Università
Ca'Foscari
Venezia

Scuola Dottorale di Ateneo
Graduate School

Dottorato di Ricerca in

Scienza e Gestione dei Cambiamenti Climatici

Ciclo XXXI, anno di discussione 2018/2019

Integrated framework for flood risk assessment in Italy

M-GGR/02; GEO/04; SECS-S/03; SPS/09

Dottorando
Mattia Amadio
825260

Supervisore
Prof. Carlo Barbante
Co-Supervisore
Prof. Jaroslav Mysiak

Coordinatore del Dottorato
Prof. Carlo Carraro

Integrated framework for flood risk assessment in Italy

Mattia Amadio

Venezia, 2018

A mia madre.

ACKNOWLEDGEMENTS

I wish to thank all the people that supported me by providing help, suggestions, food for thought, but also plain fun.

Jaro, my coach, and all my colleagues at CMCC, ANU and IVM for being always supportive and encouraging about my work.

Elco, for being a great example and for being always patient and motivating.

Takuya, for helping me as a talented programmer and for introducing me to amazing things.

My PhD mates Silvia, Remo and my chela Sepehr.

My co-authors Francesca, Annarita, Arthur, Alessio, for being a fantastic team to work with.

All my friends out of the office, because I really needed those frabjous gallops.

Anna for many reasons, but most importantly for keeping me on track.

Prof. Zanetto, if it wasn't for him I wouldn't have been writing all of this.

“If it keeps on raining, the levees’ gonna break...”

Led Zeppelin

EXECUTIVE SUMMARY

This thesis is related to climate-risk management and adaptation. Specifically, it focuses on the assessment of flood risk to preventively understand the costs of hazard scenarios in a changing climate in order to timely elaborate adequate adaptation policies. Italy is a flood-prone country that suffers the highest economic impact in all of the EU. Despite this, there is no established framework or method at national level to estimate flood risk, and existing country-wide assessments are very broad-brushed. This is critical considering that impacts from extreme meteorological events are expected to increase by 2050 in Europe.

The objective of the thesis is to develop an improved methodology for the assessment of flood risk combining the most updated and reliable data available by means of advanced spatial and statistical approaches. By reducing the uncertainty typical of simple customary methods, the improved flood risk model can be used to translate any change in flood hazard probability and magnitude into variation of Expected Annual Damage. Key for the improvement of damage modelling is the collection and analysis of empirical data from observed flood events; starting from a large dataset collected after the 2014 flood on the Secchia basin (Emilia-Romagna), the thesis shows different approaches, more and more sophisticated, to elaborate the available information into a prognostic tool that can be reliably employed for risk management. Starting from a general investigation on the performances of transferred univariable models (depth-damage curves) over different damage categories (**Ch.2**), my research proceeded with the development of an empirical-base univariable model by using a statistical calibration procedure (**Ch.3**). In my third study (**Ch.4**) I collected heterogeneous country-wide data (e.g. land use, soil sealing, population and building census, cadastral information, production value) and combined them by using a dasymetric approach in order to draw a detailed and homogeneous representation of exposure in terms of asset, population, GDP, and social vulnerability. Exposure to different hazard scenarios is then estimated in relation to JRC flood hazard modelling.

My last study (**Ch.5**) takes a far more advanced step in the identification of a tool that can be practically employed for country-wide risk assessment by validating an innovative multivariable, synthetic damage model for residential structures by means of machine learning.

Keywords: flood risk; climate hazard; economic modelling

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
BVM	Bivariable model
CGE	Computable General Equilibrium
FD	Flood Directive (2007/60/EC)
FLF-IT	Flood Loss Function for Italy
GVA	Gross Value Added (equal to GDP plus subsidies minus taxes)
GLMV	Gross Local Market Value, based on GVA
LC	Land Cover
LMA	Local Market Area (corresponds to SLL, Sistema Locale di Lavoro)
MAE	Mean Absolute Error
MBE	Mean Bias Error
MVM	Multivariable model
NUTS	Nomenclature of Territorial Units for Statistics
SDC	Stage-Damage Curve
RER	Regione Emilia-Romagna
RF	Random Forest
RMSE	Root Mean Square Error
UVM	Univariable model

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Background and motivation

Human civilisation depends on water. Since its very beginning, societies required to understand the hydrological cycle and consequently manage the water resources in order to provide a service (drinking water, irrigation, sanitation, energy) while controlling the flood hazard. This has never been an easy task, as evidenced by numerous myths and stories found in the history of mankind concerning devastating deluges and floods (Pindar 450BC; Pleins 2003; Wu et al. 2016). History also shows how we learnt to successfully control the waters through trials, errors and observations, ultimately claiming larger and larger portions of the valleys for the development of human activities. However, while population and wealth grew, the anthropic pressure on the floodplains intensified. Natural ecosystems regulate water flow and foster infiltration and natural water retention. When forests and wetlands are turned into impermeable urban areas, agricultural land or pastures, precipitation runoff is affected; water is discharged much faster through the catchments, resulting in higher peak flows that increase the probability of floods.

Today, in Europe, up to 90% of floodplains have been lost or are no longer able to provide an ecosystem service in terms of flood risk reduction (EEA 2016). In their place, river embankments and other man-made defences were put in action to control river flows and reduce (yet never eliminate) flood risk, leading to the accumulation of people and wealth within once flood-prone areas (Domeneghetti et al. 2015). At the same time, human-induced climate change is altering the water cycle, eventually increasing the variability of extreme meteorological events (e.g. intense rainfall) which are among the root causes of flood hazard and threatening to make the existing flood defences inadequate (Jonkman 2013; Alfieri et al. 2015a; Alfieri et al. 2016). Floods and inundations are today the most frequent and costliest natural disasters worldwide. According to the estimates European Environment Agency, between 1980 and 2017 floods caused losses in the EU exceeding 200 billion Euro, which is approximately 6 billion Euro of losses per year affecting 250,000 people (EEA 2010; EASAC 2018). Worse, other studies from the EEA have shown that future flood damage and losses may increase fivefold with increasing exposure and hazard (Alfieri et al. 2017). Under this perspective, a reliable

assessment of current and future flood risk is vital for designing effective disaster risk reduction and climate change adaptation policies and plans.

In the age of information, our capacity to record, measure and analyse natural phenomena has far outreached any previous effort. Remote sensing, Artificial Intelligence (AI) development, numerical modelling and other data-management and computation improvements are proving essential to tighten the knowledge gap. Still, there are many sources of uncertainty that affect the estimate of flood risk, such as the scarce number of complete empirical observation datasets that can be used to extrapolate models, the difficulties to obtain a detailed characterization of the exposed value, and the limited availability of hazard scenarios.

My thesis aims to contribute to the literature related to flood risk by investigating the performance of existing risk assessment methods while proposing further advancements based on updated data sources, improved spatial techniques and innovative modelling approaches. Starting from empirical observations collected from historical floods in Italy, my research leads to a better understanding of the complex "flood impact" process and provides insights about the economic modelling of flood damage. The improvements brought by the conducted research are relevant to the objectives of the EU Floods Directive (2007/60/EC) and to the targets of the UN's Sendai Framework for Disaster Risk Reduction 2015-2030, specifically to the reduction of economic disaster losses in relation to global Gross Domestic Product (GDP), and the reduction of disaster damage to critical infrastructure and the disruption of basic services.

1.2 Flood risk assessment

Flood impact is a complex phenomenon human, social, economic and environmental capitals. It can directly affect health and pose a threat to life, cause physical damage to households and productive activities, harm livestock and pollute land, erode social cohesion, but also interrupt critical infrastructures and supply chains (e.g. power outages), thus generating a series of indirect impacts that spread along the economic network. A multi-disciplinary approach is therefore required to assess the different dimensions of flood impact. In general terms, flood risk is expressed as a function of the hydrological hazard, the extent and value of the exposed assets, and their relative susceptibility to suffer damage from water inundation (Smith 1994; Meyer and Messner 2005; Scawthorn et al. 2006; Messner et al. 2007a; Jonkman et al. 2008a; Thieken et al. 2009; Merz et al. 2010a; de Moel and Aerts 2011a; Jongman et al. 2012a; Huizinga et al. 2017). Flood hazard is estimated by assessing the probability that a given area is flooded, as a result of an event of defined magnitude. This information is usually provided in

the form of flood scenarios referring to different frequency of occurrence and measured in terms of hazard variables (i.e. flooded extent, water depth, flow velocity and event duration). The elaboration of accurate flood scenarios involves advanced physical modelling and the thoroughly collection of detailed datasets related to land morphology, soil, hydrological features and climatic conditions. Flood risk assessment provides an estimate of the costs associated to each hazard scenario by combining them with the representation of exposed value by means of a vulnerability model. However, risk assessment approaches can vary depending on the temporal and spatial scales at which the assessment is carried on (de Moel et al. 2015), the category of impact that is investigated, the tools employed, the availability of data (de Moel and Aerts 2011a) and the type of ood (e.g. uvial, pluvial or coastal ood) (Apel et al. 2004; Thieken et al. 2006a; Messner et al. 2007b). Impacts can translate into intangible and tangible costs: intangible are those costs that are hardly measured in economic terms, such as harm to health and life, environmental degradation, or social disruption; differently, tangible costs can be quantified in economic terms. Impacts are also identified as direct or indirect. Direct tangible impacts refer to the economic costs sustained to repair or replace damaged physical asset. Indirect tangible losses, on the other hand, refer to the consequence of direct impacts affecting the production chain on a larger spatial and temporal scale (Meyer et al. 2013b; Hallegatte 2015). Indirect losses can also be referred to as flow losses (Koks et al. 2015a; Carrera et al. 2015).

1.3 Objectives and outline of the thesis

My research has contributed to an integrated framework for assessing the potential change in economic losses caused by floods accounting for both climate variability and socio-economic development. The risk assessment framework proposed in this thesis is shown in figure 1.1. The framework is built over the three customary macro-components of risk: hazard, exposure and vulnerability.

The hazard component involves the modelling of flood scenarios that are consistent with observations (when assessing past events) and with climate projections (when assessing future events). The exposure component combines a variety of heterogeneous datasets into a standardised spatial representation of socio-economic value, including population count, classification of buildings, reconstruction costs, production value per economic sector, and other variables. These two components are the inputs for the vulnerability model, which first produces an estimate of the direct economic impacts by means of spatial analysis; then, a sub-

national macro-economic model simulates the effect of the economic shock over regional production, including spill-over effects to other Regions.

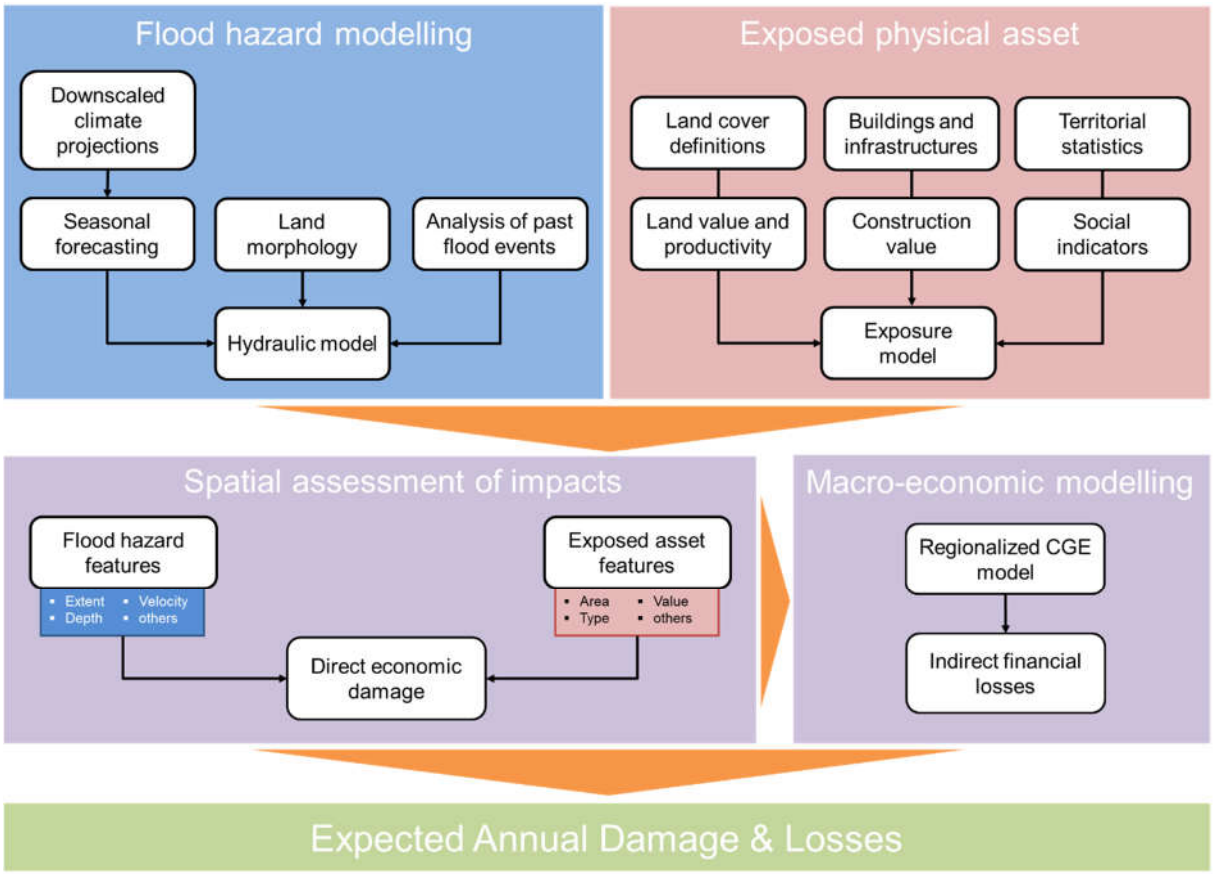


Fig. 1.1. Proposed flood risk assessment framework built over the main components of risk assessment: hazard, exposure and vulnerability.

The sum of direct economic damage to physical asset and indirect losses to production provides a measure of the expected total losses in relation to the probability of the accounted scenario; it can be multiplied by the event probability, to compute the annual average damage and losses. This is then summed for all scenarios to provide a measure of the Expected Annual Damage and Losses (EAD&L), i.e. the integral of the area under the damage probability curve. By comparing EAD&L measured from a set of climate scenarios, it is possible to produce an estimate of the costs associated with global warming.

The thesis takes the case of Italy, the EU Country with the largest population living in flood-prone areas (about 7 million people) and the largest amount of annual uninsured economic losses: around 4 billion Euro of public money were spent over a 10 years period to compensate the damage inflicted by major extreme hydrologic events countries (Associazione Nazionale fra le Imprese Assicuratrici 2015; Alfieri et al. 2016; EEA 2016; Paprotny et al. 2018).

According to ISPRA, from 2009 until 2012, the recovery funding amounted to about 1 billion Euro per year, and this is only a fraction of the real damage, estimated to be around EUR 2,2 billion (Zampetti et al. 2012). In this context, and particularly compelled by the EU Flood Directive (2007/60/EC), sound and evidence-based flood risk assessments must underpin the development and implementation of cost-effective flood risk reduction strategies and plans.

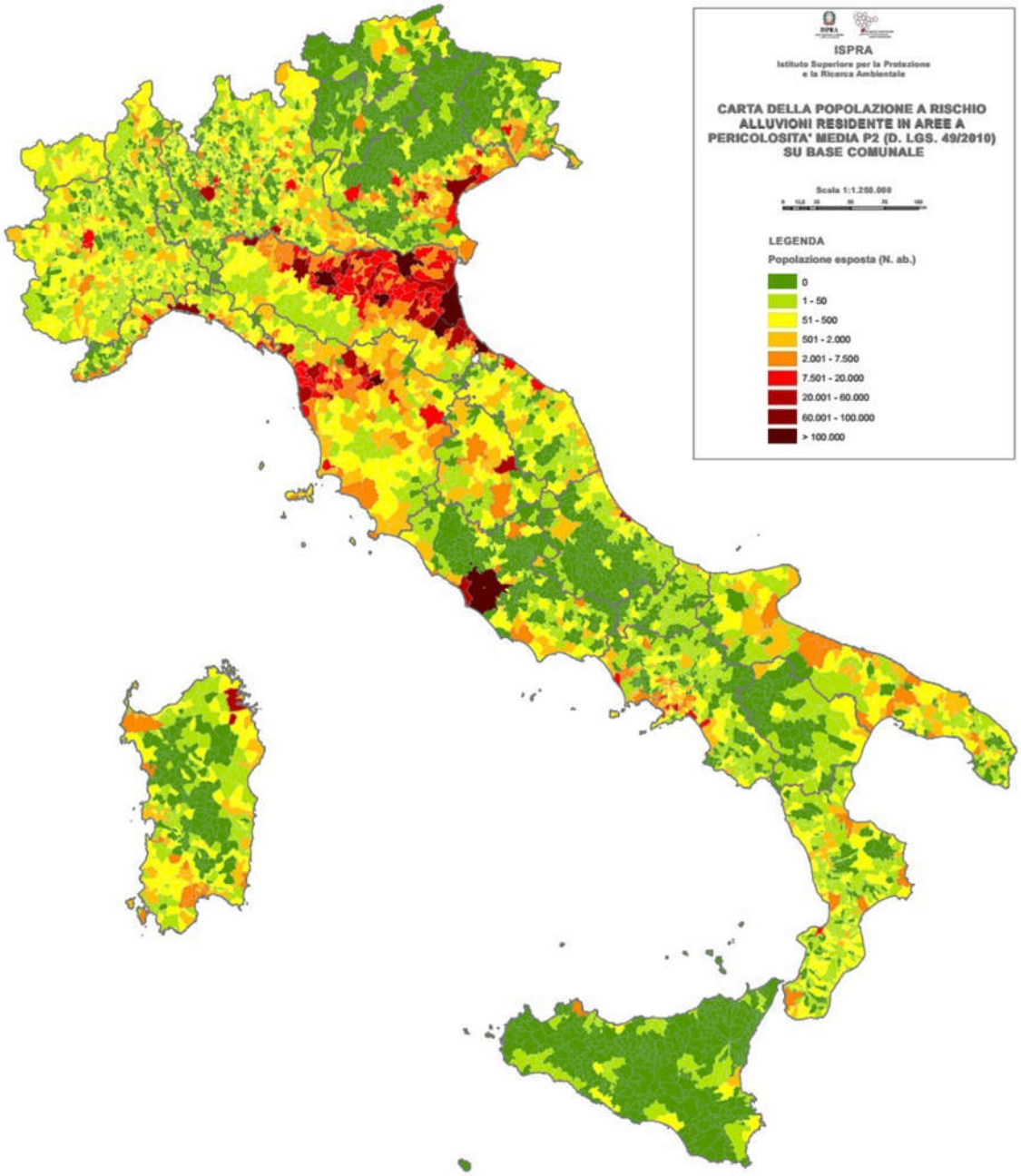


Fig.1.2. Population exposed to flood risk in Italy, map produced by ISPRA (2017) in compliance to the Flood Directive (2007/60/EC). The estimate is based on the mosaicking of flood-prone areas with a return time between 100 and 200 years (P2, medium hazard) obtained from local District Basin Authorities, and on the resident population from the 15th ISTAT Census (2011).

However, efforts to produce a country-wide flood risk assessment in Italy have been limited to the initial requirements of the Flood Directive; i.e., the current Italy-wide flood risk assessment is based on a generalized definition of hazard perimeters assembled from the information produced by local administrations (Fig. 1.2) (ISPRA 2014). In addition, the existing assessment does not include a proper damage model, which is required to connect the hydrologic and social dimensions and to translate hazard mapping into risk estimates. This issue is linked to the lack of a systematic, standardized collection of ex-post data and their methodical analysis, despite the frequency of flood events (Molinari et al. 2013; Molinari et al. 2014b).

The thesis document is organised in five chapters. The chapters from 2 to 5 represent individual research articles that have been previously accepted and published by international peer-reviewed journals, except for the last one which has been just recently submitted to the journal. Each essay represents a step forward in my understanding of the flood damage phenomenon and in the elaboration of practical instruments for flood risk management. Specifically, the four manuscripts included in this thesis focus on the modelling of direct economic costs of flood events by addressing the spatial representation of socio-economic value and the development of proper empirically validated tools to estimate the risk associated with hazard scenarios. More research has been conducted to cover all components of the proposed framework, planning to include additional chapters on hazard modelling, indirect losses and implications of risk scenarios; however, these essays could not be completed by the time of this dissertation (see par. *1.5 - Future research*).

In my first publication (**Chapter 2**), I conducted a general evaluation of flood impact over different damage categories. The transferability potential of an existing univariable model (DamageScanner) is tested using hazard and compensation records from the aforementioned Secchia case study. Additionally, the study explores both physical asset and foregone production losses, the latter measured amidst the spatially distributed gross added value (GVA).

The second article (**Chapter 3**) was developed during my visiting period at the Austrlian National University. Here I produce a new Flood Loss Function for Italian residential structures (FLF-IT) based on the empirical observations included in the Secchia dataset. The performance of the statistical model is validated for the prediction of loss ratios and absolute damage values. A three-fold cross-validation procedure is carried out over the empirical sample to measure the range of uncertainty from the actual damage data. The validation procedure shows that FLF-IT performs better in the estimation of direct damage compared to transferred third-party models.

The third article (**Chapter 4**) is dedicated to improve the spatial representation of socio-economic exposure and to measure social vulnerability. A spatially-weighted dasymetric approach based on multiple ancillary data is applied to downscale important socio-economic variables resulting in a grid dataset for Italy. This set contains multi-layered georeferenced information about physical exposure, population, GDP and social vulnerability. We test the performance of this approach compared to other spatial interpolation methods. Then, we combine the grid dataset with flood hazard scenarios to exemplify its application for the purpose of risk assessment.

The final article of the thesis (**Chapter 5**) compares the performances of expert-based and empirical (both uni- and multivariable) damage models for estimating the potential economic costs of flood events to residential buildings. Multivariable models account for a larger number of factors related to hazard and exposure and are potentially more robust and flexible when extensive input information is available. In this study the observation dataset comprises three recent major flood events in Northern Italy (Adda 2002, Bacchiglione 2010 and Secchia 2014), including hazard features (depth, velocity and duration), buildings characteristics (size, type, quality, economic value) as well as reported losses. First, the relative importance of each explanatory variable on the damage output is assessed by means of a tree-based regression approach. Then, the performance of four literature flood damage models of different nature and complexity are compared with the performance of univariable, bivariable and multivariable models empirically developed for Italy and tested at the micro scale based upon observed records. The uni- and bivariable models are produced testing linear, logarithmic and square root regression while multivariable models are based on two machine learning techniques, namely Random Forest and Artificial Neural Networks. Results identify the best fitting models and provide important insights about the choice of the approach for operational disaster risk management.

1.4 Main findings and lesson learnt

The idea behind the flood risk framework that gives the title to the thesis was to screen and collect information and data about hazard, exposure and vulnerability components, having them standardised using a solid spatial approach and then combined using a proper impact model for assessing flood risk in Italy at any scale. But as I found out soon, a single approach cannot easily provide an answer to a whole range of challenges a various spatial and temporal scales, typologies of floods and categories of impacts. After narrowing down the analysis to

measurable economic damage, thus excluding non-economic capitals (i.e. human, environmental, social), the task of producing a sound tool for risk assessment remains a difficult one due to the complexity of the phenomena and the scarcity of empirical data. As I found out from the study presented in chapter two, existing assessment tools are afflicted by uncertainties, depending for the most part on the methods employed to estimate vulnerability. Required observations about impacts are usually trackable for some risk categories, which represent a large share of the total. Specifically, in my research the direct losses suffered by residential private properties often constitute the bulk of the damage, followed by impacts on commercial and industrial activities and infrastructures. Economic impacts on agricultural production are strongly depending on seasonality and hardly verifiable, but in general they are several orders of magnitude smaller. Indirect losses are likely as important as direct ones; but in reality, they are unlikely measurable as they spread in time (and time) long after the aftermath of a flood disaster. For all these reasons, I chose to dedicate my efforts towards direct residential damage, which also offers the largest availability of observations to train impact models.

A critical stage of the research has been dedicated to build a solid dataset of empirical flood damage records, which currently represents one of the most extensive and detailed for the country. On this basis, it was possible to apply several statistical approaches to investigate the flood impact process, from rudimental to advanced, and to identify the best performing damage models. In chapter five, advanced assisted machine learning is proven to be a powerful tool to investigate flood loss data for this purpose. My results corroborate previous findings from literature about the relationship between hazard variables and estimated losses; first, water depth is identified to be the most important predictor of damage. Additional hazard variables only marginally improve the model estimates, at the cost of a major complexity and data requirements. Secondly, performance testing showed that expert-based models are better candidates for transferability in space and time compared to empirically-obtained models. Alongside, the work conducted in chapter four contributes to the practical applicability of a scalable national risk model by providing a detailed, standardised representation of exposed economic capitals in the form of a spatial georeferenced dataset which includes all the variable required by the proposed risk assessment framework (plus indicators of social vulnerability). The grid allows to perform a risk analysis at high resolution (250 meters) and obtain an estimate of losses that can be represented at any administrative level.

In addition to the scientific contribution on the field of flood risk management, the research comprised in this thesis contributed to several European research and innovation projects (e.g. ENHANCE, *Enhancing risk management partnerships for catastrophic natural hazards in Europe*, and CLARA, *Climate forecast enabled knowledge services*) and to the *Italian National Climate Change Adaptation Plan* (PNACC).

1.5 Limitations and further research

The difficulty to collect exhaustive empirical information about flood losses makes the few available datasets extremely precious for the development of flood risk modelling. This is why the flood data acquired from the Secchia flood event have been recursively analysed during my research. Indeed, more independent flood damage datasets could benefit the value of the analysis. In particular, the inclusion of damage records from other types of flood events (e.g. flash floods, coastal floods) would help to understand if a general damage model is adequate to interpret slightly different damage processes.

Moreover, the research does not provide conclusions about the measure of impacts over the assets and production of business activities. Additional efforts are required to this end both on the side of data collection and on the elaboration of proper modelling approaches.

Ultimately, the body of this thesis does not cover all the components of the risk framework shown in figure 1.1, focusing mostly on the direct impacts. However, more studies have been under production in relation to the other components. My ongoing research sets out to address:

- **Evaluation of hazard modelling approaches at different scales**
  This essay will include a comparison on the performance of different flood hazard modelling approaches at different spatial scales. Expected outputs will include a new national map of flood prone areas and associated hazard probabilities.

- **Measuring the indirect impacts of extreme events**
  In this study a macro-economic approach is used to assess the indirect costs of disasters caused by extreme meteorological events. A Computable General Equilibrium model is applied to assess the indirect financial losses on regional economies due to economic shocks from natural disasters. The proposed case study is the disastrous event of October 2018 in Veneto.

- **Disaster insurance as economic policy instrument to cope with climate change**

  The paper will include an assessment of Expected Annual Damage scenarios in relation to an ensemble of downscaled climate projections for Italy. The study builds upon the knowledge and the methodologies developed in previous publications and it should provide important conclusions on the expected upcoming losses from flood disasters.

1.6 References

Alfieri L, Bisselink B, Dottori F, et al (2017) Earth' s Future Global projections of river flood risk in a warmer world. doi: 10.1002/eft2.183

Alfieri L, Burek P, Feyen L, Forzieri G (2015) Global warming increases the frequency of river floods in Europe. Hydrol Earth Syst Sci 19:2247–2260. doi: 10.5194/hess-19-2247-2015

Alfieri L, Feyen L, Salamon P, et al (2016) Modelling the socio-economic impact of river floods in Europe. Nat Hazards Earth Syst Sci 16:1401–1411. doi: 10.5194/nhess-16-1401-2016

Apel H, Thieken a. H, Merz B, Blöschl G (2004) Flood risk assessment and associated uncertainty. Nat Hazards Earth Syst Sci 4:295–308. doi: 10.5194/nhess-4-295-2004

Associazione Nazionale fra le Imprese Assicuratrici (2015) Le alluvioni e la protezione delle abitazioni. 22.

