process, and through a reflection on the results.

Structure

It is plausible that good estimations for residential damage and exposure require a different expertise than good estimations for total damage to multiple sectors. Therefore, the experiment was separated into two parts as seen in Table 7.1. Besides the topic, number of calibration- and seed questions, and the number of experts is shown.

Table 7.1: Overview of the two parts the expert elicitation session was split into. N_{CQ},N_{TQ}, $N_{Exp.}$ relate to the number of calibration questions, target questions and experts respectively, that was used in the study.

Part	Topic	$\mathbf{N_{CQ}}$	$\mathbf{N_{TQ}}$	$\mathbf{N_{exp}}$
Average residential	Average damage and exposure of residential houses.	7	4	10
Total direct damage	Total direct damage to residential, infrastructure, companies and agriculture assets	5	3	9

For the residential part, experts were calibrated using seven calibration questions. For two questions, WTS data was used (RVO, 2022). For the other five questions, calibration was initially planned using actual damage data of the Limburg 2021 flooding event. Unfortunately this data was not available on time. Therefore, two calibration scenarios where created using damage functions from literature:

- **Scenario 1 - Benchmark:** Benchmark damage data was created out of four damage curves: 1) Dutch: SSM single family house, 2) American: HAZUS 3) British: MCM 4) German: Flemops. These curves were combined with the maximum damage of SSM single family house
- **Scenario 2 - SSM:** A Dutch damage curve for 2-4 story low rise buildings was used. This curve closely represents the function used in the GFRT (De Bruijn et al., 2015).

The following 7 calibration questions were used:

1. How do you estimate the average valuated amount per application for damage to residential **houses (structure)**, that is applied for at the WTS for the Limburg floods 2021?
2. How do you estimate the average valuated amount per application for damage to residential **inventory**, that is applied for at the WTS for the Limburg floods 2021?
3. How do you estimate the average residential damage to a single house in the area, with a maximum water depth from the house sill of **10cm**.
4. How do you estimate the average residential damage to a single house in the area, with a maximum water depth from the house sill of **35cm**.
5. How do you estimate the average residential damage to a single house in the area, with a maximum water depth from the house sill of **75cm**.
6. How do you estimate the average residential damage to a single house in the area, with a maximum water depth from the house sill of **150cm**.
7. How do you estimate the average residential damage to a single house in the area, with a maximum water depth from the house sill of **250cm**.

For the total damage section, five calibration questions were created. Three questions asked about historical flood events in the Meuse, and two concerned events in the UK and Germany. The expert answers were validated with historical data from Jongman et al. (2012) and Wind et al. (1999)

1. How do you estimate the total valuated damage around the Meuse that was applied for at the WTS for the flood in January 2011?
2. How do you estimate the total valuated damage around the Meuse that was applied for at the WTS for the flood in January 1995?
3. How do you estimate the total valuated damage around the Meuse that was applied for at the WTS for the flood in December 1993?

4. How do you estimate the total valuated damage in Eilenburg (Germany), due to the flood in August 2002?
5. How do you estimate the total valuated damage in Carlisle (UK), due to the flood in August 2005?

Experts

To answer the questions, a group consisting of 10 experts was gathered in multiple online sessions. Experts were selected based on flood damage expertise and can be divided roughly into two groups, based on their background: Model experts and damage experts.

Four participants were classified as model experts. The model experts had a scientific background and gained experience by working on flood risk studies at academic research institutions and/or governmental bodies. Some were also involved in Limburg floods, both the 2021 event as well as earlier large events in 1993 and 1995. These experts were selected because it was suspected they had a holistic view of the damage situation, across multiple sectors. However, these experts could be biased towards model thinking, and therefore unsuitable for validation purposes.

Subsequently, four experts were classified as damage experts. Damage experts have 'real' flood damage experience from the field, where they have gained expertise by determining the financial amount of damage to an asset for insurance companies or independent valuation institutions. All participating damage experts were highly involved in the 2021 floods and saw the effects that the flood had on people's direct life and property. These experts were particularly experienced in residential damage, and had less knowledge on other sectors such as infrastructure and company damage. Some had also experienced the large floods in 1993 and 1995. The two final experts were not classified as either Damage Expert or Model expert, but had broader flood experience[1].

The expert estimations were elicited in online sessions of approximately 2,5 hours. The context of the study, probabilistic training and sample questions were discussed at the start of the session. Subsequently the experts were elicited using an online form. Relevant information was provided together with the questions, and in an attached document.

Decision makers

For this experiment, 5 different Decision Makers[2] (DM) were created.One equal-weighted average of all experts, two optimized DMs based on the previously mentioned calibration scenarios, and two DMs based on expert background.

1. **DM-eq:** The equal weights DM gives equal weights to all experts, irrespective of their calibration or information scores on the calibration questions.
2. **DM-bench.**: This is an optimized[3] decision maker using calibration scenario one: a benchmark of damage models[4] as calibration data.
3. **DM-SSM**: This is an optimized decision maker using the damage model SSM - single family house as calibration data.
4. **DM-model:** This is a self-defined DM that assigns equal weights to all four experts with a modelling background.
5. **DM-de:** The damage expert DM is another self-defined DM, with equal weights to all four experts with a damage expert background.

[1] Expertise was more focused on macro-economical factors, indirect damage and overseeing of damage claims.
[2] Linear combinations of expert estimations, see Section 2.5
[3] The experts are combined in such a way that the statistical accuracy of the Decision Maker (in terms of calibration and informativeness) is maximised.
[4] The average of SSM single family house, MCM, HAZUS and Flemops was used, together with exposure value of SSM single family house

7.2. Results - Average residential

This section discusses the results of the average residential section. Here, experts were asked to give their estimate for average damage to a single building, for three locations, and for their estimate of the maximum damage that could occur to a building.

Calibration

Calibration[5] was done using two scenarios, the resulting expert scores can be seen in Table 7.2. The benchmark scenario includes expert 4 and expert 10, with a dominant weight for expert 10 due to a higher information score. The SSM scenario only includes expert 4, as expert 10 has a much lower calibration score in this scenario. All other experts are excluded in the optimized decision makers due to too low calibration scores.

Table 7.2: Expert scores and weights for estimating average residential damage, depending on the two calibration scenarios: Calibrating with benchmarked European damage functions and calibrating with the SSM damage function for Lowrise buildings. Experts are sorted on benchmark calibration score. Higher calibration score indicates better calibration. Higher information score indicates more informative experts.

Experts	**Expertise**	**Cal score**		**Info score**		**Weight**	
		Benchmark	SSM	Benchmark	SSM	Benchmark	SSM
Exp. 4	Model	0,0903	0,0903	0,4410	0,4410	0,295	1,0
Exp. 10	Damage	0,0903	0,0124	1,0241	1,0241	0,705	
Exp. 8	Damage	0,0016	3,20E-05	1,4825	1,4825		
Exp. 3	Model	4,30E-04	2,31E-05	1,0163	1,0163		
Exp. 5	Other	2,31E-05	4,30E-04	0,9560	0,9560		
Exp. 2	Model	7,87E-06	7,87E-06	1,4819	1,4819		
Exp. 7	Damage	7,87E-06	7,87E-06	2,6373	2,6373		
Exp. 9	Damage	7,87E-06	7,87E-06	1,4516	1,4516		
Exp. 1	Model	2,45E-07	2,45E-07	2,2274	2,2274		
Exp. 6	Other	2,45E-07	2,45E-07	2,1851	2,1851		

The calibration results give doubt to the robustness of the experiment due to three reasons. First, calibration scores rapidly decline after expert 4 and expert 10. This indicates that if accidentally these experts had not participated in the research, completely different results would have come out of the optimized decision makers. A robustness analysis supports this, with calibration scores dropping rapidly in both scenarios[6]. Second, the optimized decision makers have very large uncertainty bandwidths, making their applicability low. Third, the calibration is executed with data from literature. Although these functions are based on real flood events, their transferability to the Limburg region remains questionable. To increase robustness of the outcomes, we focus the conclusions of the experiment on DM-de and DM-model. These may be more reliable due to the inclusion of more expert estimations.

[5] Plots of calibration questions can be found in Appendix B Figure B.2 and Figure B.3
[6] Plots can be found in Appendix B Figure B.4

Average building estimates

The experts were asked to estimate average residential at three areas[7] with different damage characteristics. First Valkenburg, as it was a heavily damaged area with many affected houses. Second, Schin op Geul was selected as an upstream location with lower water depth. Third, an area in the municipality of Meerssen was selected. This third area has higher uncertainty due to an unkown waterdepth[8]

For each of the locations, the number of affected (ground level) houses was estimated in GIS using BAG data [9]. It was found that the area in Valkenburg consisted of 749 affected houses, Meerssen of 159, and Schin op Geul had 166 affected houses. The total residential damage calculated in the GFRT was divided by these numbers respectively.

A large spread is seen between individual expert estimations in Figure 7.1. Especially model experts indicate a large upper uncertainty. Contrarily, damage experts are more informative and give consistently lower Q-50 estimates.

Looking at decision makers, estimations of DM-de align closely with the GFRT's estimate for Schin op Geul and Meerssen, but larger error of twice the mean for Valkenburg. DM-model estimations correspond less well to the GFRT's estimate, but exhibit a similar pattern.

When comparing DM-de estimates between the three locations, it is found that the Q-50 estimate does not change significantly. Despite the communicated difference in inundation depth, experts do not expect significant damage differences. This may be explained by a difference in house surface area, as was indicated in the evaluation[10].

Moreover, all GFRT realizations are placed within the 90% confidence interval of both DM-de and DM-model which theoretically indicates their support of the model's validity. Nevertheless, these uncertainty bandwidths are 2-3 times as large as the estimated mean, which is much larger than the concluded uncertainty in Section 5.2, and the range suggested by Huizinga et al. (2017). This may have two conclusions. First, it may doubt the added value of the SEJ results as a means to reduce uncertainty. Second, it could indicate that total uncertainty is much larger than suggested in these earlier studies.

[7] Delineated using CBS defined neighbourhoods (buurten).
[8] The GFRT calculated with 0,5m depth, however this was manually inserted into the flood map.
[9] Exact procedure can be found in Appendix B
[10] Schin op Geul and Meerssen were larger, with average surface areas of 175 and 210 m^2, vs 100 m^2 in Valkenburg. Respective inundation depths where indicated as 0,5 - 0,75 and 1,0m

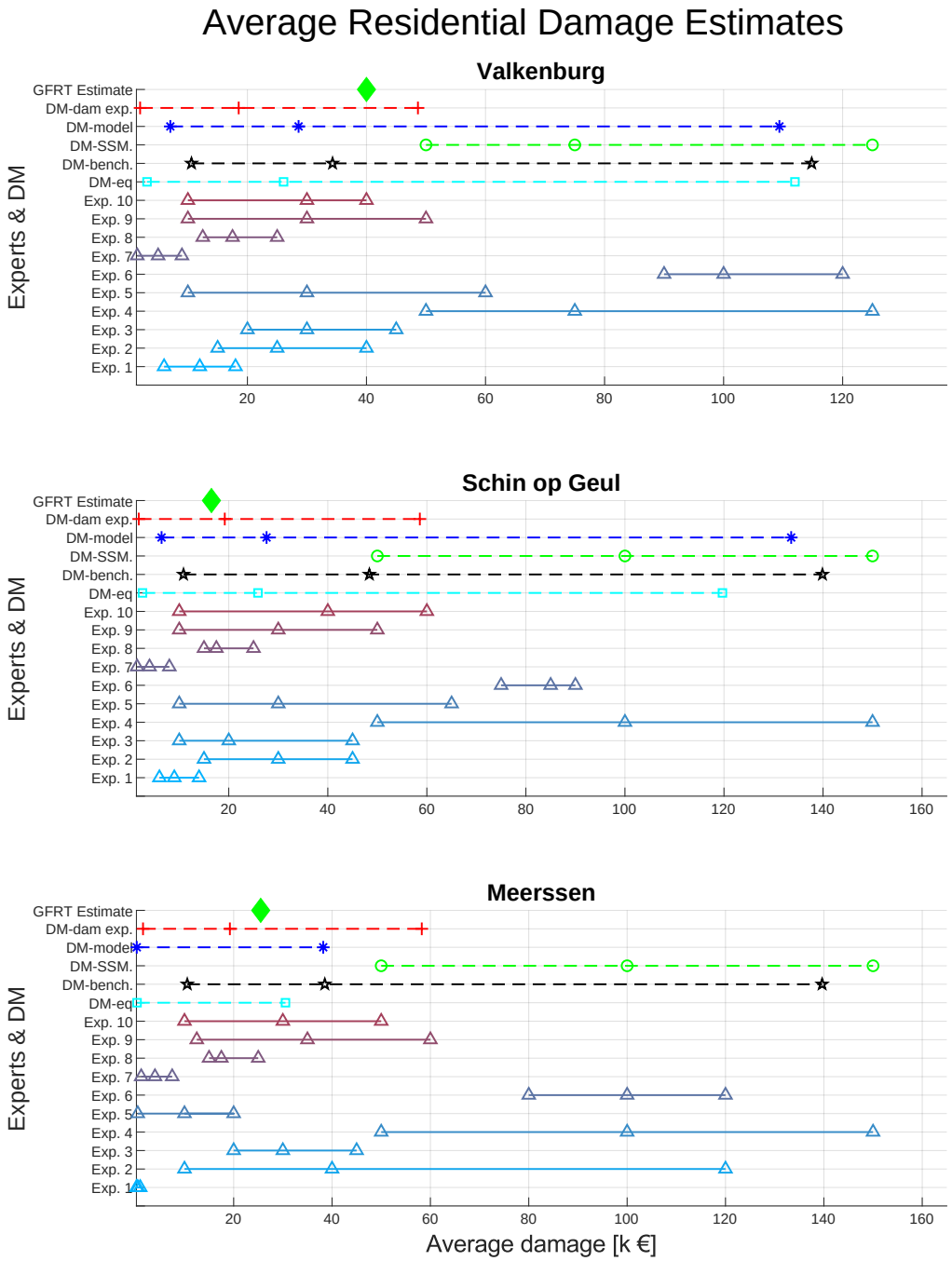


Figure 7.1: Results for the target questions on average residential damage for three locations. The GFRT estimate is shown on top, and compared with 5 DMs and the 10 single expert estimations.

Depth damage curves

Besides average damage, experts were asked to estimate damage at a specific water depth. Five questions asked about average residential damage at 10cm, 35cm, 75cm, 150cm and 250cm of inundation depth, measured from the door sill. Their answers are compared to the GFRT's vulnerability function in Figure 7.2.

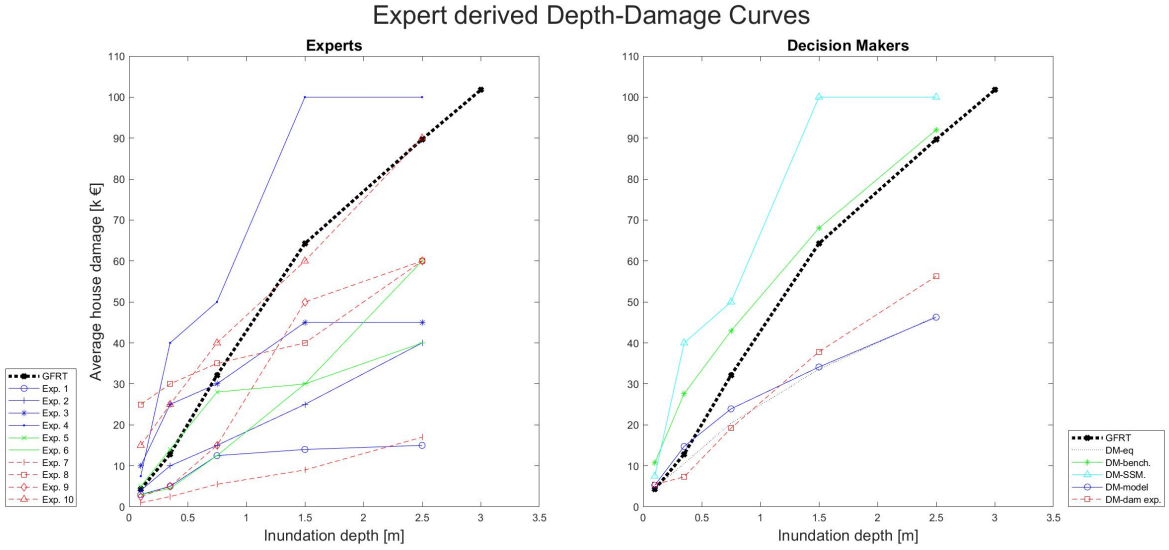


Figure 7.2: Depth Damage curves according to the 10 experts. On the left, damage experts are noted with a dashed red line (- -), whereas model experts are annotated with a blue extended line. On the right, decision makers consisting out of only model experts or only damage experts are indicated in the same method.

Looking at the left graph showing all individual expert estimations, the following three conclusions are drawn:

1. At inundation depths below 0,75m, half of the experts believe damage to be less than the used vulnerability, whereas the other half believes it to be more. At higher inundations, eight out of 10 experts believe the damage to be considerably less than modelled with the vulnerability curve. This is also confirmed through the right plot showing the equal weights decision maker.
2. It is seen that 3/4th of modellers believe that damage does not increase further after 1,5m inundation depth, whereas all other experts still indicate a significant increase between 1,5m and 2,5m inundation depth.
3. Estimates amongst damage experts are highly variable. One damage expert substantiated his deviating estimation by considering cellars. In his view, a large part of the residential damage in affected houses was attributable to rooms below ground level. 10 cm of water in the living room would mean large inundation depths on the lower levels, resulting in tens of thousands of euros in damages. Other damage experts did not indicate sub-terrain areas in their estimations, despite having seem the same affected location.

From these insights, we draw the conclusion that experts agree with the used vulnerability model for low depths up to 0,75m. Afterwards, experts believe considerably less damage will occur to houses than the model predicts. The right graph with DMs further supports this conclusion. However, the uncertainty of this conclusion should be acknowledged, given the strongly diverging individual estimations.

Maximum damage

In Section 5.1, it was found that the used exposure value for residential houses was uncertain. To validate the used data, experts were asked to estimate this exposed asset value. The value was defined as the damage that would occur in a total-loss situation[11]. Expert estimations for the exposure value of an average house in the area can be seen in Appendix B Figure B.1. Here, the exposure used in the GFRT is also shown at the top of the plot[12]

Looking at the results, it is found that both modellers as well as damage experts on average expect a higher exposure value. This coincides with the earlier found result that the exposure was underestimated by excluding inventory (Section 5.1). However, an underestimated exposure would also indicate an underestimated damage estimation. This contradicts the earlier conclusion where experts pointed at a model overestimation, in the absolute depth-damage curves.

In the evaluation, experts pointed out that this total-loss situation rarely occurred in the 2021 floods. Therefore, the used exposure value has not lead to an underestimaton. In the case that more extreme floods are modelled in future flood-risk studies, the exposure value could be altered. Correspondingly, the relative depth-damage function would require adjusting.

[11] For instance if a flood damage so large had occurred that the entire home would need to be rebuild and the entire inventory replaced.

[12] GFRT average house exposure was calculated using the same average house surface area of 100 m^2 that experts were told to assume. This resulted in a total exposed asset value of 1071 * 100 = 107.100 euro.

7.3. Results - Total damage

In the second section, experts were queried on total direct damage estimates. This included residential damage, infrastructure damage, direct business damage and agricultural damage.

Calibration

The resulting calibration scores, information scores and expert weights are shown in Table 7.3. Here, we see that expert 1 and 8 perform the best, being the only two experts in the optimized decision maker. For the global weights decision maker, Experts 3, 4 and 6 are also included for an additional weight of 11,7%. The other experts have a near zero weight based on their calibration and information scores.

Table 7.3: Expert weights for part 2: estimating total direct flood damage. Ranked on calibration score.

Experts	Expertise	Cal score	Info score	Weight	
				Global	Opt.
Exp. 1	Model	1,40E-02	1,305	0,304	0,348
Exp. 8	Damage	1,40E-02	2,302	0,571	0,652
Exp. 3	Model	1,81E-03	1,213	0,047	
Exp. 4	Model	1,81E-03	0,864	0,026	
Exp. 6	Damage	1,10E-03	2,112	0,043	
Exp. 5	Other	1,30E-04	2,748	0,006	
Exp. 2	Model	3,62E-05	1,405	0,001	
Exp. 7	Damage	1,58E-05	2,465	0,001	
Exp. 9	Damage	1,58E-05	2,226	0,001	

For the individual calibration questions[13], the two large floods in 1993 and 1995 were captured in the uncertainty of all decision makers. However, the much smaller Meuse flood in 2011 is not estimated accurately. Furthermore, also the estimates for Eilenburg (Germany) and Carlisle (UK) are far off. Despite their experience, no expert predicted the correct scale of these international floods. Most experts even estimated the damage to be lower than the Limburg floods of 1993 and 1995.

These results indicate that experts are not able to give reliable estimates for total direct damage. Although decent estimates are given for the known floods of 1993 and 1995, it can be concluded that estimates cannot be extrapolated to lower or higher extremes.

Total damage estimates

Due to the unsatisfactory calibration, the optimized and global decision maker were not used to draw conclusions. Instead, focus was placed on analysing the damage expert (DM-de) and model expert (DM-model) decision makers.

The target questions are presented in Figure 7.3. Looking at the damage experts, we see that the best estimates are 3 to 4 times lower than the GFRT realization for the Meuse and Geul. Contrarily, the Q-50 estimate for the Roer corresponds well. Moving on to the model experts, a larger uncertainty and slightly higher Q-50 estimations are seen.Both model experts as well as damage experts capture the GFRT realization within their uncertainty. However, the lower Q-50 estimates, for the Meuse, and to lesser amount the Geul indicates the expert's opinion that the GFRT currently overestimates damage for these areas.

Based on the gathered data, it may be concluded that experts agree with the model's prediction for total direct damage. However, as overall uncertainty bandwidths remain too large, this conclusion can only be of added value during initial stage validation. Furthermore, as the calibration questions indicated poor performance for unknown flood events, the method is likely unapplicable in unfamiliar terrain.

[13] Individual results are shown in Appendix B Figure B.5

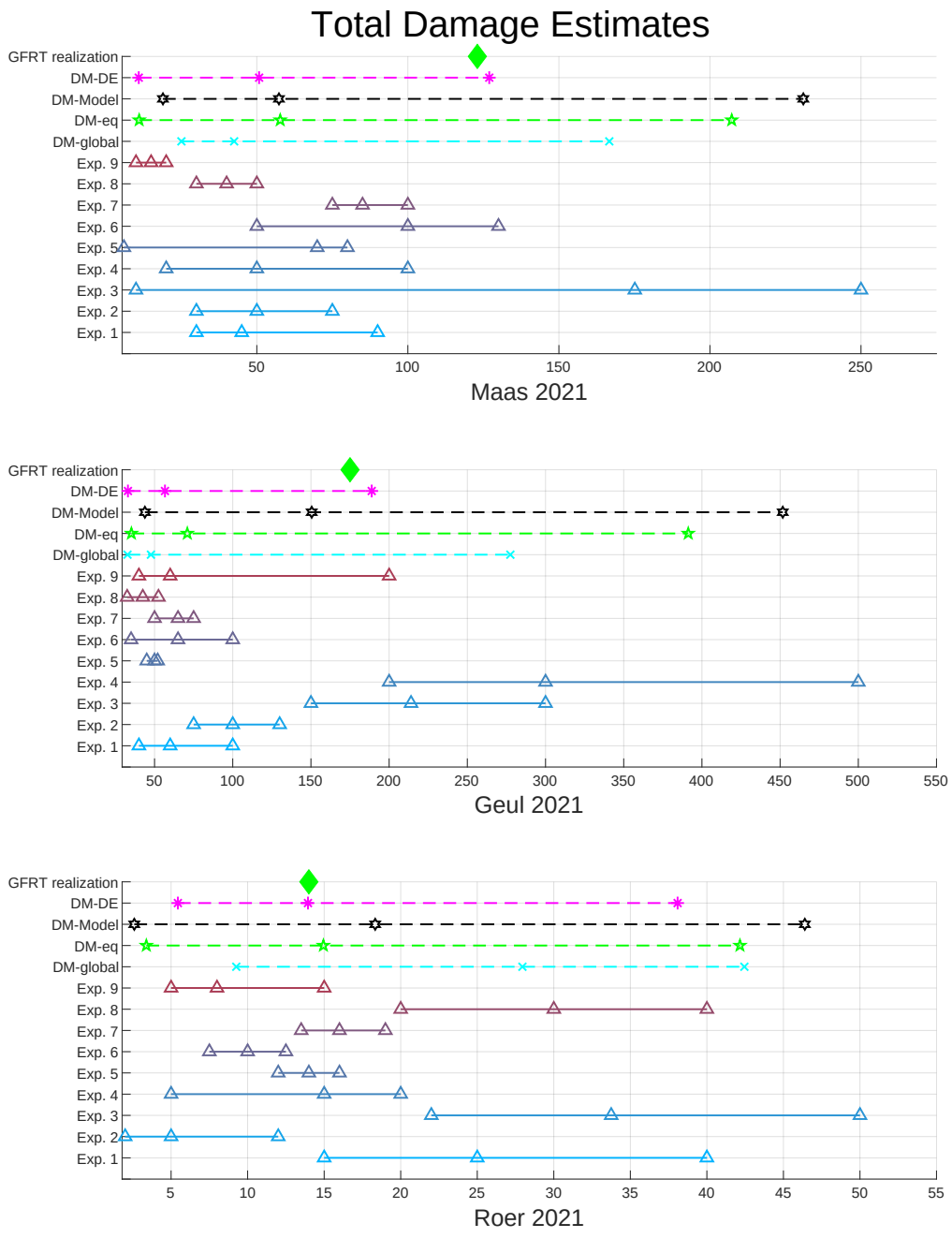


Figure 7.3: Expert estimations for total damage along the three main rivers in Limburg: Meuse, Geul and Roer. The GFRT estimate is shown on top, and compared with 5 DMs and the 10 single expert estimations.

7.4. Method evaluation

After the elicitation, experts were personally asked to evaluate the elicitation process. This explained some of the data, and may give guidance to future expert judgement experiments.

Expert feedback

The main difficulty for the residential damage section, was that experts were aware of the highly variable nature of damage to residential houses. Therefore, they found it difficult to estimate an average value.

Furthermore, experts indicated the significant influence of many uncertain variables as sill-height, flood preparation and flow velocity. Especially model-experts emphasised that these factors led them to give broad estimates. Also ambiguity in the type of house remained, with some damage experts underlining the huge difference cellars could make in an affected house. Some experts inferred that a 10cm inundation had a high probability that a cellar was completely flooded, leading to high damages.