Carrera L, Standardi G, Bosello F, Mysiak J (2015) Assessing direct and indirect economic impacts of a flood event through the integration of spatial and computable general equilibrium modelling. Environ Model Softw 63:109–122. doi: 10.1016/j.envsoft.2014.09.016

de Moel H, Aerts JCJH (2011) Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates. Nat Hazards 58:407–425. doi: 10.1007/s11069-010-9675-6

de Moel H, Jongman B, Kreibich H, et al (2015) Flood risk assessments at different spatial scales. Mitig Adapt Strateg Glob Chang 20:865–890. doi: 10.1007/s11027-015-9654-z

Domeneghetti A, Carisi F, Castellarin A, Brath A (2015) Evolution of flood risk over large areas: Quantitative assessment for the Po river. J Hydrol. doi: 10.1016/j.jhydrol.2015.05.043

EASAC (2018) Extreme weather events in Europe. Preparing for climate change adaptation: an update on EASAC's 2013 study.

EEA (2016) Flood risks and environmental vulnerability - Exploring the synergies between floodplain restoration, water policies and thematic policies.

European Environment Agency (2010) Mapping the impacts of recent natural disasters and technological accidents in Europe - An overview of the last decade.

Hallegatte S (2015) The Indirect Cost of Natural Disasters and an Economic Definition of Macroeconomic Resilience. The World Bank

Huizinga J, De Moel H, Szewczyk W (2017) Methodology and the database with guidelines Global flood depth-damage functions.

ISPRA (2014) Mappe di pericolosità idraulica e popolazione esposta a rischio alluvioni in italia.

Jongman B, Kreibich H, Apel H, et al (2012) Comparative flood damage model assessment: Towards a European approach. Nat Hazards Earth Syst Sci 12:3733–3752. doi: 10.5194/nhess-12-3733-2012

Jonkman SN (2013) Advanced flood risk analysis required. Nat Clim Chang 3:1004–1004. doi: 10.1038/nclimate2031

Jonkman SN, Bočkarjova M, Kok M, Bernardini P (2008) Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. Ecol Econ 66:77–90. doi: 10.1016/j.ecolecon.2007.12.022

Koks EE, Carrera L, Jonkeren O, et al (2015) Regional disaster impact analysis: comparing Input-Output and Computable General Equilibrium models. Nat Hazards Earth Syst Sci Discuss 3:7053–7088. doi: 10.5194/nhessd-3-7053-2015

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

Messner F, Penning-rowsell E, Green C, et al (2007a) Evaluating flood damages: guidance and recommendations on principles and methods. Flood Risk Manag Hazards, Vulnerability Mitig Meas 189.

Messner F, Penning-rowsell E, Green C, et al (2007b) Evaluating flood damages : guidance and recommendations on principles and methods principles and methods. Flood Risk Manag Hazards, Vulnerability Mitig Meas 189.

Meyer V, Becker N, Markantonis V, et al (2013) Review article: Assessing the costs of natural hazards-state of the art and knowledge gaps. Nat Hazards Earth Syst Sci 13:1351–1373. doi: 10.5194/nhess-13-1351-2013

Meyer V, Messner F (2005) National flood damage evaluation methods : A review of applied methods in England, the Netherlands, the Czech republic and Germany. 49.

Molinari D, Aronica GT, Ballio F, et al (2013) Le curve di danno quale strumento a supporto della direttiva alluvioni: criticità dei dati italiani. 10–15.

Molinari D, Menoni S, Aronica GT, et al (2014) Ex post damage assessment: an Italian experience. Nat Hazards Earth Syst Sci 14:901–916. doi: 10.5194/nhess-14-901-2014

Paprotny D, Sebastian A, Morales-Nápoles O, Jonkman SN (2018) Trends in flood losses in Europe over the past 150 years. Nat Commun 9:1985. doi: 10.1038/s41467-018-04253-1

Pindar (450BC) Deucalion and the big deluge. In: Olympian Ode 9.

Pleins JD (2003) When the great abyss opened : classic and contemporary readings of Noah's flood. Oxford University Press

Scawthorn C, Flores P, Blais N, et al (2006) HAZUS-MH Flood Loss Estimation Methodology. II. Damage and Loss Assessment. Nat Hazards Rev 7:72–81. doi: 10.1061/(ASCE)1527-6988(2006)7:2(72)

Smith D (1994) Flood damage estimation. A review of urban stage-damage curves and loss function. Water SA 20:231–238.

Thieken a. H, Ackermann V, Elmer F, et al (2009) Methods for the evaluation of direct and indirect flood losses. RIMAX Contrib 4th Int Symp Flood Def 1–10.

Thieken A, Merz B, Kreibich H, Apel H (2006) Methods for flood risk assessment: concepts and challenges. In: International Workshop on Flash Floods in Urban Areas. pp 1–12

Wu Q, Zhao Z, Liu L, et al (2016) Outburst flood at 1920 BCE supports historicity of China's Great Flood and the Xia dynasty. Science 353:579–82. doi: 10.1126/science.aaf0842

Zampetti G, Ottaviani F, Minutolo A (2012) I costi del rischio idrogeologico. Dossier Legambiente, Roma

2 IMPROVING FLOOD DAMAGE ASSESSMENT MODELS IN ITALY

2.1 Introduction

The EU Floods Directive (FD, 2007/60/EC) manifested a shift of emphasis away from structural defence approach to a more holistic risk management, with structural and non-structural measures having the same importance. The FD compels the identification of areas exposed to flood hazard and risk, and the adoption of measures to moderate flood impacts. A sound, evidence-based risk assessment should underpin public disaster risk reduction and territorial development policies. Stage-damage curves (SDCs) are a customary tool used for assessing risk arising from the physical disruption of physical tangible assets (Genovese 2006; Messner et al. 2007c; Thieken et al. 2009; Jongman et al. 2012c), typically as a function of flood characteristics (primary water depth, in some cases speed and persistence) over different land cover (LC) categories (Messner et al. 2007c; Merz et al. 2010b). SDCs are either empirically determined from observed damage events or inferred from bibliographic sources. Most flood risk assessment studies employ empirical SDC models that are developed elsewhere and neither tested nor calibrated for the specific study area (Sargent 2013). The lack of practical corroboration compromises the reliability of the model results. In addition, the SDC models are afflicted by substantial uncertainties stemming from the variability of assets value and vulnerability (Messner et al. 2007c; Merz et al. 2010b; de Moel and Aerts 2011b). To some extent, these uncertainties can be reduced if the damage models are designed to reproduce the economic conditions of households and businesses (Luino et al. 2009a; de Moel and Aerts 2011b). Different SDC models have been reported in literature, but most of them have been developed for site-specific application and are rarely tested for transferability. SDC based on empirical material from Italy are rare (Molinari et al. 2012a; Scorzini and Frank 2015a). This is despite the common practice of state compensation for household (private) losses, for which certified damage reports are collected. In addition, SDC models often assume that the potential damage is constant throughout the year. This does not hold for agricultural land, where the crop value varies depending from the crop maturity. Furthermore, SDC models address physical assets damage and hence are not able to determine output losses in terms of foregone production

that arises from impairment of economic activities until after the production process are fully recovered. Spatially distributed economic and social variables such as population density and GDP can help to estimate impact on the economic flow from natural hazards. Different methodologies are employed for this purpose, such as econometric models (Noy and Nualsri 2007; Strobl 2010; Cavallo et al. 2012), Input-Output (IO) models (Jonkman et al. 2008b; Hallegatte 2008; Henriet et al. 2012; Koks et al. 2014a) and Computable General Equilibrium (CGE) models (Jonkhoff 2009; Bosello et al. 2012; Rose and Wei 2013; Carrera et al. 2015). These are useful to estimate the impact of a hazard on the economy up to the regional level but require disaggregated data that is rarely available at lower scales. The availability of sound flood risk models appropriate for the Italian economic and social circumstances is essential for well-designed and informed flood risk management policies. In this paper, we explore ways to improve the damage and loss assessments for the sake of a better risk assessment and management. Methods such as those explored in this paper have been tested elsewhere at the national (Winsemius et al. 2013) and international scale (Ward et al. 2013).

The paper is structured as follows. First, we test the applicability and transferability of up-to-date SDCs against household damage declarations in the aftermath of the 2014 Modena flood in the Emilia-Romagna Region. Successively, we describe a detailed crop-specific model for agricultural losses, better suitable for compensation claims (Forster et al. 2008; Tapia-Silva et al. 2011; Twining 2014). Ultimately, we explore the use of Gross Value Added (GVA) as an indicator of exposure for production losses (Peduzzi et al. 2009).

2.2 Data and methods

Most commonly, *flood risk* **R** is determined as a function of *hazard probability* (**H**), *exposure* (**E**) and *vulnerability* (**V**): $\boldsymbol{R = H \times E \times V}$ (Crichton 1999; Kron 2005; Messner et al. 2007c; Barredo and Engelen 2010). Hazard is expressed as observed or modelled probability $\boldsymbol{p}$ of river discharges exceeding the holding capacity of river embankments. Exposure represents the depreciated or replacement value of the tangible physical assets in hazard-prone areas. Vulnerability is the susceptibility to damage under different levels of flood submersion. The structural damage to physical tangible assets is also termed direct impact or damage on stock (Merz et al. 2010b; Meyer et al. 2013a). When productive capital is damaged, the impacts can also be valued in terms of production *losses* or foregone flows of production. Sometimes, flow losses are equated to indirect impacts or damage. This is misleading because production losses are an alternative manifestation of material damage to productive capital assets, one that

contemplate the value of *output* (good and services) that would have been produced during the time of suspended production, rather than the depreciated value of the damaged asset. Flow losses are able to capture situations in which production is disrupted as a result of dearth of critical input with no material damage to productive capital, for example in case of lifeline disruption (Przyluski and Hallegatte 2011). Here we avoid this ambiguity by referring to damage in terms of partial or total physical asset destruction and losses in terms of foregone production flows. This is consistent with economic theory according to which the value of a stock is the discounted flow of net future returns from its operation (Rose, 2004). We estimate the flood damage both as *asset damage* using the SDC model and as *production losses* in terms of affected annual GVA (Figure 2.1). Agricultural losses are estimated using a complementary model that accounts for crop production cost and the value of yields (Thieken et al. 2009).

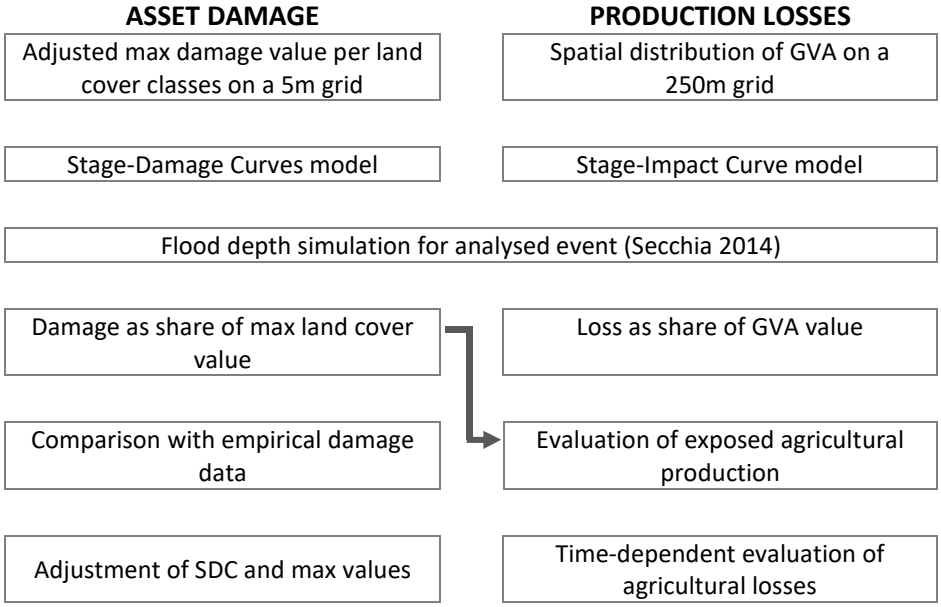


Figure 2.1. Flood damage assessment methodological approach.

2.2.1 SDC models for asset damage

Among the SDC models found in literature, two have been found performing reasonably well compared to reported empirical damage in Italy (Scorzini 2015): Damage Scanner (DS or **SDC-1**) (Klijn et al. 2007a) and JRC (**SDC-2**) (Huizinga 2007a). SDC-1 has been recently updated (de Moel et al. 2013b; de Moel et al. 2014; Koks et al. 2014a) with additional sub-classes for residential, rural and industrial damage (Tebodin 2000). This set estimates impact over buildings surface separately from other sealed areas such roads. Differently, SDC-2 aggregates the impact for mixed land cover classes: the maximum value for each of these main classes is built over the weighted sum of buildings and area, including both the structure and content.

This approach is adapted to work in conjunction with low resolution land cover maps such CLC. Depth resolution also varies among the two sets: SDC-1 takes steps of 0.1 m, while the other set has 0.5 m steps. All these curves are based on expert judgment and none of them have been validated on empirical damage data. SDC-1 and SDC-2 are the best available options up to date for transferability testing. The damage is estimated for sealed areas and agricultural land, while roads and natural areas are neglected. For residential damage we consider both the damage to physical structure of buildings and to their associated content. The model accounts also for damage to passenger vehicles based on average price from statistical registers (ACI, 2014).

With the proposed methodology, we aim to simulate the impact of the flood event which hit the province of Modena province during 2014. On January 19th, a 80 meters wide levee breach occurred on the Secchia river, spilling 200 cubic meters per second in the surrounding countryside, covering nearly 6.5 thousand ha of cultivated land (figure 2.2). Seven municipalities were affected, with the small towns of Bastiglia and Bomporto suffering the largest share of losses being flooded for more than 48 hours. The total volume of water pumped out of the inundated area was estimated to exceed 20 million cubic meters (Fotia 2014). For the purpose of this paper we used the hydrological simulation of the event produced by D'Alpaos et al. (2014). The extent of the simulated flood is nearly five thousand hectares, with an average depth of 1 meter.

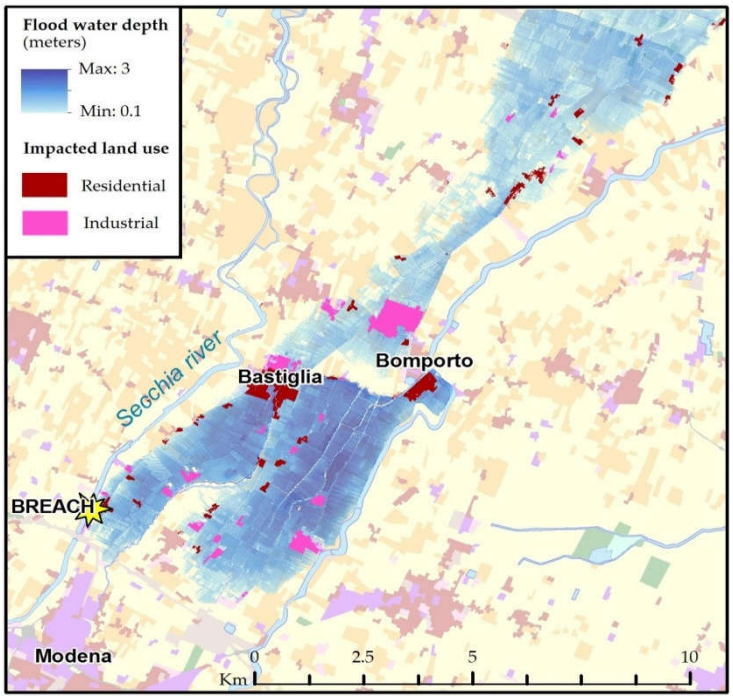


Figure 2.2. Simulated max flood depth ensuing from the Secchia levee breach in January 2014 near Modena. Impacted areas are highlighted for residential and industrial land cover.

The damage estimated through SDC is compared to households-declared damage made available by local authorities, while damage to business activity is not yet available. The damage reports distinguish between *structure*, *mobile goods* (furniture and common domestic appliances) and *registered vehicles* (private cars and motorcycles).

2.2.2 Agricultural losses

Expected losses in sparsely populated rural areas are often substantially lower than those in residential areas, since the density of exposed value is lower. For this reason, agricultural damage is often neglected or accounted for by using simple approaches with coarse estimates. Yet a thorough loss assessment is necessary in areas where agricultural production is the predominant activity (Messner et al. 2007c) as it guides compensation where compelled by liability or granted in form of state aid (Forster et al. 2008; Tapia-Silva et al. 2011; Twining 2014). Standard SDC models are suboptimal for this purpose as they hardly account for the variety in cultivated crops values, yields, and the progressive distribution of production costs. The SDC typically assumes a constant economic value throughout the year, which is not consistent with the fact that the damage depends from when a flood occurs (Ward et al. 2011). In our enhanced model, we determine the representative full crop damage per hectare D_{MAX} as a weighted average of all major crops' values in the analysed area (equation 2.1) at any time during the growing session (equation 2.2 and 2.3).

$$D_{MAX} = \sum_{i=1}^{n} P_i \times Y_i \times \frac{UAA_i}{UAA} \qquad \text{[Equation 2.1]}$$

Where:

- ***i*** = crop index
- ***P*** = producer prices (per tonnes)
- ***Y*** = yield (tonnes/hectare)
- ***UAA*** = Utilised Agricultural Area[1]

D_{MAX} at any time *t* during the growing season can be estimated either by taking into account the end-of-the-season yield and producer price of crop *i*, minus production costs not exerted until the end of the production cycle (equation 2.3); or as a sum of all production costs exerted from the beginning of the growing season up to the damaging event, plus the land rent (equation 4). The best estimate of the crop value at the harvesting time is Gross Saleable Product[2] (GSP).

$$D_{MAX}^{t} = \sum_{i=1}^{n} \left(\left[GSP_i \quad \sum_{t}^{End} DC_i \right] \times \frac{UAA_i}{UAA} \right) \qquad \text{[Equation 2.2]}$$

[1] UUA comprises total area of arable land, permanent crops and meadows.
[2] The average gross income from the sale of the yield expressed in €/ha, not inclusive of direct costs.

Where:

- DC are the direct production costs[3]
- t a defined moment of the production cycle ($0 < t <_{End}$).

$$D^{t}_{MAX} = \sum_{i=1}^{n}\left(\left[TNI_i + \sum_{0}^{t} DC_i\right] \times \frac{UAA_i}{UAA}\right) \quad \text{[Equation 2.3]}$$

Where:

- TNI = Total Net Income calculated on the previous years' average
- DC = sum of crop specific production costs exerted until the damaging event

The average yield, production cost and net income per hectare of arable and permanent crops are determined for different cultivation patterns in the Emilia-Romagna administrative region (RER) based on empirical observations (Altamura et al. 2013). The direct cost is calculated as a function of average cost of technical means (raw materials, machinery) and labour per hectare. Costs are distributed among the production year on the basis of each crops lifecycle as exemplified in figure 2.3.

	Growing season months												***Period***	***Costs***
	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	1	sow	50%
Wheat	1	1	2	2	2	2	2	2	3	4	4	4	2	growth	30%
Maize	3	4	4	4	4	4	1	1	1	2	2	3	3	harvest	20%
Alfalfa	2	2	2	2	2	1	1	2	2	2	2	2	4	renewal	0%

Figure 2.3. Allocation of production cost and the typical growing season for the most common cereal crops in the study area.

Field analysis conducted after the event (Setti 2014) highlighted that the flood occurred at a time when many field crops had not yet been planted. Wheat and alfalfa were the most commonly exposed crops, but the only physical harm reported was some occasional yellowing among crop fields. Vineyards and other permanent crops were in vegetative rest and apparently did not suffer any damage. In the end, the report on regional agricultural production for the year 2014 (OAA-RER 2014) does not revealed any substantial yield reduction. On the contrary, the average yield per hectare in 2014 were slightly higher than 2013.

2.2.3 Gross Value Added model for production losses

To estimate the production losses we use gridded *Gross Value Added* (GVA) (Peduzzi et al. 2009; Green et al. 2011) based on the statistical disaggregation of GVA at the *local market*

[3] Sum of the costs for technical means and labour, excluding subsidies.

areas[4] (in Italian S*istemi Locali di Lavoro* SLL) for three macro-economic branches: agriculture, industry and services (ISTAT 2013). We assume that within the SLL the GVA is uniformly spread, but only over the land cover classes ascribed to each specific branch of economic activities. In the case of agriculture and industry, the GVA is attributed to respectively the UAA and total industrial area distinguishable in the land use/cover data sets. The GVA generated by services is distributed proportionally to the population density. The assumption behind this is that since services are multiple and dispersed, they are proportional to the number of residents served. A population density grid is produced based on the 2011 census tracks (ISTAT 2011a) using a cell resolution of 250 meters. The expected losses as a share of GVA per cell are then calculated using a step function (Equation 2.4, Figure 2.4) (Carrera et al. 2015), inspired by literature on flood damage functions (de Moel and Aerts 2011b; De Moel et al. 2012; Jongman et al. 2012c; Saint-Geours et al. 2014). The curve assumes that the higher the water level, the more persistent is the productivity loss. This assumption is based on three principles: a) higher water-depths cause larger productive asset damage; b) larger asset damage typically requires longer recovery periods; and c) flood water retreat is a function of flood depth. The relation between water depth and persistence of the impact is likely afflicted by uncertainty, however we assume the curve suited for our purposes.

$$Impact\ on\ GVA_{S,L} = \sum_{k=1}^{n} FC_{S,L_k} \times c_k \qquad \text{[Equation 2.4]}$$

Where:

FC = flooded cell k

c = damage factor applied to each FC_k based on its water depth

N = number of cells belonging to sector S for each system L

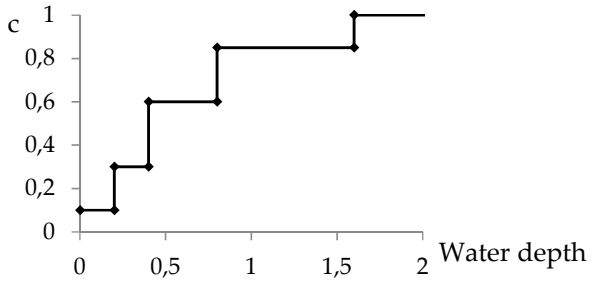


Figure 2.4. Stage-impact curve for GVA losses.

[4] Local market areas (SLL) have been devised by the Italian Statistical Bureau as continuous territorial areas in which most of the daily work activity of resident people takes place. Typically a SLL is smaller than a NUTS3 unit and larger than a municipality.

2.3 Results

2.3.1 Asset losses

The damage assessment carried out on the Modena flood using the two selected SDC models (SDC-1 and SDC-2) yield values that differ by 170 million, corresponding to one third of the SDC-2 estimate (figure 2.5, left). Besides, there is a sizeable divergence in distribution of the estimated damage across the land cover categories. The SDC-1 yields a damage that is more than two times higher than SDC-2 output for the industrial land cover category. On the contrary, SDC-1 estimated damage is lower than SDC-2 by a factor 0.7 for the residential land cover category and only one fifth for rural category.

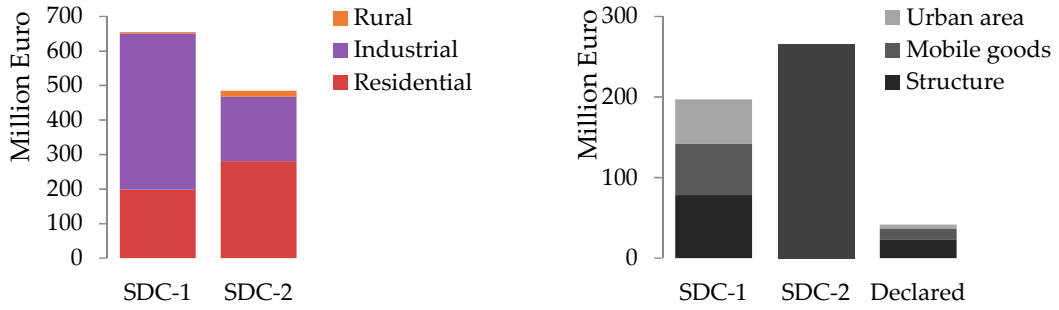


Figure 2.5. (left) damage estimates from SDC-1 and SDC-2 models for the 2014 flood event on aggregated land covers; (right) comparison of models output for residential areas against registered compensation requests from households.

Overall, SDC-1 overestimates declared damage in residential areas by a factor 4.5, but for the urban spaces outside buildings this difference peaks factor 9.2. SDC-2 results are even larger, 13 times greater than those observed. The damage shares between structure, mobile goods and private vehicles simulated by SDC-1 resemble[5] the ratios of declared damage (figure 2.5, right). For the calibration exercise, we have chosen SDC-1 over SDC-2 because it is able to disaggregate structural and content-wise damage in isolated dwellings and built-up areas. Both estimated and declared damage are geocoded and aggregated into a 250 m grid.

[5] Simulated damage: 57/33/10%. Declared damage: 60/35/5%.

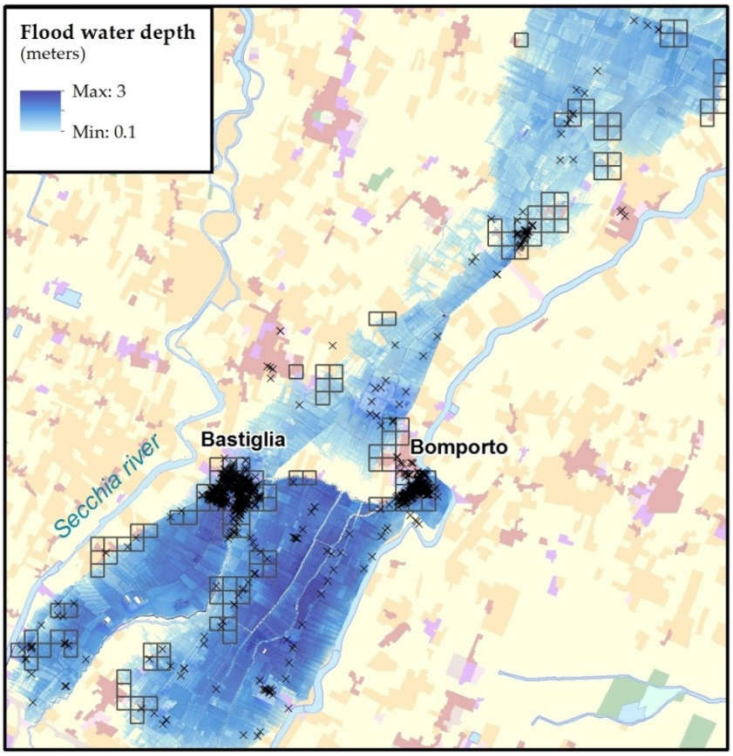


Figure 2.6. Location matching for residential land cover between empirical (black Xs) and simulated damage (aggregated to 250 meters cells).

The calibration is carried out only on matching cells using regression analysis under the hypothesis of linear relationship. There are 61 (out of 157) matching cells between simulated and empirical damage, which is less than 40% in terms of affected area but the matching cells account for 83% of simulated and 75% of the declared damage. As shown in Figure 2.6, this mismatch is caused mainly by uncertainty in the land cover data for sparsely developed areas and in the extent of the flood boundaries, but the core damage areas of Bastiglia and Bomporto match well between recorded and simulated damage. For each land cover category, the maximum damage value is individually adjusted using the *B* (slope) coefficients as scaling factor. Figure 2.7 shows the results of linear regression between SDC-1 output and empirical damage before and after calibration for *total (*A*)*, *structural* (B) and *content* (C) damage categories.

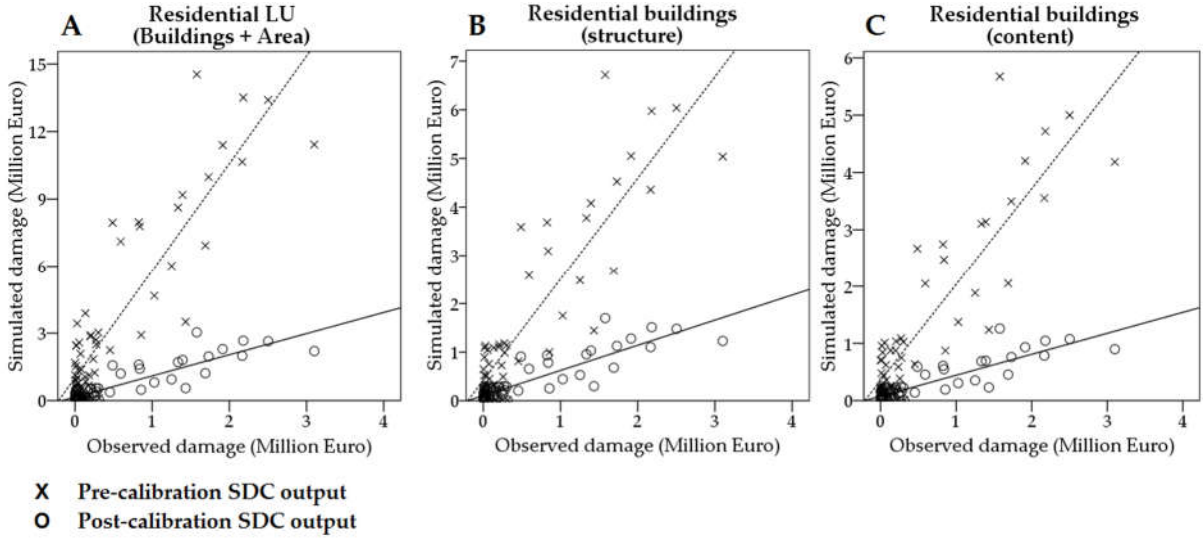


Figure 2.7. Scatterplot showing empirical damage (X axis) and SDC results (Y axis) per grid cell using original land cover values (cross indicator, dotted line) and calibrated ones (circle indicator, black line) for: A) total residential area; B) building structure; C) buildings content.

The pre-calibration output overestimated the total damage in residential areas by a factor 4.5-7 depending on the within-urban land cover category. The calibrated damage values are regressed with the observed/reported damage with good results (R^2=0.8) for all categories except for *urban area* where registered vehicles are assumed to be homogeneously distributed. This proven to be an over simplistic assumption. For buildings structure and content the coefficient (*B*) is close to 1.0, and the final output overestimate recorded residential damage by just 6% (Table 2.1).

Land cover		**Observed**	**Simulated**		
Description	Area (m^2)	Damage (million Euro)	Damage (million Euro)	R^2	B
Urban area (vehicles)	1,432,650	5,5	2,4	0.3	0.2
Buildings	234,950	36	41,9	0.8	1.0
Buildings structure		22,3	24,3	0.8	1.0
Buildings content		13,7	17,6	0.7	1.0
Total	1,667,600	41,5	44,4	0.8	0.9

Table 2.1. Observed exposed area and simulated damage inclusive of regression results for each calibrated land cover category tested against empirical data.

2.3.2 Agricultural losses

The flood extent comprises predominately rural areas (43 km^2), with a prevalent share of arable crops (81% of UUA). The typical crops include cereals, in particular soft wheat and maize (40% of arable crops) and forage (52% of arable crops). Other arable crops together cover less than 8 per cent. Vineyards and other permanent crops cover the remaining 19% of UAA. As shown in Figure 2.3, in January maize crops are fallow, while wheat is in its vegetative stage. This

means that just half of cereal production is affected. Losses for wheat crops include all the initial costs, which amounts to 50 per cent of total value. Permanent crops are affected by 20% of annual production value. The maximum damage (total loss) to cropland estimated from these share using equation [2] is 343 Euro/ha, less than half compared to the max value used by SDC-1 (790 Euro/ha). This adjusted max value leads to a maximal estimated loss by SDC-1 of 375 thousand Euro over 4.2 thousand ha of crop land. Empirical sampling on crop production suggests that the assumption of total loss for exposed crops is over-pessimistic, since crop plants shown good tolerance to inundation(Setti 2014)(Setti 2014)(Setti 2014)(Setti 2014)(Setti 2014)(Setti 2014)(Setti 2014)(Setti 2014)(Setti 2014)(Setti, 2014) (Setti, 2014). Overall, an estimate based on case-specific data should be preferred over unadjusted SDC values.

2.3.3 Production losses

The losses are calculated for each economic sector as a share of total annual production. The largest share of damage come from the industrial sector, affected for 434 million Euro, equivalent to 14% of its annual production (4.2% of total GVA for SLL Modena, see Table 2.2). The ratio between asset damage and annual GVA sheds light on the equivalence of structural damage and production losses as a function of the flood characteristics (Figure 2.8). For water depth around 1 meter, the linear trend describes an asset damage close to annual production losses (ratio of 1), similarly to the stage-impact curve assumptions in figure 2.5.

	Million Euro	Sector %	Total %
Agriculture	9,1	6,41	0,09
Industry	434.1	14,11	4,20
Services	147.2	2,07	1,42
TOTAL	590,4		5,71

Table 2.2. Modelled impact on GVA from the event of Modena 2014.

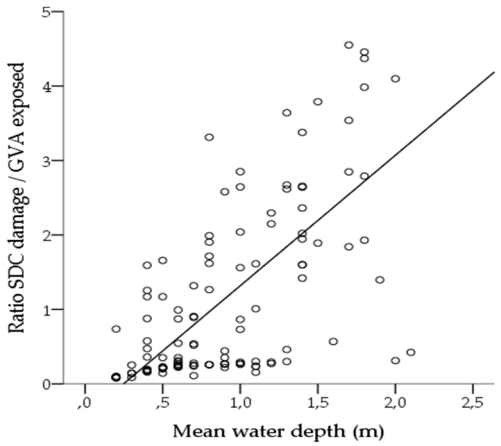


Figure 2.8. Scatterplot of mean water depth (X) and ratio of SDC damage over exposed GVA (Y).

2.3.4 Discussion

In this paper, we have presented three ways to improve the current state-of-the-art of flood risk assessment models based on SDC method. Major uncertainties in damage assessments are associated with the value of risk-exposed elements (i.e. maximum damage values) and the depth-damage curves (De Moel and Aerts 2011; Scorzini 2015). First, we have shown that by

adjusting the maximum damage values for the specific conditions of the assessment area, the consistency of the model improves substantially. Prior adjustments, the tested SDC models overestimate the reported damage by a factor 4 to 13. After calibration, the maximum damage values for residential buildings are 4 to 4.5 times smaller than the original values and the simulation of total damage is very close to empirical observations. These considerations for the Italian territory are consistent with those found by Scorzini (2015), who similarly stresses the importance of evidence-based SDC to perform a meaningful flood risk assessment.

Then, we considered the temporal variability in the agricultural sector using detailed crop yield data and local production patterns. This approach produces a different outcome compared to the conventional SDC estimate: the maximum crop-yield loss per hectare is less than a half of what is assumed by SDC-1; similarly, lower damage estimates using a time-dependent approach are found in Forster et al. (2008). Still, our estimate appears to represent a pessimistic scenario compared to available evidences of small to no damage to crops production in our case study. Lastly, we explains how the GVA approach can approximate output losses within the flooded area with relative ease, if economic data are available. We estimated that the production losses amount to around 600 million Euros, or 5.7 per cent of the annual GVA of the Modena SLL. Asset damage appears close to the annual GVA when the average water depth reaches one meter. However, these results are hardly comparable with empirical observations about production losses at the regional scale and thus cannot be properly validated.

2.4 Conclusion

Our analysis aimed at improving flood damage assessment modelling in Italy. The comparison of damage estimates made by SDC models with empirical recorded damage is key for this task. In this paper we tested two frequently used SDC models against reported flood damage after a major flood event in Northern Italy. Model calibration is proven mainly useful to improve the loss assessment in a specific event area, while it is yet to be studied how these calibrated curves can be adjusted for application in surrounding regions. The calibration here is carried out for residential land cover categories only, while empirical damage records about industrial land cover is awaited to complete the assessment in future research.

Further improvements can be achieved when a larger number of empirical damage evidences, typically collected by the Civil Protection Agency (CPA), is made accessible to academic community. Another research thread capable to improve the reliability of flood risk models by reducing the largest uncertainty in the definition of maximum damage values entails the

spatially disaggregated socioeconomic data such as population, household income, cadastral value of property. With the growing availability of digital spatial data related to these variables, their implementation in an integrated model is a advisable step to improve the representativeness and reliability of flood risk assessment.

2.5 Acknowledgment

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 308438 (*ENHANCE - Enhancing Risk Management Partnerships For Catastrophic Natural Hazards In Europe*).

We are grateful to the Department of Civil Engineering, Environment, Territory and Architecture of the University of Parma for the provision of flood simulation of the 2014 Modena flood event.

2.6 References

Altamura V, Cesena C, Aldo P, Unibo B (2013) Costi di produzione delle principali specie vegetali in Emilia-Romagna 2013.

Barredo JI, Engelen G (2010) Land Use Scenario Modeling for Flood Risk Mitigation. Sustainability 2:1327–1344. doi: 10.3390/su2051327

Bosello F, Nicholls RJ, Richards J, et al (2012) Economic impacts of climate change in Europe: sea-level rise. Clim Change 112:63–81.

Carrera L, Standardi G, Bosello F, Mysiak J (2015) Assessing direct and indirect economic impacts of a flood event through the integration of spatial and computable general equilibrium modelling. Environ Model Softw 63:109–122. doi: 10.1016/j.envsoft.2014.09.016

Cavallo E, Bank ID, Galiani S, Pantano J (2012) Catastrophic Natural Disasters and Economic Growth. Washington DC, USA

Citeau J (2003) A new flood control concept in the Oise catchment area: definition and assessment of flood compatible agricultural activities. 1–19.

Crichton D (1999) The Risk Triangle. Nat Disaster Manag 102–103.

D'Alpaos L, Brath A, Fioravante V (2014) Relazione tecnico-scientifica sulle cause del collasso dell ' argine del fiume Secchia avvenuto il giorno 19 gennaio 2014 presso la frazione San Matteo.

De Moel H, Aerts JCJH (2011) Effect of uncertainty in land use, damage models and inundation depth on flood risk estimates. Nat Hazards 58:407–425.

De Moel H, Asselman NEM, Aerts J (2012) Uncertainty and sensitivity analysis of coastal flood damage estimates in the west of the Netherlands. Nat Hazards Earth Syst Sci 12:1045–1058.

De Moel H, Bouwer LM, Aerts JCJH (2014) Uncertainty and sensitivity of flood risk calculations for a dike ring in the south of the Netherlands. Sci Total Environ 473-474:224–234. doi: 10.1016/j.scitotenv.2013.12.015

De Moel H, van Vliet M, Aerts JCJH (2013) Evaluating the effect of flood damage-reducing measures: a case study of the unembanked area of Rotterdam, the Netherlands. Reg Environ Chang 1–14. doi: 10.1007/s10113-013-0420-z

European Environment Agency (2006) CORINE Land Cover.

Forster S, Kuhlmann B, Lindenschmidt K, Bronstert A (2008) Assessing flood risk for a rural detention area.

Fotia F (2014) Esondazione del Secchia. http://www.meteoweb.eu/2014/01/esondazione-del-secchia-evacuati-20-milioni-di-metri-cubi-dacqua/255270/. Accessed 29 Aug 2014

Gazzetta di Modena (2014) Alluvione: 54 milioni di euro di danni in agricoltura.

Genovese E (2006) A methodological approach to land use-based flood damage assessment in urban areas: Prague case study. Eur. Communities, DG-JRC, Ispra, EUR

Green PC, Viavattene C, Thompson P, Green C (2011) CONHAZ Report - Guidance for assessing flood losses.

Hallegatte S (2008) An adaptive regional input-output model and its application to the assessment of the economic cost of Katrina. Risk Anal 28:779–99. doi: 10.1111/j.1539-6924.2008.01046.x

Henriet F, Hallegatte S, Tabourier L (2012) Firm-network characteristics and economic robustness to natural disasters. J Econ Dyn Control 36:150–167. doi: 10.1016/j.jedc.2011.10.001

Huizinga HJ (2007) Flood damage functions for EU member states.

ISTAT (2013) Banca dati Italia.

ISTAT (2011) 15° censimento della populazione e delle abitazioni.

Jongman B, Kreibich H, Apel H, et al (2012) Comparative flood damage model assessment: towards a European approach. Nat Hazards Earth Syst Sci 12:3733–3752.

Jonkhoff W (2009) Flood risk assessment and policy in the Netherlands. In: OECD (ed) Green Cities new approaches to confronting Clim. Chang. OECD, Las Palmas de Gran Canaria, pp 220–240

Jonkman SN, Bočkarjova M, Kok M, Bernardini P (2008) Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. Ecol Econ 66:77–90.

Klijn F, Baan P, Bruijn K de, Kwadijk J (2007) Overstromingsrisico's in Nederland in een veranderend klimaat.

Kok M, Huizinga HJ, Barendregt A (2005) Schade en Slachtoffers als gevolg van overstromingen.

Koks EE, Bockarjova M, De Moel H, Aerts JCJH (2014) Integrated direct and indirect flood risk modeling: development and sensitivity analysis. Risk Anal.

Kron W (2005) Flood Risk = Hazard x Values x Vulnerability. Water Int 30:58–68. doi: 10.1080/02508060508691837

Luino F, Cirio CG, Biddoccu M, et al (2009) Application of a model to the evaluation of flood damage. Geoinformatica 13:339–353. doi: 10.1007/s10707-008-0070-3

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

Messner F, Penning-rowsell E, Green C, et al (2007) Evaluating flood damages: guidance and recommendations on principles and methods.

Meyer V, Becker N, Markantonis V, et al (2013) Review article: Assessing the costs of natural hazards – state of the art and knowledge gaps. Nat Hazards Earth Syst Sci 13:1351–1373. doi: 10.5194/nhess-13-1351-2013

Molinari D, Aronica GT, Ballio F, et al (2013) Le curve di danno quale strumento a supporto della direttiva alluvioni: criticità dei dati italiani. 10–15.

Noy I, Nualsri A (2007) What do exogenous shocks tell us about growth theories? Santa Cruz Inst. for International Economics, Santa Cruz

Osservatorio Agro-Alimentare Emilia Romagna (2014) Il sistema agro-alimentare dell'Emilia-Romagna.

Peduzzi P, Dao H, Herold C, Mouton F (2009) Assessing global exposure and vulnerability towards natural hazards: the Disaster Risk Index. Nat Hazards Earth Syst Sci 9:1149–1159. doi: 10.5194/nhess-9-1149-2009

Przyluski V, Hallegatte S (2011) Indirect Costs of Natural Hazards - CONHAZ Report.

Regione Emilia-Romagna (2011) Coperture vettoriali dell'uso del suolo 2008. http://servizigis.regione.emilia-romagna.it. Accessed 1 Jul 2015

Rose A, Wei D (2013) Estimating the Economic Consequences of a Port Shutdown: the Special Role of Resilience. Econ Syst Res 25:212–232. doi: 10.1080/09535314.2012.731379

Saint-Geours N, Bailly J-S, Grelot F, Lavergne C (2014) Multi-scale spatial sensitivity analysis of a model for economic appraisal of flood risk management policies. Environ Model Softw 60:153–166. doi: doi.org/10.1016/j.envsoft.2014.06.012

Sargent DM (2013) Updating Residential Flood Stage-Damage Curves based on Building Cost Data. Reach. out to Reg. - SIAQ Conf.

Scorzini A, Frank E (2015) Flood damage curves : new insights from the 2010 flood in Veneto, Italy.

Setti G (2014) Recuperare i suoli dopo l'alluvione. Terra e vita 11:

Strobl E (2010) The Economic Growth Impact of Hurricanes: Evidence from U.S. Coastal Counties. Rev Econ Stat 93:575–589. doi: 10.1162/REST_a_00082

Tapia-Silva F-O, Itzerott S, Foerster S, et al (2011) Estimation of flood losses to agricultural crops using remote sensing. Phys Chem Earth, Parts A/B/C 36:253–265.

Tebodin (2000) Schadecurves industrie ten gevolge van overstroming.

Thieken a. H, Ackermann V, Elmer F, et al (2009) Methods for the evaluation of direct and indirect flood losses. RIMAX Contrib 4th Int Symp Flood Def 1–10.

Twining S (2014) Impact of 2014 Winter Floods on Agriculture in England.

UOOML, PSAL (2009) Linee guida per la sorveglianza sanitaria e la prevenzione dei rischi per la salute e la sicurezza nel settore cerealicolo.

Vacondio R, Aureli F, Mignosa P, Palù AD (2014) 2D shallow water GPU parallelized scheme for high resolution real-field flood simulations.

Ward PJ, Jongman B, Weiland FS, et al (2013) Assessing flood risk at the global scale: model setup, results, and sensitivity. Environ Res Lett 8:044019. doi: 10.1088/1748-9326/8/4/044019

Winsemius HC, Van Beek LPH, Jongman B, et al (2013) A framework for global river flood risk assessments. Hydrol Earth Syst Sci 17:1871–1892. doi: 10.5194/hess-17-1871-2013

3 FLOOD LOSS MODELLING WITH FLF-IT: A NEW FLOOD LOSS FUNCTION FOR ITALIAN RESIDENTIAL STRUCTURES

3.1 Introduction

Floods are the natural hazards that cause the largest economic impact in Europe today (European Environment Agency 2010). Italy is no exception, with about 80% of its municipalities being exposed to some degree of hydrogeological hazards (Zampetti et al. 2012). Regarding flood hazard frequency, 8% of Italy's territory and 10% of its population are exposed to a flood probability of once every 100 to 200 years (ANCE/CRESME 2012a; Trigila et al. 2015a). This issue is reflected in over a billion Euros spent from 2009 to 2012 on recovery from extreme hydrological events (Zampetti et al. 2012). Italy is, in fact, the European country where floods generate the largest economic damage per annum (Alfieri et al. 2016). This is especially worrisome considering that the frequency of extreme flood losses may be doubled at least by 2050 in Europe due to climatic change factors and urban expansion (Jongman et al. 2014). Climate variability already affects rainfall extremes and the peak volumes of discharge in rivers (Alfieri et al. 2015a; Karagiorgos et al. 2016). Relentless urban sprawl within catchments alters the water run-off speed and propagation while increasing the value of exposed land use (Barredo 2009). In order to effectively prevent massive losses, disaster risk management requires estimation well in advance of the frequency and magnitude of potential flood events, and their consequences in terms of economic damages (Kaplan and Garrick 1981; UNISDR 2004; Thieken et al. 2008; Elmer et al. 2010; Neale and Weir 2015; Hammond et al. 2015). Therefore, it is indispensable to provide decision-makers with reliable assessment tools that are able to produce such knowledge, after which an efficient risk reduction strategy can be adequately planned (Penning-Rowsell et al. 2005; Merz et al. 2010a; Emanuelsson et al. 2014; McGrath et al. 2015).

In general, flood losses are classified as marketable (tangible) or non-marketable (intangible) values, and as direct or indirect (Thieken et al. 2005; Jonkman 2007; Kreibich et al. 2010; Meyer et al. 2013b; Molinari et al. 2014a). Direct damage takes place when the floodwater physically inundates buildings and structures, whereas indirect damage accounts for the consequences of direct damage on a wider scale of space and time (Hasanzadeh Nafari et al.

2016c). The tools employed to assess flood risk consist of a variety of damage models, with differing methods depending on the type of accounted losses. While Input-Output models, Computable General Equilibrium models and other econometric tools are often used to estimate indirect economic losses (Hallegatte 2008; Koks et al. 2015a; Carrera et al. 2015), the focus of most flood damage models is still on the estimation of direct, tangible losses using stage-damage curves. Stage-damage curves or flood loss functions are used to depict a relationship between water depth and economic damage for a specific kind of structure or land cover (Messner et al. 2007b; Kreibich and Thieken 2008; Thieken et al. 2009; Merz et al. 2010a; Jongman et al. 2012a). Damage curves can be empirical or synthetic. Empirical curves are drawn based on actual data collected from one specific event. Due to the differences in flood and building characteristics, they cannot be directly employed in different times and places (McBean et al. 1986; Gissing and Blong 2004). To resolve this issue, general synthetic curves based on a valuation survey have been created for different types of buildings. Valuation surveys assess how the structural components are distributed in the height of a building (Smith 1994; Barton et al. 2003). Afterwards, the magnitude of potential flood losses is estimated based on the vulnerability of structural components and via "what-if" questions (Gissing and Blong 2004; Merz et al. 2010a). Damage functions can also be distinguished as absolute or relative. The first type states the damage directly in monetary terms, while the relative type states the damage as a percentage of the total exposed value, which can refer to the total replacement value or the total depreciated value (Kreibich et al. 2010). Relative functions have an advantage over absolute functions, namely that they are more flexible for transfer to different regions or years since the damage ratio is independent of the changes in market values (Merz et al. 2010a). Still, both types are developed on sample areas which have particular geographical characteristics that affect both the quality of the exposed value and the flood phenomena (Proverbs and Soetanto 2004; McGrath et al. 2015). Therefore, transferred models may carry a high level of uncertainty, unless they are calibrated with an empirical dataset collected from the new study area (Cammerer et al. 2013; Molinari et al. 2014b; Hasanzadeh Nafari et al. 2015).

Although Italy has seen several flood disasters in recent years, flood records do not enable development or validation of a national loss flood function because the information is still poor, fragmented and inconsistent. This issue largely depends on the lack of an established official procedure for the collection and the storage of damage data (Molinari et al. 2014b). Another obstacle is the heterogeneity across different regions of digital geographic information, which is the key to correctly represent the driving factors of exposure and vulnerability influencing

the sustained damage. Few attempts at drawing a depth-damage relation from post-disaster reports have been made (Luino et al. 2009b; Molinari et al. 2012b; Papathoma-Köhle et al. 2012; Molinari et al. 2014b; Scorzini and Frank 2015b; Amadio et al. 2016), while other uncalibrated synthetic functions have been derived from pan-European studies (Huizinga 2007b). The use of such uncalibrated functions on the Italian territory has proven troublesome (Amadio et al. 2016), showing a large degree of uncertainty.

Our research aims to calibrate and validate a new relative flood loss function for Italian residential structures (FLF-IT) based on real damage data collected from one large river flood event in the region of Emilia-Romagna at the beginning of 2014. The focus of this study is on direct tangible damage, and the spatial scale is on the order of individual buildings. This research builds on a newly derived Australian approach called FLFA (Hasanzadeh Nafari et al. 2016a, 2016b).

3.2 Case study

The region of Emilia-Romagna is located in Northern Italy, on the southern side of the Po River, the longest of all Italian rivers. This region has the greatest flood prone area both in relative and absolute terms: about 10,000 km^2, including 64% of the population are exposed to a medium flood probability (return period between 100 and 200 years), while 2,500 km^2 and 10% of the population are exposed to a high probability (return period between 20 and 50 years) (Trigila et al. 2015a). This includes more than half of the region's territory. Our empirical data comes from a flood generated by the Secchia river in 2014 near the town of Modena, in the central part of Emilia-Romagna.

3.2.1 Event description

January 2014 was a dramatic month for floods in Italy, with 110 flood events recorded over a span of 23 days due to extreme meteorological conditions. Severe precipitations hit central Emilia-Romagna between the 17th and the 19th of January, with an areal mean of 125 mm of cumulative rain over 72 hours flowing in the Secchia catchment. The increase in the river flow volumes caused heavy stress on the levees, which stand 7-8 meters over the flood plain. At around 6 am, approximately 10 meters of the eastern Secchia levee were overwashed and breached at the top by one meter, thereby starting to flood the countryside. In 9 hours, the levee section was completely destroyed for a length of 80 meters, spilling 200 m^3 per second in the surrounding plain and flooding nearly 65 km^2 of rural land (Figure 3.1) (D'Alpaos et al. 2014b). Seven municipalities have been affected, with the small towns of Bastiglia and Bomporto

suffering the largest share of losses. Both towns, including their industrial districts, remained flooded for more than 48 hours. The total volume of water inundating the area was estimated to be around 36 million m^3 (D'Alpaos et al. 2014b).

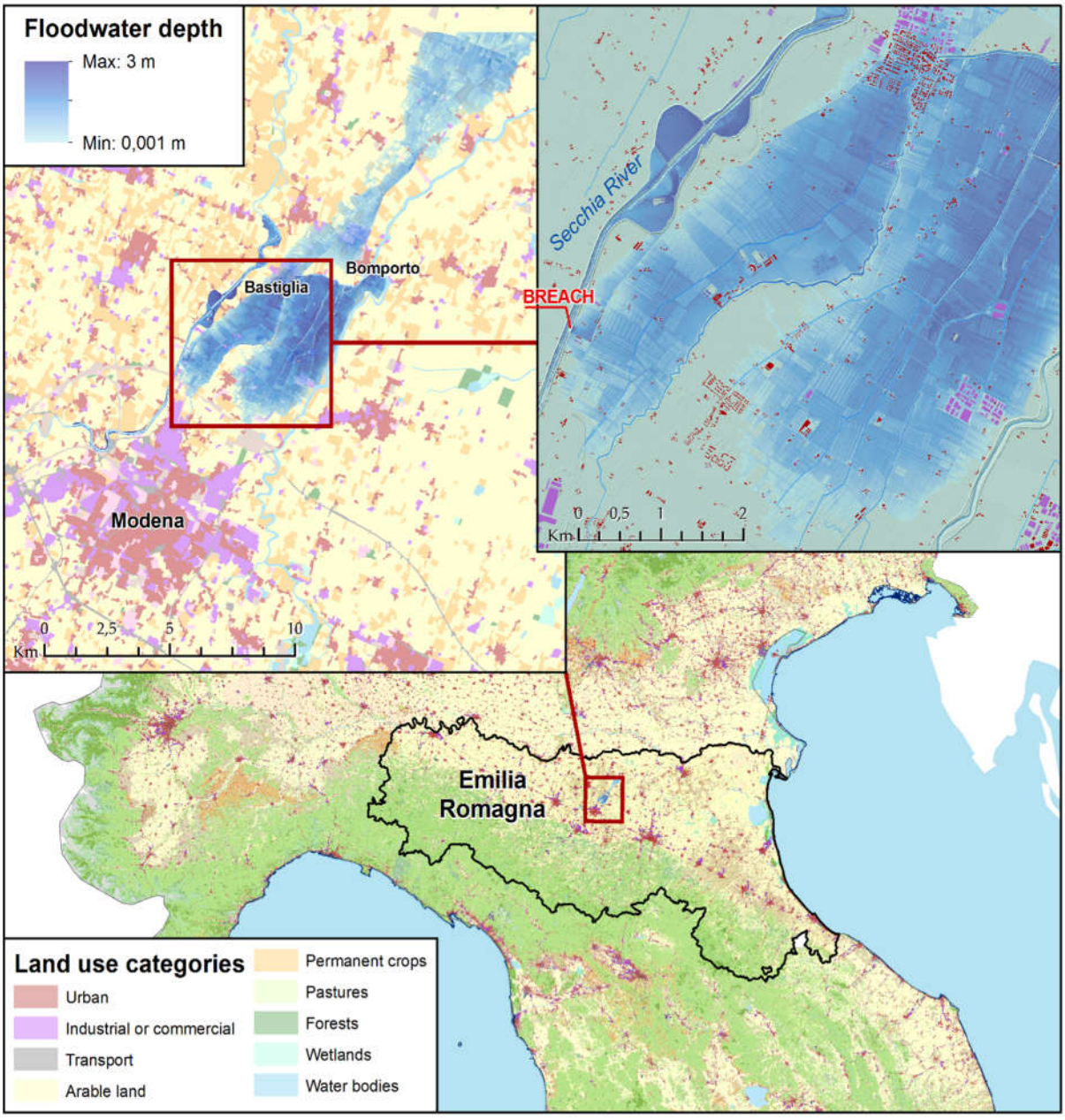


Figure 3.1. Identification of case study, flooding from the river Secchia during January 2014 in central Emilia-Romagna, Northern Italy.

3.2.2 Data description

The information about cumulative water depths comes from the hydraulic simulation of the event produced by the technical-scientific committee in the official report (D'Alpaos et al. 2014b; Vacondio et al. 2014). The extent of the simulated flood is nearly 5 km^2, with an average depth of one meter. The flow volume at the breach is calculated using the 1-D model HEC-

RAS calibrated on recorded observations from the event. The evolution of the flooding is simulated by a 2-D hydraulic model using the finite-volume method over a Digital Terrain Model (DTM) obtained by LiDAR scans at a one-meter resolution. The simulation also accounts for the gradual change in the size of the breach from 10 to 80 meters (Vacondio et al. 2014).

A database of damage declared by residential properties has been made available for this research by the local authorities. Damage records are listed by address for the three municipalities of Bastiglia (70% of the total damage), Bomporto (24%) and Modena (6%). The total damage sums up to EUR 41.5 million, of which: 54% is damage to structural parts, including installations; 33% is damage to movable contents, meaning furniture and common domestic appliances; and 13% is represented by registered vehicles, such as cars and motorcycles. For the purpose of our study, only the structural damage is considered. The recorded damage is compared to the average market values of the residential properties, as reported by the cadastral map for the semester preceding the flood event (Agenzia delle Entrate 2018). The majority of residential structures in the area share the same general characteristics: they are brick or concrete buildings built in the last 30 years, with no underground basement or parking (slab-on-ground). Houses have at least two or three floors. However, only the ground floors have been affected in this particular event.