The questions on total damage were evaluated as difficult, and experts indicated that these estimation scales were too large. Experts approached the problem by breaking down estimations in multiple sectors, and transforming the given residential damages. Here, the model experts communicated the use of theoretical approaches, by also including if the area had flood experience, and where it was likely located in the watershed. On the other hand, damage experts gave more importance to the economical factors by using GDP data and used impressions of the city The historical events were estimated as if the area would be flooded by that event today. As the damage pattern may have changed significantly due to different inventory, this may induce errors.

As the results indicate that damage experts may have higher informativeness and calibration, these experts seem to be more suitable for elicitation.

Opportunities and drawbacks of SEJ as a validation tool

On the positive side, this method of expert judgement can generate academic-quality data. By querying the experts on their opinion, without first sharing the model's result, an opinion with minimum bias can be created (O'Hagan, 2019). Furthermore, the application of the results are separated from the elicitation process itself, which also helps to minimize sources of ambiguity (Cooke & Goossens, 2000). Additionally, the personal evaluations with experts may give new insights into the damage behaviour of assets in that specific region (e.g. the previously mentioned cellars). This can highlight important differences that are difficult to identify from a theoretical perspective.

The evaluation also showed drawbacks of the applied method. The first problem that emerged was the high cost of time, both in the preparation from the researcher as well as during the elicitation from the experts. Especially the calibration questions add to the large workload of this research type. Second, the results of part two show that actual local expertise is highly important to elicit good estimations from the experts. However, the damage experts used in this case work for large insurance companies. If the insurance industry is not mature at the location of a flood damage study, then local damage experts will also not be available. It is plausible that this is the case especially in less-developed and data-sparse areas of the world. Therefore, no suitable experts may be available in data-sparse regions, where this study type is especially relevant. A final potential improvement point could be the inclusion of out-of-sample validation. When more resources are available, out-of-sample validation of optimized decision makers could strongly improve the reliability of the expert results, which could increase validity of the flood damage model.

The drawbacks and unsatisfactory results outweigh the positive points. The evaluations hint that repeating the experiment with better trained experts is unlikely to give improved results, as experts disagree inherently on the Q-50 estimations, include many other uncertain factors, and do not see the method connect well with reality. Therefore, the application of another type of expert elicitation method may be more suitable. The classical Delphi method could give higher consensus between expert estimations, however this may result in a biased answer that is not closer to reality (Hallowell & Gambatese, 2010). Furthermore, other estimation methods as without iteration rounds and calibration requirements, for instance staticized groups, could also minimize bias and reduce costs (Erffmeyer & Lane, 1984).

7.5. Conclusion & discussion

In this section, a conclusion on the SEJ results for the case study is given, followed by a conclusion on the applied SEJ method as a validation tool.

Case study conclusion

From the analysis of calibration and target questions, three conclusions were drawn for average residential damage.

First, the Q-50 estimations indicate that experts may agree with residential GFRT estimations for inundation depths below 0.75m. At higher inundation depths, the experts believe that the model could overestimate residential damage. Second, experts believe that the GFRT asset exposure is in itself underestimated by approximately 500 $€/m^2$. However, this parameter has not led to erroneous damage predictions, as experts indicated that total-loss situations occurred rarely in the 2021 flood. When more extreme flood scenarios are included in a flood risk analysis, it may be preferred to increase the exposure value. Correspondingly, the relative depth-damage curve will then need to be adjusted accordingly. Third, it was seen that expert uncertainty is too large for general applicability, and strongly questions the first two conclusions based on the Q-50 estimates. As SEJ uncertainty is larger than that of general literature, it brings up doubst on the added value of SEJ as a validation method (Huizinga et al., 2017).

For the section on total damage estimates, it was concluded that experts are not able to reliably estimate aggregated damage numbers. Although two acquainted flood events were estimated adequately, estimates were far off for three unknown floods with strongly deviating damage numbers.

SEJ as a validation method

From these results, it is concluded that Structured Expert Judgement is likely not useful for general validation procedures of flood damage estimations. As the overall uncertainty was found to be too large to increase confidence in the model, the q-50 estimations alone are of limited use. Furthermore, the evaluations indicate that experts are unlikely to improve their calibration and informativeness when repeating the experiment. Other drawbacks as the high cost, the requirement of scarcely available calibration data and lack of suitable experts (especially in data-scarce areas), further reduce the applicability of this method.

Two main benefits were found: first, the increased confidence in the model by acquiring expert estimations with minimum bias. Second, the qualitative insights on accurateness of model parameters and characteristics on the flooded area. However, these benefits can also be achieved through other, less costly, forms of expert judgement.

8

Limburg - Conclusion

The goal of this analysis was to validate the initial Global Flood Risk Tool (GFRT) estimation of 526,2 M€, for direct damage caused by the 2021 river floods in Limburg. By applying the proposed framework, a best estimate of 349,4 M€ was found, indicating an initial overestimation of 34%. Besides, an uncertainty range of 271,8 M€ (-23%) to 388,2 M€ (+11%) was concluded.

This was assessed using plausible assumptions and damage observations. Initial plausible errors were found with public spatial datasets for cropland and infrastructure, financial information and model-model comparisons with HIS-SSM. Subsequently, damage data from the WTS, damage experts and the insurance industry were used to verify the adjusted estimates.

First, plausible errors in agriculture and infrastructure sections indicated an overestimation by the model of at least 170,8 M€. A large number of floodplains and grassland were incorrectly classified as damaged agricultural land, and linear infrastructure was modelled poorly due to elevation and rasterization.

Second, estimations for residential houses were found to be highly uncertain between 62,9 to 134,3 M€. The main uncertainty source was the inundation depth, which was unknown for 51% of the affected residential flood extent. Furthermore, the exposure values and vulnerability function were not validated for this region. However, this uncertainty was not quantified due to lacking information. Total uncertainty may be reduced through the application of detailed damage data, for instance present in insurance databases. Besides the uncertainty, phase 2 assessed a plausible exposure in residential houses. Through a model-model comparison with HIS-SSM it appeared damage to inventory was excluded from the maximum damage value, leading to an underestimation of 23.7 M€. However in phase 3, no underwriting evidence of this error was found in high-level insurance data and WTS data. Therefore, it was concluded that the initial exposure value was accurate after all.

Third, damage to highly specific assets such as sluices, power- and water treatment plants was analysed. For these locations, it was concluded that the surface-based approach is too uncertain to draw any conclusions. Instead, the number of affected assets should be reported and the damage validated individually.

These results can be used as a validation base when applying the model to a probabilistic analysis. First, the modelled agriculture and infrastructure damage classes can be calibrated to reduce errors for the analysed event. Furthermore, the found uncertainty in residential assets can guide engineers in the uncertainty bandwidth they should expect for flood scenarios with different return periods. Finally, the identified assets where a surface-based damage calculation was deemed unsuitable can be assessed on a case-by-case basis, with additional focus on the indirect damage impacts.

III

Results - Beira coastal flooding

9

Beira - Case description

The second case study of this report is focused on the 2019 coastal flooding in Beira - Mozambique caused by hurricane Idai, illustrated in Figure 9.1. Similar to the previous case event, this chapter describes step 1 of the validation framework. The relevant background of the area is described, together with flood event characteristics. Finally, the applied data are summarized here.

Case background

Mozambique, is an east-African country with a total surface area of around 800.000 km^2 Today more than 40% of the population is living under the poverty threshold (Santos & Salvucci, 2016). Next to this, food insecurity and malnutrition is a major problem for the area, especially in the Northern and rural parts.

The country has a population of approximately 31 million people, and is characterised as very young and rapidly growing. 45% of the population is younger than 15 years old (CIA, 2021), and an annual growth rate of 2.8% per year is expected. Currently, more than 30% of the population is estimated to live in urban areas, which is expected to increase to 50% around 2040 (UN-Habitat, 2018). This means that all of Mozambique's cities will continue to undergo rapid growth in the near future, with an estimated total of 80.000 new households annually in urban areas. Already, around 80% of all urban households reside in an informal settlement. These settlements are characterised by inadequate infrastructure, no formal land registration and poor sanitation (UN-Habitat, 2018).

Three large urban areas are present in Mozambique: Nearby the capital Maputo in the South, around Beira in the middle of the country, and around Nampula in the North (CIA, 2021). Like these areas, most cities in Mozambique are located near rivers or the coastline, resulting in an estimated 60% of the population living in flood-prone areas (UN-Habitat, 2018). At the same time, the country is also frequently hit by other natural hazards. Tropical storms have hit the country every year on average for the past 25 years, and droughts are a common cause of widespread food insecurity and malnutrition.

Focusing on the case area, Beira is the capital of the Sofala province, and has around 500.000 inhabitants. The city has a major trade position due to the importance of its port, with a total throughput of 9.5 M tonnes of cargo (2016 estimate), including Petroleum products, containers and grain (GFDRR, 2019). Southwest of Beira, the river Buzi and Pungwe flow into the Indian ocean, cumulatively draining an area of approximately 60.000 km^2 (1/3 of the Rhine drainage basin).

Beira flooding event

On the night of 14-15 March, hurricane Idai made landfall at Beira city with winds of up to 200 km/hour. The wind itself inflicted major damage across the city and moved further into the land, causing widespread damage across Sofala province, and to much lesser effect damage in the neighbouring Manica province. After changing course and weakening, the storm re-emerged over Zambezia province. For the entire province of Sofala, initial impact estimates included 600 fatalities, 1600 injuries and 400.000 displaced people. Around 300.000 km^2 of land was flooded, including 750.000 hectares of cropland.In Beira city, the cyclone was reported to have damaged as much as 65% of the city's housing and infrastructure (Reach, 2019).

The second case study of this report, is focused on the coastal flooding in Beira caused by hurricane Idai, illustrated in the bottom figure of Figure 9.1. Due to a storm surge from the strong winds, multiple residential areas and commercial assets in Praia Nova, as well as the port in the northeast of the city were flooded by coastal water. Areas remained inundated for up to two weeks, increasing further risk to the population through disease and bad sanitation. One month later, on the 25th of April, the north of Mozambique was hit by a second cyclone Kenneth, worsening the situation in the country (GFDRR, 2019).

Main datasets

To model the coastal flood caused by hurricane Idai, datasets shown in Table 9.1 were used.

Table 9.1: Main Datasets used for the Beira 2019 case study. The top four datasets were used in GIS analysis and the GFRT. The bottom two datasets are external reports and used for validation.

Name	Description	Resolution	Author
Inundation map	Idai coastal flood inundation map created for validation of a risk analysis study. Validated using local highwater marks and aerial imagery.	10x10m	RHDHV (2019)
Land use map	A land use map with 20 different classes, including four residential zones differentiated on income, created in collaboration with local governments.	Polygon	RHDHV (2019)
Vulnerability curves	Combination of vulnerability curves and maximum damage data from Englhardt et al. (2019), Huizinga et al. (2017), and van den Berg et al. (2000).	-	RHDHV (2019)
Building footprint	Indicates geometric representation of buildings based on satellite imagery.	Polygon	Ecopia and DigitalGlobe (2017)
Mozambique PDNA	A post-disaster needs assessment for Mozambique executed in May 2019.	-	GFDRR (2019)
Beira PDNA	A PDNA executed only for the Beira urban area.	-	Municipality of Beira (2019)

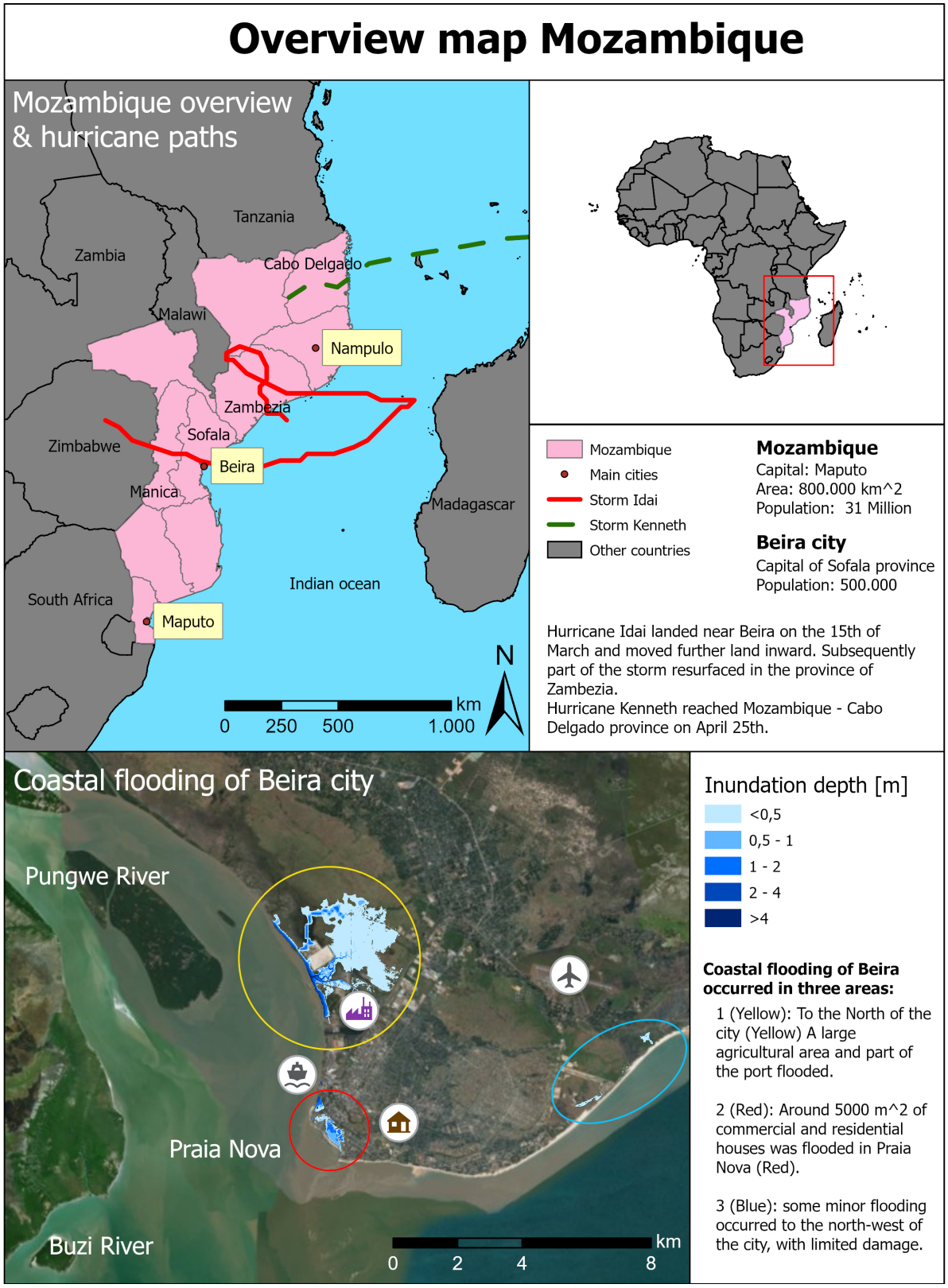


Figure 9.1: Overview of hurricane idai induced coastal flooding situation in Beira, Mozambique.

10

Beira - Model plausibility

This chapter applies steps two to four to the case study of Beira, as described in Chapter 3.

10.1. Input assessment

In this chapter the input data used for the damage modelling is considered. First, the different sources used for the input data are described. Second, the characteristics of the area are compared against the assumptions of the input data.

Hazard data

The used hazard data is sourced from an earlier project of RHDHV (2019) and was created by Deltares. Validation of the map was done internally using high-water marks and local reports. Inspection of the data itself shows an average inundation value of 0,45m, with a maximum depth of approximately 2m[1]. This indicates the section of the vulnerability curves that the damage estimation is especially sensitive to.

For this research it was also attempted to compare the map with any available public data, mainly satellite imagery and aerial pictures. Province-wide satellite-derived flood extent maps were available for multiple days during and after the flood event (Maxar, 2019; UNOSAT, 2019). These maps indicated large flood extents in Sofala, especially in the eastern Delta of the rivers Buzi and Pungwe. However, a very low agreement was seen between the applied coastal flooding map and the satellite-derived maps. These maps indicated that systematic underestimation of flood extents near dense vegetation, river sides and urban areas was likely. Therefore this validation step unfortunately did not give further information. Further validation was attempted with online aerial pictures taken during the event. Unfortunately only the locations of the images were set online, not the pictures themselves (INGC, 2019).

These attempts show the difficulty in validating the hazard aspect of flood modelling in data-scarce areas. As such, no further conclusion could be drawn on the uncertainty present in the hazard aspect.

[1] The histogram can be seen in Appendix C Figure C.1

Exposure and vulnerability

The exposure data consists of a land-use map with 17 different classes with an asset value larger than zero. Corresponding vulnerability functions can be seen in Figure 10.1.

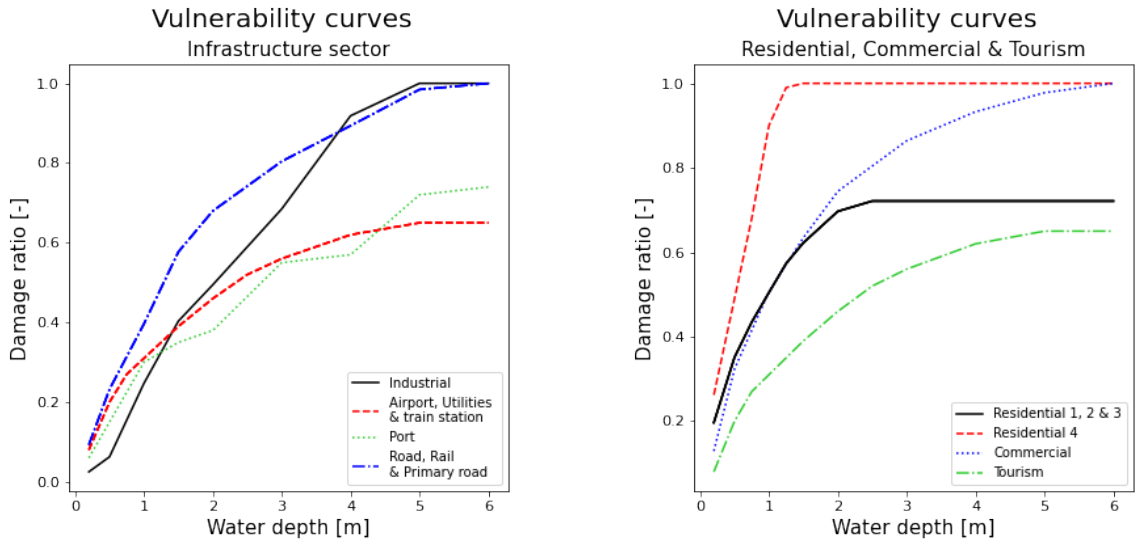


Figure 10.1: Beira vulnerability curves for the infrastructure sector (left) and general buildings (right).

Three out of four residential classes have the same vulnerability function, but differentiating exposure value. It is noted that the maximum damage ratio equals 0,72 for very high inundation depths, meaning not all asset value is lost. The fourth residential class represents informal housing, and has an exposure of approximately 20% less than the cheapest formal housing. Furthermore, a much steeper vulnerability curve is applied, and a maximum damage ratio of 0,9 from 1 meter inundation depth.

These curves are sourced from Englhardt et al. (2019) who created damage functions for an Ethiopian study based on (similar) structural properties (Rudari et al., 2017). For this application, weighted averages were taken to be aligned with the area's building properties. For the exposure value, an accurate building density factor is of major importance. Densities for Residential one to three were estimated at 0-10%, 10-20% and 20-30% respectively (RHDHV, 2019). The fourth (informal) residential class has no reported density.

Three asset classes represent the transport sector: Primary road, Road and Rail. Although exposure values vary significantly, all three assets have the same vulnerability function and are sourced from Huizinga et al. (2017). For linear infrastructure, the assumed width is important, which was found to be 20, 10 and 10 meters respectively. A single asset class corresponds to agricultural damages, with a fresh-water vulnerability curve from Huizinga et al. (2017). It is likely that these crops[2] were highly vulnerable when the cyclone hit (FAO, 2021). Finally, commercial and multiple industrial land use classes are created from the same source. Only the port asset class was created from van den Berg et al. (2000).

[2]For the entirety of Mozambique Maize, Sorghum, Sesame and Rice are reported to be of major importance for the agricultural sector (GFDRR, 2019). When the cyclone hit, these crops were in the middle of their growth stage, and therefore highly vulnerable (FAO, 2021).

From these observations, three conclusions were drawn:

1. **Source uncertainty:** although the data from Huizinga et al. (2017) is published for global applicability, large uncertainties are expected because the functions are not created specifically for this situation. Smaller uncertainty is expected in the functions of Englhardt et al. (2019) because they are validated in a similar environment, and adjusted to structural properties.
2. **Land use uncertainty:** Due to high informality in building and land use, a higher uncertainty may be expected in the land-use map.
3. **Vulnerability:** Damage ratio's of residential 1, 2 and 3 are capped at 0,72. This assumption of the non-destructible part should be aligned with the assumption for maximum damage. Furthermore, the damage ratio of the agricultural class does not assume saline water explicitly.

10.2. Model output validation

After assessing the model plausibility, this section applies steps three and four of the framework, analysing the output of the GFRT. Shown in Figure 10.2, total damage was estimated at 4,46 M$ for five damage classes. First, these classes are analysed using four sectors. Second, an analysis is done of the remaining classes, to check for underestimation.

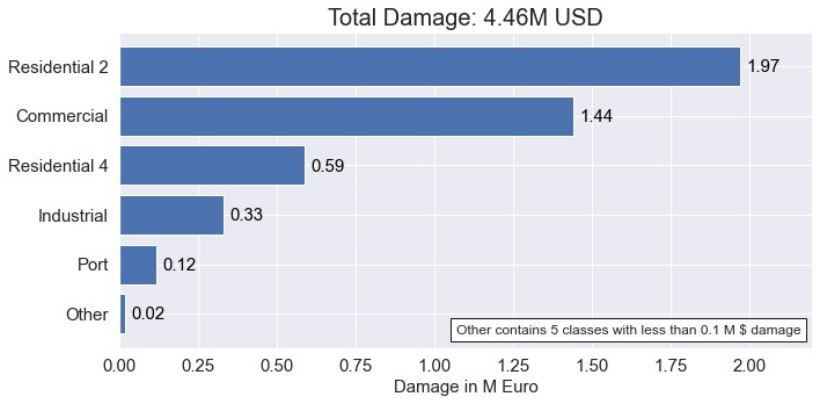


Figure 10.2: The outcome of applying the GFRT to the case area. Other consists of road, agriculture, drainage and residential 1 & 3.

Residential

Residential housing is the most damaged sector from the coastal flood, with impacts estimated at around 2,56 M$. The area is visualised in Figure 10.3. The sector is subdivided into two housing classes: Residential 2 (1,97 M$) and residential 4(0,59MM$). Residential 2 describes concrete structures with masonry filling and average exposure value, and will be referred to as formal housing. Residential 4 describes informal housing and will be referred to as such.

The formal housing is affected by an average inundation depth of 0,7m, with a maximum of approximately 1,5m. This corresponds to damage ratios of 0,42 and 0,62 respectively. Analysis of the affected buildings using a building footprint dataset indicated a total of 98 affected buildings, with an average of 167 m^2 of building surface area. Visual inspection indicated this included medium-rise building complexes, not only single-family houses. Using this data, a building density of 25% was found, which is slightly higher than the assumed 10-20% described in Section 10.1. Transforming the total damage resulted in 20.102 $/building or 120 $/$m^2$ of building. This is 14% higher than the building based damages of Huizinga et al. (2017)[3]. but falls within the general uncertainty bandwidth of -28% to + 53%.

[3] 2010 value of 134$$m^2$$ transformed with inflation factor of 1,82 to exposure of 245 $/$m^2$, then multiplied with the average damage ratio of 0,42.

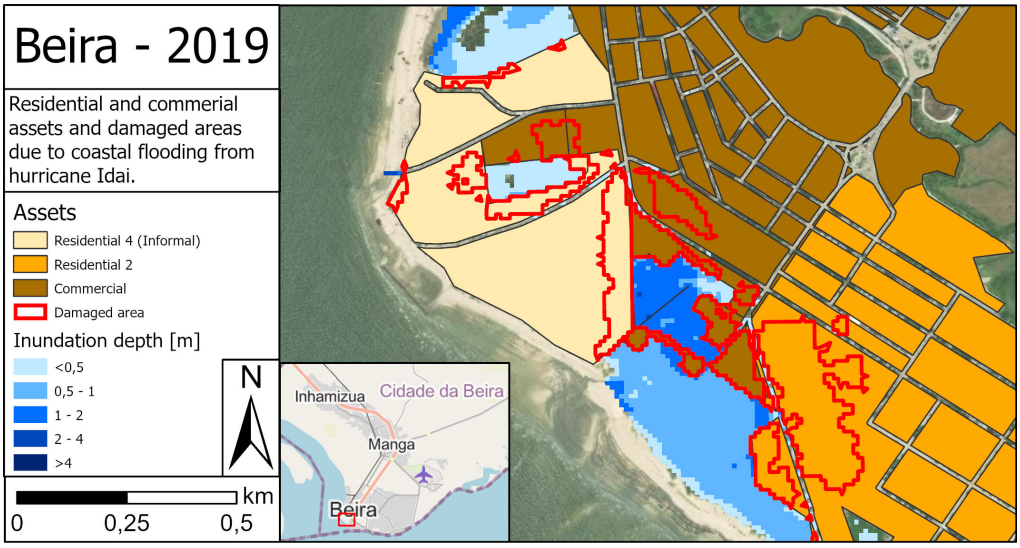


Figure 10.3: Overview map of main flooded areas in Praia Nova - Beira.

The informal housing is affected by a lower average inundation depth of 0,4m, with most houses being affected by 0 - 1m of depth. A small number of the area is affected by depths up to 2 meters. This corresponds to damage ratios of 0,41, 0,90 and 1,0 respectively. Using the building footprint dataset, a total of 310 affected houses were identified with an average surface area of 49,5 m^2. This resulted in a damage of 1806,45 \$/building or 36,49 $\$/m^2$. Building density was found to be 39%, but can not be used for validation, as no assumed building density was reported.

The damaged informal housing area is located next to the beach. Satellite imagery shows the presence of many moored fishing boats under normal circumstances. These boats are excluded from the damage calculation, but may be largely connected to the residential housing present here. Despite the exclusion of the model, large damages were likely incurred here, as the fishery sector was heavily damaged by the event and of major importance for Mozambique (GFDRR, 2019).