The information related to water depth, total structural damage and average market value is linked together at the building scale (Figure 3.2) by combining the street numbers points and residential buildings perimeters from the official regional geodatabase (Regione Emilia Romagna 2011). The mean of cumulative water depths simulated by the hydraulic model is calculated within the area of each building unit. Accordingly, each address linked to a damage record is first georeferenced as a street number point; then the points falling within the same building unit are summed into an aggregated value representing the total structural damage occurred in that building, including private dwellings and common parts. This spatial join is necessary since building perimeters do not include any information about addresses.

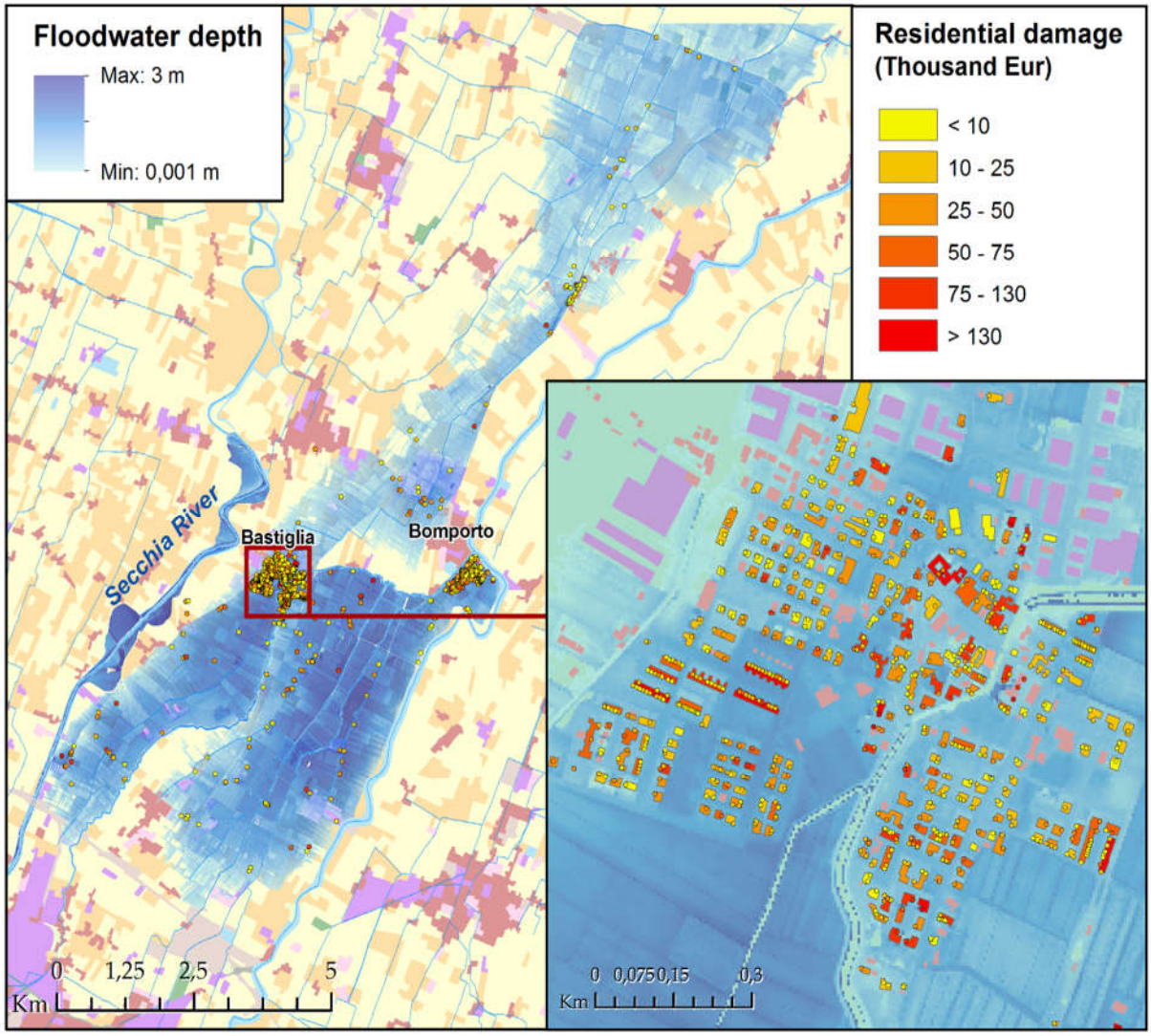


Figure 3.2. Visualisation of the empirical damage records suffered by the individual dwellings during the flood event of 2014. Records are projected to official street number points by using their "address" field. The information is then transferred from the points to the building features that contains them. The point records that fall within the same building perimeter are summed up into one aggregated damage value for each residential building. About 97% of damage records are correctly projected. The remaining 3% of damage records is discarded due to inconsistent projection, incomplete address or gaps in the record data. The colour gradient (yellow to red) indicates the magnitude of the damage for both individual points and building units.

The procedure it is performed successfully for EUR 21.7 million, corresponding to 97% of the total residential damage. The remaining 3% of records are excluded due to incomplete addresses or inconsistency with the spatial data. Percentages of damage vs. depths of water for all 613 final samples have been depicted in Figure 3.3.

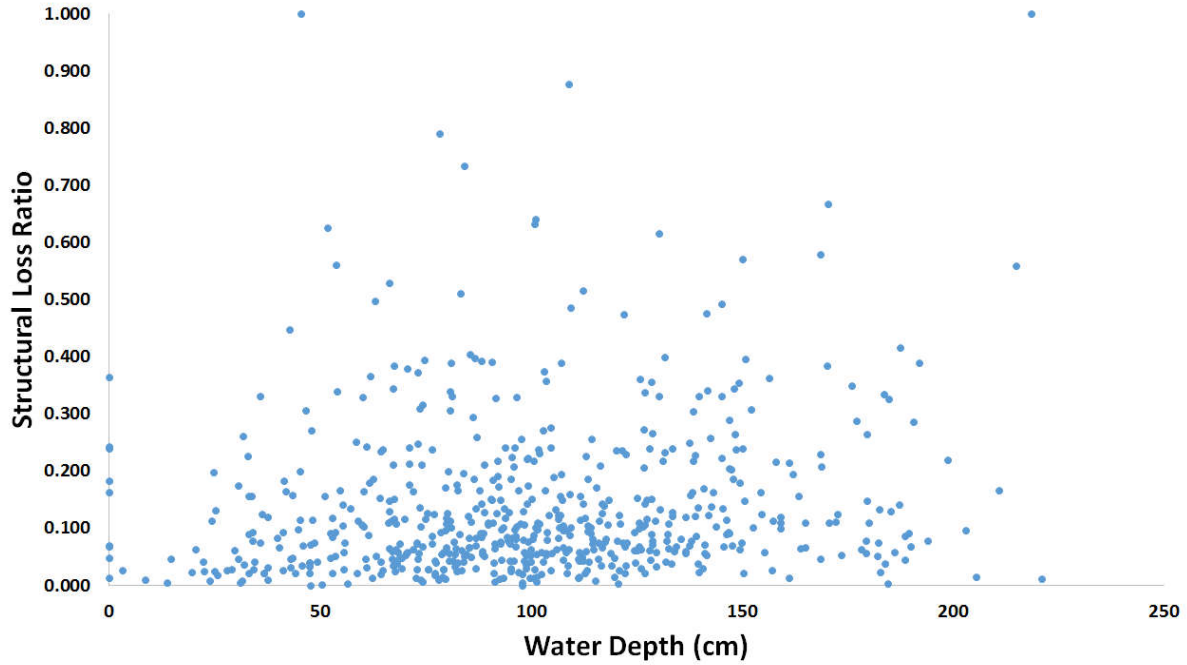


Figure 3.3. Empirical data utilised for calibrating the FLF-IT model (613 relative damage records in the original dataset).

3.3 The FLFA method

The FLFA method is based on a simplified synthetic approach called the sub-assembly method, proposed by the HAZUS technical manual (FEMA 2012). This method measures the extent of losses for each stage of floodwater and suggests a flexible curve that accounts for the variability in the characteristics of structures. In the first step, one or more representative building categories are selected from the study area. The ratio of damage for every stage of water and within each category of the building is a function of the vertical distribution of structural components (i.e., vulnerability and the total value exposed to flood) (Lehman and Hasanzadeh Nafari 2016). More specifically, each structural component starts suffering damage after a specific stage is reached. Commonly the first decimetres of water cause damage to some of the most valuable items such walls, floors, insulation and electrical wiring (FEMA 2012). Accordingly, the relationship between the damage percentage (d_h) and water depth can be described by a root function (Kreibich and Thieken 2008; Elmer et al. 2010; Cammerer et al. 2013). The following function (1) is developed by Hasanzadeh Nafari et al. (2016a) for the Australian case study:

$$d_h = \left(\frac{h}{H}\right)^{\frac{1}{r}} \times D_{max} \qquad \text{[Equation 3.1]}$$

The root (r) controls the rate of alteration in the percentage of damage relative to the growth of the water depth (h) over a total height (H) of the floor. The D_{max} is the total percentage of damage corresponding to the total height of the floor. A higher value of r means a slower increase in the rate of damage. The obtained curve is then adjusted and calibrated using the empirical data collected from the selected study area. Hence, this approach is defined as an empirical-synthetic method. Due to the inherent uncertainty in the data sample, the study has employed a bootstrapping approach, which produces three stage-damage functions (i.e. most likely, maximum and minimum damage functions) for each type of building. This range of estimate describes confidence limits around the functional parameters and represents the uncertainty that exists in the data sample. The advantages of this simplified synthetic approach include calibration with empirical data, a better level of transferability in time and space, consideration of the epistemic uncertainty of data, and the ability to change parameters based on building practices across the world.

3.4 Calibration of FLF-IT

Based on the formula represented previously, the model calibration process includes choosing the most appropriate values for the root of function and the maximum percentage of damage (i.e., r and D_{max}), with reference to the empirical dataset (Hasanzadeh Nafari et al. 2016b). The selection will be made by the chi-square test of goodness of fit, to minimise predictive errors. Also, instead of a deterministic regression analysis, this study has relied on the probabilistic relationship among the percentage of damage and other damage-related parameters (i.e. building and flood characteristics) (Hasanzadeh Nafari et al. 2016a). In this regard, a bootstrapping approach has been employed to resample the damage data 1,000 times. This method assists in exploring the confidence limits around the parameters and illustrates the epistemic uncertainty of the empirical damage data (Lehman and Hasanzadeh Nafari 2016). To be more specific, first the original dataset including 613 data points was resampled using a bootstrapping approach. For the new resample, the most appropriate value of the root function and the maximum percentage of damage were selected by the chi-square test of goodness of fit. The two previous steps were repeated 1,000 times, and 1,000 sets of parameters (i.e., r and D_{max}) were generated as the result. Finally, by the above iteration, the averages of the 1,000 calibrated parameters converged to a fixed value considered as the most likely scenario. The most likely parameters produce the smallest cumulative error compared to the actual damage data. Also, from the 1,000 sets of parameters generated above, the function that maximises the

depth-damage relationship was taken as a maximum damage curve, and the observation that created the minimum depth-damage relationship was considered for the minimum depth-damage function. Results of the model calibration are presented in Table 3.1 and Figure 3.4.

Number of Samples	Parameters	Range of parameters		
		Minimum	Most likely	Maximum
613	r	2.7	2	1.7
	D_{max}	10%	20%	40%

Table 3.1. Number of samples and range of r and D_{max} values, calculated by the bootstrap and chi-square test goodness of fit.

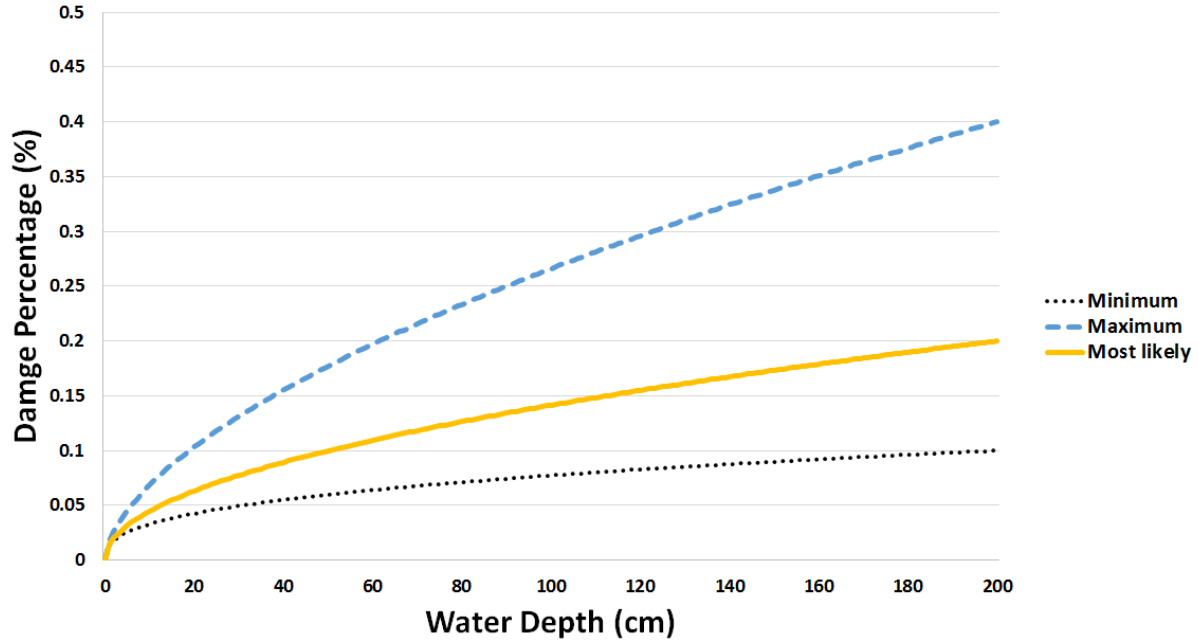


Figure 3.4. Visualisation of minimum, most likely and maximum damage functions, calculated by bootstrap and chi-square test of goodness of fit.

3.5 Model validation

3.5.1 Applied damage models

Besides FLF-IT, Damage Scanner as an uncalibrated relative model frequently used in European studies has been selected for comparison in this study. The Damage Scanner model (de Bruijn 2006; Klijn et al. 2007b) is based on depth-damage curves previously developed by the synthetic approach in the Netherlands using data from "what-if" analyses at the building scale (Kok et al. 2004). These curves estimate the magnitude of damage separately for building structure and movable content. The damage is expressed in relation to an average maximum damage value per square meter, which varies according to land cover classes (e.g. residential, industrial, agriculture, and infrastructure). The Damage Scanner model have been employed for

predictive purpose in various studies (Bouwer et al. 2010; Ward et al. 2011; de Moel et al. 2011; Aerts and Botzen 2011; Koks et al. 2012; Poussin et al. 2012), and it has been more recently updated including additional land cover subclasses (de Moel et al. 2013a; Koks et al. 2014b). The uncertainty of Damage Scanner has been investigated in comparison to other damage models (Bubeck et al. 2011; Jongman et al. 2012a), and its transferability has been evaluated for use in different areas of study such as Northern Italy (Amadio et al. 2016). Damage Scanner is, in fact, easy to tailor to land cover description available for Italy, and because it expresses damage in relative terms, it can be adapted to work on region-specific maximum values. For the purpose of comparison with FLF-IT, the curve related to residential structure damage has been selected from the Damage Scanner set and applied at building scale on the residential units using the same average market values and simulated water stages employed to produce the FLF-IT. It is worth noting that the predicted absolute damage values are calculated by multiplying the estimated loss ratio by the average market value and the area of each property.

3.5.2 Result comparison and model validation

Results of the applied damage models have been compared with the observed loss data, and their performances have been validated in contrast to real damage data. Due to the lack of an independent dataset, a three-fold cross-validation technique was employed for this purpose (Seifert et al. 2010). Accordingly, the original damage records including 613 data points were first shuffled and partitioned into three equally sized subsets. Then, three iterations of model calibration and model testing were performed. In each iteration, one subset including 204 samples was singled out for model testing, while the remaining two parts including 409 data points were used for model calibration (Refaeilzadeh et al. 2009). Model calibration in each iteration was performed based on the approach explained earlier. Eventually, the loss ratio of the held-out subset was estimated by the FLF-IT model calibrated without it, and the results were compared with the actual records. Errors including the mean bias error (MBE), the mean absolute error (MAE) and the root mean square error (RMSE) were calculated and averaged over all three iterations. The MBE illustrates the direction of the error bias (i.e. a positive MBE shows an overestimation in the predicted values, while a negative MBE depicts an underestimation); the MAE shows how close the estimates are to the actual damage ratios; and the RMSE signifies the variation of the predicted ratios from the actual records (Seifert et al. 2010; Chai and Draxler 2014). In addition to FLF-IT and for each iteration, errors of the Damage Scanner model's estimates were calculated. The results are presented in Table 3.2.

	MBE		MAE		RMSE	
	FLF-IT	**Damage Scanner**	**FLF-IT**	**Damage Scanner**	**FLF-IT**	**Damage Scanner**
Iteration 1	0.015	0.152	0.092	0.188	0.119	0.212
Iteration 2	-0.010	0.125	0.104	0.177	0.157	0.204
Iteration 3	-0.009	0.125	0.091	0.164	0.133	0.188
Average	0.00	0.13	0.10	0.18	0.14	0.20

Table 3.2. Error estimation for the performance of the FLF-IT model (MBE: Mean Bias Error; MAE: Mean Absolute Error; RMSE: Root Mean Squared Error).

This table clearly shows that FLF-IT has a better performance compared to the Damage Scanner model which is not calibrated with the local damage data. The average of the MBE over all iterations shows no bias and represents only around 1% bias in each iteration. The MAE is 10% on average, and RMSE alters between 12 and 16% (14% on average). The results of the Damage Scanner model show 13% average deviation from the validation subsets ratios; larger average values of absolute error; and higher variation of the predicted ratios from the actual records. Overall, the small value of the deviations and the low variation of the errors signify that the new model performance is accurate.

The predictive capability has also been studied for some sub-classes of water depth. By this test, the performance of the applied damage models will be evaluated for different stages of the flood. Figures 3.5 and 3.6 show the precision of the results and the number of relative damage records for seven different sub-classes of water depth.

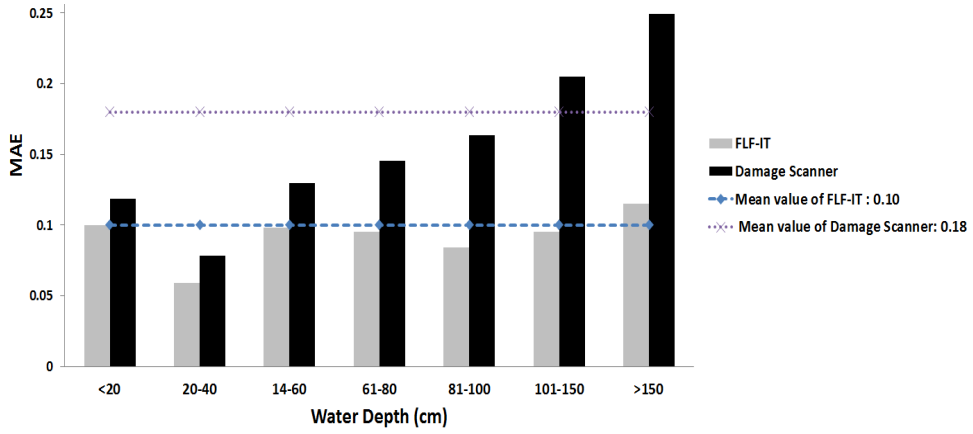


Figure 3.5. Comparison of the flood damage estimation models' precision per water-depth class (MAE: mean absolute error; Number of damage records for each sub-class of water depth, respectively, are 14, 36, 52, 96, 125, 222, and 68).

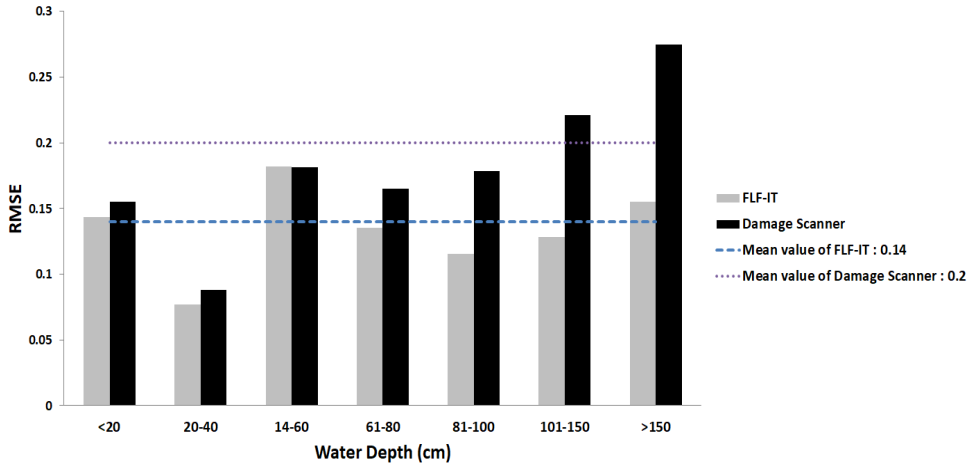


Figure 3.6. Comparison of the flood damage estimation models' precision per water-depth class (RMSE: root mean square error; Number of samples for each sub-class of water depth, respectively, are 14, 36, 52, 96, 125, 222, and 68).

These Figures clearly show that the uncertainty of FLF-IT is less than the Damage Scanner model and the results justify the overall better performance of the FLF-IT model. This test shows that the application of the Damage Scanner model using the original uncalibrated maximum damage values leads to overestimating the actual damage occurred during this flood event especially when the water depth is high. In contrast to Damage Scanner, FLF-IT performs well specifically when the flood is deep, the extent of damage is more considerable, and the prediction performance of the model is more important. The high number of samples with a depth more than 60 centimetres supports the reliability of this outcome.

In addition to the above comparison on the loss ratios, the performance of the model is also validated for predicting the absolute damage values. As stated before, the overall reported loss for the 613 cases (building fabric) amounted to EUR 21.7 million. In this regard and for each iteration, the absolute damage records are resampled using the bootstrapping approach 10,000 times, and the 95% confidence interval of the total losses was calculated. If the total damage value estimated by the models falls within the 95% confidence interval, their performance is accepted. Otherwise, it is rejected (Thieken et al. 2008; Seifert et al. 2010; Cammerer et al. 2013). By this approach, the performance of the applied damage models in terms of structural damage estimation in the area of study will be evaluated. The results are presented in Table 3.5, which shows that the results of all iterations of the FLF-IT model with the most likely functional parameters r and D_{max} lie within the 95% confidence intervals and the FLF-IT model has an acceptable performance. However, results of Damage Scanner do not lie within the confidence intervals of the mean loss ratios, and its performance is rejected in this area of study.

	95% confidence interval	Estimated damage values (in 10^6 EUR)			
		FLF-IT	Within 95% interval	Damage Scanner	Within 95% interval
Iteration 1	4.88-6.8	6.5	Yes	16.2	No
Iteration 2	5.81-7.8	7.7	Yes	15.6	No
Iteration 3	8.07-10.4	10.1	Yes	21.8	No
All records	19.94-24.5	24.3	Yes	53.7	No

Table 3.3. Comparison of total absolute losses estimated by FLF-IT with the 95% confidence interval of the resampled damage records.

Results of these validation tests illustrate the importance of model calibration, especially when the water depth is the only hydraulic parameter taken into account (McBean et al. 1986; Chang et al. 2008; Cammerer et al. 2013). In other words, flood damage, being a complicated process, could be dependent on more damage influencing parameters than those considered here (Fuchs et al. 2011; Merz et al. 2013; Grahn and Nyberg 2014; Schröter et al. 2014; Hasanzadeh Nafari et al. 2016c). However, by calibrating the loss function with an actual damage dataset and providing an empirically-based model, the function estimations are good (i.e. low predictive error, low variation and acceptable reliability in results) and its performance is validated for use in flood events with the same geographical conditions (i.e. flood characteristics and building specifications) as the area of study (McBean et al. 1986; Hasanzadeh Nafari et al. 2016a). While the FLF-IT model is shown to be more accurate, there are still some limitations that can be the subject of new research. Model validation in this study was based on random samples which were not independent of the data used for model calibration, and this test does not give information about the transferability of the FLF-IT model. Hence, improvements can be made by considering more influencing factors of hazard, exposure and vulnerability; validation with more actual damage records from other study areas in Italy; considering other types of structure, and taking into account more variations of residential buildings.

3.6 Conclusion

Floods are frequent natural hazards in Italy, triggering significant negative consequences on the economy every year. Their impact is expected to worsen in the near future due to socio-economic development and climate variability. To be able to reduce the probability and magnitude of expected economic losses and to lessen the cost of compensation and restoration, flood risk managers need to be correctly informed about the potential damage from flood hazards on the territory. A loss function that can reliably estimate the economic costs based on available data is the key to achieving this objective. However, despite a significant number of

flood disasters hitting Italy every year, few attempts at developing a flood damage model from post-disaster reports have been made.

Flood loss functions are an internationally accepted method for estimating direct flood damage in urban areas. Flood losses can be classified as marketable or non-marketable values, and as direct or indirect damages. This study focused on direct, marketable damage due to riverine floodwater inundation. We employed a newly derived Australian approach (FLFA) with empirical damage data from Italy to develop a synthetic, relative flood loss function for Italian residential structures (FLF-IT). The FLFA approach takes data of damage and depth, stratified by building classifications, and uses the chi-square test of goodness of fit to fix a parameterized function to compute depth-damage estimates. Parameters include the height of the stories, maximum damage as a percentage of their total value, and the internal elevation of ground floors. Additionally, FLFA illustrates a bootstrapping approach to the empirical data to assist in describing confidence limits around the parameterized functional depth damage relationship.

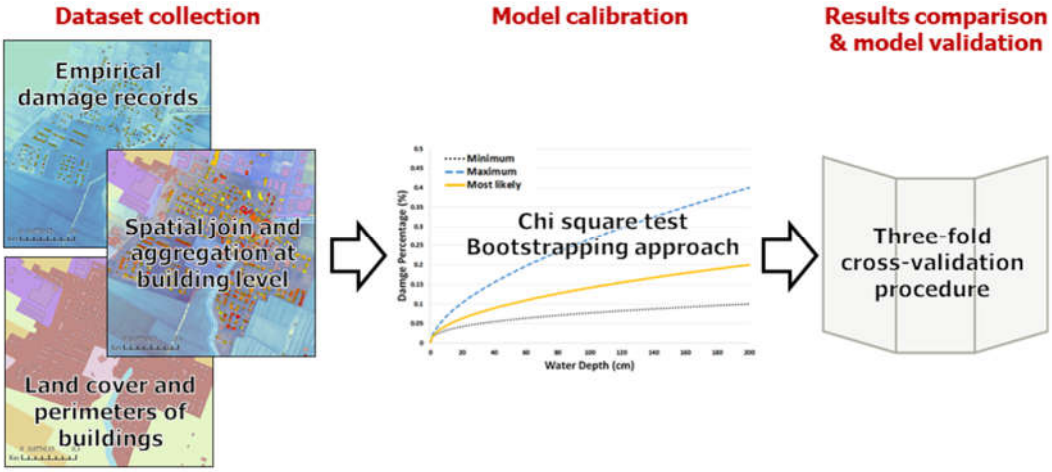


Figure 3.7. Workflow of the analysis.

Accordingly, the advantages of the new model (FLF-IT) include calibration with empirical data, consideration of the epistemic uncertainty of data and the ability to change parameters based on building practices across Italy (see Figure 3.7). After model calibration, its performance was also validated for predicting the loss ratios and absolute damage values. Also, the performance of the new model in comparison to the empirical data has been contrasted with an uncalibrated relative model frequently used in Europe. In this regard, a three-fold cross-validation procedure and the usual bootstrap approach were applied to the empirical sample to measure the range of uncertainty from the actual damage data. This validation test was selected to compensate for the lack of comparable data from an independent flood event. Finally, the predictive capability has also been studied for some sub-classes of water depth. The validation procedure shows that

estimates of FLF-IT are good (no bias, 10% mean absolute error and 14% root mean square error) especially when the flood is deep, and its performance is acceptable. However, the application of the Damage Scanner model using the original uncalibrated maximum damage values leads to overestimating the actual damage occurred during this flood event.

Results of these validation tests depict the importance of model calibration, especially when the water depth is the only hydraulic parameter considered. In other words, by calibrating the loss function and providing an empirically-based model, the function performs well (i.e. low predictive error, low variation and acceptable reliability) and its performance is validated for use in events with the same geographical conditions as the area of study. Awareness of these issues is necessary for decision-making in flood risk management. Further research will be aimed at considering some additional parameters that may govern the significance of the damages for a given depth. An independent dataset is required to evaluate the predictive capacity and transferability of the model.

3.7 Acknowledgment

The authors would like to acknowledge the ongoing financial contribution and support from the Bushfire & Natural Hazards Cooperative Research Centre. The research leading to this paper has also received funding from project ENHANCE (Grant Agreement N° 308438) and the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7 (Grant Agreement number 609642). The authors want to thank the local administrations of Emilia-Romagna for providing the official damage records of the 2014 flood event.

3.8 References

Aerts, J.C.J.H., Botzen, W.J.W., 2011. Climate change impacts on pricing long-term flood insurance: A comprehensive study for the Netherlands. Glob. Environ. Chang. 21, 1045–1060. doi:10.1016/j.gloenvcha.2011.04.005

Agenzia delle Entrate, 2014. Osservatorio del Mercato Immobiliare - Quotazioni zone OMI [WWW Document]. URL http://www.agenziaentrate.gov.it/wps/content/Nsilib/Nsi/Documentazione/omi/Banche+dati/Quotazioni+immobiliari/ (accessed 7.1.15).

Alfieri, L., Burek, P., Feyen, L., Forzieri, G., 2015. Global warming increases the frequency of river floods in Europe. Hydrol. Earth Syst. Sci. 19, 2247–2260. doi:10.5194/hess-19-2247-2015

Alfieri, L., Feyen, L., Salamon, P., Thielen, J., Bianchi, A., Dottori, F., Burek, P., 2016. Modelling the socio-economic impact of river floods in Europe. Nat. Hazards Earth Syst. Sci. 16, 1401–1411. doi:10.5194/nhess-16-1401-2016

Amadio, M., Mysiak, J., Carrera, L., Koks, E., 2016. Improving flood damage assessment models in Italy. Nat. Hazards 82, 2075–2088. doi:10.1007/s11069-016-2286-0

ANCE/CRESME, 2012. Lo stato del territorio italiano 2012. Insediamento e rischio sismico e idrogeologico 284 p.

Barredo, J.I., 2009. Normalised flood losses in Europe: 1970–2006. Nat. Hazards Earth Syst. Sci. 9, 97–104. doi:10.5194/nhess-9-97-2009

Barton, C., Viney, E., Heinrich, L., Turnley, M., 2003. The Reality of Determining Urban Flood Damages, in: NSW Floodplain Management Authorities Annual Conference. Sydney.

Bouwer, L.M., Bubeck, P., Aerts, J.C.J.H., 2010. Changes in future flood risk due to climate and development in a Dutch polder area. Glob. Environ. Chang. 20, 463–471. doi:10.1016/j.gloenvcha.2010.04.002

Bubeck, P., de Moel, H., Bouwer, L.M., Aerts, J.C.J.H., 2011. How reliable are projections of future flood damage? Nat. Hazards Earth Syst. Sci. 11, 3293–3306. doi:10.5194/nhess-11-3293-2011

Cammerer, H., Thieken, a. H., Lammel, J., 2013. Adaptability and transferability of flood loss functions in residential areas. Nat. Hazards Earth Syst. Sci. 13, 3063–3081. doi:10.5194/nhess-13-3063-2013

Carrera, L., Standardi, G., Bosello, F., Mysiak, J., 2015. Assessing direct and indirect economic impacts of a flood event through the integration of spatial and computable general equilibrium modelling. Environ. Model. Softw. 63, 109–122. doi:10.1016/j.envsoft.2014.09.016

Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. Geosci. Model Dev. 7, 1247–1250. doi:10.5194/gmd-7-1247-2014

Chang, L.F., Lin, C.H., Su, M.D., 2008. Application of geographic weighted regression to establish flood-damage functions reflecting spatial variation. Water SA 34, 209–216.

D'Alpaos, L., Brath, A., Fioravante, V., 2014. Relazione tecnico-scientifica sulle cause del collasso dell' argine del fiume Secchia avvenuto il giorno 19 gennaio 2014 presso la frazione San Matteo.

de Bruijn, K.M., 2006. Bepalen van schade ten gevolge van overstromingen voor verschillende scenario's en bij verschillende beleidsopties.

de Moel, H., Aerts, J.C.J.H., Koomen, E., 2011. Development of flood exposure in the Netherlands during the 20th and 21st century. Glob. Environ. Chang. 21, 620–627. doi:10.1016/j.gloenvcha.2010.12.005

de Moel, H., van Vliet, M., Aerts, J.C.J.H., 2013. Evaluating the effect of flood damage-reducing measures: a case study of the unembanked area of Rotterdam, the Netherlands. Reg. Environ. Chang. 14, 895–908. doi:10.1007/s10113-013-0420-z

Elmer, F., Thieken, a. H., Pech, I., Kreibich, H., 2010. Influence of flood frequency on residential building losses. Nat. Hazards Earth Syst. Sci. 10, 2145–2159. doi:10.5194/nhess-10-2145-2010

Emanuelsson, M. a E., Mcintyre, N., Hunt, C.F., Mawle, R., Kitson, J., Voulvoulis, N., 2014. Flood risk assessment for infrastructure networks. J. Flood Risk Manag. 7, 31–41. doi:10.1111/jfr3.12028

European Environment Agency, 2010. Mapping the impacts of recent natural disasters and technological accidents in Europe - An overview of the last decade, EEA Technical report. doi:10.2800/62638

FEMA, 2012. Multi-Hazard Loss Estimation Methodology, Flood Model: Hazus-MH Technical Manual. Washington, D.C.

Fuchs, S., Kuhlicke, C., Meyer, V., 2011. Editorial for the special issue: vulnerability to natural hazards, the challenge of integration. Nat. Hazards 58(2), 609–619. doi:10.1007/s11069-011-9825-5

Gissing, A., Blong, R., 2004. Accounting for variability in commercial flood damage estimation. Aust. Geogr. 35, 209–222. doi:10.1080/0004918042000249511

Grahn, T., Nyberg, R., 2014. Damage assessment of lake floods: Insured damage to private property during two lake floods in Sweden 2000/2001. Int. J. Disaster Risk Reduct. 10, 305–314. doi:10.1016/j.ijdrr.2014.10.003

Hallegatte, S., 2008. An adaptive regional input-output model and its application to the assessment of the economic cost of Katrina. Risk Anal. 28, 779–99. doi:10.1111/j.1539-6924.2008.01046.x

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

Hasanzadeh Nafari, R., Ngo, T., Lehman, W., 2016a. Calibration and validation of FLFArs – a new flood loss function for Australian residential structures 16, 15–27. doi:10.5194/nhess-16-15-2016

Hasanzadeh Nafari, R., Ngo, T., Lehman, W., 2016b. Development and evaluation of FLFAcs – A new Flood Loss Function for Australian commercial structures. Int. J. Disaster Risk Reduct. doi:10.1016/j.ijdrr.2016.03.007

Hasanzadeh Nafari, R., Ngo, T., Lehman, W., 2015. Results comparison and model validation for flood loss functions in Australian geographical conditions. Nat. Hazards Earth Syst. Sci. Discuss. 3, 3823–3860. doi:10.5194/nhessd-15-1-2015

Hasanzadeh Nafari, R., Ngo, T., Mendis, P., 2016c. An Assessment of the Effectiveness of Tree-Based Models for Multi-Variate Flood Damage Assessment in Australia. Water 8, 282. doi:10.3390/w8070282

Huizinga, H.J., 2007. Flood damage functions for EU member states. Technical report implemented in the framework of the contract # 382441-F1SC awarded by the European Commission - Joint Research Centre.

Jongman, B., Hochrainer-stigler, S., Feyen, L., Aerts, J.C.J.H., Mechler, R., Botzen, W.J.W., Bouwer, L.M., Pflug, G., Rojas, R., Ward, P.J., 2014. Increasing stress on disaster-risk finance due to large floods. Nat. Clim. Chang. 1–5. doi:10.1038/NCLIMATE2124

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. Nat. Hazards Earth Syst. Sci. 12, 3733–3752. doi:10.5194/nhess-12-3733-2012

Jonkman, S.N., 2007. Loss of life estimation in flood risk assessment; theory and applications. Delft University of Technology.

Kaplan, S., Garrick, B.J., 1981. On the quantitative definition of risk. Risk Anal. 1, 11–27.

Karagiorgos, K., Heiser, M., Thaler, T., Hübl, J., Fuchs, S., 2016. Micro-sized enterprises: vulnerability to flash floods. Nat. Hazards 84, 1091–1107. doi:10.1007/s11069-016-2476-9

Klijn, F., Baan, P.J.A., De Bruijn, K.M., Kwadijk, J., 2007. Overstromingsrisico's in Nederland in een veranderend klimaat: verwachtingen, schattingen en berekeningen voor het project

Nederland Later. Delft Hydraulics, Deltares (WL).

Kok, M., Huizinga, H.J., Barendregt, A., 2004. Standard Method 2004. Damage and Casualties Caused by Flooding. Client: Highway and Hydraulic Engineering Department.

Koks, E.E., Bockarjova, M., De Moel, H., Aerts, J.C.J.H., 2014. Integrated direct and indirect flood risk modeling: development and sensitivity analysis. Risk Anal. 35(5), 882–900.

Koks, E.E., Carrera, L., Jonkeren, O., Aerts, J.C.J.H., Husby, T.G., Thissen, M., Standardi, G., Mysiak, J., 2015. Regional disaster impact analysis: comparing Input-Output and Computable General Equilibrium models. Nat. Hazards Earth Syst. Sci. Discuss. 3, 7053–7088. doi:10.5194/nhessd-3-7053-2015

Koks, E.E., Moel, H. de, Bouwer, L.M., 2012. Effect of spatial adaptation measures on flood risk in the coastal area of Flanders. Valorisation Report 10b: CcASPAR.

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

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

Lehman, W., Hasanzadeh Nafari, R., 2016. An Empirical, Functional approach to Depth Damages, in: FLOODrisk 2016, the 3rd European Conference on Flood Risk Management. Lyon, France.

Luino, F., Cirio, C.G., Biddoccu, M., Agangi, a., Giulietto, W., Godone, F., Nigrelli, G., 2009. Application of a model to the evaluation of flood damage. Geoinformatica 13, 339–353. doi:10.1007/s10707-008-0070-3

McBean, E., Fortin, M., Gorrie, J., 1986. A critical analysis of residential flood damage estimation curves. Can. J. Civ. Eng. 13, 86–94. doi:10.1139/l86-012

McGrath, H., Stefanakis, E., Nastev, M., 2015. Sensitivity analysis of flood damage estimates: A case study in Fredericton, New Brunswick. Int. J. Disaster Risk Reduct. 1–9. doi:10.1016/j.ijdrr.2015.09.003

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

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

Messner, F., Penning-rowsell, E., Green, C., Tunstall, S., Veen, A. Van Der, Tapsell, S., Wilson, T., Krywkow, J., Logtmeijer, C., Fernández-bilbao, A., Geurts, P., Haase, D., 2007. Evaluating flood damages : guidance and recommendations on principles and methods principles and methods. Flood Risk Manag. Hazards, Vulnerability Mitig. Meas. 189.

Meyer, V., Becker, N., Markantonis, V., Schwarze, R., Van Den Bergh, J.C.J.M., Bouwer, L.M., Bubeck, P., Ciavola, P., Genovese, E., Green, C., Hallegatte, S., Kreibich, H., Lequeux, Q., Logar, I., Papyrakis, E., Pfurtscheller, C., Poussin, J., Przyluski, V., Thieken, a. H., Viavattene, C., 2013. Review article: Assessing the costs of natural hazards-state of the art and knowledge gaps. Nat. Hazards Earth Syst. Sci. 13, 1351–1373. doi:10.5194/nhess-13-1351-2013

Molinari, D., Ballio, F., Berni, N., Pandolfo, C., 2012. Implementing the Floods Directive: the case of the Umbria Region, in: FLOODRisk 2012. Rotterdam, pp. 1–7.

Molinari, D., Ballio, F., Handmer, H., Menoni, S., 2014a. On the modeling of significance for flood damage assessment. Int. J. Disaster Risk Reduct. 10, 381–391. doi:10.1016/j.ijdrr.2014.10.009

Molinari, D., Menoni, S., Aronica, G.T., Ballio, F., Berni, N., Pandolfo, C., Stelluti, M., Minucci, G., 2014b. Ex post damage assessment: an Italian experience. Nat. Hazards Earth Syst. Sci. 14, 901–916. doi:10.5194/nhess-14-901-2014

Neale, T., Weir, J.K., 2015. Navigating scientific uncertainty in wildfire and flood risk mitigation: A qualitative review. Int. J. Disaster Risk Reduct. 13, 255–265. doi:10.1016/j.ijdrr.2015.06.010

Papathoma-Köhle, M., Keiler, M., Totschnig, R., Glade, T., 2012. Improvement of vulnerability curves using data from extreme events: debris flow event in South Tyrol. Nat. Hazards 64(3), 2083–2105. doi:10.1007/s11069-012-0105-9

Penning-Rowsell, E., Johnson, C., Tunstall, S., Morris, J., Chatterton, J., Green, C., Koussela, K., Fernandez-bilbao, A., 2005. The Benefits of Flood and Coastal Risk Management: a Handbook of Assessment Techniques. Middlesex Univ. Press 89. doi:10.1596/978-0-8213-8050-5

Poussin, J.K., Bubeck, P., Aerts, J.C.J.H., Ward, P.J., 2012. Potential of semi-structural and non-structural adaptation strategies to reduce future flood risk: case study for the Meuse. Nat. Hazards Earth Syst. Sci. 12, 3455–3471. doi:10.5194/nhess-12-3455-2012

Proverbs, D.G., Soetanto, R., 2004. Flood Damaged Property; A Guide to Repair. Blackwell Publishing Ltd.

Refaeilzadeh, P., Tang, L., Liu., H., 2009. "Cross-Validation.," in: Encyclopedia of Database Systems. Springer US, pp. 532–538.

Regione Emilia Romagna, 2011. Coperture vettoriali dell'uso del suolo [WWW Document]. URL http://servizigis.regione.emilia-romagna.it/ctwmetadatiRER/metadatoISO.ejb?stato_IdMetadato=iOrg01iEnP1idMetadato76868 (accessed 7.1.15).

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

Scorzini, A.R., Frank, E., 2015. Flood damage curves: new insights from the 2010 flood in Veneto, Italy. J. Flood Risk Manag. n/a-n/a. doi:10.1111/jfr3.12163

Seifert, I., Kreibich, H., Merz, B., Thieken, A.H., 2010. Application and validation of FLEMOcs – a flood-loss estimation model for the commercial sector. Hydrol. Sci. J. 55, 1315–1324. doi:10.1080/02626667.2010.536440

Smith, D., 1994. Flood damage estimation. A review of urban stage-damage curves and loss function. Water SA 20, 231–238.

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 Resour. Res. 41, 1–16. doi:10.1029/2005WR004177

Thieken, A.H., Olschewski, A., Kreibich, H., Kobsch, S., Merz, B., 2008. Development and evaluation of FLEMOps – a new Flood Loss Estimation MOdel for the private sector. Flood Recover. Innov. Response, WIT Press 315–324.

Thieken, a. H., Ackermann, V., Elmer, F., Kreibich, H., Kuhlmann, B., Kunert, U., Maiwald, H., Merz, B., Müller, M., Piroth, K., Schwarz, J., Schwarze, R., Seifert, I., Seifert, J., 2009. Methods for the evaluation of direct and indirect flood losses, in: RIMAX Contributions

at the 4th International Symposium on Flood Defence (ISFD4). pp. 1–10.

Trigila, A., Iadanza, C., Bussettini, M., Lastoria, B., Barbano, A., 2015. Dissesto idrogeologico in Italia: pericolosità e indicatori di rischio. ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale.

UNISDR, 2004. Terminology. Basic terms of disaster risk reduction. United Nations Int. Strateg. Disaster Reduct. (UNISDR), Geneva 1–30.

Vacondio, R., Dal Palù, A., Mignosa, P., 2014. GPU-enhanced finite volume shallow water solver for fast flood simulations. Environ. Model. Softw. 57, 60–75. doi:10.1016/j.envsoft.2014.02.003

Ward, P.J., de Moel, H., Aerts, J.C.J.H., 2011. How are flood risk estimates affected by the choice of return-periods? Nat. Hazards Earth Syst. Sci. 11, 3181–3195. doi:10.5194/nhess-11-3181-2011

Zampetti, G., Ottaviani, F., Minutolo, A., 2012. I costi del rischio idrogeologico. Dossier Legambiente, Roma.

4 MAPPING SOCIO-ECONOMIC EXPOSURE FOR FLOOD RISK ASSESSMENT IN ITALY

4.1 Introduction

In an increasingly urbanized world, a reliable estimation of social, economic and environmental capitals which are located in hazard-prone areas is crucial for devising effective disaster risk reduction and climate change adaptation strategies. Maps of population and asset exposure are very important for helping planners in the design of risk mitigation policies (Bajat et al. 2011), but also for responding properly during emergencies (Moon and Farmer 2001). Risk assessment typically combines various dimensions (structural, social, economic, institutional and environmental) of vulnerability (Fuchs 2009; Kienberger et al. 2009; Fernandez et al. 2016). An integrated approach requires a quantitative evaluation of each individual component's characteristics (e.g. buildings, people, productive facilities) in order to determine their overall vulnerability (Garcia et al. 2016). This means that a number of different datasets must be collected and combined so that an integrated risk assessment can be carried out (Koks et al. 2015b). Since the quality of the data used to represent these indicators directly affects the outcomes of risk assessment (Thieken et al. 2006b), its improvement is directly relevant for risk management and spatial planning. However, collecting and harmonizing all the required datasets can be challenging. While hazards are typically modelled in a spatially explicit way (e.g. raster grids), exposure data such as population density is often available for administrative or survey units (Thieken et al. 2006b), where a single value homogeneously represents the entire unit's extent. This is not only an oversimplification of reality, but also the reason behind spatial mismatches when trying to combine and compare information projected on statistical units that are heterogeneous in their shapes and sizes (Zandbergen and Ignizio 2010). To overcome this problem, various areal interpolation methods have been tested for transferring spatial information from irregular to uniform units (e.g. regular square grid). The most straight-forward method is *areal weighting*, in which the variable from the source is spatially apportioned to the target layer, based on how much of the source area falls within the target area. A more advanced set of methods is referred to as *dasymetric mapping* (Wright 1936). In using this approach, ancillary data are employed to gain additional information about the spatial distribution of socio-economic variables within the source units. This downscaling process will result in

smaller units of homogeneity that more closely resemble the analysed phenomena, rather than arbitrary administrative boundaries (Mrozinski and Cromley 1999; Wu et al. 2005). Then, dasymetric zones can be generalised into a regular grid by using areal weighting. This step is necessary in order to reduce the size of the dataset, making it possible to run the risk assessment on very large areas (e.g. countries) while maintaining the advantages of the dasymetric approach and an acceptable level of detail. Population and buildings maps can ultimately serve as a proxy to geographically project correlated socio-economic variables, for example GDP (Eicher and Brewer 2001; Nordhaus et al. 2006).

We describe a high-resolution, multi-layer exposure grid for Italy, including information about exposure and social vulnerability that can be employed to perform risk analysis at the national or local scale. The dataset is built by means of a dasymetric approach based on official census records about population and buildings, and additional ancillary data related to land cover, buildings, demographics and GDP. The ensuing gridded exposure was used for the purpose of national climate change adaptation planning in Italy (Mysiak et al. 2018). Our downscaling approach is tested within one Italian Region against global state-of-the-art population grids to compare their performances and reliability. Then, our grid is combined with hazard scenarios produced at a European scale for six different return periods (Alfieri et al. 2014; Alfieri et al. 2015c) in order to give an example of its application for the purpose of risk assessment.

4.2 Study area

Because of its peninsular and mountainous morphology, Italy is susceptible to flood hazard and disaster risk. Some 8% of its territory and 10% of its population are reportedly located in areas flooded with a chance of 1 in 100-200 in a single year (ANCE/CRESME 2012b; ISPRA 2014; Trigila et al. 2015b). In absolute terms, Italy's annual expected damage under current climate is the largest in the European Union (Alfieri et al. 2016). The damage and recovery costs inflicted by major flood events over the period 2009-2012 amounted to about one million Euros per day (Zampetti et al. 2012). Flood risk is expected to increase due to changes in risk-driving factors: on the one side, growing anthropization of natural environments has increased the value of asset along the floodplains in the last 50 years, requiring the adoption of higher hazard-defence standards; on the other side, climate change affects the frequency and magnitude of extreme precipitation events (Rojas et al. 2013; Alfieri et al. 2015c; Alfieri et al. 2015b). We chose Emilia Romagna Region (RER), located in Northern Italy, to compare the performance of interpolation methods (figure 4.1, left). According to the Hydrological System Plan (PAI),

RER is the most flood-prone region in the country (ISPRA 2014) (figure 4.1, right), both in relative and absolute terms: 2.5 thousand km^2 and 10% of its population are exposed to a high flood probability (once every 20 to 50 years) and up to 10 thousand km^2 and 64% of the population to a medium probability (once in 100 of 200 years) (ISPRA 2014).

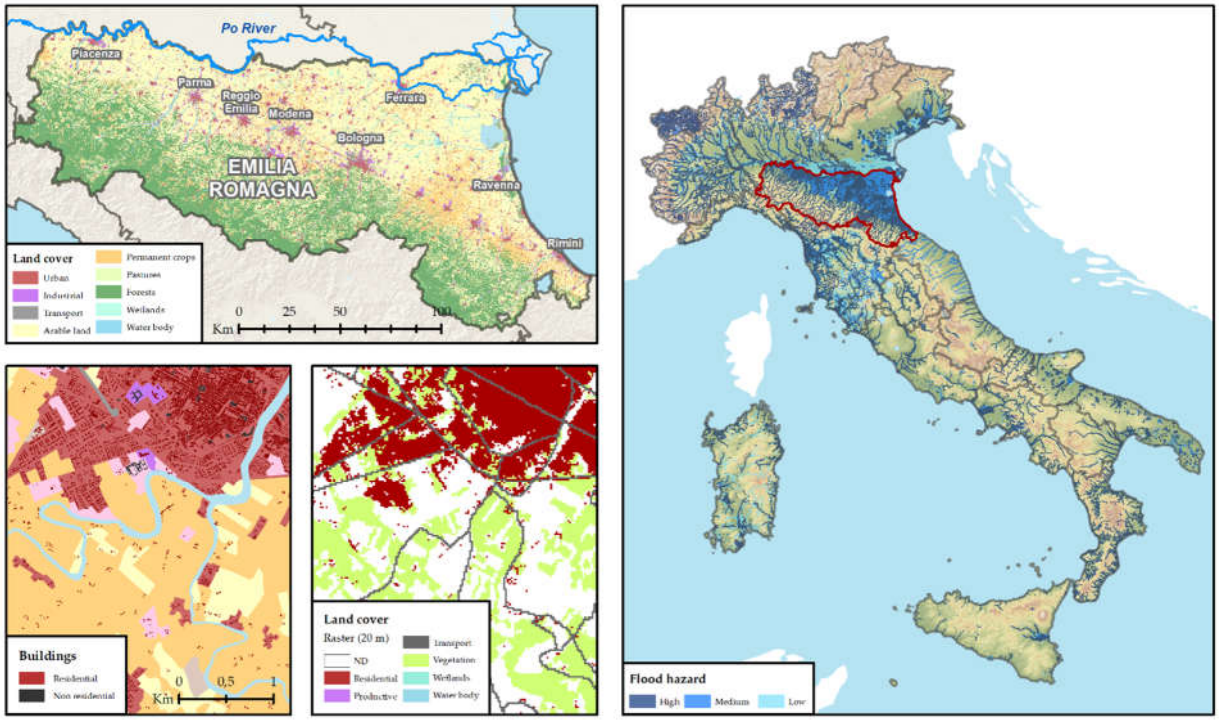


Figure 4.1. (left) Emilia-Romagna Region, available land cover datasets; (right) official definition of flood-prone hazard zones in Italy (FD 2007/60/EC).

4.3 Methodology

4.3.1 Data description

We used the most recent census data of the Italian Statistical Office (ISTAT) to obtain variables related to demographics (residential population, gender, age structure, education, employment, rate of foreigners), buildings (type, use, age, quality, number of floors) and dwellings (type of occupants, area). The size of a census tract is variable and ranges between a hundred square meters to hundreds of square kilometers, with an average size of ~750 square meters. Residential buildings account for 84% of the total stock and half of them are single-dwelling units. More than three-quarters of the dwellings included in the census are occupied by primary residents. Industrial activities, retail assets and offices account for 3%, 2.5% and 1.8% of buildings (ISTAT 2014). To define the value of assets we used real estate market values, available for ~76% of the stock, aggregated to sub-urban functional zones: center and semi-center, periphery, extra-urban or sub-urban (Agenzia delle Entrate 2014). Land cover data and

building perimeters are acquired as shapefiles from the digital Regional Technical Map (RER 2011), although in most Italian municipalities this information is comparable to that available from Open Street Map (OSM) (Geofabrik GmbH 2012). We also used the country-wide land cover dataset, made available by the Copernicus Land Monitoring Services (Munafò et al. 2015), which includes imperviousness, forests and marshes as a raster at a 20-meter resolution. This information is combined with the Corine Land Cover dataset (2012) and the transport network map (Ministry of Environment, 2008) to mask non-residential built-up areas. A similar approach was used by other authors to refine land cover characterization (Steinnocher et al. 2011; Batista e Silva et al. 2013). The analysis of exposure is performed by means of flood hazard simulations produced by LISFLOOD at a European-scale (Dottori et al. 2016a) for six flood hazard scenarios of different intensity and probability (return period 10, 20, 50, 100, 200 and 500 years). The same data were previously used for assessing flood hazard exposure and risk across Europe (Alfieri et al. 2015c).

4.3.2 The dasymetric approach

Dasymetric mapping uses one or more geographical proxies to downscale the spatial information usually provided at the scale of administrative units. As for population, housing density is an ideal covariate, as residential floor area is closely correlated (Naroll 1962; Brown 1987; Wu et al. 2017). Roads, industrial districts and other non-habitable artificial areas are identified and masked out, and population is distributed over the remaining built-up areas (*binary method*) (Eicher and Brewer 2001). The resulting *dasymetric zones* are nested within the census tracts and the land cover polygons (Zandbergen and Ignizio 2010), as exemplified in Figure 4.2.

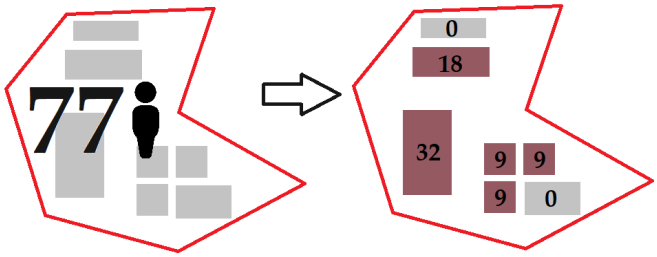


Figure 4.2. Example of dasymetric distribution of population count from the census tract unit to individual residential buildings based on their area extent.