Commercial

Commercial is the second-highest damaged sector, with an estimated total of 1.45M\$ in damages. Shown in Figure 10.3, the damaged area is located between the formal and informal residential areas.

The modelled average inundation depth was 0,8m, with a maximum of 1,6m. This corresponds to damage ratios of 0,42 and 0,65. The building footprint data indicates a total of 93 affected buildings, with largely varying surface areas. Around 70 buildings have an area of less than 100 m^2, whereas the largest building has an area of 1041 m^2. From this data and visual inspection, it is assumed that the damaged area consists mostly of smaller stores and offices, and a few larger commercial stores.

Aerial pictures were taken most recently two years before and one year after the event, and can be seen in Appendix C Figure C.4. From these images, it was concluded that the land use of part of the area is uncertain. Shown in Figure C.4, the location below the road is classified as commercial, however visually this looks similar to the informal housing on the left of it. The commercial class has an exposure value of three times the informal housing exposure value. However, the informal housing damage ratio is twice as high as the more resilient commercial structures in this location. As the overestimating effect of the higher exposure is larger than the underestimating effect of the vulnerability, this may have led to an overestimation of damages. However, the area further south of that location is classified as derelict, whereas visual inspection shows more informal housing. The exclusion of these houses may lead to an underestimation, which may cancel out the previous overestimation.

No certain conclusion can be drawn from this data, except that the southern commercial area is highly uncertain and requires further investigation.

Industrial

The following section describes the third most damaged class: Industry. On the western coast of Beira, North of Praia Nova, flooding of two areas resulted in a modelled damage of approximately 330.000 $. A total area of 7.050 m^2 including a shipbuilding drydock and an area adjacent to a river outlet was flooded[4]. Due to the dry dock's low placement very high inundation depths of up to 8m were modelled, with maximum damage ratios of 1.

This estimation is likely incorrect due to two reasons: First, the used global depth damage function based from Huizinga et al. (2017), assumes many different buildings (e.g. warehouses) to be in the industrial class, whereas currently a very specific area is flooded related to shipbuilding. Second, the flooded area is of such small scale and specifically affects areas designed for inundation, that a surface-area-based calculation seems unsuitable. The actual amount of damage is highly uncertain and largely dependent on the current state of the area at the time of inundation relating to flood preparedness.

More significant was the modelling decision to exclude the port and quay area. Although debatable if damage to these assets is created by wind or water damage, high asset value and damage are expected to the fishing industry and their boats. A storm surge can heavily damage boats through high waves and strong mooring forces, resulting in collisions (AIR Worldwide, 2012). Furthermore, multiple container cranes, warehouses and buildings are spotted to be (partly) excluded from the land use map. Excluding these assets from the model can be valid, but should be communicated clearly as damage to the fishing industry was reported to be of major importance for Mozambique's economy and this event (GFDRR, 2019)

Port

The port section concerns a large area on the North-West side of Beira. Exposed assets are container terminals along the coast, silos and warehousing near the city, and a large petrochemical area on the northern side. Modelled damage totals around 100.000 $ due to two inundated areas.

The first area is close to the coast, where an elongated tongue of water has flown into the port where silos and containers are located. Damage is calculated only at part of the area, because the vulnerability curves assume that damage starts at inundation depths larger than 20 cm. The second area is located north of the petrochemical area. Using satellite pictures was found that port areas here were still under development. Two satellite images were available, from April 2017 and January 2020. As the area was not yet finished in the 2020 image, it was inferred that the area was under construction during the flood event. From this information, no quantitative conclusion was drawn. However again the uncertainty of the situation related to land use uncertainty was pointed out.

Roads and Agriculture

Roads and agriculture can be heavily impacted by flooding (Jongman et al., 2012), however in the current model they appear to be affected minimally. Roads account for approximately 14.000$ in the modelled damage, being affected in the Praia Nova area. An average inundation depth of 0,5m resulted in a damage ratio of 0,23. Images of the area indicate that this concerns low-speed paved roads. However, no further conclusions can be made on the accuracy of this estimation, and require further empirical data insights.

Agriculture accounts for even less damage of approximately 3.000 $. A large area to the north of Beira[5] was flooded with a low inundation depth of approximately 0,25m. However, due to the vulnerability function, damage is calculated only above 0,20m inundation depth. This results in only part of the area being modelled as damaged. An important error here is found in the assumptions, as the flood concerns saline water, whereas the vulnerability curve does not explicitly include this. This may result in an underestimation of the damages (Brémond & Grelot, 2013). However, even with a higher vulnerability, the total amount is likely negligent compared to the other damage sectors. An increase of 10 times would still result in total damages of only 30.000 $, which is much less than the other sectors.

[4] The area is visualised in Appendix C Figure C.2.
[5] Shown in Appendix C Figure C.3

10.3. Conclusion

The concluded errors and largest uncertainty sources are summarised in Table 10.1. Uncertain components are indicated with ↔, GFRT overestimations are indicated with ↑, and GFRT underestimations are indicated with ↓. For each class, a qualitative conclusion is drawn in the right column, considering all indicated uncertainty components.

For the three most affected damage classes, no conclusion could be drawn on over- or underestimation due to large uncertainties in applied vulnerability curves, maximum asset value and land use. For Residential 2, a minor underestimating error in surface density was found, however this is expected to be of minor importance compared to overall uncertainty. On the border between Commercial and Residential 4 assets, land-use uncertainty was found. Furthermore, the exclusion of fishery boats in the damage calculation was highlighted as an important assumption.

For the industrial area, an underestimation was concluded due to more expensive assets being inundated than assumed in the vulnerability curve. However, high uncertainty depending on the flood preparedness of the area remains a much larger, and unknown issue. Also damage around the port area is uncertain, mostly due to found differences between the area and the modelled vulnerability and exposure models.

Finally, agriculture damage is likely underestimated by the occurrence of a salt-water flood, to which crops are more vulnerable. It was validated that the flood occurred during the growing season and will therefore inflict heavy damage. Even when considering this, the class is insignificant for the modelled area. Besides the mentioned error roots, an unknown amount of uncertainty in the hazard analysis remains present.

Although this chapter gives important insights into uncertainty factors and errors, it remains a qualitative conclusion due to the lack of data. In the next chapter, a PDNA damage report is connected to the GFRT estimates to verify modelled damage and quantify possible errors.

Table 10.1: Overview of major error/uncertainty roots for each damage class. Symbols are used in the following definition: ↔Uncertain component. ↑The error results in an GFRT overestimation. ↓The error results in an GFRT underestimation. A qualitative conclusion on underestimation or overestimation of the damage class is presented in the final column, together with important assumptions. DR Denotes damage ratio.

Class	**Vulnerability**		**Exposure**		**Conclusion and Assumptions**
	Temporal	DR	Dam_{Max}	Assets	
Residential 2		↔	↔	↓	**Uncertain** Dam. ratio and Max. Dam. outweigh minor asset density underestimation.
Commercial			↔	↔	**Uncertain:** global vulnerability curves and uncertain S southern land use area.
Residential 4		↔		↓	**Uncertain:** Dam. Ratio and Max. Dam. outweigh extra assets in derelict area. • Fishery boats are excluded.
Industrial	↔	↑	↓	↓	**Underestimation:** Higher Max. Dam. expected due to machines and assets. Uncertainty depending on flood preparation • Fishery boats in the port are excluded
Port			↔	↔	**Uncertain:** Model assumes breakbulk terminal while the flood affects petroleum assets under construction.
Roads	-	-	-	-	-
Agriculture	During growth	↓			**Underestimation:** Model assumes freshwater whereas a saltwater flood occurs. • Flood occurs during growth season

11

Beira - Verification using damage observations

This chapter applies steps five and six of the proposed framework to the Beira case: Empirical data analysis and verification (Figure 3.1). Here, empirical data is transformed to enable a comparison between the model's outcomes and real-world observations.

11.1. Analysis approach

After hurricane Idai made landfall in 2019, a team led by the government of Mozambique started assessing all damages through a Post Disaster Needs Assessment PDNA methodology (GFDRR, 2008, 2019). Members of central, provincial, district and municipal levels were trained in this methodology to estimate damages of the hazard, which aims at reducing bias. Furthermore, quality control is assumed as the report forms the basis of international financial aid and lending.

The PDNA estimated the total damage of the hurricane to be over $ 1,4 Billion. This amount is divided over several sectors, the largest one being the transport sector with an estimated $ 442 M of damages. The second most damaged sector is housing, with $411 M. Subsequently, industry & commerce was affected with $ 140 M in damages. Other damages were inflicted in the energy-, environmental- and agriculture sectors[1].

The report was created by the Mozambique government, and supported by the World Bank, The United Nations and the European Union. The report indicates damages per sector for the provinces of Zambezia, Tete Manica, Sofala, and Inhambane. The exact damage reporting format differs per sector and province. For Sofala, generally the number of affected assets (houses, km of road) was mentioned, together with a qualitative description of the damaged components. This damage is attributable to strong winds, pluvial-, fluvial- and coastal flooding.

As only damage by coastal flooding is modelled, the damage numbers were transformed using two approaches. First, a comparison between Sofala-wide damages and damage in Beira was enabled by transforming the total amounts to $/damaged asset. Second, a consideration was made on how the reported damage would relate to the modelled flood damage. This was done either by comparing wind and flood fragility curves or by using the qualitative descriptions in the report. Construction-cost data could be compared to modelled damages for some sectors. As these data were generally older, a price transformation was required using GDP-deflator-based inflation data from the World Bank (2022). Currency conversions were executed using an exchange rate of 1:64 $ to MZN (Mozambique Meticais) (GFDRR, 2019). To exchange euro to US Dollar, exchange rates of Macrotrends (2022) were used.

[1] Besides damages, another $1,39 Billion in losses was reported, mainly in the agricultural and industry & commerce sectors ($513 M and $470 M respectively). As this research focuses only on direct damages, only numbers from the damage categories were used for the analysis.

11.2. Residential damage

This section will describe the analysis of the housing sector in Beira. First, PDNA damage data reports actual damages to the housing sector. Second, this damage data is disaggregated to acquire flood damage for informal houses. This is done in three steps: 1) Comparing wind and flood vulnerability curves, 2) applying a scenario analysis for hurricane idai, and 3) disaggregating the final data. Third construction data is used to verify Residential 2 assets. Fourth, a synthesis of these analyses resulted in the creation of two uncertainty bandwidths, shown in Table 11.3.

PDNA damage data

In the top two rows of Table 11.1 reported PDNA damage is shown, besides modelled GFRT damage in the bottom two rows. The two PDNA reports indicate total damages of 410 M$ for the whole Sofala province, and 184,4 M$ in the Beira area (GFDRR, 2019; Municipality of Beira, 2019). This translates to average damages of 1.708 and 2.111 $/house respectively.

Although the raw compound damage data in the PDNA report is not similar enough to our case study to draw any conclusions, an initial comparison is made. The informal class Residential 4 compares closely to the PDNA data. On a damage/unit basis the modelled 1806 $ differs +5% and -17% from the reported numbers of for Sofala and Beira. Differences become larger on a damage/ house surface area basis, with the model giving a +20% and + 35% higher estimation. Contrarily, the modelled damage of formal housing classified as Residential 2 seems strongly overestimated, with modelled damage being four times higher than the reported damage[2].

Table 11.1: Overview of found damage data by hurricane Idai and modelled numbers by the GFRT. Dam_{unit} is calculated as dam_{total} divided by the number of assets. Surface damage refers to the damage per m^2 of house.

Source	Asset	$\mathbf{Dam_{Total}}$ M$	Assets -	$\mathbf{Dam_{Unit}}$ $	$\mathbf{Area_{Unit}}$ m^2	$\mathbf{Dam_{Surface}}$ $\$/m^2$
PDNA	Sofala[3]	410	240.000	1.708	72,5[4]	23,6
	Beira[5]	184,4	87.339	2.111	72,5	29,12
GFRT	Residential 2	1,97	98	20.102,04	170	120,39
	Residential 4	0,56	310	1.806,452	50	36,49

Comparing vulnerability curves

To analyse the relative contribution of wind and flood damage to the reported damage numbers, fragility curves made by RHDHV and Celsius Pro (2020), were compared for the case situation in Figure 11.1. Both curves calculate damage as a fraction of total building value, meaning object exposure is equal for flood and storm damage.

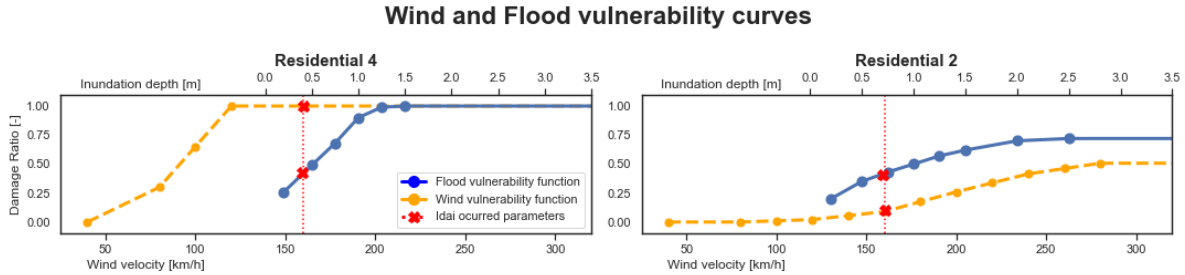


Figure 11.1: Fragility curves for flood and storm hazards, made by RHDHV and Celsius Pro (2020).

[2] On a damage per surface area basis. The difference on a building level is even higher, but this comparison is illogical as the GIS model showed these numbers concern highrise building complexes, not individual houses.
[3] Source: GFDRR (2019)
[4] Found by using a weighted average of 98 Residential 2 buildings and 310 Residential 4 buildings
[5] Source: Municipality of Beira (2019)

Looking at informal housing on the left of Figure 11.1, both wind and flood damage curves can reach the maximum damage ratio of 1,0,In this case, complete building value is lost and both hazards damage the same components. Therefore it is argued that DR_{Flood} and DR_{Wind} can be related to each other and are additive with a maximum of 1,0, as shown in Equation 11.1.

$$\begin{aligned} &\textbf{Informal housing} \\ &\text{Independent flood damage} = DR_{Flood} * Exposure_{Object} \\ &\text{Independent storm damage} = DR_{Wind} * Exposure_{Object} \\ &\text{Compound damage} = Min(DR_{Flood} + DR_{Wind}; 1) * Exposure_{Object} \end{aligned} \tag{11.1}$$

Describing formal housing compound damage requires three assumptions: First, the maximum damage ratio will always be lower than 1. The asset class consists of reinforced components, that are assumed to be indestructible. Second, a compound event will have a higher maximum damage ratio than a separate hazard. As strong winds generally damage the top of the structure whereas floods damage the bottom, it can be inferred that both hazards affect different building components (Baradaranshoraka et al., 2017). Third, the combined maximum damage ratio depends on the correlation between the exposed building components, as elements can not be destroyed twice.

The resulting compound damage ratio is described in Equation 11.2, but is uncertain due to the unknown correlation. The damage ratio likely ranges between the maximum of each single event damage ratio and either the sum of both separate damage ratios, or a ratio lower than 1,0.

$$\begin{aligned} &\textbf{Formal housing} \\ &\text{Compound DR} = DR_{Flood} + DR_{Wind} \\ &\qquad - DR_{Correlated}(Flood, Wind) - DR_{Undamageable} \\ &\text{Lower bandwidth DR} = Max(DR_{Flood}; DR_{Wind}) \\ &\text{Upper bandwidth DR} = Min(DR_{Flood} + DR_{Wind}; < 1) \end{aligned} \tag{11.2}$$

Flood & wind scenario analysis

By using Equation 11.1 and Equation 11.2, the compound event damage ratio for the damaged areas is estimated in Table 11.2. This is done by fixing the occurred inundation depth and wind speed in a scenario. Average inundation depths of 0,4m and 0,7m were extracted from the GIS model. Reported wind speeds are uncertain, with sources indicating values from "Larger than 120 km/h" to 180-200 km/h (Probst & Annunziato, 2019; RHDHV, 2019). Based on this, an average wind speed of 160 km/h was assumed. The corresponding damage ratios were extracted from fragility curves in Figure 11.1.

Table 11.2: Comparison of wind and flood damage to formal and informal housing classes using the scenario analysis. Residential 1 and 3 only experienced pluvial flooding and wind damage. DR denotes Damage Ratio.

Class	Wind velocity km/h	Inundation depth m	DR_{Wind} -	DR_{Flood} -	$DR_{Compound}$ -	Destroyed -
Informal housing	160	0,4	1,0	0,41	1,0	Partially
Formal housing		0,7	0,1	0,41	0,41 - 0,51	Completely
Residential 1 & 3	160	Pluvial	0,1	-	> 0,1	Partially

For formal housing, the maximum damage ratio of 1 is already reached at wind speeds of 100 km/h, and is therefore insensitive to wind speed uncertainty. At the assumed inundation depth, the flood damage ratio is 0,41. Therefore, it is concluded that all buildings are completely destroyed, and that wind damage has a higher impact than flood damage for informal structures. Contrarily, formal houses are only partially destroyed by wind and flood hazards due to their stronger structure. Furthermore, flood damages are more impactful than strong winds in the case area, as the theoretical DR_{wind} is lower than DR_{flood}. Other formal housing classes (Residential 1 and Residential 3) share the same vulnerability function as Residential 2, and are therefore also classified as partly destroyed.

Disaggregating flood & wind damage

Subsequently, the found state of informal and formal housing assets is used to disaggregate the damage data. The PDNA made by the Municipality of Beira (2019), estimates that 23.833 houses are completely destroyed and 63.506 are partly damaged in the Beira area, resulting in 184,4 M$ of damages. The damage is disaggregated in Equation 11.3, Equation 11.4, and Equation 11.5. First, the total damage is split up into damage for completely destroyed buildings, and damage for partially damaged buildings.

$$\begin{aligned} \text{Total PDNA damage} &= \text{Damage completely destroyed} \\ &\quad + \text{Damage partially damaged} \\ \text{Total PDNA damage} &= N_{Destroyed} * DR_{Destroyed} * Exposure_{Destroyed} \\ &\quad + N_{Damaged} * DR_{Damaged} * Exposure_{Damaged} \end{aligned} \tag{11.3}$$

Assumptions made to fill in this formula are given in Equation 11.4. First, completely destroyed houses are assumed to be informal houses (indicated with subscript $_{Destroyed}$), thereby equalling informal exposure values. Second, the assumption is made that partially damaged houses are formal houses (indicated with the subscript $_{Damaged}$). Third, The object-based exposure of informal and formal houses differs by a factor 6,6[6]. Fourth, the uncertain damage ratio of the partially destroyed buildings is varied as a sensitivity analysis, within the likely $DR_{Compound}$ range found in Table 11.2 .

$$\begin{aligned} &\text{Assumption:} && Exposure_{Destroyed} = Exposure_{Informal} \\ &\text{Assumption:} && Exposure_{Damaged} = Exposure_{Formal} \\ &\text{Assumption:} && Exposure_{Destroyed} * 6,6 = Exposure_{Damaged} \\ &\text{Sensitivity:} && DR_{Damaged} = [0,10; \quad 0,25; \quad 0,50] \end{aligned} \tag{11.4}$$

Subsequently the assumptions and sensitivity parameters were filled into Equation 11.5, and subsequently solved for the exposure value of informal housing. This resulted in disaggregated exposure values extracted from the PDNA. The found exposure value in $/house is transformed to an exposure value in $/m^2 of LU area, using average building surface area and density found in Section 10.1.

$$\begin{aligned} 184,4M\$ &= 23.833 * 1,0 * Exposure_{Informal} \\ &\quad + 63.506 * Exposure_{Informal} * 2.5 * DR_{partial} \\ \text{Exposure}_{\text{Informal}} &= [2.809; \quad 1.444; \quad 792]\$/House \\ &= [56,75; \quad 29,02; \quad 16,00]\$/m^2 building \\ \text{Building density} &= 39\% \\ \text{Exposure}_{\text{Informal}} - \text{LU} &= [22,13; \quad 11,32; \quad 6,24]\$/m^2 \end{aligned} \tag{11.5}$$

When comparing the disaggregated exposure values to the model exposure value of 26,29, it is found that the assumption is outside of the upper bandwidth acquired from the PDNA report. This indicates that the model may have overestimated informal housing damage. The lower bandwidth is 63% lower than the used input data and is therefore used as the lower limit. These values are used to create a new damage estimation in Table 11.3.

[6] Land use exposure values of residential 1-3 and residential 4 differ a factor 2.5 (Englhardt et al., 2019; RHDHV, 2019). Combining this with found object densities of 15% and 39% indicates 2,5 * 0,39/0,15 = 6,6

Residential construction data

Construction data is used for the formal housing class, as a large uncertainty remains around its damage estimation. CAHF (2019) reported a minimum house construction price of $ 55.000 from official contracting companies. When multiplying this value with the same damage ratio of 0,41, and dividing over the reported house surface area of 65 m^2, damage of 347 $\$/m^2$ is found. This is much higher than the reported damage of 120,30 $/m^2. However, this is likely an overestimation as the structure has an un-damageable part, and is a depreciated asset instead of a new construction (Huizinga et al., 2017). Therefore, it is used as a suitable upper bandwidth. The average PDNA reported housing damage of 29,12 $\$/m^2$ is used as a lower bandwidth.

Total damage estimates are created for both classes by scaling initial damage, summarized in Table 11.3. Formal housing is scaled using the found damage/building surface number, whereas informal housing is scaled using exposure values. A lower- and upper bandwidth was created for formal housing, whereas for informal housing, a decrease in damage is suggested.

Table 11.3: Uncertainty bandwidths for modelled damage of formal and informal housing

Class	Estimation	$\mathbf{Dam}$ [7]	Exposure	Difference	$\mathbf{Dam_{total}}$	Comment
	-	$\$/m^2$	$\$/m^2$	-	$M\$$	-
Formal housing	Original	120,30		-	1,97	
	Lower	29,12		-76%	0,48	Beira average PDNA
	Higher	347,00		+282%	5,56	Official contracting cost.
Informal housing	Original		26,29	-	0,56	
	Lower		6,24	-76%	0,13	Disaggregated PDNA
	Higher		22,13	-16%	0,47	Disaggregated PDNA

11.3. Commercial damage

The PDNA indicates data of 311 damaged businesses disaggregated over 10 sub-sectors. For each subsector the number of businesses, number of affected employees and total estimated damage are reported. The largest part of the reported damage was attributable to wind, with damages to roofs of warehouses and offices, with only the industry sector reporting flooded materials and machinery.

Table 11.4: PDNA reported damage for ten subsectors in Sofala. Ranked on damage per business.

Subsector	$\mathbf{N_{business}}$	$\mathbf{N_{employees}}$	Damage	$\frac{\text{Employees}}{\text{Business}}$	$\frac{\text{Damage}}{\text{Business}}$	$\frac{\text{Damage}}{\text{Employee}}$
	-	-	k $	-	k $	k $
Agro business	17	735	28.493	43	1.676	38,8
Industry	28	2.745	31.249	98	1.116	11,4
Fishery	6	2.314	6.311	386	1.052	2,7
Transport	35	2.560	14.291	73	408	5,6
Commerce	53	2.397	17.166	45	324	7,2
Construction	36	1.590	6.337	44	176	4,0
Hospitality	29	589	4622	20	159	7,8
Services	**64**	**2.152**	**9.885**	**34**	**154**	**4,6**
Poultry	35	143	803	4	23	5,6
Planing[8]	8	292	147	37	18	0,5
Total average	**311**	**15.517**	**119.304**	**50**	**384**	**7,7**

The extracted data can be seen in Table 11.4. The first conclusion from the data is the large spread of damage across the subsectors. The least damaged companies in the Planing industry report average damages of 20.000 $, whereas the most damaged companies in the Industry and Agro sectors report average damages of 11,1 and 16,8 M $. A factor of 93 difference between the minimum and maximum.

[7] Damage per building surface area
[8] Woodworking industry, e.g. Sawmills

It is expected that the flooded companies in the modelled area are most comparable to the services sector (GFDRR, 2019). There are two main differences between the data and the modelled assets. First, the PDNA services data concerns mostly wind damage, which is expected to be lower than flood damage as it was seen in Figure 11.1 that wind damage is expected to be lower than flood damage for formal buildings.

Second, as the companies hire an average of 50 employees, it was concluded that this sample only contains medium to large companies. Contrarily, it is expected that the flooded companies are much smaller, due to their central location. Therefore, a damage comparison on a unit basis is unsuitable. Instead, a methodology based on the estimated number of employees was applied. This is done more frequently in flood damage assessments, for instance by Arrighi et al. (2013) and Tezuka et al. (2014) and for economic loss-assessment by Herath et al. (2003).

The companies in the GFRT were analysed using GIS and a building footprint dataset (Ecopia & DigitalGlobe, 2017). From this analysis, it was found that 93 different buildings were affected, with surface areas of 10 to 1000 m^2. An estimated number of employees was assigned to each feature, based on its surface area[9], which resulted in an estimated range of 478 to 1003 affected employees. Subsequently in Equation 11.6, these numbers were multiplied by the average damage per employee reported in the services sector, giving an estimated commercial damage between 2,2 and 4,61 M$.

$$\begin{aligned} Damage_{Total} &= N_{employees} * \frac{Damage_{Services}}{Employee} \\ \text{Lower bandwidth} &= 478 * 4600 = 2,20M\$ \\ \text{Upper bandwidth} &= 1003 * 4600 = 4,61M\$ \end{aligned} \tag{11.6}$$

This bandwidth exceeds the GFRT estimate by 52% to 220%, which could be due to multiple reasons:

1. **Estimated number of employees:**The largest assumption in this calculation is the estimated amount of employees. The calculated damage scales linearly with this assumption, and therefore has high sensitivity. Unfortunately, it was not possible to get local data on this estimate for further validation.
2. **Reported number of employees:** No background information on the definition of the reported number of employees in the PDNA is given. Therefore it is uncertain which criteria were chosen to determine this number, which could significantly affect the average damage per employee.
3. **Average damage per employee:** The average damage per employee for the large reported business is now applied immediately to the modelled businesses. It is plausible that larger companies have more advanced (and valuable) assets and materials, than smaller companies (GFDRR, 2019). Therefore, the amount of damage per employee could be lower for smaller enterprises, explaining the difference.

[9]Buildings were divided into 4 classes based on their surface area. Buildings with less than 25 m^2 were assigned 1 employee, buildings with less than 50 m^2 were assigned 5, less than 250 m^2 received 10, and more than 250 m^2 were assigned 25. Assuming an average of 10 - 20 m^2 per employee resulted in a bandwidth of 478 - 1003 employees, supported by industry standards in US (Matheney, 2022; van Ramshorst, 2019)

11.4. Other damage classes

Industrial

Reported industrial damages can be seen in Table 11.4. The sector is indicated as one of the most severely impacted, with an average damage of 1,1 M$ per business. According to the report, this damage is mostly attributable to flooded machinery and work equipment and therefore comparable to the modelled case.