Dasymetric zones are the source for areal interpolation to different aggregation units, such as regular grid cells. To this end, the dasymetric zones are overlaid with the target areas (cells) and population is spatially apportioned based on areal weighting. This approach assumes population density to be uniform within each dasymetric zone, but because dasymetric zones

are smaller than the original census tract area, the result is more accurate than a simple areal weighting (Wu et al. 2005; Maantay et al. 2007). In fact, interpolation from small source areas to large target areas entails smaller errors than the other way round (Langford 2006; Zandbergen and Ignizio 2010). Several spatial proxies can be employed for this purpose, but imperviousness and land cover data have proven to be more suitable proxies than other ancillary data, such as road density or night-time lights (Xie 1996; Ong and Houston 2003); still, if the definition of census units and residential areas is too coarse, small scattered settlements may not be captured correctly. This may lead to biases in low-density areas (Steinnocher et al. 2011). This issue can be addressed by using different urban density classes based on sample areas (Langford et al. 1991; Goodchild et al. 1993; Bielecka 2005) and by using an average number of dwelling units and dwellers per unit within sub-areas (Wu et al. 2005). However, all methods relying on a 2-D land cover definition lack the information about the height of buildings; the same building footprint area can then be associated with a different number of floors, resulting in different estimates of population density. For this reason, the linear relationship between population and footprint area entails an error and leads to an underestimation of the residential population in taller buildings within the same census unit. Most studies apply a corrective factor (pycnophylactic constraint) for the total population to coincide with the source data (Thieken et al. 2006b; Gallego 2010).

Several global population grids are available, including GeoStat (EFGS 2011), LandScan (Bhaduri et al. 2002) and the GWP (Balk and Yetman 2004). The state-of-the-art of global population grids today is the Global Human Settlement Layer (GHSL), which combines remote sensing, ancillary census data and spatial modelling into a global dataset, with a resolution of 250 meters (Ouzounis et al. 2013; Pesaresi et al. 2014; Freire et al. 2016). Although there is no single best dasymetric method, multiple ancillary data are expected to improve the results of the downscaling compared to the single data-source method. The precision and accuracy of the input data likely plays a more important role than the method used to combine them (Martin et al. 2000). To test this hypothesis in our case study, we have applied and compared three methods to transfer the information from the recent population and dwellings census to the new exposure grid:

A) simple interpolation based on areal weighting;
B) dasymetric method based on imperviousness data;
C) dasymetric method based on detailed land cover and building data.

In method A, population density (number of inhabitants per square meter) is directly transferred from each census unit into a regular square grid with cell size 250x250 meters, by means of the share of interpolated cell area as a weight. In method B and C, population density is calculated within each census tract *i* over residential land cover. The difference between B and C lies in the quality and type of land cover data employed to identify residential areas; method B uses imperviousness data from Copernicus at a 20 m resolution, while method C relies on polygon geometries of land cover and buildings. Population (*P*) is assumed proportional to housing density (*Hd*) and allocated over residential land cover classes (*R*) only. This assumption is supported by the good correlation found between normalized dwelling area and normalized population within each census tract (R^2=0.95). Hd_i is calculated as the ratio between tract population and residential area as identified within each tract boundary. Areal interpolation is used to transfer the information from the source layers to the dasymetric zones of residential land cover. Then, these are aggregated into regular grid cells by using area weights (A_w) calculated as the share of residential area in dasymetric zones over the total area of the cell *j* (Figure 4.3 and Equation 4.1).

$$P_j = \sum_i \left(A_{W\,ij} \times Hd_i \right) \times R_j \qquad \text{[Equation 4.1]}$$

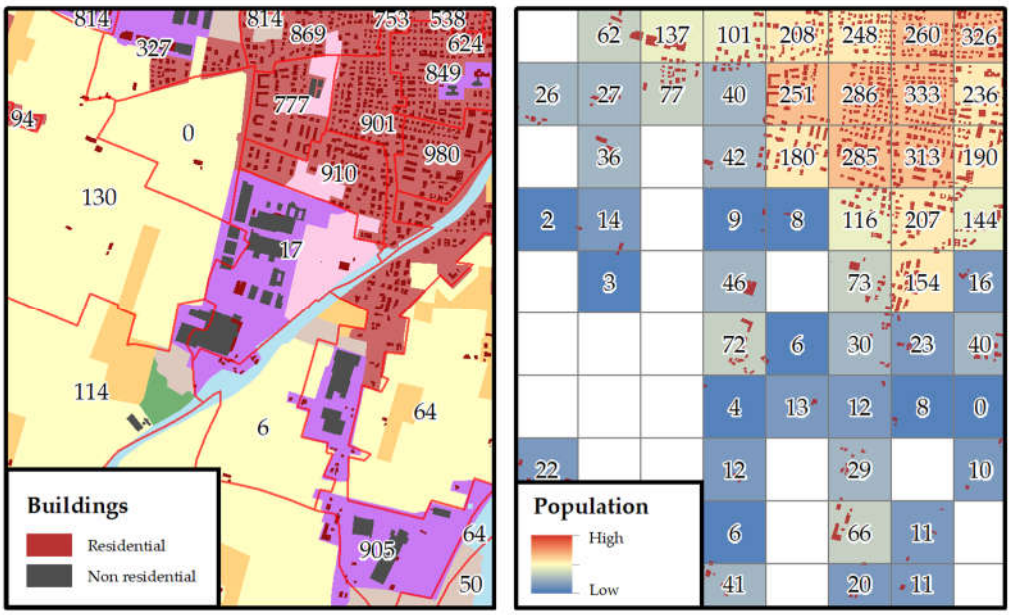


Figure 4.3. Transferring of population information from the census units to a regular square grid (cell size of 250 meters) by using residential buildings as dasymetric units.

4.3.3 Mapping economic activity

The quality of the population grid has a direct influence on the derived grid of Gross Domestic Product (GDP). The dataset produced by G-Econ for example (Nordhaus et al. 2006) is drawn

from a global population grid with a resolution of about 100 km; the population density in each grid cell is multiplied by the Gross Regional Product per capita (GRP), available at NUTS3 level[6] (EUROSTAT 2013). A much finer global GRP dataset was produced by the UN for 2010 using a 1 km resolution (UNEP 2012) based on LandScan population (Bright et al. 2006). This method distinguishes between rural and urban population, assuming the latter to have a higher GRP per capita. Other researchers (Huang et al. 2014) projected GDP as disaggregated measures per macro-sectors (primary, secondary and tertiary sector production) on the basis of land cover. Validation against empirical data showed a small degree of uncertainty for this method. For Italy, disaggregated GDP[7] measures are available for Labour Market Areas (LMA[8]) for the three main macro-economic branches: agriculture, industry and services. Within each branch, Gross Labour Market Value (GLMV) is divided by the respective productive area as identified by land cover classes, and proportionally distributed over them. Namely, GLMV density is calculated for each LMA: 1) over the total LMA area for agriculture product, excluding steep slopes and quotas over 600 meters; 2) over the total industrial area for industry product; 3) over the total population for services product. GLMV density per sector is finally distributed on each cell by using as weight, respectively, LMA area, industry LC area and population density. We assume here an average value per square meter for agriculture and industry products within each LMA. Contrarily, services are scattered over the territory; we assume that their value is proportional to the number of people served, so population is used as a proxy for mapping the distribution of their value.

4.3.4 Social Vulnerability Index

Many risk assessment studies focus on the vulnerability of physical assets by accounting for potential damage to structures and, occasionally, their content (Jongman et al. 2012b; Jongman et al. 2014; Koks et al. 2014a; Roder et al. 2017). Social variables are often neglected, although they can help for understanding how the expected burden is distributed. The capacity to cope with disaster damage and losses has been previously analysed by using a number of socio-economic characteristics such as wealth, age, ethnicity, and quality of dwellings, which are

6 Classification of Territorial Units for Statistics developed and regulated by the European Union. NUTS3 correspond to the administrative level “Region” in most EU countries.

7 The measure provided is in fact the Gross Value Added (GVA), which is GDP minus taxes plus subsidies on products. For simplicity, it will be referred as GDP.

8 Labour Market Areas refers to ISTAT’s Local Labour Systems (in Italian “Sistemi Locali di Lavoro”, SLL).

employed to define a vulnerability index (Cutter et al. 2003; Fekete 2009; Willis et al. 2014; Zhou et al. 2014; Koks et al. 2015b; Bakkensen et al. 2016; Fernandez et al. 2016; Frigerio et al. 2016; Roder et al. 2017). The choice of variables used in such studies is typically influenced by the availability of spatially explicit data (Balica et al. 2009). We select and combine indicators which are linked to exposure and vulnerability of households from most recent censuses (ISTAT 2011b) into an aggregated index of social vulnerability to flood hazard (SVI). It includes data on population (P), buildings (B) and dwellings (D). Four metrics related to P are selected among those that describe socio-economic status, as this is assumed to be a main determinant of social vulnerability (Cutter 1996; Fekete 2009). First, the Total Age-Dependency rate expresses the share of the dependent (younger than 15 and older than 65 years) over the non-dependent population (Hewitt 1997; Koks et al. 2015b). Second, the unemployment rate is used as a proxy of labour economic status. Third, the rate of foreigners over total population is assumed to be an expression of accessibility and comprehension of emergency indications. Lastly, the rate of population commuting daily outside their municipality reveals if the owner can promptly activate hazard mitigation when the event strikes. Information related to residential buildings (B) includes the number of floors and the preservation state, or quality, of the structure. During floods, ground floors are those mostly affected by water; therefore, a high share of 1-floor buildings over the total increases vulnerability. Similarly, a high rate of buildings in poor preservation status is used as a proxy for higher chances of damage, as well as an indication of limited access to finance. Finally, statistics about dwellings include number, total area and vacancy rate. Unoccupied dwellings are less likely to be protected from hazards by private risk-reduction actions.

Most frequently, such indexes are aggregated linearly by addition assigning equal weights to all indicators (Cutter et al. 2003; Koks et al. 2015b). With this approach, the lower performance of one indicator can be compensated by some higher performance of another indicator. In some cases a non-compensatory multi-criteria approach may be more appropriate (OECD 2008); we use a fuzzy inference, non-compensatory aggregation that accounts for uncertainty as an intrinsic component of the model (Martins et al. 2012). Fuzzy logic makes it possible to deal with complex or ambiguous concepts, which are not (easily) quantifiable but blurry, or “fuzzy” (Phillis and Andriantiatsaholiniaina 2001; Andriantiatsaholiniaina et al. 2004; Hefny et al. 2013). Indicators are translated into linguistic values (e.g. high, medium or low) via membership function (fuzzification), and then decision rules are applied to combine them. Rules reflect the importance assigned to each indicator. The hierarchical structure of our

framework is shown in Figure 4.4. The intrinsic indicators related to the three main components P, B and D are multiplied by their extrinsic count (number of P, B and D elements) and classified into three classes (low, medium, high) using a fuzzy membership function. In doing so, we avoid abrupt changes and can discriminate between indicators of primary and secondary importance. We assume that, for example, a high rate of commuters does not affect the vulnerability of households as much as a high unemployment rate could. Primary indicators are distributed into three vulnerability classes (high, medium and low), while secondary indicators are split into two classes (high and low). The criteria for the aggregation of each component of the SVI are defined by a set of decision rules. The choice of the indicators, criteria and rules is one out of the many possible, and inevitably reflects the expert knowledge of the authors; our purpose is to provide an example of application of the available data, but decision makers may tailor the index on the basis of largely-shared preferences and expert opinions. The index is calculated for each census unit and then distributed on the grid by using the dasymetric approach in relation to residential density, similarly to population.

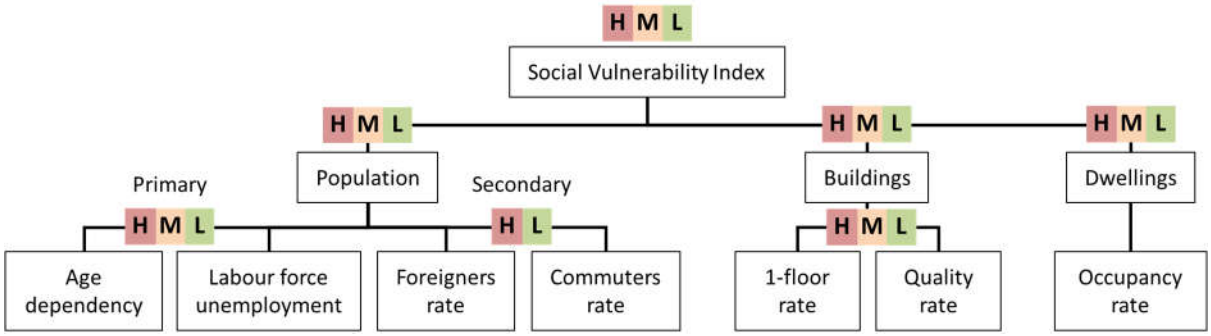


Figure 4.4. Hierarchical representation of criteria and vulnerability classes employed to produce the fuzzy aggregation of the SVI.

4.4 Results

4.4.1 Population

The outputs from the three downscaling methods A, B and C are all consistent with their source data when comparing total population within provinces (NUTS3). We use a multi-scale approach (Ouzounis et al. 2013) to evaluate the quality of these outputs (figure 4.5, 4.6 and 4.7). Simple interpolation (method A) largely underestimates population density in residential areas, since it distributes population uniformly from either small/dense and large/scattered sections. This method is excluded from further analysis. The GHSL grid overcomes this issue by using remote sensing data to project population only within built-up areas. Both methods B and C provide a better output by masking non-residential built-up areas and accounting for

sparse residential buildings in rural areas. They both perform well in this region, underestimating total population by a small share (1.2% and 0.5% respectively) compared to the source census data. The error in method B is due to the coarser definition of the land cover layer, which does not identify residential areas within some inhabited census units. This bias is larger in some regions, such as Basilicata (4,3%), Calabria (3.6%) and Umbria (3.5%). Within the RER, we measured the error of relative population distribution for both methods B and GHSL (2015) using regression analysis and assuming method C as the best fit to represent the baseline (figure 4.5).

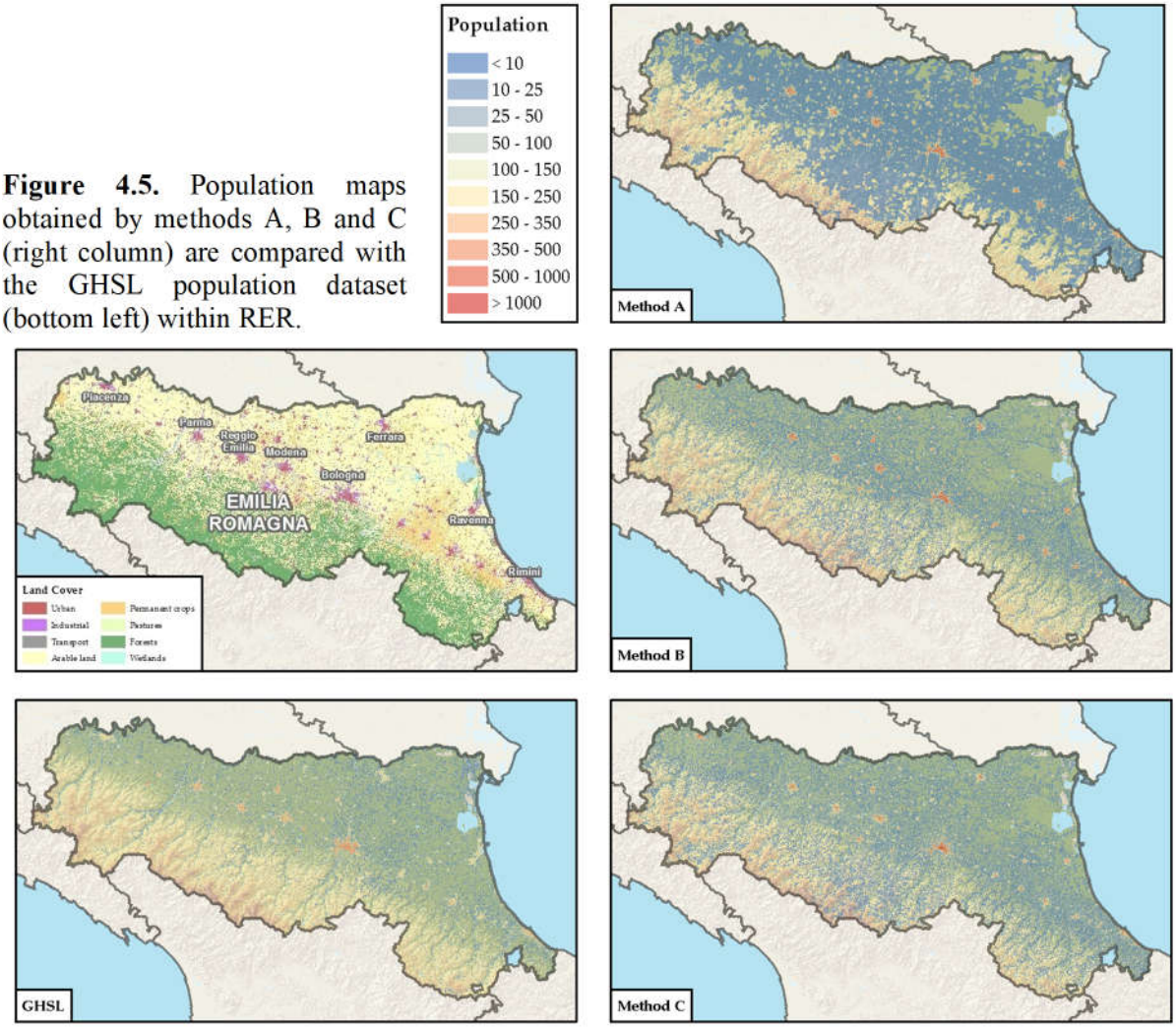


Figure 4.5. Population maps obtained by methods A, B and C (right column) are compared with the GHSL population dataset (bottom left) within RER.

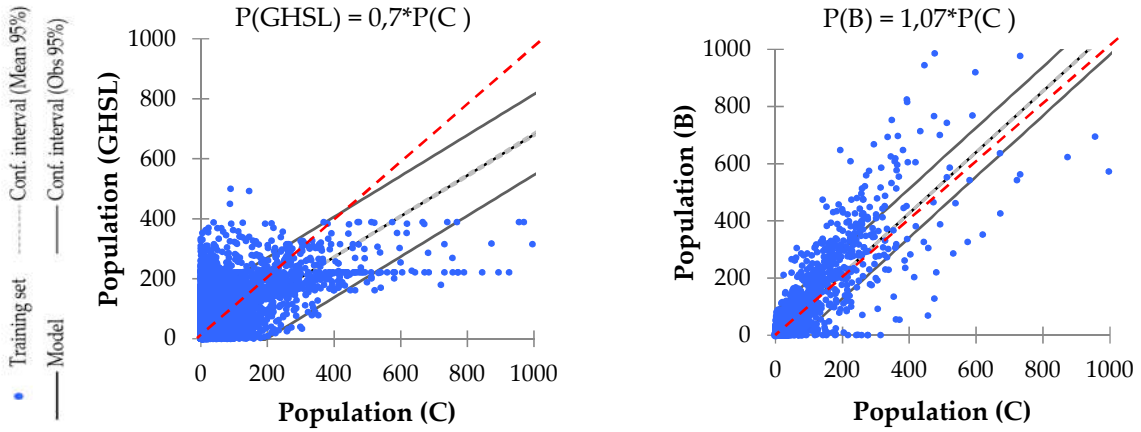


Figure 4.6. Scatterplot for regression analysis comparing on the Y-axis population cells from GHSL (left) and from method B (right) to population from method C, on the X-axis. The red dashed line represents 1:1. The test covers the RER area and assumes set C as the most reliable. The R^2 coefficient for set B is equal to 0.7. The GHSL set has a R^2 coefficient of 0.2 against set C. Both models are significant within 95%.

Overall, dasymetric models B and C perform similarly (figure 4.6, right), projecting population only where residential land cover is found. The GHSL grid performs decently (figure 4.5, bottom left), showing a comparable relative error for the whole country (±1.3%); however, the scatterplot shows a much larger dispersion and a lower correlation score (figure 4.6, left). This seems to confirm that the combination of multiple ancillary data sources and the use of smaller source units have a major importance for the quality of the dasymetric mapping output. By narrowing the comparison of method B, C and the GHSL to the urban scale (figure 4.7) we can find overall a close agreement between the three datasets.

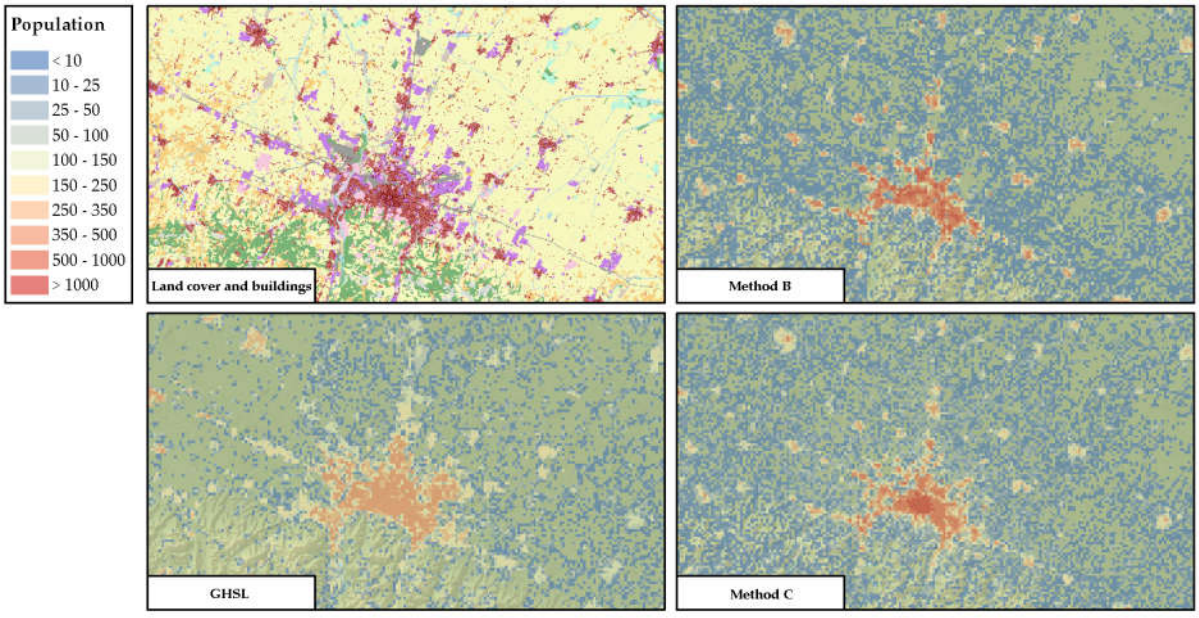


Figure 4.7. Comparison at the urban scale (city of Bologna) between methods.

The GHSL correctly identifies the main urban settlements and some scattered urban areas, while ignoring smaller settlements and sparse dwellings. Moreover, it does not discriminate residential areas from non-residential impervious surfaces such as industrial districts and transport infrastructures, thus distributing population on an area much larger than the one effectively occupied by dwellings. Dasymetric methods B and C take advantage of more detailed information from the census and project it exclusively within residential units. As a further example of the importance of having a reliable description of land use, see figure 4.8: the top row shows the ancillary data employed to distribute population density, while the bottom row shows the resulting dasymetric maps according to the three methods. The interport in the center of the map is a large transport infrastructure which includes few dwellings along its extent, for a total of 256 residents. The GHSL populates the whole area with 4,467 residents (bottom left). Method B identifies the area as generic non-residential, and projects 245 residents only within the few cells including residential units (bottom mid). Method C correctly identifies the interport as transport infrastructure, and projects a total of 404 residents within the few residential building surfaces (bottom right).

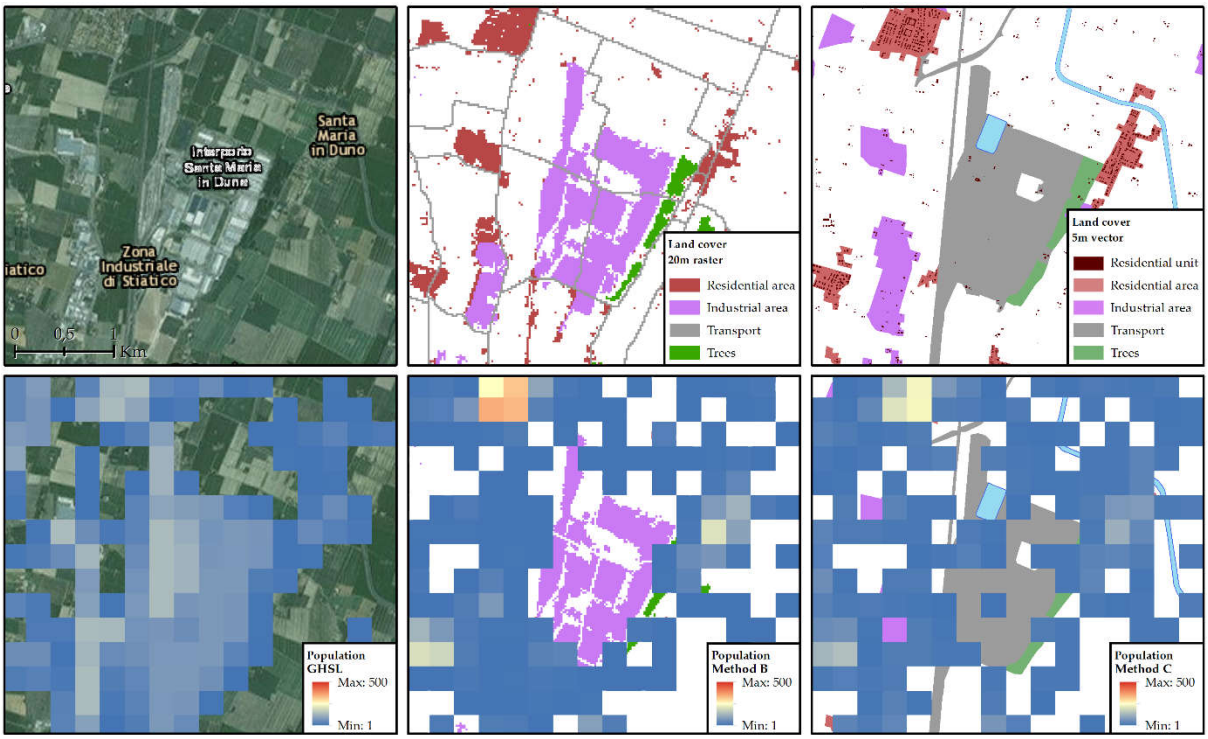


Figure 4.8. Sample from Bologna province (RER) including industrial zones, transport network infrastructures and sparse residential land cover. Comparison between the land cover information (top row) and the resulting population grids (bottom row) according to GHSL (left), method B (middle) and method C (right).

The error metrics from the comparison of the three approaches within the RER are summarised in table 4.1, by assuming that method C is the most reliable for baseline comparison. The linear regression is calculated by selecting cells where at least one of the two datasets is populated. Dataset B accounts for about twice the populated area compared to the GHSL and it fits much better to dataset C (0.72). The Mean Absolute error is lower for B, however the Mean Absolute Percentage Error shows that large residuals exist also in this case. This is caused by the mismatch between the two land cover datasets, which in some cases can lead to large differences in the cell-by-cell distribution of population (figure 4.9).

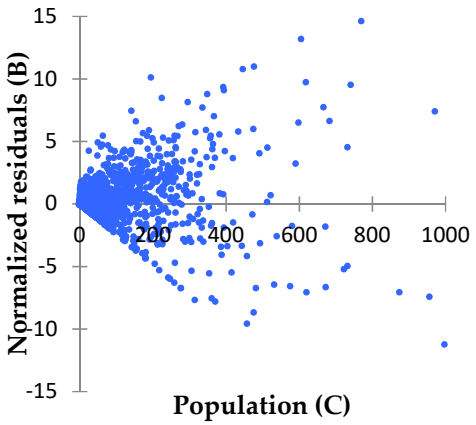


Figure 4.9. Scatterplot of normalised residuals from linear regression of dataset B on C.

Dataset	Populated area	Population	Relative error	Populated cells	MAE	RMSE	R^2
(C)	44%	4,321,494	-0.5%	Regression against C dataset			
(B)	39%	4,287,982	-1.2%	168,093	0.19	43.6	0.72
(GHSL)	22%	4,452,335	+2.5%	80,681	14.0	68.5	0.24

Table 4.1. Error metrics within Emilia-Romagna comparing method B, C and GHSL datasets. Dataset C is selected as baseline for the estimation of error in set B and GHSL.

4.4.2 Economic value

The GDP is projected from the LMA data by using both population and land cover/use as a proxy (figure 4.10). The resulting dataset is compared to the GDP dataset produced by the UN for 2010. To do so, the UN grid is first re-projected to the same resolution and the GDP is converted from 2000USD constant value to 2005EUR (Figure 4.10, left). A comparison of the two grids at the scale of LMA units shows a negligible difference (R^2=0.92), while on a cell-by-cell basis the coincidence is rather poor (R^2=0.20); this partially depends on the re-projection and resampling procedure. Overall, the UN dataset depicts GDP distribution adequately for the purpose of country-wide studies; our method offers a better alternative where sub-regional, regional and multi-regional studies are concerned. There is one important limitation inherent to this approach: GDP is measured according to taxes paid by people and enterprises; however, the headquarters of large economic activities individually account for a production value that

is in fact generated over a larger territory. Therefore, some urban units in major cities such Milan and Rome show extremely large GDP values, since large corporations are legally represented there. This can lead to a misestimation of relative exposure.

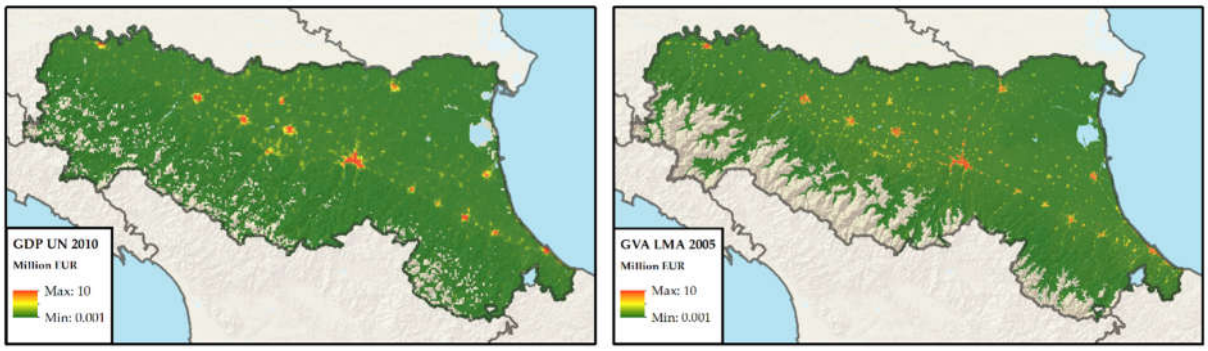


Figure 4.10. Comparison between GDP datasets: the UN grid for 2010 (left), although coarser, is coherent with the dataset we produced using dasymetric method B and LMA values of 2005 (right).

Table 4.2 shows the regional share of GDP per each macro-sector in relation to population and residential area. The three macro-sectors differ consistently in terms of relative contribution to GDP: agriculture accounts for about 2%, industry for 27% and services 71% of the total. That means that more than two thirds of GDP value are distributed over residential areas.

Region	Population	Residential area	GDP			
			Agricult.	Industry	Services	Total
Piedmont	6.8%	6.5%	5.7%	9.1%	7.8%	8.1%
Aosta Valley	0.2%	0.2%	0.2%	0.2%	0.3%	0.2%
Lombardy	16.5%	17.1%	10.6%	26.2%	19.3%	21.0%
Trentino South-Tyrol	1.8%	1.7%	3.4%	1.9%	2.1%	2.1%
Veneto	8.3%	8.9%	8.6%	12.3%	8.5%	9.5%
Friuli	2.1%	2.5%	1.7%	2.3%	2.4%	2.3%
Liguria	2.7%	2.0%	2.2%	1.8%	3.2%	2.8%
Emilia Romagna	7.4%	8.4%	9.9%	10.8%	7.9%	8.7%
Tuscany	6.2%	5.9%	6.1%	6.8%	6.7%	6.7%
Umbria	1.5%	1.5%	1.4%	1.4%	1.4%	1.4%
Marches	2.6%	2.4%	2.3%	3.1%	2.3%	2.5%
Lazio	9.3%	7.8%	5.9%	6.1%	13.1%	11.0%
Abruzzo	2.2%	2.9%	2.4%	2.0%	1.7%	1.8%
Molise	0.5%	0.6%	0.7%	0.4%	0.4%	0.4%
Campania	9.8%	7.7%	8.1%	4.4%	6.9%	6.3%
Apulia	6.9%	6.2%	9.2%	3.9%	4.6%	4.5%
Basilicata	0.9%	1.0%	1.8%	0.7%	0.7%	0.7%
Calabria	3.2%	2.9%	5.6%	1.4%	2.4%	2.2%
Sicily	8.5%	9.7%	10.7%	3.6%	6.1%	5.5%
Sardinia	2.8%	4.2%	3.6%	1.6%	2.3%	2.2%

Table 4.2. Distribution of population and productive macro-sectors across Italian Regions.

4.4.3 Exposure to flood scenarios

The country-wide grid dataset produced by using method B is employed to estimate social and economic exposure to hazard according to six different hazard probability scenarios (Dottori et al. 2016a). The non-compensatory version of the SVI is chosen to identify the exposed areas with higher vulnerability. The change in the amount of population and GDP exposure according to the six RP scenarios is plotted in figure 4.11, left. Exposure grows almost proportionally with the increase of magnitude. On the right, exposure is classified according to ten water-depth stages. Much of the exposure is related to a water depth between 0.1 and 1.5 meters, although the relative depth of a cell increases with the event magnitude.

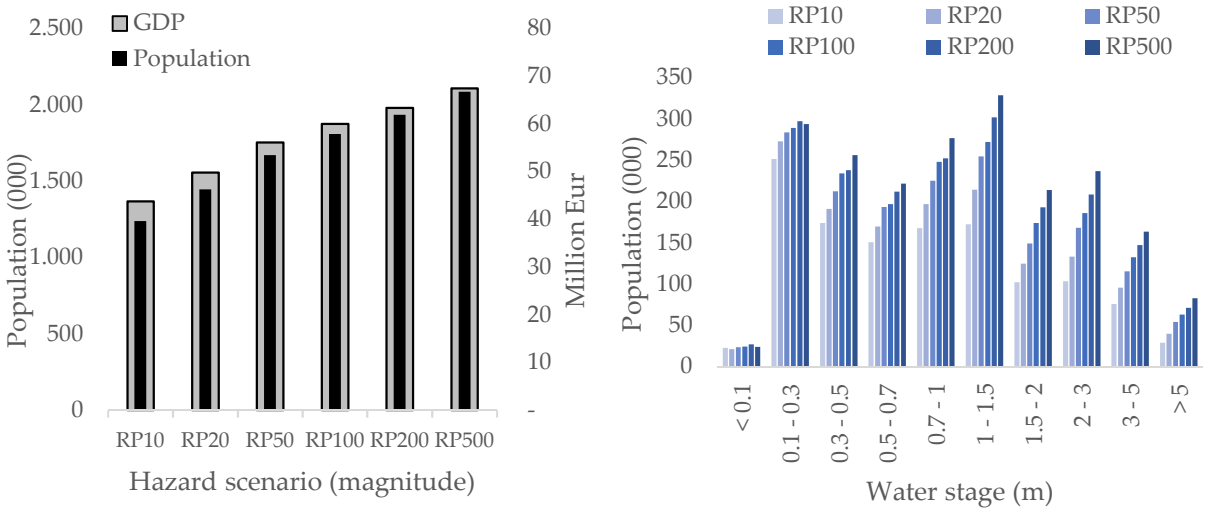


Figure 4.11. (left) Total Population and GDP exposed to flood hazard for six Return Period scenarios; (right) exposure of Population and GDP cells to ten stages of floodwater depth from each RP scenario.

Table 4.3 shows population and GDP exposure as a share of their total. Exposure increases almost linearly for both indicators with the growth of hazard magnitude (extent and depth). Population exposure ranges from 2.1% to 3.6% of the total. Among GDP sectors, agriculture is the one that is expected to be most exposed in relative terms, from 6.6% to 9.2% of its production value. However, agriculture production accounts for just about 2% of the Italian economy, thus in absolute terms its exposure ranges from 0.1 to 0.2% of total GDP. Industry, being mostly scattered over rural land and peripheral areas, shows similar relative shares of exposure, but its weight over the GDP is much more important (27%), thus exposure of industrial production ranges from 1.7% to 2.4% of total GDP. Services exposure is concentrated in urban areas and is the lowest in relative terms, growing from 2.3 to 3.8%; however, because services contribute for about 71% of Italian economy, their exposure accounts for 1.6% to 2.7% of the domestic product.

Scenario	Population	GDP Agriculture		GDP Industry		GDP Services		GDP Total
		Sector %	Total %	Sector %	Total %	Sector %	Total %	
RP10	2.1%	6.6%	0.1%	6.2%	1.7%	2.3%	1.6%	3.4%
RP20	2.5%	7.4%	0.2%	7.0%	1.9%	2.7%	1.9%	3.9%
RP50	2.9%	8.1%	0.2%	7.7%	2.1%	3.1%	2.2%	4.4%
RP100	3.1%	8.5%	0.2%	8.1%	2.2%	3.3%	2.4%	4.7%
RP200	3.3%	8.8%	0.2%	8.5%	2.3%	3.5%	2.5%	5.0%
RP500	3.6%	9.2%	0.2%	8.9%	2.4%	3.8%	2.7%	5.3%

Table 4.3. Shares of exposed population and GDP according to six return period scenarios. GDP exposure is expressed relatively to sector production and to GDP total.

Figure 4.12 displays the exposed cells for population and GDP in relation to scenario RP 100, which will be referred to as "medium frequency". Labels display the share of exposure in relation to regional totals. Results show how the northern regions are the most exposed to flood hazard, both in terms of population and GDP. Tuscany shows the highest share of population exposure within the Florence metropolitan area. In its history, Florence has been flooded several times by the Arno, with the last tragic event of 1966 causing 35 deaths and millions in losses. In terms of GDP exposure, the largest amount is located in Veneto; a large part of the productive asset and residential buildings in this Region is spread over floodplains that are exposed to hazard triggered by the rivers flowing from the Eastern Alps (Adige, Piave and Brenta-Bacchiglione catchments) involving the main towns of Vicenza, Padova and Verona already from the RP10 scenario. In 2010, central Veneto was hit by massive floods involving more than 260 municipalities and causing losses for about 500 million Eur. The lower section of the Po River, shared by Veneto and the RER, is the area where the largest relative amount of agricultural land is exposed to flood hazard. On the northern side of the Po river, the rural Province of Rovigo (Polesine) was hit in 1951 by the largest flood event in recent history for Italy, causing around 100 deaths and 200 million Eur of damage, and causing people and investments to move in other Provinces. Table 4.4 summarises the results of the share of exposure at the regional scale.

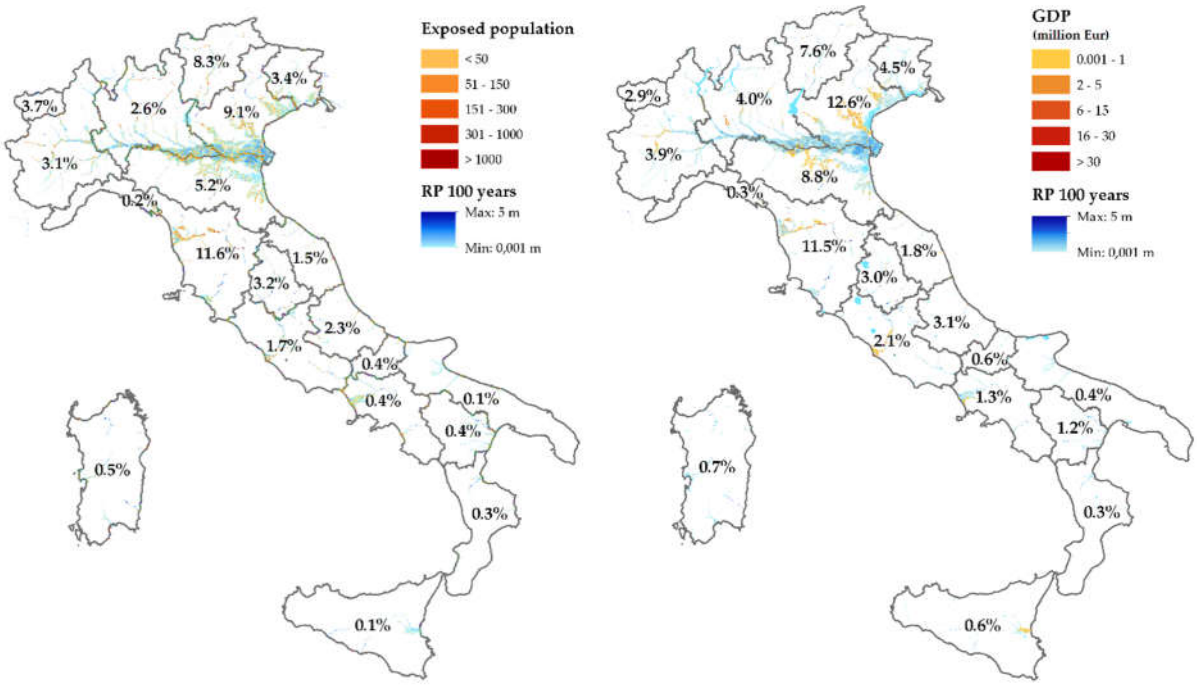


Figure 4.12. Exposed population (left) and GDP (right) according to scenario RP100 (medium frequency). Percentage value is relative to the regional total.

Region	Population		GDP							
			Agriculture		Industry		Services		Total	
	REG%	IT%	REG%	IT%	REG%	IT%	REG%	IT%	REG%	IT%
Piedmont	3.1%	0.2%	6.1%	0.3%	6.5%	0.6%	2.8%	0.2%	3.9%	0.3%
Aosta Valley	3.7%	0.0%	0.8%	0.0%	0.8%	0.0%	3.7%	0.0%	2.9%	0.0%
Lombardy	2.6%	0.4%	18.8%	2.0%	7.5%	2.0%	1.9%	0.4%	4.0%	0.8%
Trentino South-Tyrol	8.3%	0.1%	3.6%	0.1%	3.5%	0.1%	9.1%	0.2%	7.6%	0.2%
Veneto	9.1%	0.8%	26.7%	2.3%	18.1%	2.2%	9.2%	0.8%	12.6%	1.2%
Friuli	3.4%	0.1%	10.7%	0.2%	9.0%	0.2%	2.7%	0.1%	4.5%	0.1%
Liguria	0.2%	0.0%	0.7%	0.0%	0.7%	0.0%	0.2%	0.0%	0.3%	0.0%
Emilia Romagna	5.2%	0.4%	21.2%	2.1%	16.8%	1.8%	4.3%	0.3%	8.8%	0.8%
Tuscany	11.6%	0.7%	4.1%	0.2%	6.7%	0.5%	13.6%	0.9%	11.5%	0.8%
Umbria	3.2%	0.0%	2.6%	0.0%	2.9%	0.0%	3.1%	0.0%	3.0%	0.0%
Marches	1.5%	0.0%	1.4%	0.0%	2.0%	0.1%	1.8%	0.0%	1.8%	0.0%
Lazio	1.7%	0.2%	2.0%	0.1%	3.9%	0.2%	1.8%	0.2%	2.1%	0.2%
Abruzzo	2.3%	0.0%	1.5%	0.0%	1.7%	0.0%	3.7%	0.1%	3.1%	0.1%
Molise	0.4%	0.0%	1.2%	0.0%	1.4%	0.0%	0.4%	0.0%	0.6%	0.0%
Campania	0.4%	0.0%	5.8%	0.5%	4.6%	0.2%	0.4%	0.0%	1.3%	0.1%
Apulia	0.1%	0.0%	1.7%	0.2%	1.1%	0.0%	0.1%	0.0%	0.4%	0.0%
Basilicata	0.4%	0.0%	3.9%	0.1%	2.5%	0.0%	0.4%	0.0%	1.2%	0.0%
Calabria	0.3%	0.0%	1.0%	0.1%	0.7%	0.0%	0.2%	0.0%	0.3%	0.0%
Sicily	0.1%	0.0%	1.2%	0.1%	3.0%	0.1%	0.1%	0.0%	0.6%	0.0%
Sardinia	0.5%	0.0%	1.6%	0.1%	1.9%	0.0%	0.4%	0.0%	0.7%	0.0%
ITALY		3.1%		8.5%		8.1%		3.3%		4.7%

Table 4.4. Percentage of exposed population and GDP per macro-sectors over total regional (left column) and national (right column) amount according to a flood scenario defined by a 100 years Return Period.

The analysis of the SVI highlights the cells with the highest social vulnerability scores. For comparison, the SVI is shown in the form of three discrete classes (low, medium and high). Figure 4.13 displays the SVI for scenario RP100 by focusing on the lower section of the Po river basin, where most of the exposure is located. SVI values are higher within small rural settlements located along the river network, and peripheral areas of larger towns such as Padua or Verona. The Po delta region poses hazards to small coastal settlements along the freeway from Venice to Ravenna, many of which are characterized by below-average wealth, and higher age dependency and unemployment rates. Depressed areas where SVI has high scores are expected to be more vulnerable in the aftermath of a disaster and to have less capacity to cope with restoration costs.

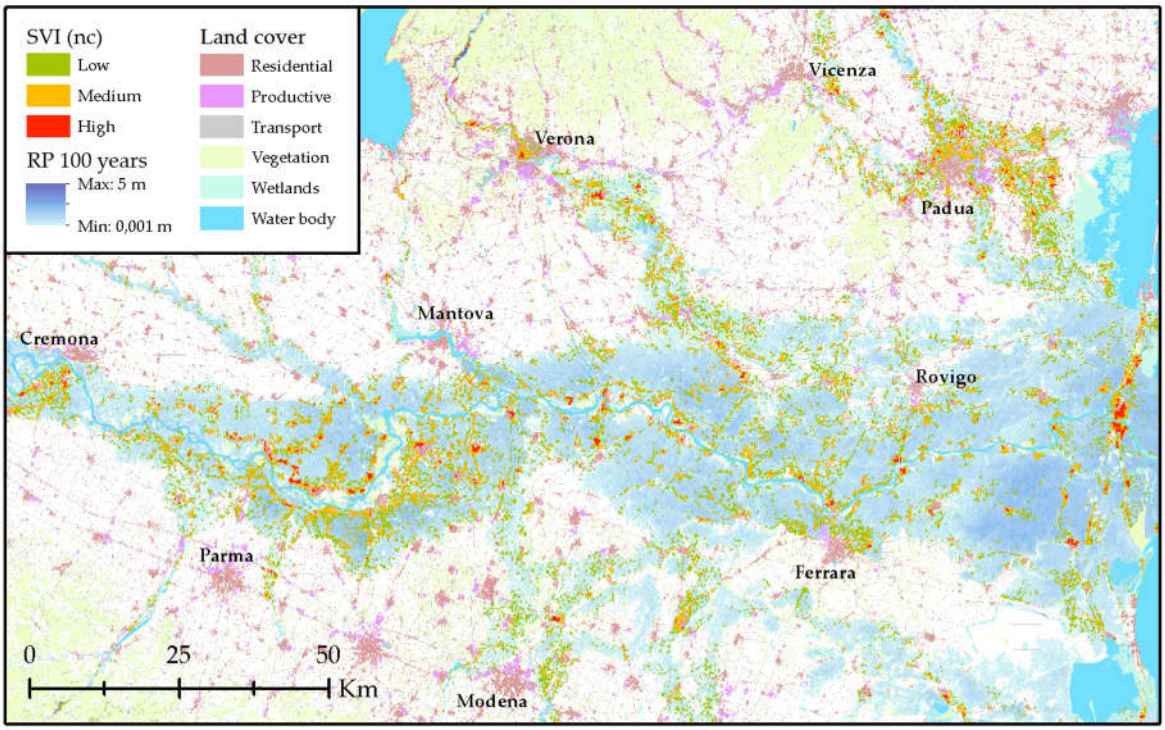


Figure 4.13. SVI for exposed population according to scenario RP100 (medium frequency).

4.5 Conclusion

In this paper we describe a method for producing a high-resolution exposure and vulnerability grid to be used for the purpose of flood risk assessment. The new grid dataset makes it possible to account for exposure and social vulnerability much better than the original unprocessed, non-homogeneous data. A similar approach can be followed for other hazard phenomena, provided that exposure and vulnerability indicators are chosen according to the specific framework of the risk analysis. Our dataset is structured in the form of a regular grid, with a resolution of 250 meters, which includes information about physical asset, population, domestic product and

social vulnerability. Information about demography, buildings, dwellings and GDP macro-sectors is collected from disaggregated sources and then projected on a grid by means of a dasymetric approach. The performance of the dasymetric method employed to build the exposure and vulnerability grid at a national scale (method B) is tested on one specific Region (the RER). The uncertainty of this approach depends in larger part on the detail of the source data providing the statistical information and the quality of the spatial proxy employed to project the socio-economic variables. The size of the statistical unit is the most important factor, the second being the quality of the georeferenced information about land cover and buildings. When a less detailed land cover classification is used, the dasymetric method performs well but has larger uncertainty, confirming the findings of Thieken et al. (2006). For analysis carried out at NUTS2 or NUTS3 scale, method B and C are comparable. The uncertainty of method B tends to grow when shifting to the municipal scale, although unevenly. Additional uncertainty is added by the discrepancy of scale and time references between the source ancillary data, and the assumptions made to proxy the indicators. In comparison, globally-produced population and GDP maps based on similar dasymetric approaches carry a much larger error rate due to coarser basedata, thus their use for spatial analysis within the national scale should be held as an option only when lacking country-specific sources.

The grid is employed to estimate socio-economic exposure to flood hazard according to six probability scenarios; by doing so, it is possible to identify the distribution of population, asset and sectorial domestic product in relation to the magnitude and frequency of potential flood events. This information is aimed to help decision-makers when running risk assessments and planning of mitigation measures. The high resolution of the grid (250x250 meters) allows to investigate the effects locally as well as at the preferred administrative level. The SVI offers a supplementary view on the variability of the social dimension. Inevitably, the composition of the SVI is also prone to an amount of uncertainty because of the subjective choices in the criteria selected for its production. A large survey over expert knowledge and a detailed analysis of sensitivity would be required to justify the use of some rules and criteria over others.

4.6 Acknowledgment

The research leading to this paper has received funding from the European Union FP7 ENHANCE projects (Grant Agreement N° 308438) and ECOCEP (Grant Agreement N° 609642).

4.7 References

Agenzia delle Entrate, 2014. Osservatorio del Mercato Immobiliare - Quotazioni zone OMI [WWW Document]. URL http://www.agenziaentrate.gov.it/wps/content/Nsilib/Nsi/Documentazione/omi/Banche+dati/Quotazioni+immobiliari/ (accessed 7.1.15).

Alfieri, L., Burek, P., Feyen, L., Forzieri, G., 2015a. Global warming increases the frequency of river floods in Europe. Hydrol. Earth Syst. Sci. 19, 2247–2260. https://doi.org/10.5194/hess-19-2247-2015

Alfieri, L., Feyen, L., Dottori, F., Bianchi, A., 2015b. Ensemble flood risk assessment in Europe under high end climate scenarios. Glob. Environ. Chang. 35, 199–212. https://doi.org/10.1016/j.gloenvcha.2015.09.004

Alfieri, L., Feyen, L., Salamon, P., Thielen, J., Bianchi, A., Dottori, F., Burek, P., 2016. Modelling the socio-economic impact of river floods in Europe. Nat. Hazards Earth Syst. Sci. 16, 1401–1411. https://doi.org/10.5194/nhess-16-1401-2016

Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., Feyen, L., 2014. Advances in pan-European flood hazard mapping. Hydrol. Process. 28, 4067–4077. https://doi.org/10.1002/hyp.9947

ANCE/CRESME, 2012. Lo stato del territorio italiano 2012: Insediamento e rischio sismico e idrogeologico. Rapporto ANCE/CRESME 2012.

Andriantiatsaholiniaina, L.A., Kouikoglou, V.S., Phillis, Y.A., 2004. Evaluating strategies for sustainable development: Fuzzy logic reasoning and sensitivity analysis. Ecol. Econ. 48, 149–172. https://doi.org/10.1016/j.ecolecon.2003.08.009

Bajat, B., Krunic, N., Bojovic, M., Kovacevic, Z., 2011. Population Vulnerability Assessment in Hazard Risk Management: a Dasymetric Mapping Approach. Rijeka.

Bakkensen, L.A., Fox-Lent, C., Read, L.K., Linkov, I., 2016. Validating Resilience and Vulnerability Indices in the Context of Natural Disasters. Risk Anal. 37, 982–1004. https://doi.org/10.1111/risa.12677

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

Balk, D., Yetman, G., 2004. The global distribution of population: evaluating the gains in resolution refinement. New York Cent. Int. Earth Sci. Inf. Netw. (CIESIN), Columbia Univ.

Batista e Silva, F., Lavalle, C., Koomen, E., 2013. A procedure to obtain a refined European land use/cover map. J. Land Use Sci. 8, 255–283. https://doi.org/10.1080/1747423X.2012.667450

Bielecka, E., 2005. A dasymetric population density map of Poland.

Bright, E.A., Coleman, P.R., King, A.L., 2006. LandScan 2005.

Brown, B.M., 1987. Population Estimation From Floor Area: a Restudy of "Naroll's Constant." Cross-Cultural Res. 21, 1–49. https://doi.org/10.1177/106939718702100101

Cutter, S.L., 1996. Vulnerability to environmental hazards. Prog. Hum. Geogr. 20, 529–539.

Cutter, S.L., Boruff, B.J., Shirley, W.L., 2003. Social vulnerability to environmental hazards. Soc. Sci. Q. 84, 242–261.

Dottori, F., Alfieri, L., Salamon, P., Bianchi, A., Feyen, L., Lorini, V., 2016. Flood hazard maps for Europe, 10-20-50-100-200-500 year return period. Eur. Comm. Jt. Res. Cent. [Dataset].

Eicher, C.L., Brewer, C. a., 2001. Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation. Cartogr. Geogr. Inf. Sci. 28, 125–138. https://doi.org/10.1559/152304001782173727

European Forum for GeoStatistics, 2011. Testing and quality assessment of pan-european population grids, in: GEOSTAT 1B Final Report, Appendix 3. pp. 1–8.

EUROSTAT, 2013. Eurostat statistics [WWW Document]. URL http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes (accessed 2.14.14).

Fekete, A., 2009. Validation of a social vulnerability index in context to river-floods in Germany. Nat. Hazards Earth Syst. Sci. 9, 393–403. https://doi.org/10.5194/nhess-9-393-2009

Fernandez, P., Mourato, S., Moreira, M., 2016. Social vulnerability assessment of flood risk using GIS-based multicriteria decision analysis. A case study of Vila Nova de Gaia (Portugal). Geomatics, Nat. Hazards Risk 7, 1367–1389. https://doi.org/10.1080/19475705.2015.1052021

Freire, S., MacManus, K., Pesaresi, M., Doxsey-Whitfield, E., Mills, J., 2016. Development of new open and free multi-temporal global population grids at 250 m resolution, in: AGILE 2016. Helsinki.

Frigerio, I., Strigaro, D., Mattavelli, M., Mugnano, S., De Amicis, M., 2016. Costruzione di un indice di vulnerabilità sociale in relazione a pericolosità naturali per il territorio italiano. Rend. Online della Soc. Geol. Ital. 39, 68–71. https://doi.org/10.3301/ROL.2016.49

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

Gallego, F.J., 2010. A population density grid of the European Union. Popul. Environ. 31, 460–473. https://doi.org/10.1007/s11111-010-0108-y

Garcia, R.A.C., Oliveira, S.C., Zêzere, J.L., 2016. Assessing population exposure for landslide risk analysis using dasymetric cartography. Nat. Hazards Earth Syst. Sci. Discuss. 2016, 1–21. https://doi.org/10.5194/nhess-2016-202

Geofabrik GmbH, 2012. OpenStreetMap data extracts [WWW Document]. URL http://download.geofabrik.de/ (accessed 3.30.17).

Goodchild, M.F., Anselin, L., Deichmann, U., 1993. A framework for the areal interpolation of socioeconomic data. Environ. Plan. A 25, 383–397.

Hefny, H.A., Elsayed, H.M., Aly, H.F., 2013. Fuzzy multi-criteria decision making model for different scenarios of electrical power generation in Egypt. Egypt. Informatics J. 14, 125–133. https://doi.org/10.1016/j.eij.2013.04.001

Hewitt, K., 1997. Regions of risk: a geographical introduction to disasters. Longman.

Huang, Y., Jiang, D., Fu, J., 2014. 1 km grid GDP data of China (2005, 2010) 69, 140–143. https://doi.org/10.3974/Global

ISPRA, 2014. Mappe di pericolosità idraulica e popolazione esposta a rischio alluvioni in italia.