In the GFRT, an industrial port area north of Praia Nova was flooded. It was concluded that the situation was highly uncertain due to the specific flooding of a dry dock. A total flooded area of 7050 m^2 resulted in a calculated damage of 330.000 $. Aerial pictures indicate that the flood affects a single large company that operates the drydock[10].

Although many uncertainties remain on the flood preparation and exposed assets of this specific situation, the data points out that damage is likely underestimated. The average reported industrial damage number of 1,1 M$ is 3,3 times as large as the modelled number.

Port

Total damage of 12.1 M $ is reported due to damages in the coal terminal fuel terminal and four quays. Further moderate damage was sustained to lightning towers and warehouses, alongside operational facilities and dredgers. From the description, this damage seems mostly attributable to wind damage (GFDRR, 2019).

In the GFRT port damage was the fourth largest damage contributor, with areas located to the north of Beira. Here damage of 120.000 $ was calculated on a flooded petroleum site. As a complication, aerial pictures showed this site was under construction at the moment of flooding (Section 10.1).

The PDNA reported damage is difficult to disaggregate and mostly attributable to wind damage. Nevertheless, the reported damages are significantly larger than the modelled damage. Even if only 10% of the PDNA reported damage was caused by flooding, the data would still point out a factor ten underestimation. Therefore, it is concluded that the GFRT likely underestimates this situation.

Road infrastructure

Modelled damage to roads is minimal, whereas the PDNA reports it as a main damage sector. The applied verification consists of three parts. First, total damage reported in the PDNA is disaggregated to find road damage per lane length. Second, this damage is compared to the damage found in the GFRT. Third, benchmark construction cost data from African road-building projects are compared to exposure values.

Damage to the transport sector is estimated at 442 M$. Of this amount, 310 M$ describes damages to the road network. Damage to the road network is subsequently divided into four infrastructure types: 1962 km of national roads, 90 culverts, 15 bridges and 24 drifts. Shown at the top of Table 11.5, multiple assumptions on the percentage of damage that is attributable to the 1962 km of roads were tested. Subsequently, these amounts were transformed to obtain damage in $ per m of road[11] Lower in the table, the two road classes modelled in the GFRT are shown. All damaged sections are classified as ‘road’, whereas a single ’main road’ section exists in the city but was unaffected by the flood.

[10] Shown in Appendix C Figure C.2.

[11] In this analysis, it is assumed that roads are not susceptible to wind damage, and that therefore all damage is attributable to flooding.

Table 11.5: Analysis of reported PDNA road damage, GFRT road damage and benchmarked African road construction costs.

Reference	Total dam. M$	Assumption M$	Road dam. M$	Dam. / m $ / m
PDNA Damage (GFDRR, 2019)	310	10%	31,01	15,80
		25%	77,51	39,50
		50%	155,03	79,00
		75%	232,54	118,50

Reference	Class -	Exposure $ / m	Dam. ratio -	Dam. / m $ / m
GFRT	Road	21,20	0,23[12]	4,90
	Main road	227,80	-	NA
Construction Cost (AfdB, 2014)	Rehabilitation	169,0		38,90
		362,10		83,28
	Construction	295,70	0,23	68,00
		458,70		105,50
(Africon, 2008)	Rehabilitation	603,00[13]		138,80

Comparing the PDNA damage with GFRT damage shows that reported damages are much higher than the modelled road damage. Depending on the assumption, reported damages per m length are 3 to 23 times as large as the modelled damage. It could be hypothesized that the PDNA-reported roads may be of higher quality than the damaged road sections, possibly explaining the difference. However, the PDNA mentions that 75% of the national road network is unpaved. Visual inspection of images indicated that all affected roads in Beira were paved. Therefore it is concluded that the Beira roads are not of significantly lower quality than the reported PDNA roads and are therefore comparable with each other. This indicates that the model strongly underestimates road damages.

Besides, large damages to the road network are described qualitatively (Beira city government, 2019; GFDRR, 2019). Contrarily, low damage ratios of 0,23 were found in the model (Section 10.1). This may also indicate that the vulnerability curves are underestimated.

Construction cost data is presented at the bottom of Table 11.5. Benchmark data was created through the analysis of different road-building projects in Africa. The data indicate large variability between projects, but also depending on the size and quality of the project. Nevertheless, also here the data indicates an underestimation of the model, as the benchmarked construction costs are significantly higher than assumed exposure values.

To quantify the underestimation, the 'conservative' assumption is taken that the percentage of total PDNA damage that is attributable to road damage lies between 10% and 50%. This indicates that damage numbers lie between 15,80 and 79,00 $ / m damaged road, versus the modelled 4,90$ per m damaged road. This increases the initial estimate of 13.841 $ by a factor of 3,2 and 16,2.

$$\text{Lower bandwidth:} \quad 13841 * 3,2 \quad = 44.291\$$$
$$\text{Higher bandwidth:} \quad 13841 * 16,2 = 224.224\$$$

[12] With a corresponding inundation depth of 0,5m, extracted from GIS.
[13] Values from 2006 were transformed using inflation factor of 2,01

11.5. Conclusion

The resulting sector-based conclusions and new quantitative damage estimations are given in Table 11.6. Overall, the original estimate is concluded as an underestimation, with the new best estimate quantified as 8,1 M$. An best estimate uncertainty range between 5,2 and 13,2 M$ is given.

This conclusion was taken due to underestimations in commercial and road assets, where the analysis indicated modelling errors with high reliability. These sectors may require calibration in future flood risk studies. The causes for underestimations in industrial and port assets are less clear, making calibration difficult. Finally, formal housing was concluded as accurate but highly uncertain.

Contrarily, damage estimates for informal housing were considered overestimated. The remaining damage estimates for agriculture and other housing classes were considered of insignificant impact to be included specifically in this estimation. Nevertheless, indicated uncertainties may be relevant for flood risk studies including larger return periods.

Table 11.6: Conclusion of Beira estimate accuracy using empirical data. All values are indicated in k $. When only a lower and upper bound estimate was given, the new best estimate is taken as the average of the lower and upper bound, and indicated with ↔

Asset	Original estimate	Conclusion	Lower bound	New estimate	Upper bound	$\frac{\text{New}}{\text{Original}}$
Residential 2	1.970	Uncertain	480	1.970	5.560	-
Commercial	1.440	Underestimation	2.200	3.405 ↔	4.610	+140%
Residential 4	585	Overestimation	130	300↔	470	-46%
Industrial	328	Underestimation		1.116		+240%
Port	120	Underestimation		1.210		+908%
Road	14	Underestimation	44	134 ↔	224	+857%
Other	6	-	-	-	-	
Total [k $]	**4.460**		**5.180**	**8.135**	**13.190**	**+82%**

12

Beira - Conclusion

The goal of this study was to validate the estimated 4,46 M$ coastal flood damage caused by hurricane Beira in March 2019. A new best estimate of 8,1 M$ is proposed, indicating a model underestimation of 82%. Furthermore, an uncertainty range of 5,2 M$ (-36%) to 13,2 M$ (+62%) was concluded.

This conclusion was quantified using disaggregated compound damage data and construction cost data, applied in phase 3 of the framework. Large underestimations were found in infrastructure, industrial and commercial assets, together with an overestimation for informal residential assets. Due to a lack of data informativeness, no conclusion on the root cause of these errors could be made. Less information was acquired from phase 2 - model plausibility assessment. Only two underestimations in industrial, and agricultural damage classes were expected from phase 2. For the industrial asset, both errors on the vulnerability and exposure parameters were expected. Contrarily, agriculture was underestimated solely through the vulnerability function due to the excluded salinity effect of the inundation.

The quantified uncertainty stems mostly from formal housing and commercial assets. Furthermore, too little information was available to quantify direct damage uncertainty for critical infrastructure (industrial and port assets). Informal housing, road, commercial and agricultural assets are expected to also be underestimated when modelling other return period hazards. Therefore accuracy may be increased through a calibration process. However, the damage observations are not informative enough for each class to indicate which parameter should be altered, and how. Furthermore, Industrial, port, and (to a lesser extent) formal housing assets have high uncertainty, that cannot be reduced further through the framework. These areas can be highlighted and the uncertainty in their risk reduction should be taken explicitly into account by decision-makers. The validated event estimates and assumptions can then be used as a basis for a probabilistic flood risk analysis.

IV

Framework Discussion & Conclusion

13

Discussion

In this chapter, the research relevance, reflection on case study results, framework limitations and recommendations are discussed.

Relevance

Current flood risk literature calls for more and better flood risk validation, particularly by collecting more and better flood event damage data (Jongman et al., 2012; Molinari et al., 2019; Sayers et al., 2016). Furthermore, the importance of assumption sharing and uncertainty communication is emphasized (Merz et al., 2008; Sayers et al., 2016). The proposed framework integrates these requirements, structures the validation process and highlights the main uncertainty sources.

For both case studies, verification with observational data led to a significantly altered damage estimate, demonstrating the necessity of validation. New data transformation procedures were required due to data heterogeneity, involving employee numbers, comparing vulnerability curves and construction costs. These new methods allowed a more precise comparison of estimated damage with observations, thereby enabling error and uncertainty quantification of specific damage categories. This has increased benefits compared to general literature, where often only model-model comparisons or aggregated observational comparisons are made (Herath et al., 2003; Jongman et al., 2012).

Case study reflection

Reflecting on the case studies, the analysis results depended on the availability of two information types: The availability of damage data and the availability of general information. Shown on the x-axis in Figure 13.1, the damage data availability is rated from having no data at all to having high-detail data that indicates a spatial- or asset-wise damage distribution (such as WTS or Post Disaster Needs Assessment (PDNA) data). On the y-axis, general information availability can range from only having satellite imagery, to having high amounts of information such as economic asset value estimates, spatial land use information, and construction data.

The Limburg and Beira case studies are characterised by their data availability in Figure 13.1 A. For Limburg, both a detailed amount of damage data (Fact finding report, WTS, insurance data) was available, as well as many information sources. For the Beira case study, the PDNA damage supplied damage information, but limited general information was available, resulting in a lower y-axis placement.

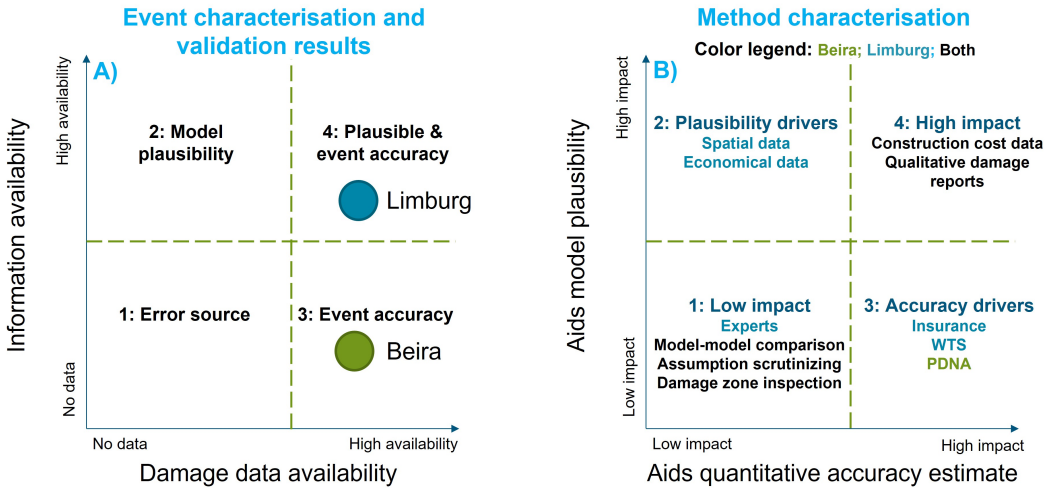


Figure 13.1: A) Characterisation of validation events on damage data availability and information availability (such as economical data, construction cost data, spatial data on land use). B) Overview of applied analyses characterised by 1) how much they support quantitative accuracy estimates and 2) underwrite the plausibility that the model simulates the damage processes well under multiple return periods.

Figure 13.1 B subsequently characterises the methods used in the case studies on two similar axes: On the x-axis, the amount a method aids the creation of a quantitative accuracy estimate, and on the y-axis, the amount a method aids the plausibility of model application under different return periods. Methods are ranked on having a low or high impact for the two axes. The graphs are similar because it was noted that many general information sources had a strong contribution to model plausibility, but a low contribution to accuracy estimates. Correspondingly, damage data sources were found to be very useful for a quantitative accuracy estimate, but aided little to model plausibility for other return periods. Finally, construction cost data and damage reports were found to be effective for both metrics.

Both analyses were initiated in quadrant 1. Here, low-impact methods such as assumption scrutinizing gave insights into where possible errors in the estimate may have occurred. Although these methods were low-effort to execute (except for expert judgement), their overall informativeness was low in comparison with later methods. Subsequently for the Limburg case, general information sources strongly increased the plausibility of the model's exposure parameter. This indicated that the model's internal workings were likely plausible, but a quantitative difference between estimates and damage observations was unknown. Finally, the application of insurance and WTS data allowed further error quantification (mostly applied to the residential sector). Therefore the case study was moved into quadrant 4, where both the plausibility of model performance as well as a quantitative accuracy estimate was reached. This allows confident model application for risk studies.

Contrarily for the Beira case study, little general information was available. This resulted in only a marginal model plausibility. After the low-impact methods, the PDNA report was extremely useful to quantify the event accuracy. Construction cost data aided both the accuracy estimation as well as model plausibility for the exposure parameter. However, as information lacked for many asset types, the overall plausibility was deemed significantly lower than in the Limburg case study. This quadrant only allowed a conclusion on the estimate accuracy for the simulated event, thus the model's application to other hazard events remains questionable. This results in much higher uncertainty when applying the model to risk studies, which needs to be communicated and mitigated in the flood-risk reduction investment case.

Limitations

Limitations of the research are rooted in the information sources information, the subjectivity of validity and uncertainty.

First, it was found that information sources limit the validation process through availability, their level of detail and reliability. In Figure 13.1, it was noticed that non-damage data aided most to the plausibility of the exposure parameter. However, little focus was placed on the validation of the hazard parameter, as no additional data was available for the process. The framework gives limited recommendations on how to validate this parameter, despite recognizing it as a large uncertainty source (Ward et al., 2013). To reduce this, additional sensitivity analyses may be considered in the validation approach. Furthermore, validation of the vulnerability parameter could only be done by damage data, but was severely restricted by the level of detail. For depth-damage curve validation, object-based damage data is required (Hasanzadeh Nafari et al., 2016; Kellermann et al., 2020), which is still barely made accessible by insurance institutes.

The level of detail in the data further limits the calibration process, as the current detail does not indicate the root inaccurate parameter value. Therefore, calibration decisions on parameter alteration, can not be supported by data. This effect can be seen in literature. Mediero et al. (2021) calibrated coupled vulnerability and exposure parameters using a single event. Subsequently, the calibrated model overestimated all the smaller events. It is plausible that this was caused by the calibration of coupled variables, whereas a specific calibration to a single parameter could have improved model accuracy across all events. Contrarily, Amadio et al. (2016) calibrated only the exposure parameter, whereas the vulnerability parameter was assumed accurate. This resulted in increased estimate accuracy for the calibration event, however, the question may be posed how well this translates to other flood events.

Finally, the reliability of the information limits its application. The case studies found significant biases in the data, such as the PDNA business damage that only included large companies, and WTS valuations biased by payout limits. Bias cannot be avoided, but many studies apply data without analysing bias effects, which can be a severe limitation to the concluded model accuracy (Kron et al., 2012; Molinari et al., 2019).Therefore, additional emphasis on finding bias effects before data usage should be placed.. The reliability may be further decreased by the application of new methods to each dataset, due to their heterogeneity. Although these methods may be plausible, their new untested application adds additional uncertainty.

The second limitation of the framework (and general validation studies) remains the subjective definition of validity (Apel et al., 2009). The framework and Figure 13.1 argue that additional data can increase the level of validity. However, no generally applicable quantitative measure that defines when validity is reached can be made. Therefore, this will in practice remain up to engineering judgement and be guided by time and data constraints. An example of this can be found in the case study validation of the hazard parameter. As no additional data sources were found to compare the applied hazard analysis with, the uncertainty was mentioned but not further elaborated on.

The third limitation of the framework is the large remaining uncertainty. For residential assets, the found hazard and exposure parameter uncertainty was larger than generally reported in other literature (Huizinga et al., 2017; Wagenaar et al., 2016). This could mean a severe restriction on the use of these ‘globally’ applicable vulnerability curves from literature, as it is shown they are not widely applicable. For critical infrastructure the limitation was more severe, as the available PDNA damage data provided insufficient information to quantify the amount of direct damage uncertainty. This shows that additional damage data exclusively is not always able to solve the validation problem, where the main literature advice lies (Molinari et al., 2019). Furthermore, it questions a standard ‘engineering’ approach of calculating indirect damage as a factor of direct damage, as done in (Aerts et al., 2014; de Ruig et al., 2019). If direct damages are already highly uncertain, a multiplication will result in even higher uncertainty for indirect damage estimates (Jongman et al., 2012). For these critical assets, the direct damage estimate may simply be of inferior importance for flood mitigation strategies, as indirect effects can be much higher (De Bruijn et al., 2016; Habermann & Hedel, 2018). Therefore an improvement can be an increased focus on other methods, such as economic input-output models, to estimate indirect damage for these critical assets (Koks et al., 2015).

Recommendations

Recommendations are given on the topics of validation procedures and additional data gathering.

For future validation studies, Figure 13.1 indicates recommended methods, as well as the conclusions that may be drawn depending on the available amount of information. First, it is recommended to start the validation process with high-impact methods such as damage reports and construction costs, as these were deemed most effective for the case studies. Second, the validation status of the event can be placed in Figure 13.1 to analyse what further information is required to move the case towards quadrant 4. For instance for Beira, engineers could focus on finding additional non-damage data to increase the plausibility of the model.

Regarding future data gathering, three recommendations are made. First, it is advised to gather future damage data in a more transparent way to minimize bias and ambiguity, as was found in both the PDNA and WTS data. Second, to increase the amount of object-based damage collection. It is understandable that not the entire disaster area can be reported in an object-based manner, but a small amount of representative data could already give strong indications for the calibration process. It is estimated that around 20 damage valuations each, for a low-, medium- and high- inundation depth area could already significantly increase support of hazard and vulnerability parameters. Added benefits may be gained, for instance, by combining damage reporting with high watermarks, which could immediately aid validation of the hazard parameter which remains a troublesome task. Third, the analysis using WTS data was severely limited due to an unknown coverage factor. Estimations of this factor resulted in strong Pearson's correlations similar to other literature (Amadio et al., 2016), which indicates the potential of the dataset. Further insights on the total direct incurred can be acquired on a (very valuable) high spatial level, if this factor was known. Therefore the importance of gathering loss-data coverage information is emphasized, next to the request for additional detail.

Future research

Other limitations of the proposed framework may be reduced by additional research on four subjects: expert judgement, additional data usage, compound damage disaggregation, and calibration procedures.

First, the study attempted to create scientific data by expert elicitation, but concluded unsatisfactory results. General application of expert judgement in flood risk model validation includes consensus and open discussions, which may lead to large biases. Instead, the applied elicitation method was able to create results with minimum bias. However, this also resulted in too large uncertainty to be useful in the validation process. For further research regarding expert judgement, a focus on balance between increasing expert informativeness while minimizing bias is advised.

Second, the application of global damage databases was tested in Appendix D, for a hypothetical case where no local damage data was available. Despite inadequate results of the analysis, this situation still occurs frequently in flood risk modelling. Further research on applying these global datasets can support researchers by enabling quick high-level damage numbers, but should put considerable effort into ensuring data reliability. This may be especially interesting as earlier research by Ward et al. (2013) did manage to acquire suitable results.

Third, the research indicated a novel way to disaggregate compound wind and flood damage as part of data transformation procedures. High dependency on structural characteristics and inundation depth was shown, which agrees with earlier case study research by Baradaranshoraka et al. (2017). Contrarily, other data-based research using the NatCatSERVICE simply concluded a 50-50 split between wind and flood damages (Kron et al., 2012). Applying this detailed disaggregation method to hindcast compound damage data may give additional insights into compound damage processes.

Fourth, no calibration procedure is defined in the proposed validation framework. As calibration can be done by altering multiple parameters, most notably the exposure data and the vulnerability curves, a suitable method can further guide experts to increase model accuracy. Here, the risk of overfitting on a single flood event (as discussed in the limitations section with a study of Mediero et al. (2021) should be balanced with the calibration possibilities that the data availability enables.

14

Conclusion

During this research, a validation framework was developed that flood-risk experts can use to validate their models. The proposed framework applies four steps to an event-based damage estimation: 1) initial situation assessment, 2) model plausibility assessment, 3) verification using empirical data, and 4) concluding on the next steps. The first two steps were found to be important to gain insight into the main model uncertainties and increase confidence in the local applicability of assumptions. Subsequently applying empirical data was required to quantify these uncertainties and the difference between the model and observations, and to support assumptions.

Application of the framework to the Limburg river flood (2021, The Netherlands) and the Beira coastal flood (2019, Mozambique) resulted in a new damage estimate of -34% and +82% respectively. In both cases, large errors were found in the modelling of linear infrastructure, due to rasterization and parameter errors. Damage estimates of the major residential classes were concluded accurate but highly uncertain. For agricultural assets, a lacking crop differentiation between resilient grass and vulnerable crops was a major error source as well as disregarding the salinity of a coastal flood. Finally, large uncertainties were found in critical infrastructure, for which an area-based direct damage calculation may be unsuitable. Instead, it is advised to focus on the effects of indirect damage.

For Limburg, the largest found errors and uncertainties were attributed to the exposure parameter, followed by the hazard and the vulnerability parameter. This conclusion may be biased due to lacking object-based damage data, which limited further insight into the uncertainty of the vulnerability parameter. For Beira, it was not possible to indicate root parameters of errors due to limited information. The damage report merely allowed to indicate the difference with model estimates.

The results indicate four conclusions. First, properly verifying flood damage model results requires a damage-class-based comparison. In this way, both under- and over-estimating errors, as well as uncertainty for each asset type can be shown. The case studies revealed that data bias may be a significant factor to consider. Second, data transformation assumptions such as ‘crop damage is not plausible at inundated grassland’ and estimate uncertainties remain in the validation process. Therefore, transparency and assumption communication is concluded as an essential value. This still lacks in (flood) modelling (Beven et al., 2016; Tennøy et al., 2012). Third, although differences between observations and estimations were shown, insights for calibration were limited due to data detail. To aid calibration, future data reporting may increase their depth-damage or spatial information. Fourth, risk-based accuracy should be increased by calibrating exposure and vulnerability parameters to approach the assessed event estimation, in such a way that the parameter alterations are plausible across all hazard scenarios. If this is not possible, the remaining uncertainties should be clearly communicated to decision-makers. Otherwise, altering parameters only to align estimations with a single event’s damage observations may result in decreased accuracy for non-assessed scenarios.

Through these steps, the framework for event-based validation can increase the validity of probabilistic flood risk models. Nevertheless, insights remain limited by the availability of local information.

V

Appendices

Appendix A: Limburg analysis

A.1. Model input plausibility

Figure A.1: Depth distribution analysed in GIS, an increased amount of half meter pixels can be seen in the distribution.

Figure A.2: Waterlevel of the Maas: Measured vs used as input in the inundation analysis.

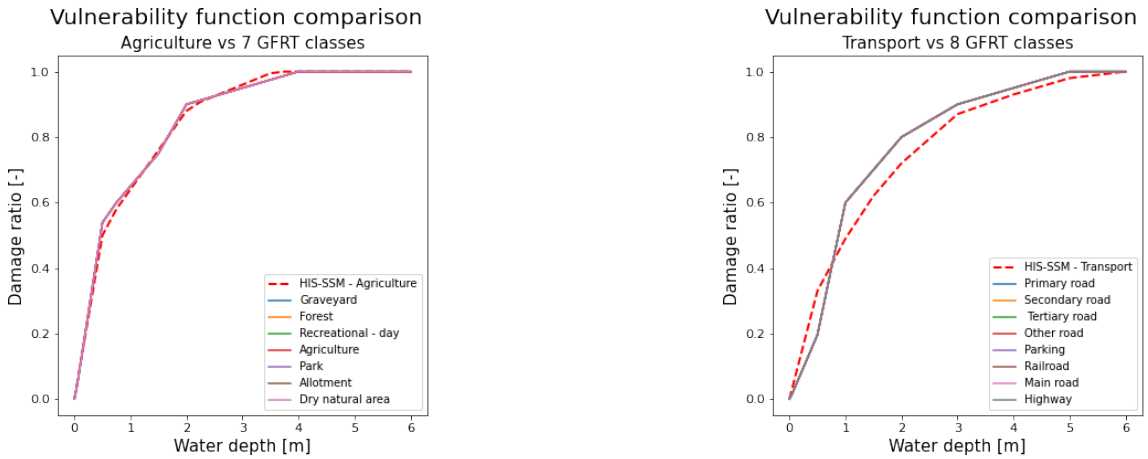


Figure A.3: Model-model comparison of other vulnerability classes. The red-dashed HIS-SSM line is compared to multiple GFRT vulnerability curves.

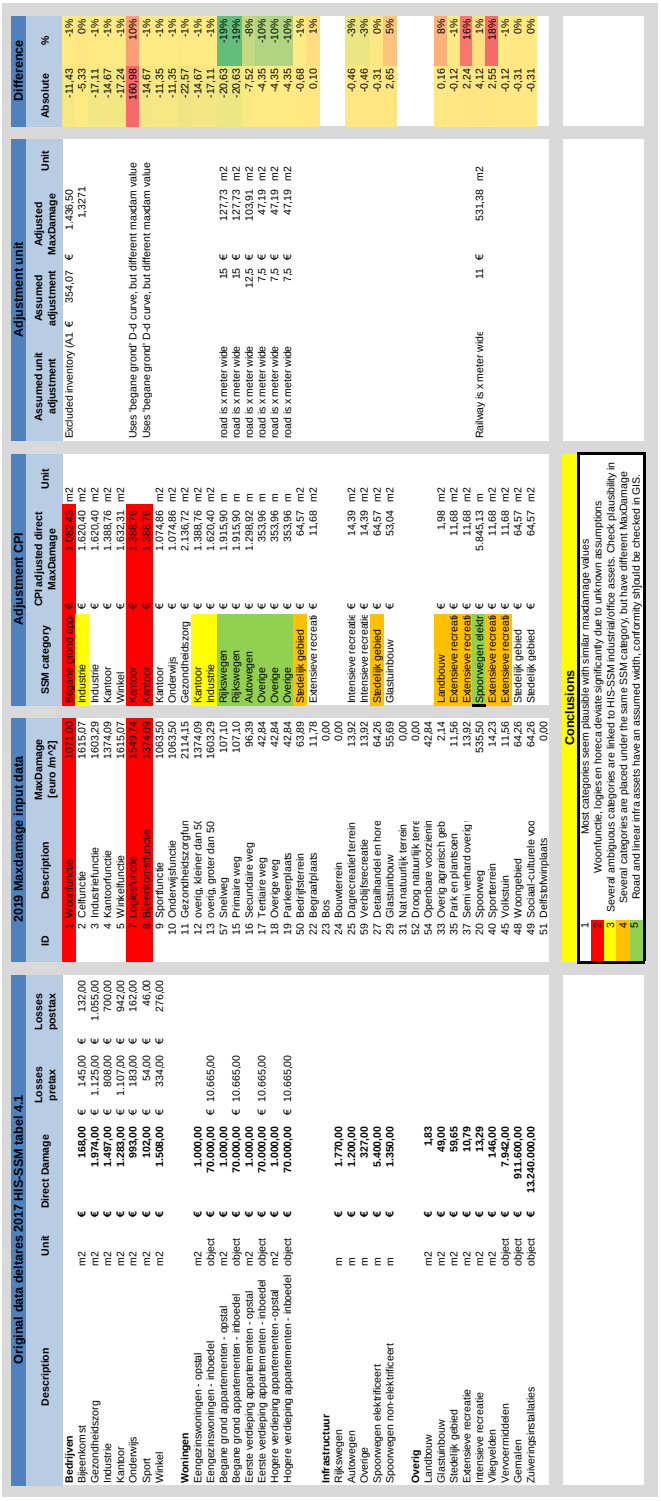


Figure A.4: Model-model comparison between HIS-SSM maxdamage values and the GFRT input maxdamage values. The left of the table shows HIS-SSM values, the second column indicates 2019 GFRT values. The GFRT values are subsequently adjusted for inflation and adjusted for units/assumptions. These assumptions and conclusions (on the bottom) are scrutinized in GIS. Lines without comparison either have no asset value or no comparable HIS-SSM category.

A.2. Geographic damage distribution: GFRT initial estimate and fact finding low and high

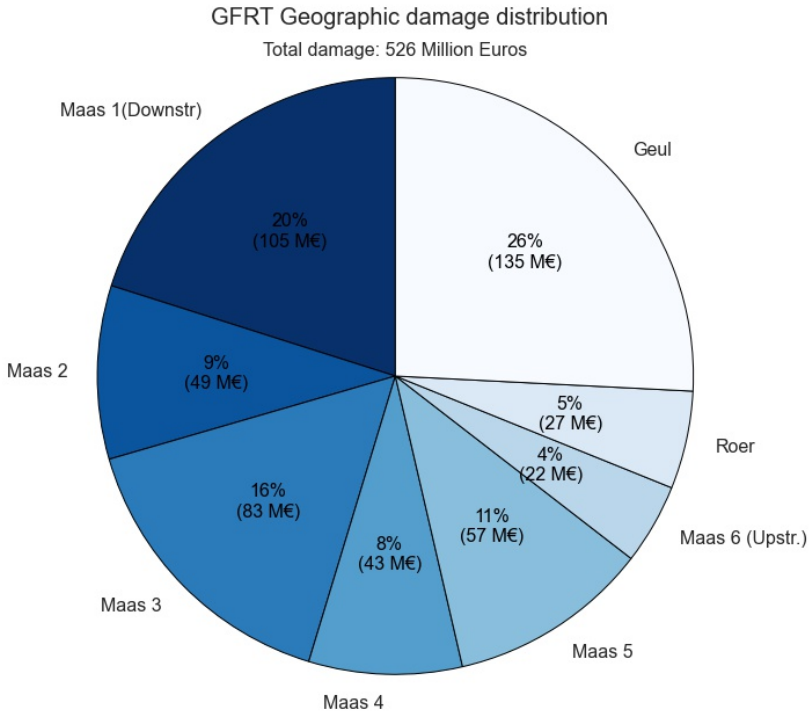


Figure A.5: Geographic damage distribution of the GFRT estimate prior to validation phase 2.

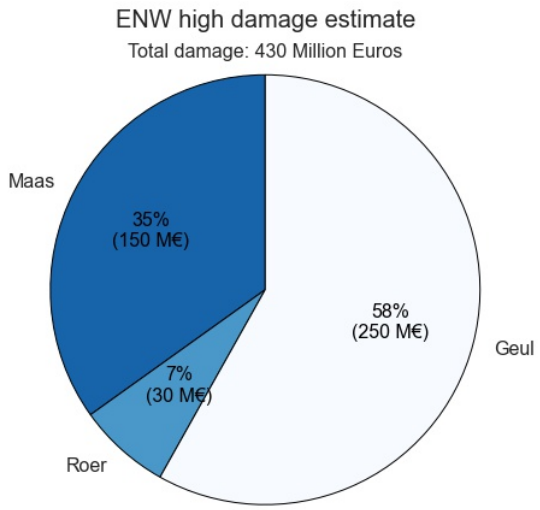


Figure A.6: Geographical damage estimate from the fact finding report including physical damages and company losses - high estimate (Expertise Netwerk Waterveiligheid, 2021).

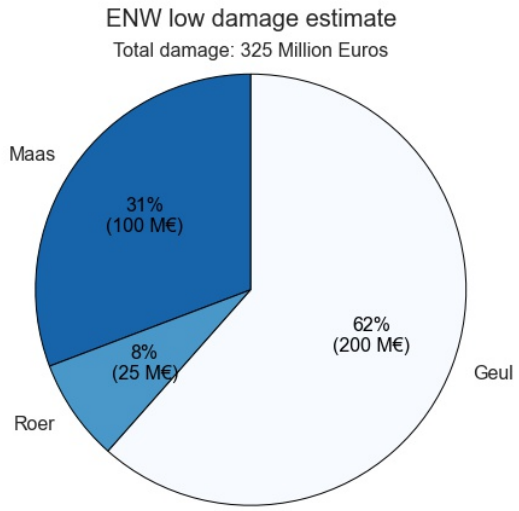


Figure A.7: Geographical damage estimate from the fact finding report including physical damages and company losses - low estimate (Expertise Netwerk Waterveiligheid, 2021).

A.3. GFRT initial and fact finding damage estimate comparison for the Meuse, Geul and Roer

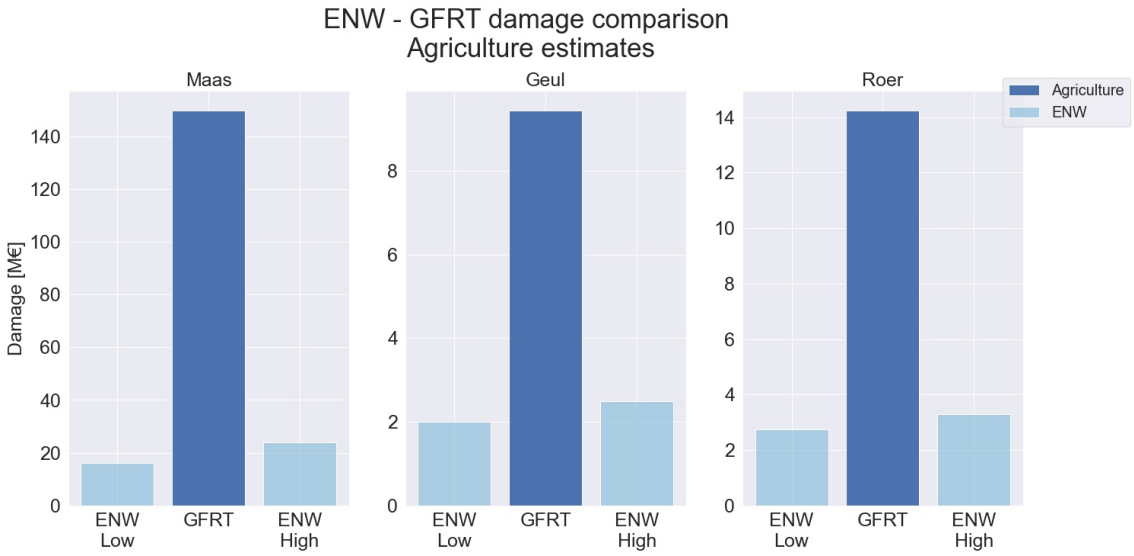


Figure A.8: Comparison of initial GFRT estimate and fact finding estimate for agriculture (Expertise Netwerk Waterveiligheid, 2021).

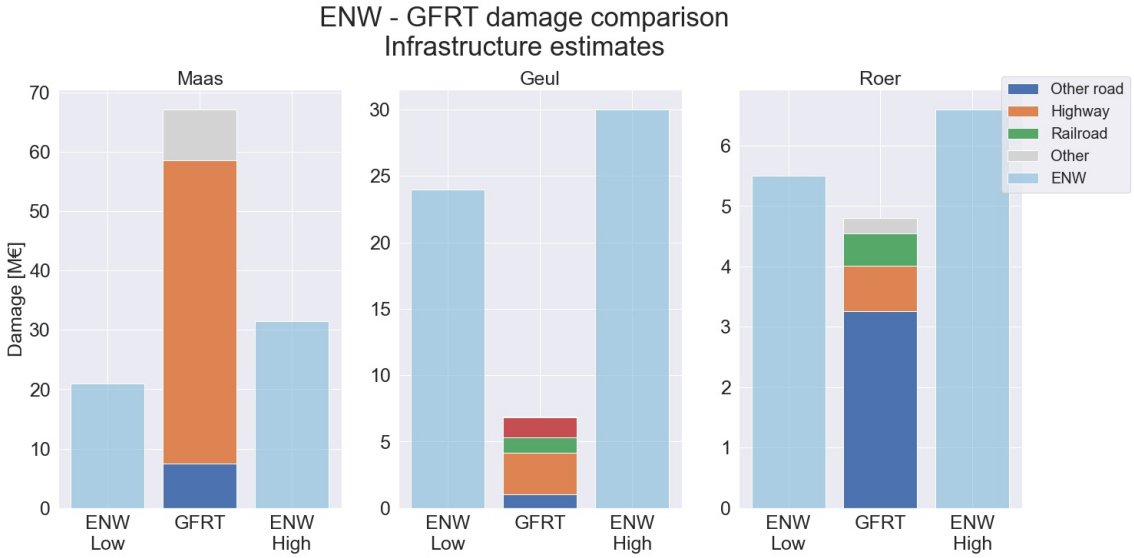


Figure A.9: Comparison of initial GFRT estimate and fact finding estimate for infrastructure (Expertise Netwerk Waterveiligheid, 2021).

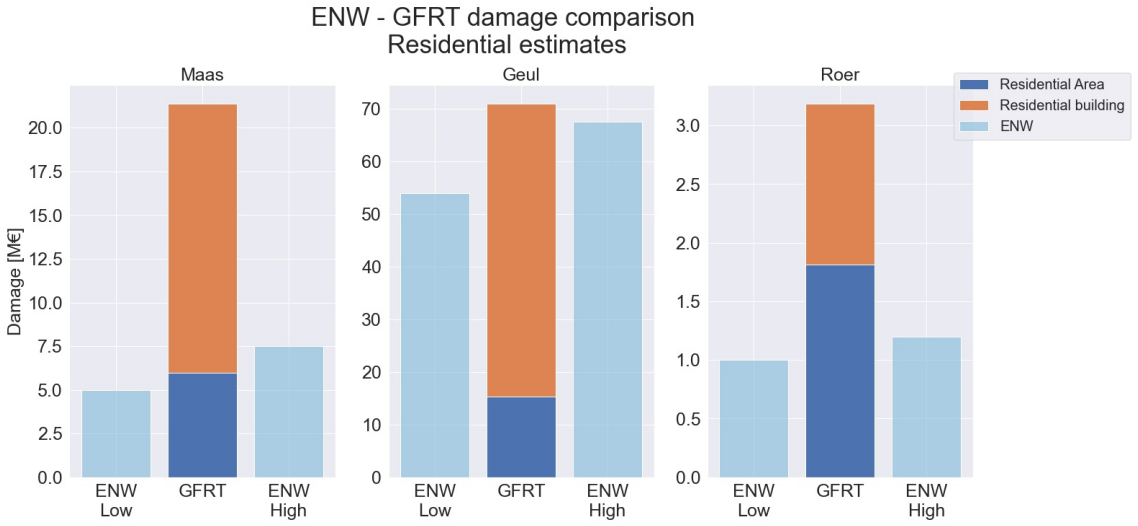


Figure A.10: Comparison of initial GFRT estimate and fact finding estimate for residential (Expertise Netwerk Waterveiligheid, 2021).

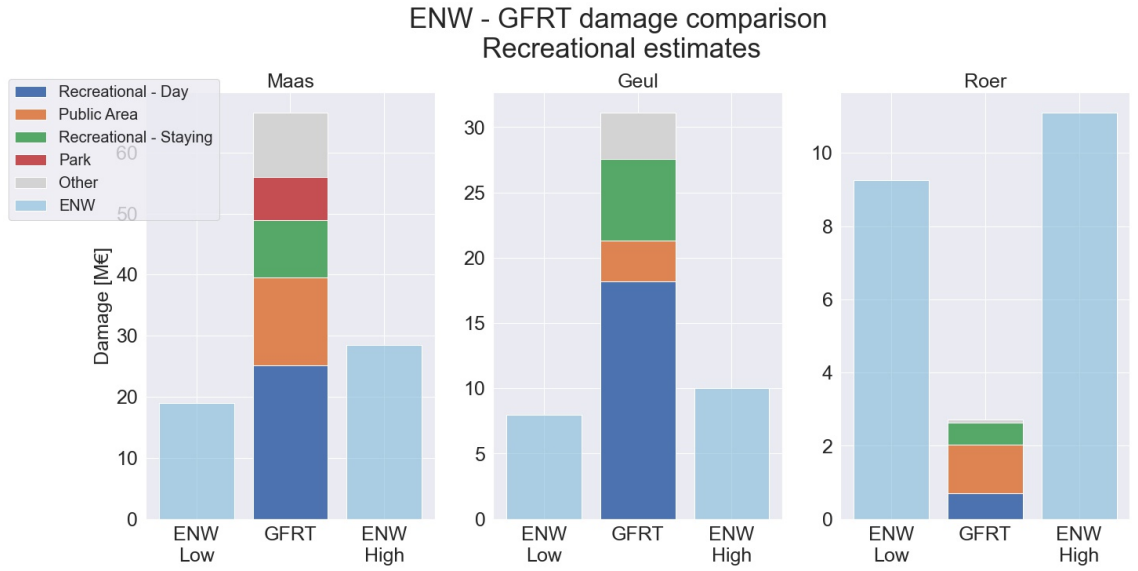


Figure A.11: Comparison of initial GFRT estimate and fact finding estimate for recreational (Expertise Netwerk Waterveiligheid, 2021).

A.4. Limburg model output validation supporting section

Figure A.12: Calculated Railroad damage at an elevated bridge and elevated berms at a river crossing.

linear road approach

To validate the minor road sections, the damaged road sections were represented by a line instead of gridcells. This line was created by clipping the NWB dataset with the damaged gridcells from the GFRT. Subsequently, along every meter of the line, the inundation depth was extracted from the inundation raster. This dataset was then exported to python, where a length based damage calculation was performed. Each water depth was transformed to a damage ratio using the same vulnerability curve, and subsequently multiplied with a maximum damage value $/m$ of road section taken from HIS-SSM in Figure 5.5 (Slager & Wagenaar, 2017)

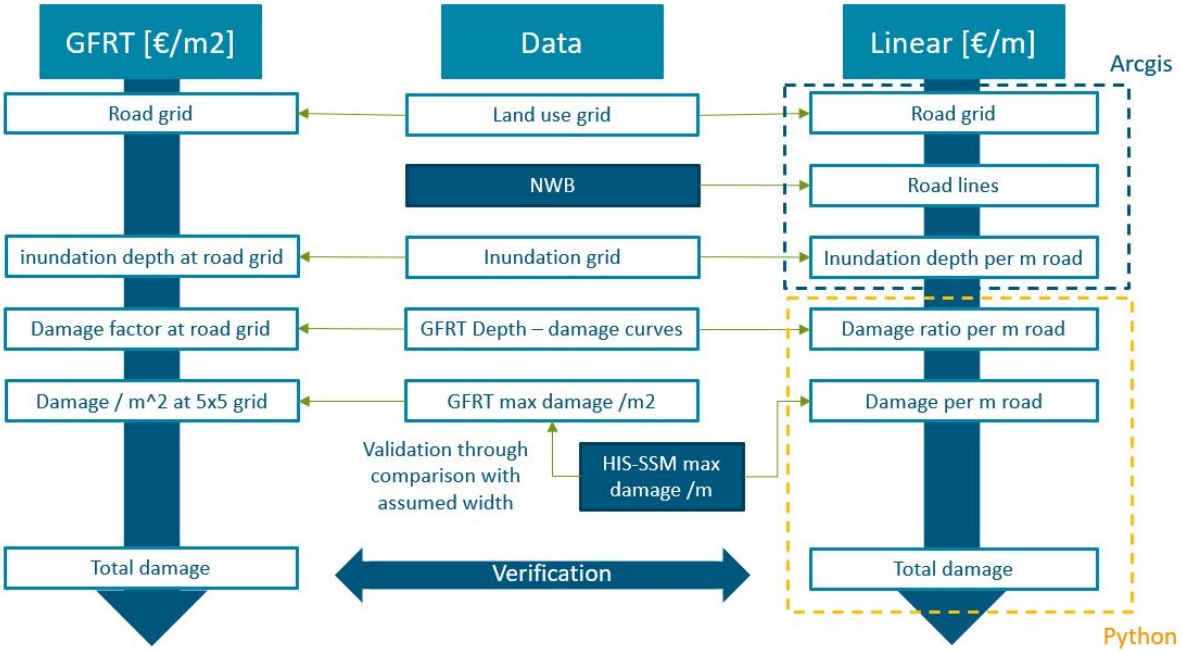


Figure A.13: Visualisation of GFRT and linear approach. Two new datasets used for the linear approach are shown in dark blue.

N1: Rasterization caused surface area error

The spatial resolution of gridcells can result in an erroneous geometrical representation of assets, especially for long thin assets like roads and railroads. de Moel et al. (2015) analysed the impact of different spatial scales and concluded that the error size depends on situational characteristics such as grid resolution and orientation, asset geometry and flood characteristics. To validate this aspect, the difference in surface area between the polygons and gridcells was compared for three minor road classes. The results can be seen in Table A.1. As the surface area ratios are near 1, it was concluded that in this particular case the gridcell based approach is valid and that no further uncertainty was caused by the spatial resolution.

N2: Analyse sensitivity to specific inundation depth

To check the sensitivity of a land use class to a specific waterdepth and corresponding damage ratio, a histogram of occurring waterdepths was analysed.

For instance for the infrastructure sector, the hypothesis could be that many roads are only slightly inundated and that therefore a large amount of predicted damage was from a small inundation. The depth-Damage curve can then be validated locally, by re-evaluating how much damage this low inundation actually caused to a road.

Results from the histogram in **??** indicate that there is no increased sensitivity for low inundation events. However, a small peak around the 0.5m mark can be seen. This indicates the effect of uncertainty in the hazard analysis mentioned in Section 5.1.

Table A.1: Results of surface area comparison between road representation by polygons and gridcells.

Class	Total floodextent			Geul floodextent		
	Polygon area m^2	Gridcell area m^2	$\frac{\text{Polygon}}{\text{Grid}}$	Polygon area m^2	Gridcell area m^2	$\frac{\text{Polygon}}{\text{Grid}}$
Other road	2276500	2293128	0.992	176625	177553	0.994
Tertiary road	1313600	1317257	0.997	85625	86108	0.994
Secondary road	198975	201018	0.989	26450	26273	1.007

A.5. WTS analysis

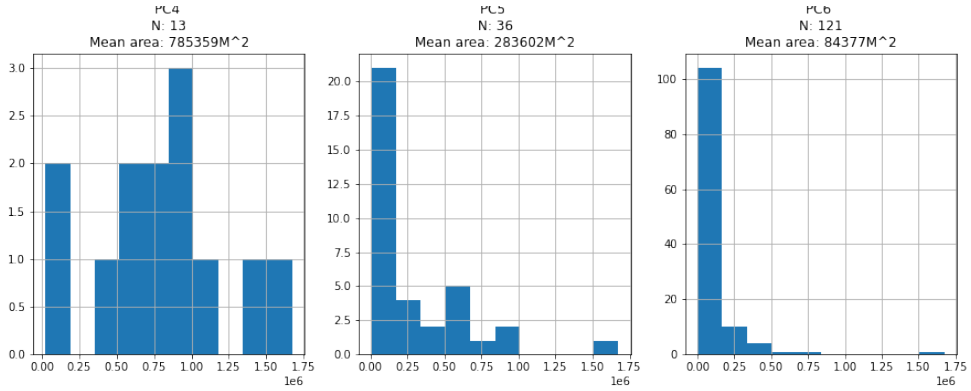


Figure A.14: Distribution of postal code areas for PC4, PC5 AND PC6.

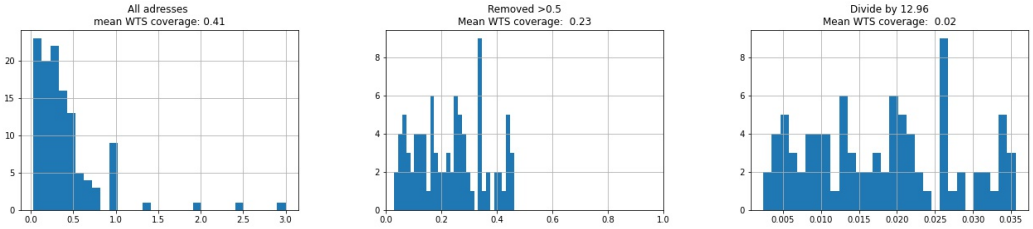


Figure A.15: Histograms of coverage factor during data transformation steps

Figure A.16: This business area has much lower reported damages than was reported on average in comparison with the GFRT. This indicates that very little flooding may have actually occurred here.

Figure A.17: Residential outliers in the municipality of Brommelen. The red area has much higher reported damage than on average in comparison with the GFRT, whereas the green areas have lower damages.

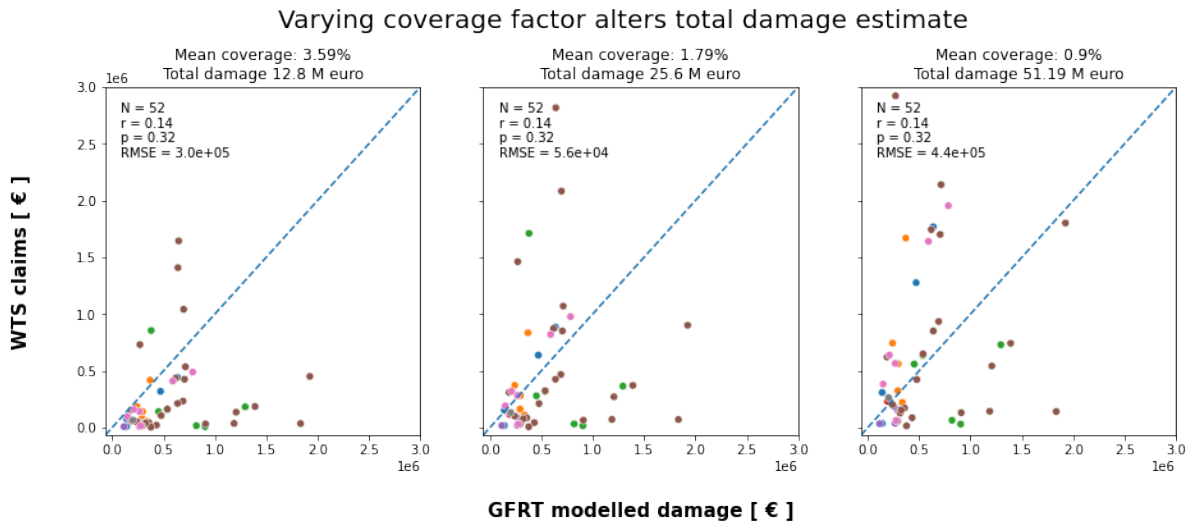


Figure A.18: Total damage estimates with three different mean coverage factors.

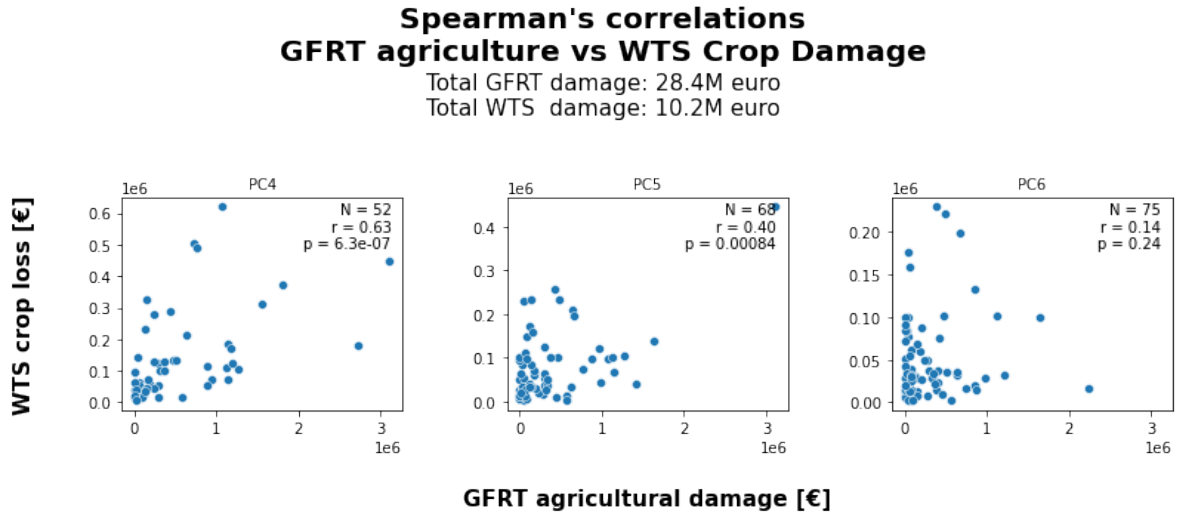


Figure A.19: Spearman correlations between WTS reported crop loss claims and GFRT agriculture estimates

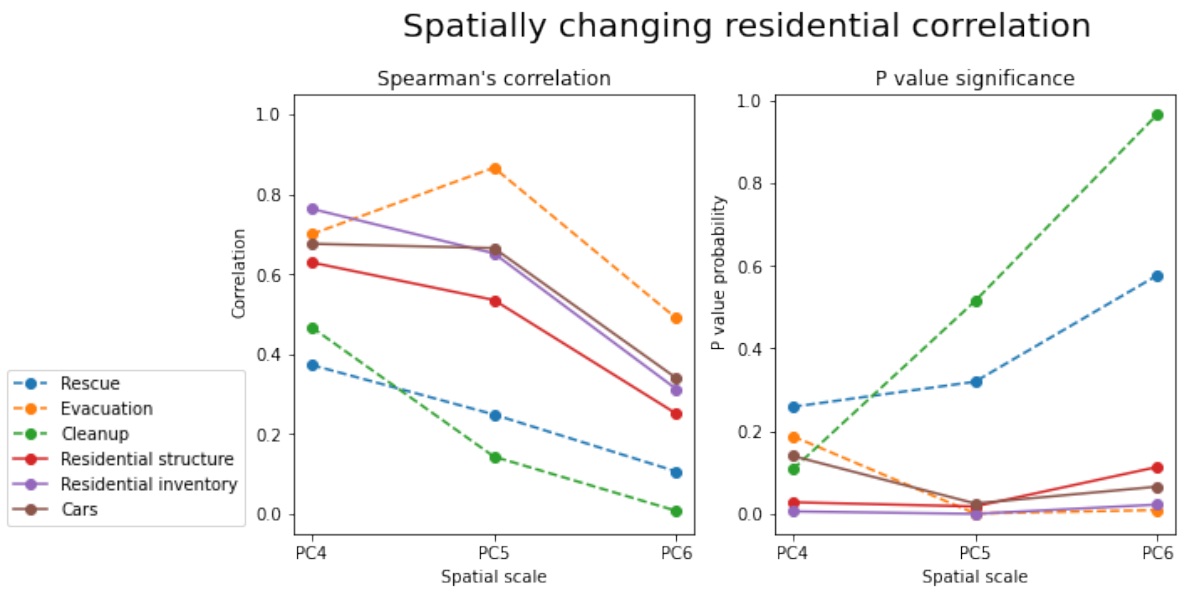


Figure A.20: Spearman correlations between WTS reported residential damage and GFRT residential house estimates

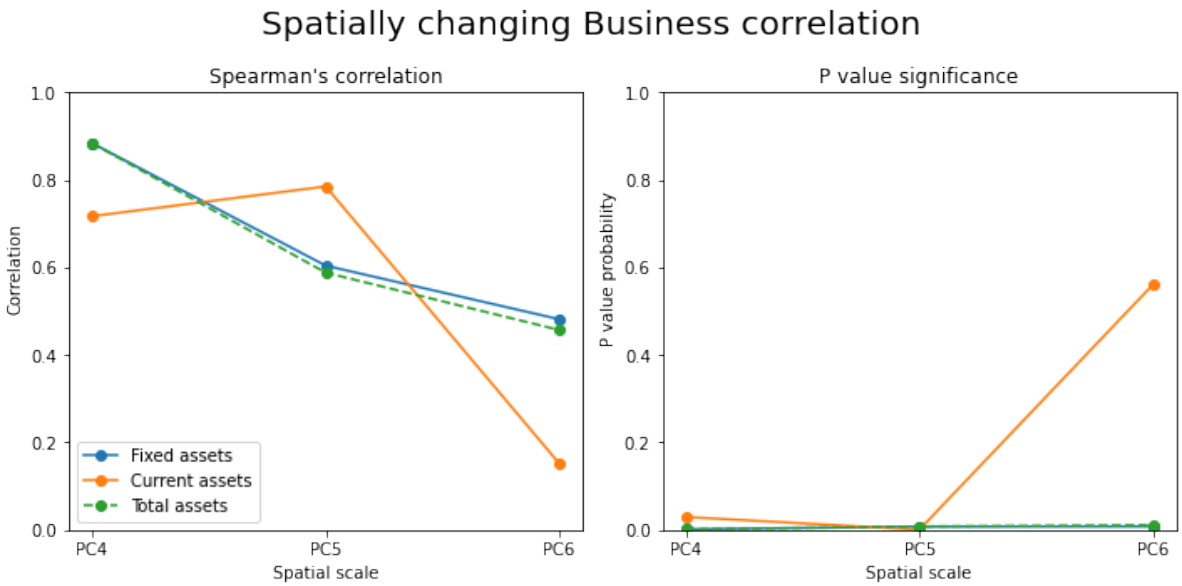


Figure A.21: Spearman correlations between WTS reported company claims and GFRT business estimates

Table A.2: Extended overview of case study methods from the discussion

Quadrant	Method	Limburg application	Beira application
1	Experts	SEJ experiment was high-cost and resulted in some qualititative insights and uninformative quantitative estimates	ND
	Model-model	Comparison with HIS-SSM or linear-based model (for roads) indicated small differences. This provided insight, but is not as concrete as verification with damage data.	comparison with Global depth-damage functions of Huizinga increased plausibility but large uncertainty remained.
	Assumption scrutinizing	Comparison of road width and residential inventory inclusion in exposure value indicated plausible errors	Not including salinity in agriculture depth-damage functions may lead to underestimation, but unclear how large the effect is
	Damage zone inspection	False damage zones could be identified and their impact summed aiding damage quantification	Zones with uncertain damage estimation could be pointed out, but no definitive conclusions or parameter insights could be obtained
2	Spatial data	BRP data indicated were agriculture damage was plausible	Housing footprint data allowed further insight in the type of buildings, in density assumptions and to compare damage values per surface area of house.
	Economic data	Average revenue for cropland and public school costs increased exposure value plausibility.	ND
3	Insurance	1) Allowed quantitative estimate of total residential damage 2) insurance payouts and reconstruction costs could be compared to maximum damage values	ND
	WTS	1) Allowed a quantitative estimate of total damage, but the unknown coverage factor resulted in large uncertainty. 2)Spatial data allowed to indicate deviating areas and possible errors in the hazard or vulnerability parameter	ND
	PDNA	ND	Very useful for verification data as information on various asset types was available. Disaggregation was plausible but possibly added uncertainty
4	Construction cost	Public school construction cost increased exposure value slightly.	Some residential construction costs were available, but too large uncertainty was present to give confident estimations
ND	Qualitative damage reports		Qualitative damage descriptions in the PDNA hinted at a wrong vulnerability curve for road infrastructure. Quantification was not possible.

B

Appendix B: Structured Expert Judgement figures

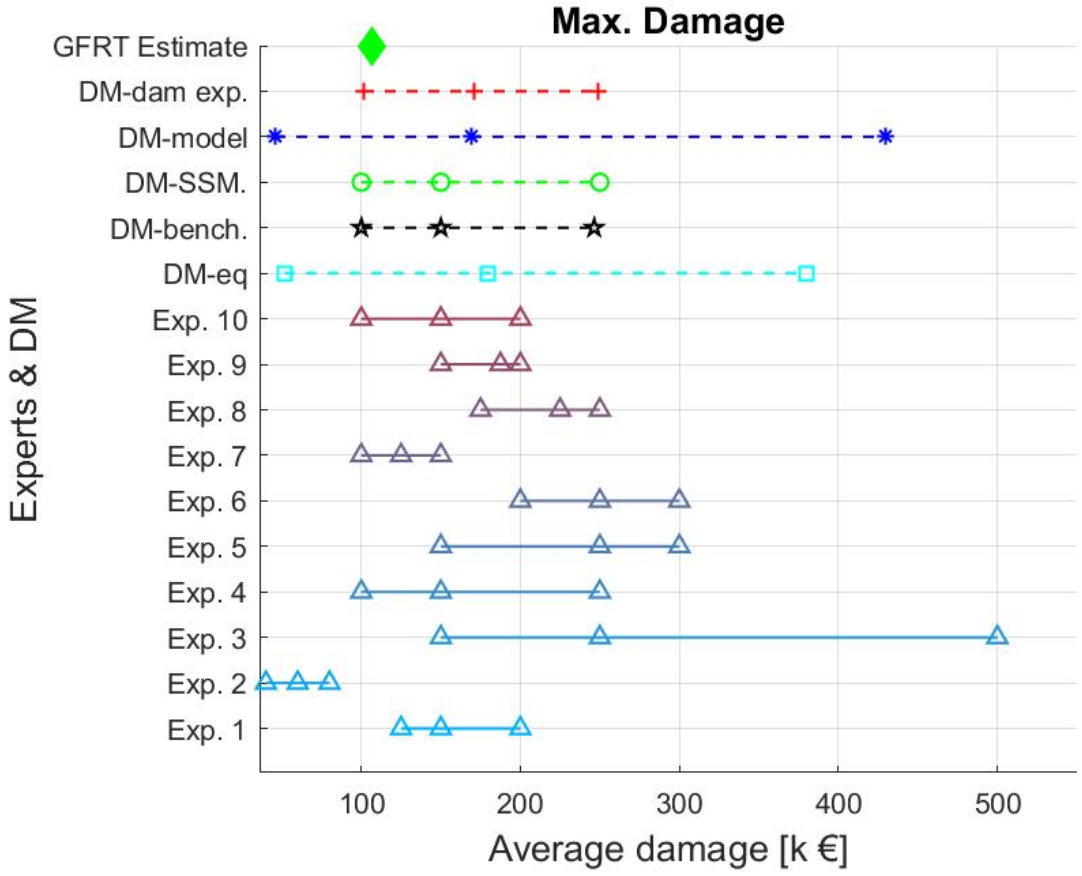


Figure B.1: Residential maximum damage estimations (Target question 1) of experts and their corresponding decision makers.

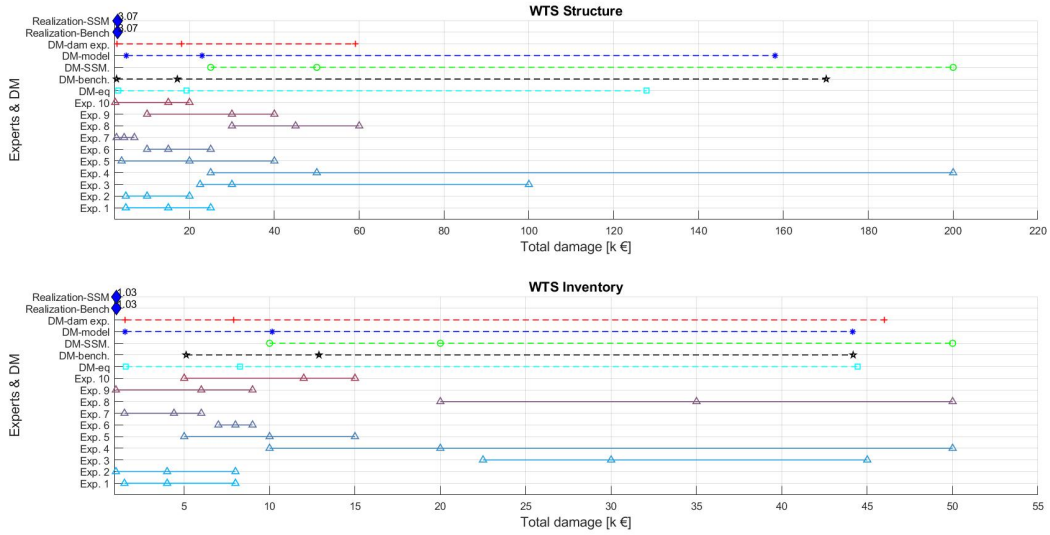


Figure B.2: Calibration questions 1 and 2 for residential damage

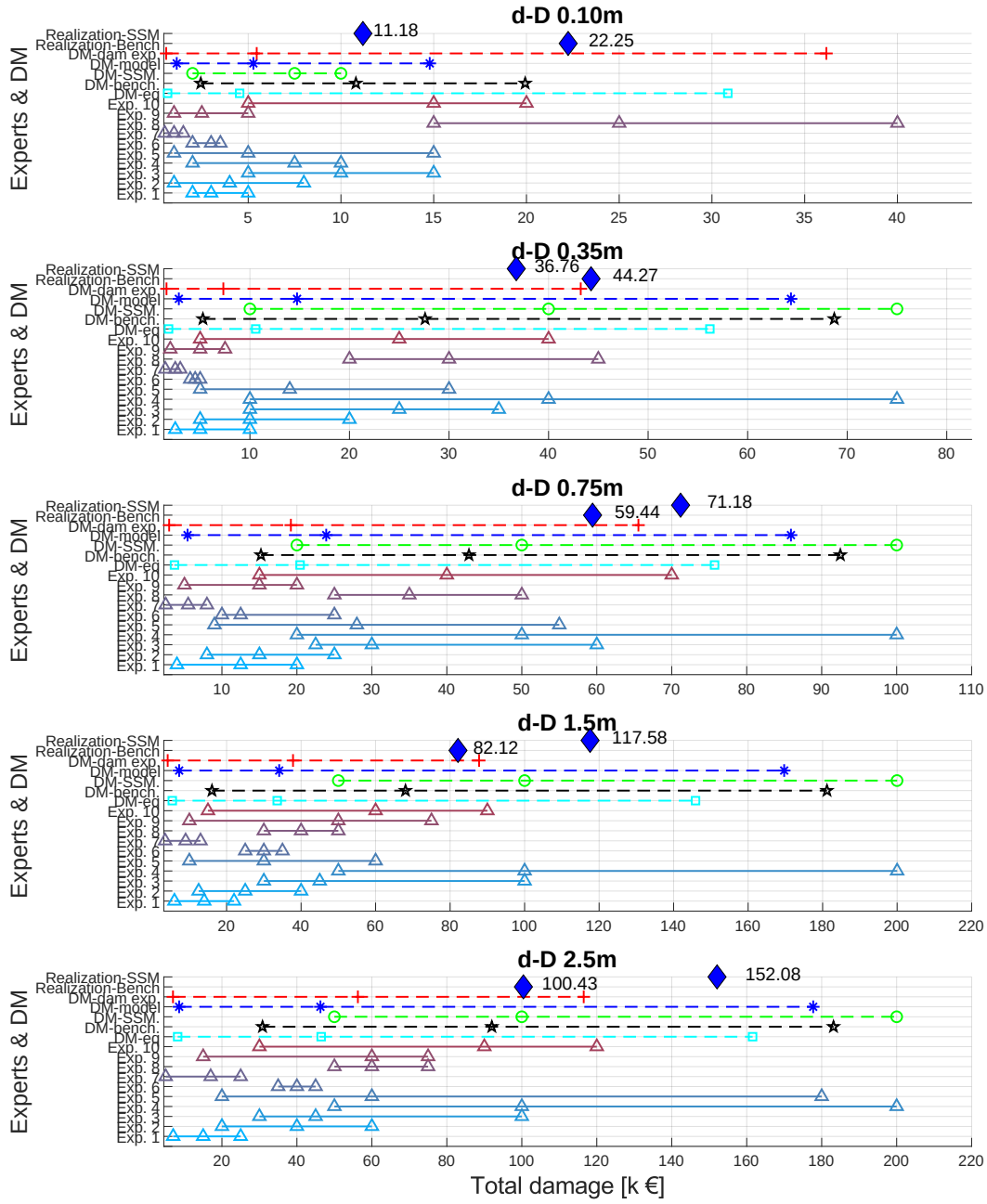


Figure B.3: Calibration questions 2 to 7 for residential damage

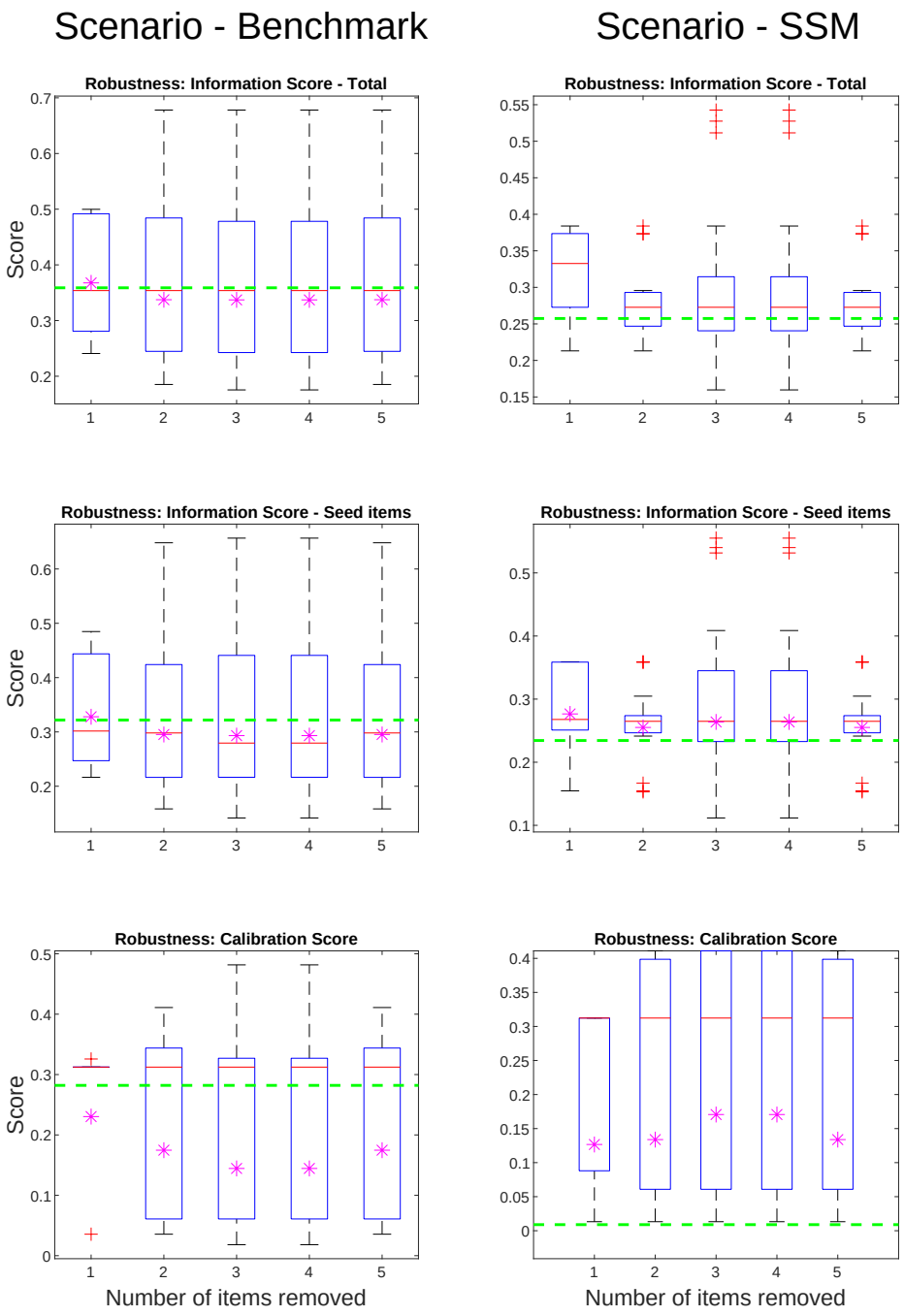


Figure B.4: Robustness plots of removed calibration questions for each calibration scenario. The green dashed horizontal line indicates the initial value before removing items. The magenta star indicates the geometric mean of the variations. The red line and plus markers indicate the median and outliers of the boxplot.

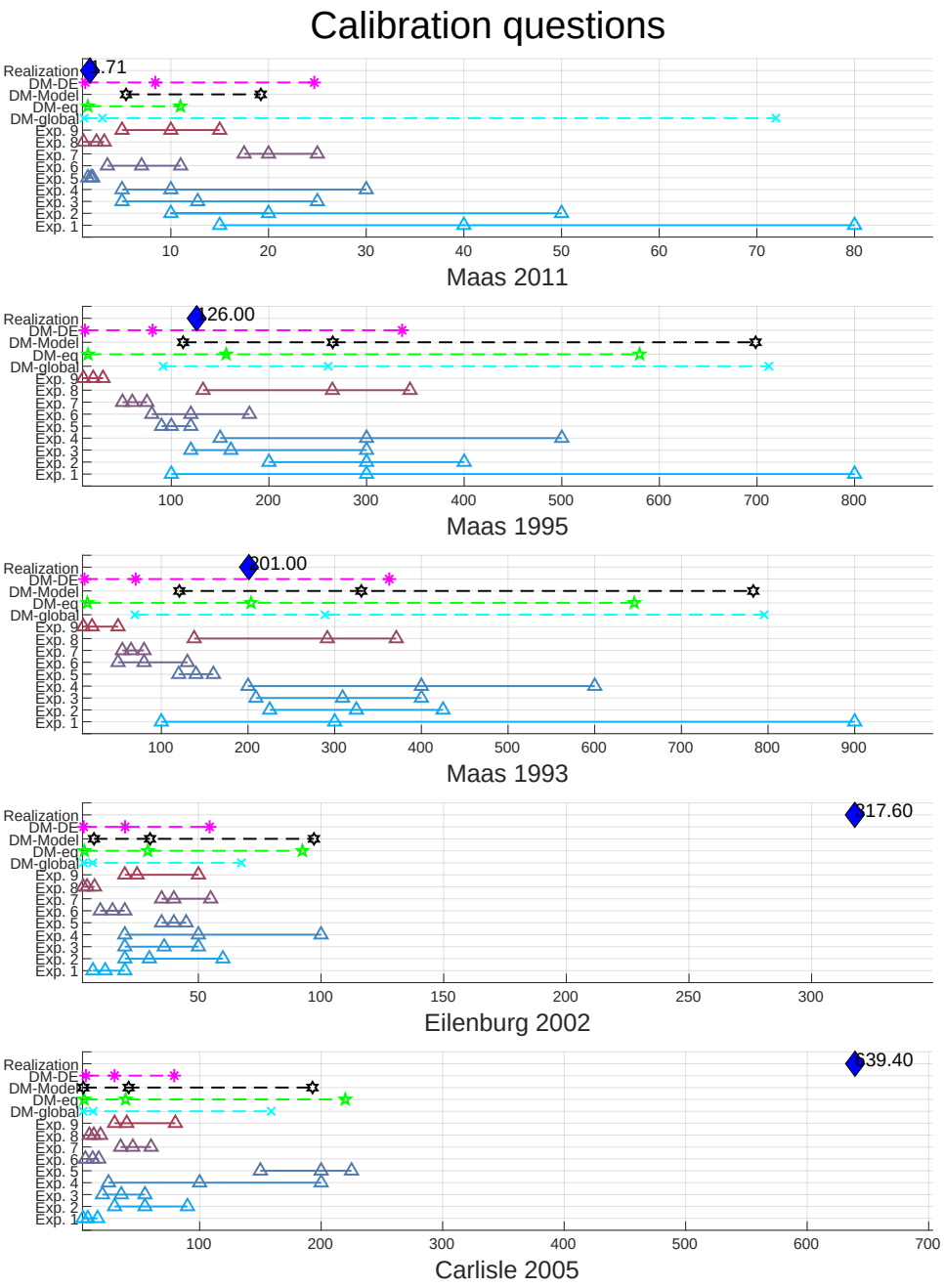


Figure B.5: Calibration questions for total damage estimations

C

Appendix C: Beira supporting figures

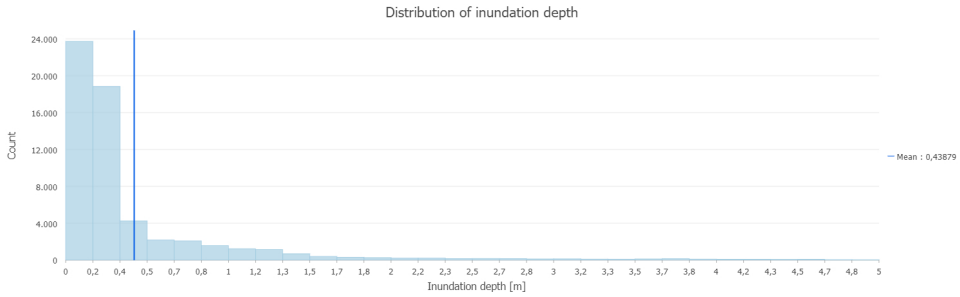


Figure C.1: Distribution of inundation depth at Beira.

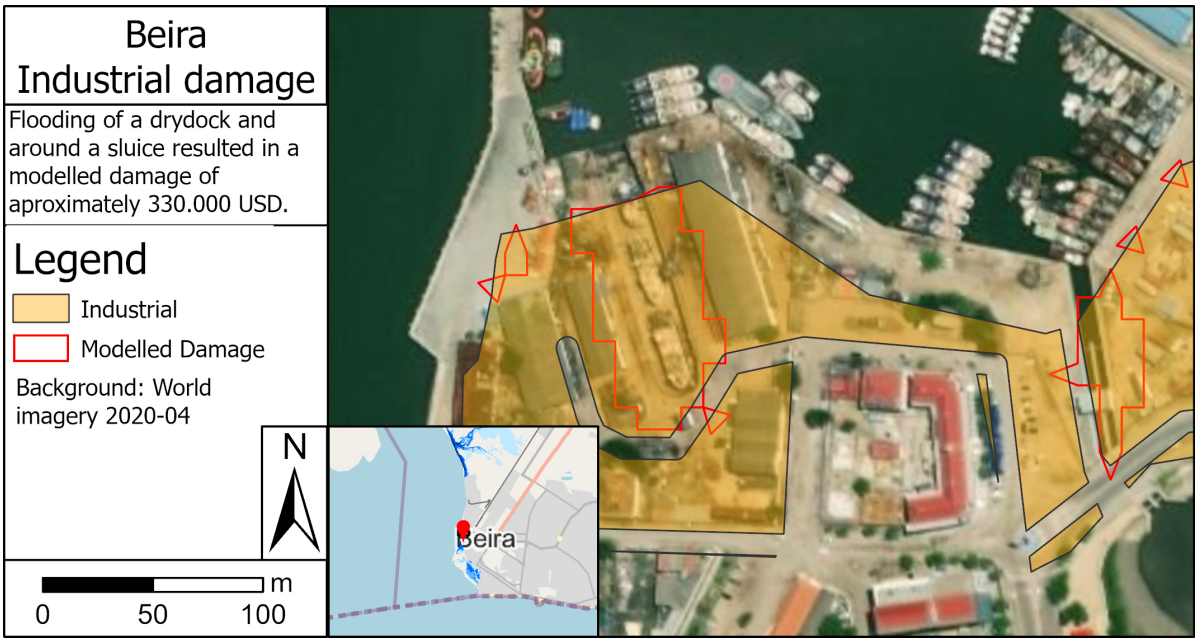


Figure C.2: Flooded industrial area north of Praia Nova

Figure C.4: Aerial pictures of the damaged commercial area in Beira

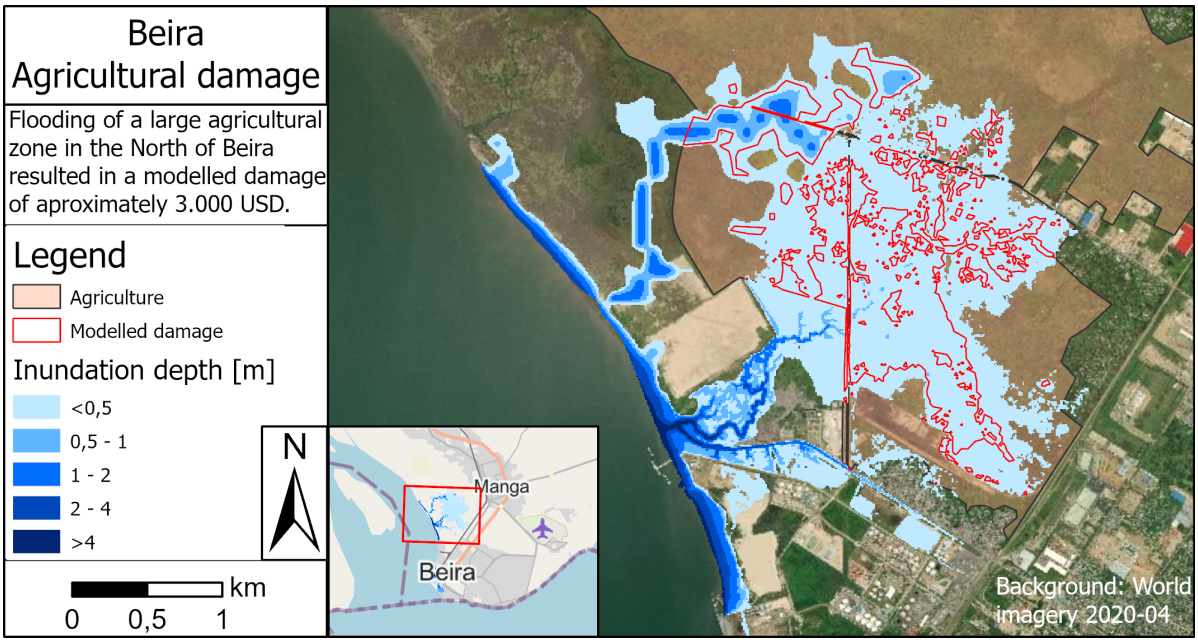


Figure C.3: Flooded agricultural areas to the North of Beira

D

Global damage database approach

D.1. Problem statement

For both case studies, damage data was present in some form. However, the situation can occur, when no suitable historical damage data is present for a case area. This means no verification process is possible with local damage data. However, global damage databases exist, which may be able to fill this gap. These databases track the impacts of multiple hazard types on a global scale. Examples are EM-DAT (2022), NatCatService (Munich Re, 2022) or Desinventar (2022). These databases give high-level event impact data, such as: damage, losses, affected people and casualties.

In this section, it was attempted to use these data to acquire insights into historical flood impacts. These insights may then be used for verification in the case that no local damage data is available. The section consists of two goals: First, it was attempted to gain insights by statistically describing the impact data. Second, spatial data on floodextents, affected landuse and affected socio-economic data was used to find relationships between affected and reported data.

D.2. Methodology

For this approach the data summarized in Table D.1 was used. The main damage data was taken from EM-DAT (2022). The size and global applicability of the dataset make it a valuable database. For this research, Flood events on the african continent between 1985 and 2020 were used, resulting in up to 2871 floods. However, after further investigation, drawbacks became present. The first main drawback is the reliability of the data, as it is unclear what the exact procedure is. Most data is taken from non-academic sources such as various forms of news reporting. Other data may be estimated by models. The second drawback is the fact that not all events report the same amount of data, with many events missing the total damage, insured damages or number of casualties. Only events where flood damage was reported were taken, which left 118 flood events.

Floodextents are taken from Tellman et al. (2021), who recently mapped 913 flood extents accurately using sattelite imagery. Unfortunately, the intersection between the Emdat events that reported damage and the floodextents resulted in only six useful flood events. For landuse, a 20m resolution global dataset was taken from ESA (2017). Socio-economic data on GDP and population was taken from

Table D.1: Data overview used for the global damage data base analysis

Name	Description	Resolution	Period	Reference
Emdat	Global hazard impact database. Used events are African hazards between 1985-2010	-	1985-2010	EM-DAT (2022)
Floodextents	Sattelite derived floodextent maps between 2000 to 2018	250m	2000 - 2018	Tellman et al. (2021)
Corine LCL	Global landuse data	20m	2016	ESA (2017)
Socio-economic	Shared Socio-economic Pathways (SSP) base level data. Global grids with estimated GDP and population	30 Arcsec	2010	O'Neill et al. (2017)

To find the affected GDP and Population for each floodextent, the workflow visualised in Figure D.1 was used. The used data is shown in the left box 1, which was elaborated on earlier in Table D.1. The socio economic data was presented in a grid with non-constant gridcell size[1]. To create constant gridcells, they were fixed locally for each floodextent. To enable the creation of snapped rasters later in the process, a fishnet was created for each locally fixed raster file.

To work with the Landuse data, the assumption was made that all economic activity and population was centered in the landuse type Urban. This was used in step 2: extraction to create two files per flood (So 12 files in total). In the first file, all urban raster cells were extracted for each floodextent. For the second file, all flooded urban raster cells were extracted.

Subsequently in step 3 two raster files were created for each floodextent. The first raster contains the total urban surface area per gridcell, whereas the second file contains the total flooded urban surface area. These files were created by first clipping the extracted rasters from step 2 with a fishnet. Second, the surface area was calculated for each singlepart polygon. Third, the features were transformed to points to enable accurate summation[2]. Fourth, the points were transformed to a raster, that was snapped to the fishnet file. Therefore, this step results in two raster files with the same orientation and resolution as the socio-economic raster per floodextent.

In step 4, the affected GDP and Population was calculated for each floodextent. This was done by first calculating the affected fraction by dividing the affected urban surface area over the total urban surface area for each gridcell. Sbusequently, this fraction was multiplied with the total GDP or Population present in that gridcell to acquire the affected GDP or Population. Finally, all gridcells in a floodextent were summed to find the total amount of affected GDP and affected population per flood.

[1] The data had a constant size in Arcseconds, which varies with distance from the equator
[2] Summing polygons gave inaccurate results due to overlap of features.

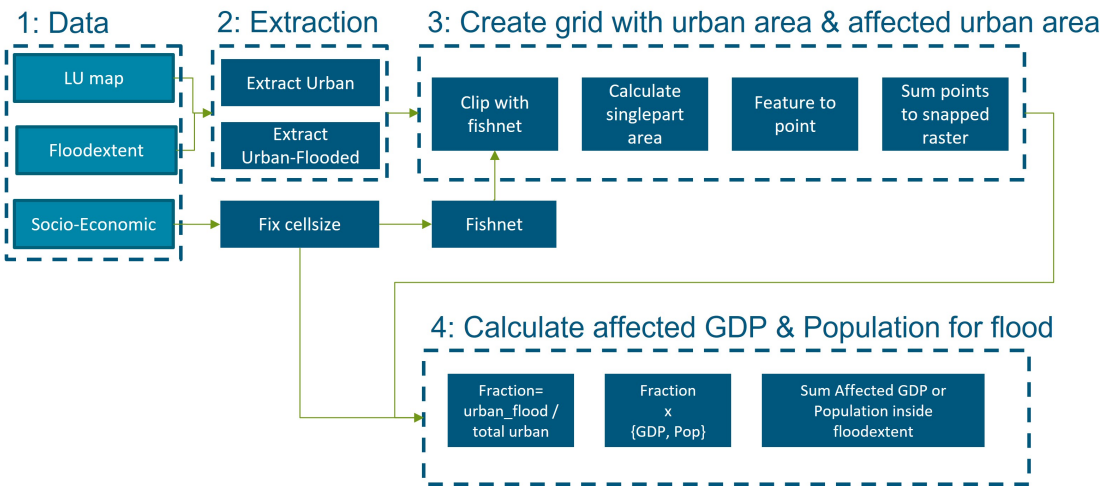


Figure D.1: GIS workflow with data sources to find affected GDP and Population per floodextent

D.3. Results - descriptive statistics

The initial analysis focused on the statistical description of the Emdat database, shown in Figure D.2. A total of 118 events reported flood damages between 1985 and 2010, with most reports being filed after 1995. The distribution of adjusted damages is subsequently shown on the right. The mean damage is 92.4 M USD, adjusted to 2022 price levels. The maximum of 1,8 Billion USD contains a South-African river flood from 1987. Although this was indeed a major flood, large price adjustment due to the age of the flood also played a major role in reaching this maximum[3]. Besides this maximum, other major flood events reach amounts of 750 - 1.000 M USD.

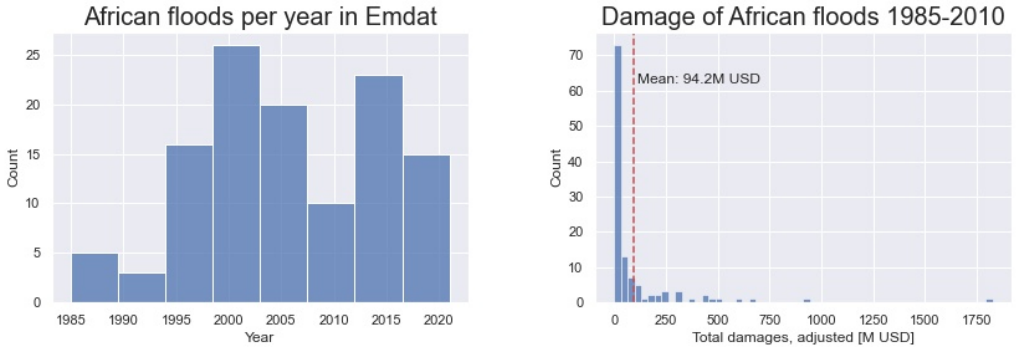


Figure D.2: Statistical overview of EMDAT reported floods for Africa 1985-2010

Besides the total numbers in the dataset, also the Pearson's correlation between reported damage and three variables were analysed across Africa in Table D.2. It was found that reported damage has a significant positive correlation in East, Southern and Western Africa. Much lower correlations were found in Northern Africa, which may be due to the different arid climate. Contrarily, a strong positive correlation was only found in Eastern and Western Africa, but not in Southern Africa. Finally, no relationship was found between damages and GDP/capita, as two areas report positive correlations of 0,20, whereas Eastern and Northern Africa report negative correlations of -0,09 and -0,11.

Looking at flooding types at the bottom of Table D.2, it was found that these relationships only stem from riverine floods, as flash floods report no correlations. As only two coastal floods were reported, not enough data was available to describe correlations[4].

[3] The 1987 estimate was 765 M USD
[4] another group of events existed where the flooding type was unknown

Table D.2: Emdat correlation between adjusted reported damages and one of the three variables Deaths, Affected or GDP/Capita. The correlation is presented for the four sections Southern-, Northern-, Eastern- and Western Africa. Middle africa had only five events, which was too few to reliably calculate correlation. N indicates the total amount of flood events in the area.

Region	**Southern Africa N = 20**			**Northern Africa N = 18**		
Variable	Deaths	Affected	GDP/Capita	Deaths	Affected	GDP/Capita
Correlation	0,87	-0,02	0,20	0,16	0,20	-0,11

Region	**Eastern Africa N = 49**			**Western Africa N = 26**		
Variable	Deaths	Affected	GDP/Capita	Deaths	Affected	GDP/Capita
Correlation	0,54	0,66	-0,09	0,77	0,91	0,20

Flood type	**Flash flood N = 22**			**Riverine flood N = 68**		
Variable	Deaths	Affected	GDP/Capita	Deaths	Affected	GDP/Capita
Correlation	-0,02	0,01	0,21	0,49	0,28	0,27

D.4. Results - Analysed flood events

The intersection of the flood event database (Tellman et al., 2021) and EM-DAT (2022) events with reported damages resulted in six flood events, shown in Table D.3. The left section indicates flood details regarding the Darthmouth Flood Observatory identifier, the main affected country and the year of occurrence. Subsequently, the GIS-found affected GDP, -population and -area are shown. On the right, the affected population, -casualties and reported damage in the emdat database is shown.

Subsequently in Figure D.3, the GIS estimated values are plotted on the x-axis, against the reported values on the y axis. Unfortunately, with three out of the four analysis, no significant correlations were found. On the right top, affected area versus affected population has a significant Pearson's correlation of 0,83. This may be caused by the fact that the reported affected population was also estimated using a spatial model. As GIS-estimated population versus EMDAT population has a lower correlation, the problem may be caused by the spatial resolution of our SSP population dataset.

Table D.3: Data of the six analysed events that intersected between EMDAT (With reported damage) and Tellman et al. (2021).

Event details			**GIS affected estimate**			**Emdat reported data**		
ID[5]	**Country**	**Year**	**Population**	**GDP** [k$]	**Area** [km^2]	**Population**	**Casualties**	**Damage** [k$][6]
2345	Nigeria	2003	2.310	4.041	0,53	210.000	16	3.786
2649	Ethiopia	2005	21.422	24.763	0,34	235.418	156	6.938
2825	South Africa	2006	835	9.067	0,12	4.160	6	95.435
3170	Uganda	2007	9.477	9.397	7,4	71.8045	29	93
4023	Mozambique	2013	47814	349.041	0,49	240.000	119	34.895
4219	Malawi	2015	4.068	3.136	3,3	638.645	278	445.868

[5]Darthmouth Flood Observatory identifier
[6]Adjusted to 2022

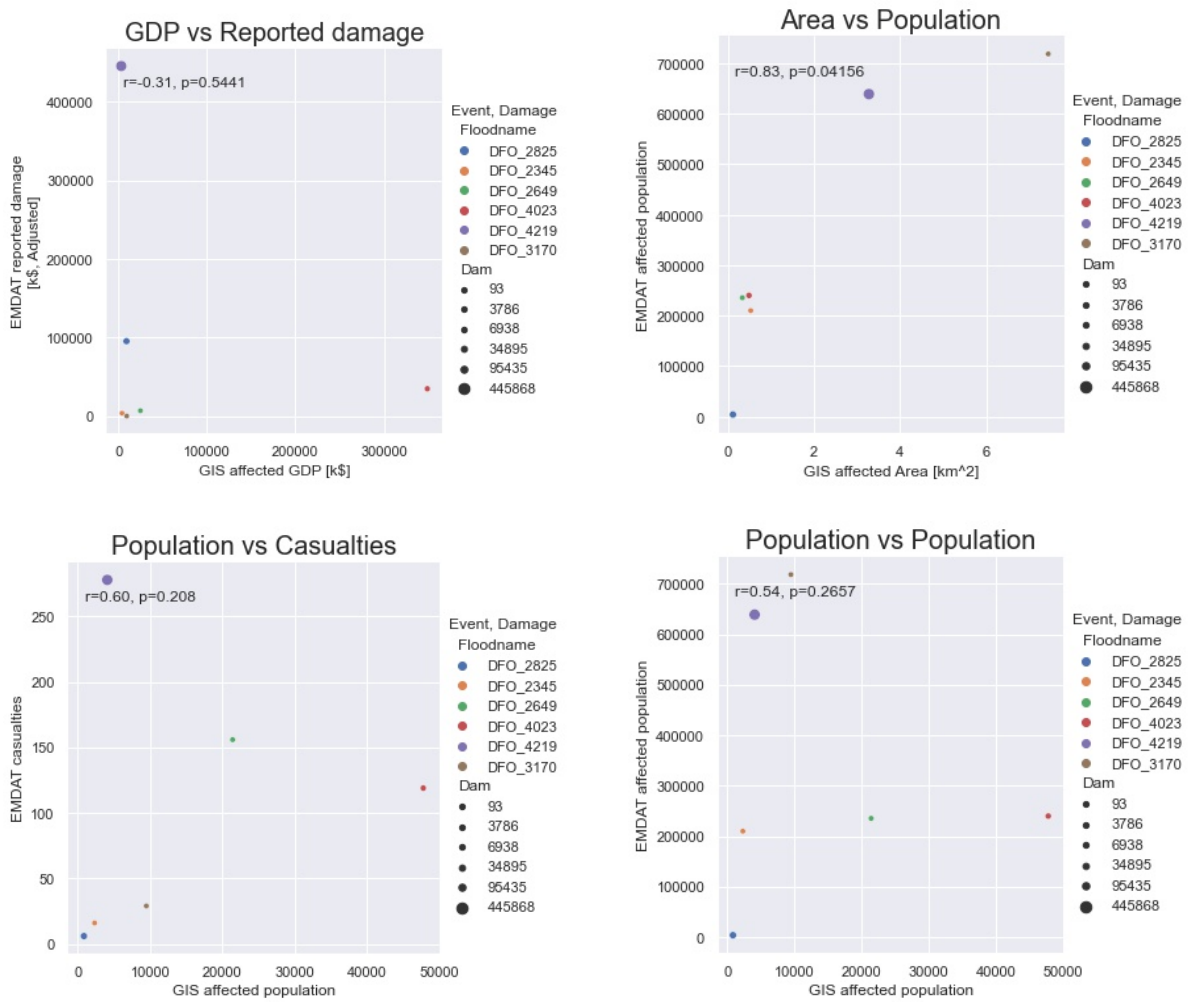


Figure D.3: Results of comparing EMDAT data with GIS analysis. As shown, no significant Pearson's correlation was found.

D.5. Discussion

The descriptive statistics presented in the previous section may be beneficial to acquire an initial overview of flood damages in Africa. However, the method is limited by various points. First and foremost the source of the data and its reliability should be considered. The exact source in EM-DAT (2022) differs per entry, but the major part is based upon sources such as newspapers and articles, of which the reliability can be strongly questioned. Second, these databases may be biased, as they are not comprehensive on all ocurring hazards. For instance, research by Moriyama et al. (2018) and Ward et al. (2013) indicated large geographical differences in coverage between multiple hazard databases. Finally, also the complex modelling process is prone to errors due to the handling of large datasets, the relatively large spatial resolution, and complex coordinate transformations.

References

Achmea. (2022). Dataset - total damage to private residential houses including inventory - valued by Achmea damage experts. Communicated by employee through email.

Aerts, J. C., Botzen, W. J., Emanuel, K., Lin, N., De Moel, H., & Michel-Kerjan, E. O. (2014). Climate adaptation: Evaluating flood resilience strategies for coastal megacities. *Science*, *344*(6183), 473–475. https://doi.org/10.1126/SCIENCE.1248222/SUPPL{_}FILE/1248222.AERTS.SM.PDF

AfdB. (2014). *Study on Road Infrastructure Costs- Analysis of Unit Costs and Cost Overruns of Road Infrastructure Projects in Africa* (tech. rep.). African Development Bank. www.afdb.org

Africon. (2008). *Africa Infrastructure Country Diagnostic* (tech. rep.). Africon, commissioned by the World Bank. www.infrastructureafrica.org.

AIR Worldwide. (2012). Modeling U.S. Hurricane Risk to Pleasure Boats | AIR Worldwide. https://www.air-worldwide.com/publications/air-currents/2012/Modeling-U-S--Hurricane-Risk-to-Pleasure-Boats/

Alfieri, L., Feyen, L., Dottori, F., & Bianchi, A. (2015). Ensemble flood risk assessment in Europe under high end climate scenarios. *Global Environmental Change*, *35*, 199–212. https://doi.org/10.1016/J.GLOENVCHA.2015.09.004

Amadio, M., Mysiak, J., Carrera, •. L., & Koks, E. (2016). Improving flood damage assessment models in Italy. *Natural Hazards*, *82*, 2075–2088. https://doi.org/10.1007/s11069-016-2286-0

Apel, H., Aronica, A. G. T., Kreibich, A. H., Thieken, A. A. H., Aronica, G. T., & Thieken, A. H. (2009). Flood risk analyses-how detailed do we need to be? *Natural Hazards*, *49*, 79–98. https://doi.org/10.1007/s11069-008-9277-8

Apel, H., Merz, B., & Thieken, A. H. (2008). Quantification of uncertainties in flood risk assessments. *International Journal of River Basin Management*, *6*(2), 149–162. https://doi.org/10.1080/15715124.2008.9635344

Arnell, N. W., & Gosling, S. N. (2014). The impacts of climate change on river flood risk at the global scale. *Climatic Change 2014 134:3*, *134*(3), 387–401. https://doi.org/10.1007/S10584-014-1084-5

Arrighi, C., Brugioni, M., Castelli, F., Franceschini, S., & Mazzanti, B. (2013). Urban micro-scale flood risk estimation with parsimonious hydraulic modelling and census data. *Hazards Earth Syst. Sci*, *13*, 1375–1391. https://doi.org/10.5194/nhess-13-1375-2013

Arrighi, C., Ballio, F., Carisi, F., Castelli, F., Domeneghetti, A., Gallazzi, A., Galliani, M., Grelot, F., Kellermann, P., Kreibich, H., Molinari, D., Mohor, G. S., Mosimann, M., Natho, S., Richert, C., Schröter, K., Scorzini, A. R., Thieken, A. H., & Zischg, A. P. (2021). A comparative analysis of flood damage models: lessons learnt and future challenges. *FLOODrisk 2020 - 4th European Conference on Flood Risk Management*, null–null. https://doi.org/10.3311/FLOODRISK2020.9.15

Balci, O. (1994). Validation, verification, and testing techniques throughout the life cycle of a simulation study. *Annals of Operations Research*, *53*, 121–173. https://doi.org/https://doi.org/10.1007/BF02136828

Bales, J., & Wagner, C. (2009). Sources of uncertainty in flood inundation maps. *Journal of Flood Risk Management*, *2*(2), 139–147. https://doi.org/10.1111/J.1753-318X.2009.01029.X

Baradaranshoraka, M., Asce, S. M., Pinelli, J.-P., Asce, M., Gurley, K., Peng, ; X., & Zhao, M. (2017). Hurricane Wind versus Storm Surge Damage in the Context of a Risk Prediction Model. *Journal of Structural Engineering*, *143*(9), 04017103. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001824

Barneveld, H., Boersema, M., Schuurman, F., & de Vriend, H. (2021). *Het verhaal van het sediment - Rijkswaterstaat Rapportendatabank* (tech. rep.). Rijkswaterstaat. https://puc.overheid.nl/rijkswaterstaat/doc/PUC_630145_31/1/

Beira city government. (2019). *Beira Municipal Recovery and Resilience Plan* (tech. rep.).

Beven, K., Koutsoyiannis, D., & Weijs, S. (2016). Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, *61*(9), 1652–1665. https://doi.org/10.1080/02626667.2015.1031761

Bisschop, C., de Groot, B., Zethof, M., & Leenders, J. (2015). *Overstromingsrisico Maasdal Limburg* (tech. rep.). Rijkswaterstaat.

Booysen, H., Viljoen, M. F., & Villiers, G. D. T. (1999). Methodology for the calculation of industrial flood damage and its application to an industry in Vereeniging. *25*, 41–46.

Bos, M., Peters, G., Driessen, T., Voesenek, D., & Huting, R. (2020). Comprehensive flood risk analysis at your fingertip through cloud-computing and standardized approach. *FLOODRisk 2020*. https://doi.org/10.3311/FLOODRisk2020.9.21

Botzen, W. J. W., & Bergh, J. C. J. M. V. D. (2008). Insurance Against Climate Change and Flooding in the Netherlands: Present, Future, and Comparison with Other Countries. *Risk Analysis*, *28*(2), 413–426. https://doi.org/10.1111/J.1539-6924.2008.01035.X

Brémond, P., & Grelot, F. (2013). Review Article: Economic evaluation of flood damage to agriculture – review and analysis of existing methods. *Natural Hazards and Earth System Sciences*, *13*(10), 2493–2512. https://doi.org/10.5194/NHESS-13-2493-2013

Budiyono, Y., Aerts, J., Brinkman, J. J., Marfai, M. A., & Ward, P. (2015). Flood risk assessment for delta mega-cities: a case study of Jakarta. *Natural Hazards*, *75*(1), 389–413. https://doi.org/10.1007/S11069-014-1327-9/FIGURES/7

CAHF. (2019). *Understanding Mozambique's housing finance market* (tech. rep.). CAHF - Center for Affordable Housing Finance in Africa. http://housingfinanceafrica.org/app/uploads/Mozambique-May-2019.pdf

Carley, K. M. (2017). *Validating Computational Models* (tech. rep.). Institute for Software Research - Carnegie Mellon University. Pittsburgh. http://reports-archive.adm.cs.cmu.edu/anon/anon/home/ftp/usr0/ftp/isr2017/CMU-ISR-17-105.pdf

CBS. (2015). Dataset: CBS Bestand Bodemgebruik. https://www.pdok.nl/introductie/-/article/cbs-bestand-bodemgebruik

CBS. (2021). Landbouw; economische omvang naar omvangsklasse, bedrijfstype. https://www.cbs.nl/nl-nl/cijfers/detail/80785ned

Chen, Y. R., Yeh, C. H., & Yu, B. (2011). Integrated application of the analytic hierarchy process and the geographic information system for flood risk assessment and flood plain management in Taiwan. *Natural Hazards*, *59*(3), 1261–1276. https://doi.org/10.1007/S11069-011-9831-7/FIGURES/12

CIA. (2021). Mozambique - The World Factbook. https://www.cia.gov/the-world-factbook/countries/mozambique/#people-and-society

Colson, A. R., & Cooke, R. M. (2017). Cross validation for the classical model of structured expert judgment. https://doi.org/10.1016/j.ress.2017.02.003

Cooke, R. M., & Goossens, L. H. J. (2000). *Procedures guide for structured expert judgment*. European Commission.

Cooke, R. M., & Goossens, L. L. H. J. (2008). TU Delft expert judgment data base. *Reliability Engineering and System Safety*, *93*, 657–674. https://doi.org/10.1016/j.ress.2007.03.005

Davis, S., & Skaggs, L. (1992). Catalog of residential depth-damage functions used by the army corps of engineers in flood damage estimation. https://apps.dtic.mil/sti/citations/ADA255462

De Bruijn, K. M., Cumiskey, L., Ní Dhubhda, R., Hounjet, M., & Hynes, W. (2016). Flood vulnerability of critical infrastructure in Cork, Ireland. *E3S Web of Conferences*, *7*. https://doi.org/10.1051/E3SCONF/20160707005

De Bruijn, Wagenaar, Slager, K., De Bel, M., & Burzel, A. (2015). *Updated and improved method for flood damage assessment: SSM2015 (version 2)* (tech. rep.). Deltares. Delft.

de Moel, H., & Aerts, J. C. J. H. (2010). Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates. *Natural Hazards 2010 58:1*, *58*(1), 407–425. https://doi.org/10.1007/S11069-010-9675-6

de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E., & Ward, P. J. (2015). Flood risk assessments at different spatial scales. *Mitigation and Adaptation Strategies for Global Change*, *20*(6), 865–890. https://doi.org/10.1007/S11027-015-9654-Z

de Ruig, L. T., Barnard, P. L., Botzen, W. J., Grifman, P., Hart, J. F., de Moel, H., Sadrpour, N., & Aerts, J. C. (2019). An economic evaluation of adaptation pathways in coastal mega cities: An

illustration for Los Angeles. *Science of the Total Environment*, *678*, 647–659. https://doi.org/10.1016/J.SCITOTENV.2019.04.308

Desinventar. (2022). United Nations DesInventar Open Source Initiative - Official Website. https://www.desinventar.net/

DNB. (2021). Individuele gegevens verzekeraars solo (Jaarlijks). https://www.dnb.nl/statistieken/data-zoeken/#/details/individuele-gegevens-verzekeraars-jaar/dataset/d66feb58-e89b-4c73-b0c7-cab81188ca77

Downton, M. W., & Pielke, R. A. (2005). How Accurate are Disaster Loss Data? The Case of U.S. Flood Damage. *Natural Hazards 2005 35:2*, *35*(2), 211–228. https://doi.org/10.1007/S11069-004-4808-4

Ecopia, & DigitalGlobe. (2017). *DigitalGlobe Building Footprints powered by GBDX in partnership with Ecopia* (tech. rep.). Bill & Melinda Gates foundation.

EM-DAT. (2022). EM-DAT public flood data of African floods between 1985-2020. https://public.emdat.be/data.

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

Erffmeyer, R. C., & Lane, I. M. (1984). Quality and acceptance of an evaluative task: The effects of four group decision-making formats. *Group & Organization Studies*, *9*(4), 509–529.

ESA. (2017). ESA & Copernicus CCI LAND COVER – S2 prototype Land Cover 20m map of Africa 2016. https://2016africalandcover20m.esrin.esa.int/download.php

Expertise Netwerk Waterveiligheid. (2021). *Hoogwater 2021 Feiten en Duiding* (tech. rep.). https://www.enwinfo.nl/publicaties/

FAO. (2021). FAO GIEWS Country Brief on Mozambique. https://www.fao.org/giews/countrybrief/country.jsp?code=MOZ&lang=FR

Förster, S., Kuhlmann, B., Lindenschmidt, K.-E., & Bronstert, A. (2008). Assessing flood risk for a rural detention area. *Hazards Earth Syst. Sci*, *8*, 311–322. www.nat-hazards-earth-syst-sci.net/8/311/2008/

GFDRR. (2008). *Post-Disaster Needs Assessments Guidelines* (tech. rep.). The european commission, The United Nations development group, The World Bank.

GFDRR. (2019). *Mozambique Cyclone Idai Post Disaster Needs Assessment Conference Version* (tech. rep.). GFDRR - Global Faility for Disaster Reduction and Recovery.

Grahn, T., & Nyberg, L. (2017). Assessment of pluvial flood exposure and vulnerability of residential areas. *International Journal of Disaster Risk Reduction*, *21*, 367–375. https://doi.org/10.1016/J.IJDRR.2017.01.016

Green, I. R. A., & Stephenson, D. (1986). Criteria for comparison of single event models. *Hydrological Sciences Journal*, *31*(3), 395–411. https://doi.org/10.1080/02626668609491056

Habermann, N., & Hedel, R. (2018). Damage functions for transport infrastructure. *International Journal of Disaster Resilience in the Built environment*, *9*(4/5), 420–434. https://doi.org/10.1108/IJDRBE-09-2017-0052

Hallowell, M. R., & Gambatese, J. A. (2010). Qualitative Research: Application of the Delphi Method to CEM Research. *Journal of Construction Engineering and Management*, *136*(1), 99–107. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000137

Hasanzadeh Nafari, R., Ngo, T., & Lehman, W. (2016). Calibration and validation of FLFArs-A new flood loss function for Australian residential structures. *Natural Hazards and Earth System Sciences*, *16*(1), 15–27. https://doi.org/10.5194/NHESS-16-15-2016

Herath, S., Dutta, D., & Musiake, K. (2003). *Flood damage estimation of an urban catchment using remote sensing and GIS.* (tech. rep.). United Nations University Japan.

Hess, T. M., & Morris, J. (1988). Estimating the Value of Flood Alleviation on Agricultural Grassland. *Agricultural Water Management*, *15*, 141–153.

Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S., Marzeion, B., Fettweis, X., Ionescu, C., & Levermann, A. (2014). Coastal flood damage and adaptation costs under 21st century sea-level rise. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3292–3297. https://doi.org/10.1073/PNAS.1222469111/-/DCSUPPLEMENTAL

Huizinga, J., de Moel, H., & Szewczyk, W. (2017). *Global flood depth-damage functions - Publications Office of the EU* (tech. rep.). Joint Research Centre (European Commission). https://doi.org/10.2760/16510

ICEYE. (2022). Flood monitoring | Get accurate data on water-caused damage to the insured property during floods, in any location on Earth.| ICEYE. https://www.iceye.com/solutions/flood-monitoring

INGC. (2019). Cyclone Idai - Aerial Survey - Humanitarian Data Exchange - INGC (National Disaster Management Authority of Mozambique). https://data.humdata.org/dataset/50e37b45-25a1-4ef4-8c83-8b2fdf923773

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

Jonkman, Bočkarjova, M., Kok, M., & Bernardini, P. (2008). Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. *Ecological Economics*, *66*(1), 77–90. https://doi.org/10.1016/J.ECOLECON.2007.12.022

Kadaster. (2021a). Dataset: Basisregistratie Adressen en Gebouwen (BAG) | Data overheid. https://data.overheid.nl/dataset/basisregistratie-adressen-en-gebouwen--bag-

Kadaster. (2021b). Dataset: Basisregistratie Topografie (BRT) TOPNL. https://www.pdok.nl/downloads/-/article/basisregistratie-topografie-brt-topnl

Kellermann, P., Schröter, K., Thieken, A., Haubrock, S.-N., & Kreibich, H. (2020). The object-specific flood damage database HOWAS21. *Natural Hazards and Earth System Sciences*, 1–30. https://doi.org/10.5194/NHESS-2019-420

Klügl, F. (2008). A validation methodology for agent-based simulations. *Proceedings of the ACM Symposium on Applied Computing*, 39–43. https://doi.org/10.1145/1363686.1363696

Kok, M., Vrouwenvelder, A., & van den Braak, W. (2005). *Standaardmethode 2005 - Schade en Slachtoffers als gevolg van overstromingen* (tech. rep.). HKV Lijn in Water en TNO Bouw.

Koks, E. E., Bočkarjova, M., de Moel, H., & Aerts, J. C. (2015). Integrated Direct and Indirect Flood Risk Modeling: Development and Sensitivity Analysis. *Risk Analysis*, *35*(5), 882–900. https://doi.org/10.1111/RISA.12300

Kron, W., Steuer, M., Löw, P., & Wirtz, A. (2012). How to deal properly with a natural catastrophe database-analysis of flood losses. *Hazards Earth Syst. Sci*, *12*, 535–550. https://doi.org/10.5194/nhess-12-535-2012

Kron, W. (2005). Flood Risk = Hazard • Values • Vulnerability, Water International. *Water International*, *30*(1), 58–68. https://doi.org/10.1080/02508060508691837

Ludwig, R., & Schneider, P. (2006). Validation of digital elevation models from SRTM X-SAR for applications in hydrologic modeling. *ISPRS Journal of Photogrammetry and Remote Sensing*, *60*(5), 339–358. https://doi.org/10.1016/J.ISPRSJPRS.2006.05.003

Macrotrends. (2022). Euro Dollar Exchange Rate (EUR USD) - Historical Chart | MacroTrends. https://www.macrotrends.net/2548/euro-dollar-exchange-rate-historical-chart

Mannakkara, S., & Wilkinson, S. (2014). Re-conceptualising “Building Back Better” to improve post-disaster recovery. *International Journal of Managing Projects in Business*, *7*(3), 327–341. https://doi.org/10.1108/IJMPB-10-2013-0054/FULL/PDF

Matgen, P., Martinis, S., Wagner, W., Freeman, V., Zeil, P., & McCormick, N. (2020). Feasibility assessment of an automated, global, satellite-based flood-monitoring product for the Copernicus Emergency Management Service. *Luxembourg: Publications Office of the European Union*.

Matheney, T. (2022). How Much Office Space Do I Need? - Aquila real estate. https://aquilacommercial.com/learning-center/how-much-office-space-need-calculator-per-person/

Maxar. (2019). Cyclone Idai flood map | Maxar. https://www.maxar.com/open-data/cyclone-idai

McBean, E., Fortin, M., & Gorrie, J. (1986). Critical Analysis of Residential Flood Damage Estimation Curves. *Canadian journal of civil engineering*, *13*(1), 86–94. https://doi.org/10.1139/L86-012

Mediero, L., Garrote, L., Santillán, D., Cueto-Felgueroso, L., & Soriano, E. (2021). Discharge-damage curve and EAD estimates in the Pamplona metropolitan area (Spain) by using insurance data. *FLOODrisk 2020 - 4th European Conference on Flood Risk Management*, null–null. https://doi.org/10.3311/FLOODRISK2020.7.3

Merz, B., Kreibich, H., & Apel, H. (2008). Flood risk analysis: uncertainties and validation. *Österreichische Wasser- und Abfallwirtschaft 2008 60:5*, *60*(5), 89–94. https://doi.org/10.1007/S00506-008-0001-4

Merz, B., Kreibich, H., Schwarze, R., & Thieken, A. (2010). Review article "assessment of economic flood damage
. *Natural Hazards and Earth System Science*, *10*(8), 1697–1724. https://doi.org/10.5194/NHESS-10-1697-2010

Merz, B., Kreibich, H., Thieken, A., & Schmidtke, R. (2004). Estimation uncertainty of direct monetary flood damage to buildings. *Natural Hazards and Earth System Science*, *4*(1), 153–163. https://doi.org/10.5194/NHESS-4-153-2004

Messner, F., Penning-Rowsell, E., Green, C., Meyer, V., Tunstall, S., Van Der Veen, A., Tapsell, S., Wilson, T., Krywkow, J., Logtmeijer, C., Fernández-Bilbao, A., Geurts, P., Haase, D., & Parker, D. (2007). Evaluating flood damages: guidance and recommendations on principles and methods. *Integrated Flood Risk Analysis and Management Methodologies*.

Middelmann-Fernandes, M. H. (2010). Flood damage estimation beyond stage-damage functions: An Australian example. *Journal of Flood Risk Management*, *3*(1), 88–96. https://doi.org/10.1111/J.1753-318X.2009.01058.X

Moel, H. d., Bouwer, L., & Aerts, J. (2014). Uncertainty and sensitivity of flood risk calculations for a dike ring in the south of the Netherlands. *Science of the Total Environment*, *473-474*(1), 224–234. https://doi.org/10.1016/J.SCITOTENV.2013.12.015

Molinari, D., De Bruijn, K. M., Castillo-Rodríguez, J. T., Aronica, G. T., & Bouwer, L. M. (2019). Validation of flood risk models: Current practice and possible improvements. *International Journal of Disaster Risk Reduction*, *33*, 441–448. https://doi.org/10.1016/J.IJDRR.2018.10.022

Molinari, D., Rita Scorzini, A., Arrighi, C., Carisi, F., Castelli, F., Domeneghetti, A., Gallazzi, A., Galliani, M., Grelot, F., Kellermann, P., Kreibich, H., Mohor, G. S., Mosimann, M., Natho, S., Richert, C., Schroeter, K., Thieken, A. H., Paul Zischg, A., & Ballio, F. (2020). Are flood damage models converging to "reality"? Lessons learnt from a blind test. *Natural Hazards and Earth System Sciences*, *20*(11), 2997–3017. https://doi.org/10.5194/NHESS-20-2997-2020

Moriyama, K., Sasaki, D., & Ono, Y. (2018). Comparison of global databases for disaster loss and damage data. *Journal of Disaster Research*, *13*(6), 1007–1014. https://doi.org/10.20965/JDR.2018.P1007

Munich Re. (2021). Record hurricane season and major wildfires – The natural disaster figures for 2020 | Munich Re. https://www.munichre.com/en/company/media-relations/media-information-and-corporate-news/media-information/2021/2020-natural-disasters-balance.html

Munich Re. (2022). NatCatSERVICE | Munich Re. https://www.munichre.com/en/solutions/for-industry-clients/natcatservice.html

Municipality of Beira. (2019). *Beira Municipal Recovery and Resilience Plan* (tech. rep.). Municipality of Beira.

Municipality Utrecht. (2021). Nadere regels vergoedingen voor schoolgebouwen.

NAIC. (2022). National Association of Insurance Commissioners - Rebuilding After a Storm: Know the Difference Between Replacement Cost and Actual Cash Value When It Comes to Your Roof. https://content.naic.org/article/rebuilding-after-storm-know-difference-between-replacement-cost-and-actual-cash-value-when-it-comes

Nogal, M. ; & Nogal, A. (2021). Sensitivity method for extreme-based engineering problems. https://doi.org/10.1016/j.ress.2021.107997

O'Hagan, A. (2019). Expert Knowledge Elicitation: Subjective but Scientific. *THE AMERICAN STATISTICIAN*, *2019*(S1), 69–81. https://doi.org/10.1080/00031305.2018.1518265

O'Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., van Ruijven, B. J., van Vuuren, D. P., Birkmann, J., Kok, K., Levy, M., & Solecki, W. (2017). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change*, *42*, 169–180. https://doi.org/10.1016/J.GLOENVCHA.2015.01.004

Pakoksung, K., & Takagi, M. (2015). Digital elevation models on accuracy validation and bias correction in vertical. *Modeling Earth Systems and Environment 2015 2:1*, *2*(1), 1–13. https://doi.org/10.1007/S40808-015-0069-3

Paprotny, D., Morales-Nápoles, O., & Jonkman, S. (2018). HANZE: a pan-European database of exposure to natural hazards and damaging historical floods since 1870. *Earth System Science Data*, *10*(1), 565–581. https://doi.org/10.5194/ESSD-10-565-2018

Paprotny, D., Vousdoukas, M. I., Morales-Nápoles, O., Jonkman, S. N., & Feyen, L. (2018). Compound flood potential in Europe. *Hydrology and Earth System Sciences Discussions*. https://doi.org/10.5194/HESS-2018-132

Parodi, M. U., Giardino, A., Van Dongeren, A., Pearson, S. G., Bricker, J. D., & Reniers, A. J. (2020). Uncertainties in coastal flood risk assessments in small island developing states. *Natural Hazards and Earth System Sciences*, *20*(9), 2397–2414. https://doi.org/10.5194/NHESS-20-2397-2020

Patankar, A., & Patwardhan, A. (2015). Estimating the uninsured losses due to extreme weather events and implications for informal sector vulnerability: a case study of Mumbai, India. *Natural Hazards 2015 80:1*, *80*(1), 285–310. https://doi.org/10.1007/S11069-015-1968-3

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

Pistrika, A. K., & Jonkman, S. N. (2010). Damage to residential buildings due to flooding of New Orleans after hurricane Katrina. *Natural Hazards*, *54*(2), 413–434. https://doi.org/10.1007/S11069-009-9476-Y/FIGURES/20

Pivot, J. M., & Martin, P. (2002). Farms adaptation to changes in flood risk: a management approach. *Journal of Hydrology*, *267*(1-2), 12–25. https://doi.org/10.1016/S0022-1694(02)00136-1

Probst, P., & Annunziato, A. (2019). *Tropical Cyclone IDAI: analysis of the wind, rainfall and storm surge impact* (tech. rep.). European Commission Joint Research Centre. http://www.gdacs.org/Knowledge/models_tc.aspx

Reach. (2019). *Impact of Idai on Beira* (tech. rep.). REACH initiative. Geneva. https://reachinitiative.maps.arcgis.com/apps/Cascade/index.html?appid=de5e77cf935a4007812abbd7eecb9a00

RHDHV. (2019). *Beira Flood and Erosion Assessment report* (tech. rep.). Amersfoort.

RHDHV, & Celsius Pro. (2020). *Design of a Risk Metric to Quantify the Impacts of TCs in Mozambique. Component 1: Research and quantification of the financial impacts of TCs on GoM. Development of vulnerability modules for inclusion in the catastrophe risk model.* (tech. rep.). Requested by the world bank.

Rijkswaterstaat. (2021a). Dijkversterkingen Maaswerken gereed | Rijkswaterstaat. https://www.rijkswaterstaat.nl/nieuws/archief/2021/02/dijkversterkingen-maaswerken-zijn-gereed

Rijkswaterstaat. (2021b). Nationaal Wegenbestand. https://nationaalwegenbestand.nl/

Rijkswaterstaat. (2021c). Waterstanden Maas - Rijkswaterstaat data.

Romanowicz, R., & Beven, K. (2003). Estimation of flood inundation probabilities as conditioned on event inundation maps. *Water Resources Research*, *39*(3), 1073. https://doi.org/10.1029/2001WR001056

Rudari, R., Campo, L., & Trasforini, E. (2017). *Final report tropical cyclone risks RISK Computation* (tech. rep.). Global Facility for Disaster Reduction and Recovery | CIMA - Deltares.

RVO. (2021a). Dataset: Basisregistratie Gewaspercelen (BRP).

RVO. (2021b). Types of damage, costs and payouts (in dutch) Soorten waterschade, kosten en de maximale vergoeding. https://www.rvo.nl/subsidies-financiering/waterschaderegelingen-juli-2021/wts/schade-en-kostensoorten#waterschade-bedrijven

RVO. (2021c). Veelgestelde vragen waterschade mest | Rijksdienst voor ondernemend nederland. https://www.rvo.nl/onderwerpen/agrarisch-ondernemen/waterschade-limburg/veelgestelde-vragen-mest

RVO. (2022). Dataset: WTS data agregated on postal code 6 level for 12 different categories.

Santos, R., & Salvucci, V. (2016). *UNU-WIDER : Policy Brief : Poverty in Mozambique* (tech. rep.). United Nations. https://www.wider.unu.edu/publication/poverty-mozambique

Sayers, P., Lamb, R., Panzeri, M., Bowman, H., Hall, J., Horritt, M., & Penning-Rowsell, E. (2016). Believe it or not? The challenge of validating large scale probabilistic risk models. DOI:%2010.1051/e3sonf/20160711004

Slager, K. (2021). Personal interview regarding floodmap of Limburg 2021 floods created with the fact finding report.

Slager, K., & Wagenaar, D. (2017). *Standaardmethode 2017 Schade en slachtoffers als gevolg van overstromingen* (tech. rep.). Deltares. Delft.

Smith, D. I. (1994). Flood damage estimation-A review of urban stage-damage curves and loss functions. *Water SA*, *Volume 20*(3). https://journals.co.za/doi/abs/10.10520/AJA03784738_1124

STOWA. (2013). *WaterSchadeSchatter (WSS); gebruikershandleiding.* (tech. rep.). STOWA. Amersfoort.

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., & Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature 2021 596:7870*, *596*(7870), 80–86. https://doi.org/10.1038/s41586-021-03695-w

Tennøy, A., Kværner, J., & Gjerstad, K. I. (2012). Uncertainty in environmental impact assessment predictions: the need for better communication and more transparency. *http://dx.doi.org/10.3152/147154606781765*
24(1), 45–56. https://doi.org/10.3152/147154606781765345

Tezuka, S., Takiguchi, H., Kazama, S., Sato, A., Kawagoe, S., & Sarukkalige, R. (2014). Estimation of the effects of climate change on flood-triggered economic losses in Japan. *International Journal of Disaster Risk Reduction*, *9*, 58–67. https://doi.org/10.1016/J.IJDRR.2014.03.004

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

Trigg, M. A., Birch, C. E., Neal, J. C., Bates, P. D., Smith, A., Sampson, C. C., Yamazaki, D., Hirabayashi, Y., Pappenberger, F., Dutra, E., Ward, P. J., Winsemius, H. C., Salamon, P., Dottori, F., Rudari, R., Kappes, M. S., Simpson, A. L., Hadzilacos, G., & Fewtrell, T. J. (2016). The credibility challenge for global fluvial flood risk analysis. *Environmental Research Letters*, *11*(9), 094014. https://doi.org/10.1088/1748-9326/11/9/094014

Undersander, D. (2016). Recovering Flooded Forages. https://fyi.extension.wisc.edu/forage/recovering-flooded-forages/

UNDRR. (2019). *Investing in disaster risk reduction: Scale and effect of investment in flood protection in Asia | UNDRR* (tech. rep.). United nations Office for Disaster Risk Reduction. https://www.undrr.org/publication/investing-disaster-risk-reduction-scale-and-effect-investment-flood-protection-asia

UN-Habitat. (2018). United Nations Human Settlements Programme 151 Macombe UN-HABITAT COUNTRY PROGRAMME MOZAMBIQUE.

UNOSAT. (2019). Cyclone Idai in Mozambique. https://disasterscharter.org/es/web/guest/disaster-types/-/asset_publisher/TC3LharmzylW/content/cyclone-in-mozambique-activation-598-;jsessionid=B50CF7678A20E1951546B977005E6008.jvm1?redirect=https%3A%2F%2Fdisasterscharter.org%2Fen%2Fweb%2Fguest%2Fdisaster-types%3Bjsessionid%3DB50CF7678A20E1951546B977005E6008.jvm1%3Fp_p_id%3D101_INSTANCE_TC3LharmzylW%26p_p_lifecycle%3D0%26p_p_state%3Dnormal%26p_p_mode%3Dview%26p_p_col_id%3Dcolumn-1%26p_p_col_count%3D1%26_101_INSTANCE_TC3LharmzylW_struts_action%3D%252Fasset_publisher%252Fview

van den Berg, K., Sluijs, L., Snuverink, M., & Wiertz, A. (2000). *Schadecurves industrie ten gevolge van overstroming - Rijkswaterstaat Rapportendatabank* (tech. rep.). Tebodin. https://puc.overheid.nl/rijkswaterstaat/doc/PUC_131055_31/

van Berchum, E. C., van Ledden, M., Timmermans, J. S., Kwakkel, J. H., & Jonkman, S. N. (2020). Rapid flood risk screening model for compound flood events in Beira, Mozambique. *Natural Hazards and Earth System Sciences*, *20*(10), 2633–2646. https://doi.org/10.5194/NHESS-20-2633-2020

van Ramshorst, J. (2019). Number of square meters of office space per employee - Flexas office space. https://www.flexas.com/blog/how-much-office-space-per-person

VvV. (2016). *Verbond van Verzekeraars - Verzekerd van Cijfers* (tech. rep.). Verbond van Verzekeraars. https://www.verzekeraars.nl/media/3544/verzekerd-van-cijfers-2016-nl.pdf

Wagenaar. (2020). *Capturing Complexity: Transferable flood impact models with Machine Learning* (Doctoral dissertation). Vrije Universiteit Amsterdam (VU). Amsterdam. https://scholar.google.nl/citations?view_op=view_citation&hl=nl&user=KLf71PoAAAAJ&citation_for_view=KLf71PoAAAAJ:_FxGoFyzp5QC

Wagenaar, De Bruijn, K. M., Bouwer, L. M., & De Moel, H. (2016). Uncertainty in flood damage estimates and its potential effect on investment decisions. *Natural Hazards and Earth System Sciences*, *16*(1), 1–14. https://doi.org/10.5194/NHESS-16-1-2016

Ward, P. J., Jongman, B., Weiland, F. S., Bouwman, A., Van Beek, R., Bierkens, M. F., Ligtvoet, W., & Winsemius, H. C. (2013). Assessing flood risk at the global scale: model setup, results, and sensitivity. *Environmental Research Letters*, *8*(4), 044019. https://doi.org/10.1088/1748-9326/8/4/044019

Ward, P. J., Winsemius, H. C., Kuzma, S., Bierkens, M. F., Bouwman, A., Moel, H. D., Loaiza, A. D., Eilander, D., Englhardt, J., Erkens, G., Gebremedhin, E. T., Iceland, C., Kooi, H., Ligtvoet, W., Muis, S., Scussolini, P., Sutanudjaja, E. H., Beek, R. V., Bemmel, B. V., … Luo, T. (2020). Aqueduct Floods Methodology. https://www.wri.org/research/aqueduct-floods-methodology

Waterschap Limburg. (2021). Dataset: Waterlevels Geul & Roer - Waterschap Limburg.

Wind, H. G., Nierop, T. M., Blois, C. J. d., & Kok, J. L. d. (1999). Analysis of flood damages from the 1993 and 1995 Meuse Floods. *Water Resources Research*, *35*(11), 3459–3465. https://doi.org/10.1029/1999WR900192

Wing, O. E. J., Pinter, N., Bates, P. D., & Kousky, C. (2020). New insights into US flood vulnerability revealed from flood insurance big data. *Nature Communications 2020 11:1*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-15264-2

Winter, B., Schneeberger, K., Huttenlau, M., & Stötter, J. (2017). Sources of uncertainty in a probabilistic flood risk model. *Natural Hazards 2017 91:2*, *91*(2), 431–446. https://doi.org/10.1007/S11069-017-3135-5

Winteraeken, H. J., & Spaan, W. P. (2010). A new approach to soil erosion and runoff in South Limburg-The Netherlands. *Land Degradation and Development*, *21*(4), 346–352. https://doi.org/10.1002/LDR.1009

World Bank. (2021). Economics for Disaster Prevention and Preparedness SUMMARY REPORT Investment in Disaster Risk Management in Europe Makes Economic Sense. www.worldbank.org

World Bank. (2022). Inflation, GDP deflator (annual %) - Mozambique | Data. https://data.worldbank.org/indicator/NY.GDP.DEFL.KD.ZG?end=2020&locations=MZ&start=1992&view=chart

Wünsch, A., Herrmann, U., Kreibich, H., & Thieken, A. H. (2009). The Role of Disaggregation of Asset Values in Flood Loss Estimation: A Comparison of Different Modeling Approaches at the Mulde River, Germany. *Environmental Management 2009 44:3*, *44*(3), 524–541. https://doi.org/10.1007/S00267-009-9335-3