ISTAT, 2014. Censimento edifici e abitazioni 12.

ISTAT, 2011. 15° Censimento generale della popolazione e delle abitazioni - manuale della rilevazione.

Jongman, B., Hochrainer-stigler, S., Feyen, L., Aerts, J.C.J.H., Mechler, R., Botzen, W.J.W., Bouwer, L.M., Pflug, G., Rojas, R., Ward, P.J., 2014. Increasing stress on disaster-risk finance due to large floods. Nat. Clim. Chang. 1–5.

https://doi.org/10.1038/NCLIMATE2124

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. Nat. Hazards Earth Syst. Sci. 12, 3733–3752. https://doi.org/10.5194/nhess-12-3733-2012

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

Koks, E.E., Bockarjova, M., de Moel, H., Aerts, J.C.J.H., 2014. Integrated direct and indirect flood risk modeling: development and sensitivity analysis. Risk Anal.

Koks, E.E., Jongman, B., Husby, T.G., Botzen, W.J.W., 2015. Combining hazard, exposure and social vulnerability to provide lessons for flood risk management. Environ. Sci. Policy 47, 42–52. https://doi.org/10.1016/j.envsci.2014.10.013

Langford, M., 2006. Obtaining population estimates in non-census reporting zones: An evaluation of the 3-class dasymetric method. Comput. Environ. Urban Syst. 30, 161–180. https://doi.org/10.1016/j.compenvurbsys.2004.07.001

Langford, M., Maguire, D., Unwin, D., 1991. The areal interpolation problem: Estimating population using remote sensing in a GIS framework, in: Masser, I., Blakemore, M. (Eds.), Handling Geographic Information: Methodology and Potential Applications. Longman, Harlow, Essex, England, pp. 55–77.

Maantay, J.A., Maroko, A.R., Herrmann, C., 2007. Mapping Population Distribution in the Urban Environment: The Cadastral-based Expert Dasymetric System (CEDS). Cartogr. Geogr. Inf. Sci. 34, 77–102. https://doi.org/10.1559/152304007781002190

Martin, D., Tate, N.J., Langford, M., 2000. Refining Population Surface Models: experiments with Northern Ireland Census Data. Trans. GIS 4, 343–360. https://doi.org/10.1111/1467-9671.00060

Moon, Z.K., Farmer, F.L., 2001. Population density surface: A new approach to an old problem. Soc. Nat. Resour. 14, 39–49.

Mrozinski, R.D., Cromley, R.G., 1999. Singly- and Doubly-Constrained Methods of Areal Interpolation for Vector-based GIS. Trans. GIS 3, 285–301. https://doi.org/10.1111/1467-9671.00022

Munafò, M., Assennato, F., Congedo, L., Luti, T., Marinosci, I., Monti, G., Riitano, N., Sallustio, L., Strollo, A., Tombolini, I., Marchetti, M., 2015. Il consumo di suolo in Italia.

Naroll, R., 1962. Floor Area and Settlement Population. Am. Antiq. 27, 587–589. https://doi.org/10.2307/277689

Nordhaus, W., Azam, Q., Corderi, D., Hood, K., Makarova, N.V., Mukhtar, M., Miltner, A., Weiss, J., 2006. The G-Econ database on gridded output: methods and data. Yale Univ. New Haven 1–30.

OECD, 2008. Handbook on Constructing Composite Indicators: methodology and user guide.

Ong, P., Houston, D., 2003. Draft Socioeconomic Report for 2003 Air Quality Management Plan. South Coast Air Quality Management District, Diamond Bar, California.

Ouzounis, G.K., Syrris, V., Pesaresi, M., 2013. Multiscale quality assessment of Global Human Settlement Layer scenes against reference data using statistical learning. Pattern Recognit. Lett. 34, 1636–1647. https://doi.org/10.1016/j.patrec.2013.04.004

Pesaresi, M., Ehrlich, D., Freire, S., 2014. The Global Human Settlement Layer (GHSL) - New Tools and Geodatasets for Improving Disaster Risk Assessment and Crisis Management,

in: Hiltz, S.R., Pfaff, M.S., Plotnick, L., Robinson, A.C. (Eds.), 11th International ISCRAM Conference. University Park, Pennsylvania, USA.

Phillis, Y. a, Andriantiatsaholiniaina, L. a, 2001. Sustainability: an ill-de ned concept and its assessment using fuzzy logic. Ecol. Econ. 37, 435–456.

Regione Emilia-Romagna, 2011. Coperture vettoriali dell'uso del suolo 2008 [WWW Document]. URL http://servizigis.regione.emilia-romagna.it/ctwmetadatiRER/metadatoISO.ejb?stato_IdMetadato=iOrg01iEnP1idMetadato76868 (accessed 7.1.15).

Roder, G., Sofia, G., Wu, Z., Tarolli, P., Roder, G., Sofia, G., Wu, Z., Tarolli, P., 2017. Assessment of social vulnerability to floods in the floodplain of Northern Italy. Weather. Clim. Soc. WCAS-D-16-0090.1. https://doi.org/10.1175/WCAS-D-16-0090.1

Rojas, R., Feyen, L., Watkiss, P., 2013. Climate change and river floods in the European Union : Socio-economic consequences and the costs and benefits of adaptation. Glob. Environ. Chang. - Press. https://doi.org/10.1016/j.gloenvcha.2013.08.006

Steinnocher, K., Köstl, M., Weichselbaum, J., 2011. Grid-based population and land take trend indicators - New approaches introduced by the geoland2 Core Information Service for Spatial Planning. NTTS Conf. Feb 2011.

Thieken, A.H., Müller, M., Kleist, L., Seifert, I., Borst, D., Werner, U., 2006. Regionalisation of asset values for risk analyses. Nat. Hazards Earth Syst. Sci. 6, 167–178. https://doi.org/10.5194/nhess-6-167-2006

Trigila, A., Iadanza, C., Bussettini, M., Lastoria, B., Barbano, A., 2015. Dissesto idrogeologico in Italia: pericolosità e indicatori di rischio.

UNEP, 2012. Global Risk Data Platform [WWW Document]. URL http://preview.grid.unep.ch/index.php?preview=data&events=socec&evcat=1 (accessed 4.10.15).

Willis, I., Gibin, M., Barros, J., Webber, R., 2014. Applying neighbourhood classification systems to natural hazards: a case study of Mt Vesuvius. Nat. Hazards 70, 1–22. https://doi.org/10.1007/s11069-010-9648-9

Wright, J.K., 1936. A method of mapping densities of population with Cape Cod as an example. Geogr. Rev. 26, 103–110.

Wu, J., Wang, X., Koks, E., 2017. Building Asset Value Mapping in Support of Flood Risk Assessment: a Case Study of Shanghai, China. Nat. Hazards Earth Syst. Sci. Discuss. 1–16. https://doi.org/10.5194/nhess-2017-17

Wu, S.S., Qiu, X., Wang, L., 2005. Population estimation methods in GIS and remote sensing: a review. GIScience Remote Sens. 42, 80–96.

Xie, Y., 1996. The overlaid network algorithms for areal interpolation problem. Comput. Environ. Urban Syst. 19, 287–306.

Zampetti, G., Ottaviani, F., Minutolo, A., 2012. I costi del rischio idrogeologico. Dossier Legambiente, Roma.

Zandbergen, P. a., Ignizio, D. a., 2010. Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates. Cartogr. Geogr. Inf. Sci. 37, 199–214. https://doi.org/10.1559/152304010792194985

Zhou, Y., Li, N., Wu, W., Wu, J., Shi, P., 2014. Local Spatial and Temporal Factors Influencing Population and Societal Vulnerability to Natural Disasters. Risk Anal. 34, 614–639. https://doi.org/10.1111/risa.12193

5 ESTIMATING FLOOD DAMAGE IN ITALY: EMPIRICAL VS EXPERT-BASED MODELLING APPROACH

5.1 Introduction

Among all natural hazards, floods historically cause the highest economic losses in Europe (EEA 2010; EASAC 2018). In Italy alone, a country with the largest absolute uninsured losses among EU countries (Alfieri et al. 2016; EEA 2016; Paprotny et al. 2018), around EUR 4 billion of public money were spent over a 10 years period to compensate the damage inflicted by major extreme hydrologic events (ANIA 2015). From 2009 until 2012, the recovery funding amounted to about EUR 1 billion per year; a fraction of the total estimated damage of around EUR 2,2 billion (Zampetti et al. 2012). In this context, and particularly compelled by the EU Flood Directive (2007/60/EC), sound and evidence-based flood risk assessments should provide the means to support the development and implementation of cost-effective flood risk reduction strategies and plans.

Several different approaches of varying complexity exist to estimate potential losses from floods, depending mainly on the category of damage (e.g. direct impacts or secondary effects, tangible or intangible costs, etc.) and the scale of application (i.e. macro, meso or micro scale) (Hallegatte 2008; Apel et al. 2009; de Moel et al. 2015; Koks et al. 2015a; Carrera et al. 2015). Direct tangible damages to assets are typically assessed using simple univariable models (UVMs) that rely on deterministic relations between a single descriptive variable (typically maximum water depth) and the economic loss mediated by the type/value of buildings or land cover directly affected by a hazardous event (Smith 1994; Meyer and Messner 2005; Scawthorn et al. 2006; Messner et al. 2007a; Jonkman et al. 2008a; Thieken et al. 2009; Merz et al. 2010a; de Moel and Aerts 2011a; Jongman et al. 2012a; Huizinga et al. 2017). Empirical, event-specific damage models are developed from observed flood loss data. A major drawback of empirically-based damage models relies on its low transferability to other study areas or regions, as significant errors are often verified when these are used to infer damage in other regions than those for which they were built to (Apel et al. 2004; Merz et al. 2004; Jongman et al. 2012a; Scorzini and Frank 2015b; Amadio et al. 2016; Wagenaar et al. 2016; Scorzini and Leopardi 2017; Hasanzadeh Nafari et al. 2017; Carisi et al. 2018). Synthetic models, on the other hand, are based on ‘‘what-if analyses’’, relying on expert-based knowledge in order to generalise the relation between the magnitude of a hazard event and the resulting economic damage. An

advantage of synthetic models over empirically-based models relies on the fact that the first are less sensitive to the input data, thus being better suited for both temporal and spatial transferability (Smith 1994; Merz et al. 2010; Dottori et al. 2016).

Both empirical and synthetic models can be configured as uni- or multivariable. The vast majority of univariable flood damage models account for water depth as the only explanatory variable to explain the often complex relation between the magnitude of a flood event and the resulting damages; however, a non-exhaustive literature search shows that other parameters may influence the flood damage process, such as flow velocity (Kreibich et al. 2009), flood duration, and water contamination (Thieken et al. 2005; Molinari et al. 2014b), to name just a few. In addition, a large number of other non-hazard factors can be significantly different from one place to another, such as type and quality of buildings, presence of basements, density of dwellings, early warning systems and precautionary measures (Smith 1994; Penning-Rowsell et al. 2005; Kreibich et al. 2005; Thieken et al. 2008; Pistrika and Jonkman 2010; Cammerer et al. 2013; Merz et al. 2013; Schröter et al. 2014; Wagenaar et al. 2017a; Carisi et al. 2018; Figueiredo et al. 2018). Therefore, multivariable models (MVMs) are potentially better-suited alternatives to describe the complex flood-damage relation (Apel et al. 2009; Elmer et al. 2010). Common techniques applied in a context of MVM are machine learning (e.g., Artificial Neural Networks and Random Forests) (Merz et al. 2013; Spekkers et al. 2014; Kreibich et al. 2017, Carisi et al. 2018), Bayesian networks (Vogel et al. 2013), and Tobit estimation (Van Ootegem et al. 2015). Moreover, some MVMs support probabilistic analyis of damage (Dottori et al. 2016b; Essenfelder 2017; Wagenaar et al. 2017b). MVMs need to be validated against empirical records in the region of the model application in order to produce reliable estimates (Zhou et al. 2013; Molinari et al. 2014b; Scorzini and Frank 2015b; Molinari et al. 2017; Hasanzadeh Nafari et al. 2017). However, greater sophistication of MVMs requires more detailed hazard, exposure and losses description. Due to the lack of consistent and comparable observed flood data, this kind of models are still seldom applied. This is why it is necessary to compile comprehensive, multivariable datasets with detailed catalogue of flood events and their impacts (see Amadio et al., 2016, Molinari et al., 2012 and 2014, and Scorzini and Frank, 2015).

Our study contributes to this end by assembling detailed data on three recent flood events in Northern Italy. For each event, our dataset comprises the following building-scale data: 1) hazard characterization derived from observational data and/or hydraulic modelling, 2) high-resolution exposure in terms of location, size, typology, economic value, etc. obtained from multiple sources, and 3) declared costs per damage categories. Building upon this extensive

dataset, we employ supervised learning algorithms to explore the parameters of hazard, exposure and vulnerability and their influence on damage magnitude. We test linear, logarithmic and square root regression to select the best fitting Uni-Variable (UVM) and Bi-Variable (BVM) models, and two machine learning techniques, namely Random Forest (RF) and Artificial Neural Networks (ANN) for training and testing the empirical MVMs. The models developed on the three considered case studies provide a benchmark to test the performance of four literature models of different nature and complexity, specifically developed for Italy. The results of this study provide important insights to understand the feasibility and reliability of flood damage models as practical tools for predictive flood risk assessments in Italy.

5.2 Study area

With an extent of 46,000 km2, the Po Valley is the largest contiguous floodplain in Italy. It extends from the Alps in the north to the Apennines in the south-west, and the Adriatic Sea to the east. It comprises the Po river basin, the eastern lowlands of Veneto and Friuli, and the south-eastern basins of Emilia-Romagna. The Po Valley is one of the most developed and populated areas in Italy, generating about half of the country's gross domestic product (AdBPo 2006). In the lower part of the Po river, flood-prone areas are protected by a complex system of embankments and hydraulic works that are part of the flood defence system in the Po Valley, extending for almost 3,000 km as a result of centuries-long tradition of river embanking (Govi and Turitto 2000; Lastoria et al. 2006; Masoero et al. 2013). However, flood protection structures generate a false sense of safety and low risk awareness among the floodplain residents (Tobin 1995). As a result, exposure has steadily increased in flood prone areas of the Po Valley (Domeneghetti et al. 2015). Records of past flood events (1950-1995) maintained by the National Research Council (Cipolla et al. 1998) show that more than 3,300 individual locations were affected by approximately 1,000 flood events within the Po Valley.

Three of the most recent flood events within the Po Valley (figure 5.1) have been chosen as case studies for this analysis: the 2002 Adda flood that affected the province of Lodi (1); the 2010 Bacchiglione flood which involved the area of Vicenza (2); and the 2014 Secchia flood in the province of Modena (3). All three locations have been subject to frequent flooding between 1950 and 2000 according to the historical catalogue. A short description of these three events is provided hereinafter to understand the dynamics and the impacts of each flood.

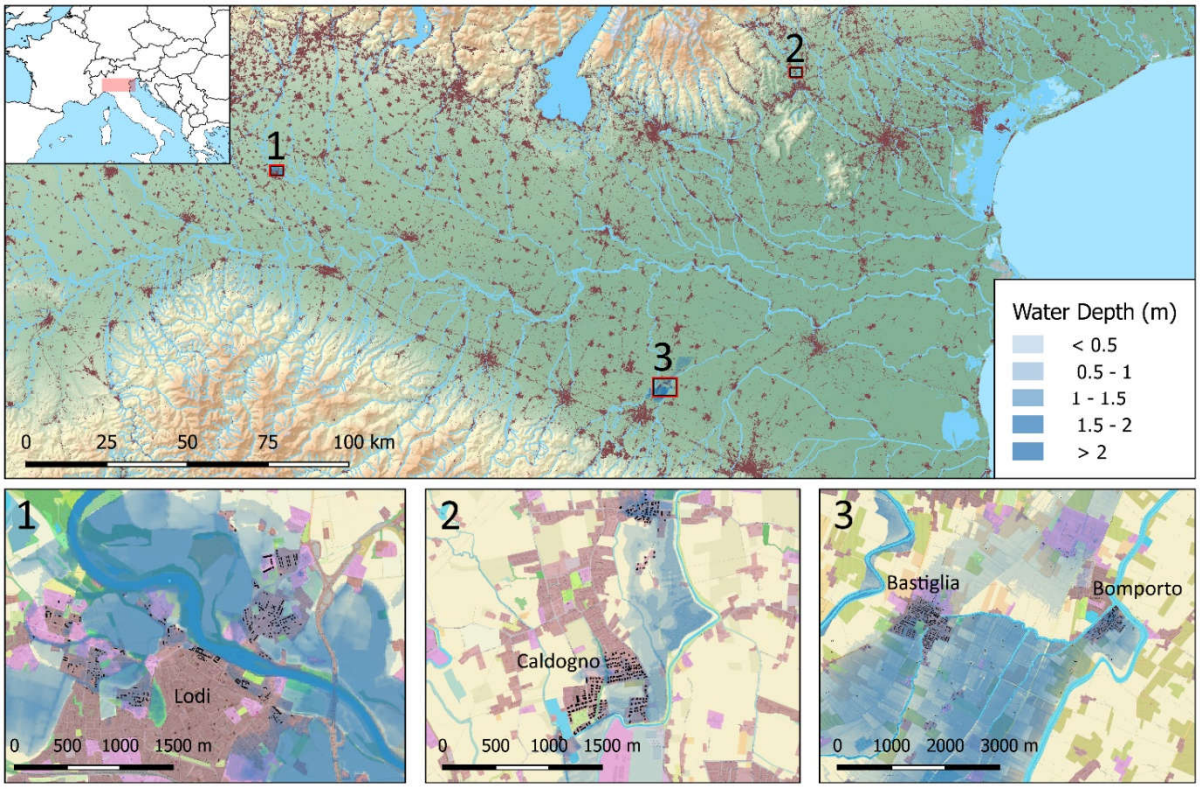


Figure 5.1. Case studies in Northern Italy (Po Valley). 1: Adda river flooding the town of Lodi, 2002; 2: Bacchiglione river flooding the province of Vicenza, 2010; 3: Secchia river flooding the province of Modena, 2014. Flooded buildings for which damage records are available are shown in black.

5.2.1 Adda 2002

On the 27th of November 2002, the province of Lodi (Lombardy) was struck by a flood caused by the overflow of the Adda river. The flood-wave reached a record discharge of about 2,000 m^3/s, corresponding to a return period of 100 years (Rossetti et al. 2010). The river overtopped the embankments and flooded the rural area first, later reaching the residential and commercial areas within the capital town of the province, Lodi. The low-speed flood waters rose up to 2.5-3m. The inundation lasted for about 24 hours and affected a large area of the Adda floodplain, including 5.5 ha of residential buildings. There were no reported casualties, but several families were evacuated during the emergency and important service nodes such as hospitals were severely affected. The reported damage to residential properties, commercial assets and agriculture summed up to EUR 17.7M, out of which EUR 7.8M relate to residential buildings.

5.2.2 Bacchiglione 2010

From the 31st of October to the 2nd of November 2010, persistent rainfall affected the pre-Alpine and foothill areas of Veneto region exceeding 500 mm in some locations (ARPAV 2010). As a result, about 140 km^2 of land were flooded, involving 130 municipalities (Regione del Veneto

2011). The Bacchiglione river, in the province of Vicenza, was particularly negatively affected. Heavy precipitation events and early snow melting increased the hydrometric levels of the Bacchiglione river and its tributaries, surpassing historical records (Belcaro et al. 2011). On the morning of November 1st, the water flowing at 330 m^3/s opened a breach on the right levee of the river, flooding the countryside and the settlements of Caldogno, Cresole and Rettorgole with an average water depth of 0.5 m (ARPAV 2010). Then the river overflowed downstream, towards the chief-town of Vicenza, which was inundated up to its historical center. The inundation lasted for about 48 hours, and its extent was about 33 Ha, of which 26 Ha consisted of agricultural land and 7 Ha were urban areas. The total damage including residential properties, economic activities, agriculture and public infrastructures was estimated to be around EUR 26M, while dwellings alone accounted for EUR 7.5 M (Scorzini and Frank 2015b).

5.2.3 Secchia 2014

In January 2014 severe rainfall endured for two weeks on the central part of Emilia-Romagna region, discharging the annual average amount of rain in just a few days. On the 19th, at around 6 AM, the water started to overtop a section (10 m) of the of the right levee of the Secchia river, which stands 7-8 meters over the flood plain. Later in the morning the levee breached at the top by one meter, flooding the countryside. After 9 hours, the levee section was completely destroyed for a length of 80 meters, spilling 200 m^3/s and flooding around six thousand hectares of rural land (D'Alpaos et al. 2014b). Seven municipalities were affected, with the small towns of Bastiglia and Bomporto suffering the largest impact. Both towns, including their industrial districts, remained flooded for more than 48 hours. The total volume of water inundating the area was estimated to be around 36 million m^3, with an average water depth of 1 meter (D'Alpaos et al. 2014b). The economic cost inflicted to residential properties according to damage declaration amounts to EUR 36M.

5.3 Materials and methods

5.3.1 Data description

The dataset we compiled for this analysis comprises:

- Detailed hazard data, including the flood extent, depth, persistence, and flow velocity.
- High-resolution spatial exposure data, including type, location and value of buildings.
- Comprehensive vulnerability data, including the characteristics of building and dwellings in terms of material, quality and age.
- Reported costs of reparation or replacement of damaged goods.

The main hazard features (extent, depth, flow velocity and duration) are obtained from flood maps produced by 2D hydraulic models based on observations performed during and after the events. In detail, the hydraulic simulation for the Adda river has been produced by means of River2D model (Steffler and Blackburn 2002) using a 5m resolution digital terrain model and high-resolution LiDAR data for the description of the floodplains obtained from the river basin district authority. The Bacchiglione flood have been simulated using the 1D/2D model Infoworks RS (Beta Studio 2012). The 1D river network geometry comes from a topographic survey of cross-sections, while the 2D floodplain morphology (5 m resolution) is obtained from LiDAR data produced by the Italian Ministry of Environment (Molinari et al. 2018). The reliability of the simulations for the Adda and Bacchiglione floods was verified using hydrometric data, aerial surveys of inundated areas and photos/videos from the affected population (Rossetti et al. 2010; Scorzini and Frank 2015b). The Secchia flood event has been simulated using an innovative, time-efficient approach (Vacondio et al. 2016) which integrates both river discharge and floodplain characteristics in a parallel computation. The simulation was performed on a 5 meters grid and its results were validated against several field data and observations, including a high-resolution radar image (Vacondio et al. 2014; Vacondio et al. 2017). The information needed for the characterization of exposure is collected from a variety of sources and then spatially projected to have a homogeneous, georeferenced dataset for each case study. External buildings perimeter and area are obtained from the Open Street Map database (Geofabrik GmbH 2018) and associated with official street-number points containing addresses. The land cover is freely available as perimeters classified by the CORINE legend (4^{th} level of detail) (Feranec, J. Ot'ahel' 1998) obtained from Regional Environmental Agencies databases. Land cover information is used to discriminate housing from other buildings (industrial, commercial, etc.). In addition, indicators for building characteristics (Table 1) have been selected from the database of the official Italian Census of 2011 (ISTAT). Construction and restoration costs as EUR/m^2 are obtained for the case study areas from the CRESME database (CRESME/CINEAS/ANIA 2014). They are used to convert the absolute damage values into relative damage shares. Empirical damage records have been collected by local administrations after the flood events in relation to households' street numbers. The records falling outside the simulated flood extents are filtered from the dataset. Each record includes: claimed; verified; and refunded damage to residential buildings. Since actual compensation often covers only a fraction of the damage costs, claimed damage is preferred in order to measure the economic impact (see Carisi et al. 2018). We restricted our analysis on direct

monetary damage to the structure of residential buildings, excluding furniture and vehicles. Economic losses, building values and construction costs for the three events have been scaled to EUR2015 inflation value.

5.3.2 Damage models overview

Empirical damage models are drawn based on actual data collected from specific events (e.g. Luino et al. 2009; Hasanzadeh Nafari et al. 2017); in some regions they represent the only knoweldge base for the assessment of flood risk. However, they carry a large uncertainty when employed in different times and places (McBean et al. 1986; Gissing and Blong 2004). Differently, synthetic models are based on a valuation survey which assesses how the structural components are distributed in the height of a building (Smith 1994; Oliveri and Santoro 2000; Barton et al. 2003). In such expert-based models, the magnitude of potential flood loss is estimated based on the vulnerability of structural components via "what-if" analysis and in the evaluation of the corresponding damage based on building and hazard features (Gissing and Blong 2004; Merz et al. 2010a). Most empirical and synthetic models are UVMs based on water depth as the only predictor of damage; yet recent studies (see e.g. Dottori et al., 2016 and Merz et al. 2013) suggest that MVMs developed using expert-based or machine-learning approaches outmatch the performances of customary univariable regression models. However, the development of MVMs requires a comprehensive set of data in order to correctly identify complex relationships among variables.

5.3.2.1 Models from literature

There are few models in the literature that are dedicated to the economic assessment of flood impacts over Italian residential structures (see e.g. Oliveri and Santoro 2000; Huizinga 2007; Luino et al. 2009; Dottori et al. 2016). These models have been developed independently using different approaches, assumptions and base data. The first model we selected for testing (Luino et al. 2009b) is an empirical UVM based on the impact data collected after the flash-flood event of May 2002 in the Boesio Basin, in Lombardy. One stage-damage curve was generated for structural damage to the most common building type in the area using loss data measured after the flood combined with estimates of water depth from a 1D hydraulic model. In this model, the estimation of building value is based on its geographical location, use and typology, based on market value quotations by the official real estate observatory of Italy (Agenzia delle Entrate 2018). Market values of residential stocks for specific areas. The second model (OS - Oliveri and Santoro 2000) is a synthetic UVM developed for a study performed in the city of Palermo

(Sicily). The model is based on water depth and consists of two curves, one for buildings with 2 floors and one for those with more than 2 floors. It considers water stage steps of 0.25 m; for each stage, the model computes the overall replacement cost as the result of damage over different components (internal and external plaster, fixtures, floors and electric appliances) plus the expenses for dismantling the damaged components. The third model we included in our analysis is part of a stage-damage curve database developed by the EU Joint Research Centre (Huizinga 2007b; Huizinga et al. 2017) on the basis of an extensive literature survey. Damage curves are provided for a variety of assets and land use classes on the global scale by normalising the maximum damage values in relation to country-specific construction costs. These are obtained by means of statistical regressions with socioeconomic development indicators. The JRC curves are suggested for application at the supra-national scale but can be a general guide to carry on assessments at the meso-scale in countries where specific risk models are not available. We select the curve provided for Italy (JRC-IT) to be tested on our dataset.

The fourth model considered is INSYDE, *In-depth Synthetic Model for Flood Damage Estimation* (Dottori et al. 2016), which is a synthetic MVM developed for residential buildings and released as open source R script. Repair or replacement costs are modelled by means of analytical functions describing the damage processes for each component as a function of hazard and building characteristics, using an expert-based "what-if" approach. Hazard features include physical variables describing the flood event at the building location, e.g. water depth, flood duration, presence of contaminants and sediment load.

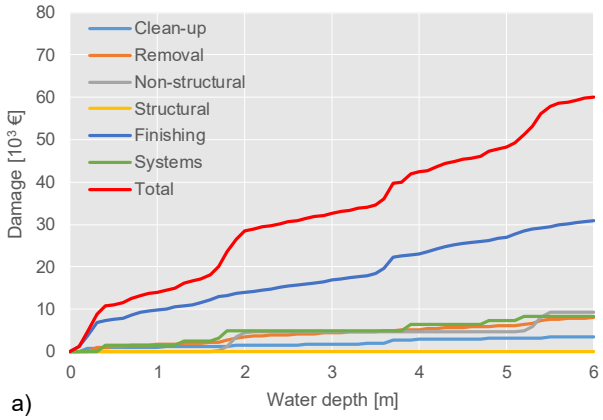


Figure 5.2. Examples of damage curves in relation to water depth produced by INSYDE for riverine floods in relation to a building with FA=100 m2, NF=2, BT=3, BS=2, FL=1, YY=1990, CS=1.

Exposure indicators include building characteristics such geometry and features. Building features affect costs estimation either by modifying the damage functions or by affecting the unit prices of the building components by a certain factor. Damage categories include clean-up and removal costs, damage to finishing elements, windows, doors, wirings and installations

(Figure 5.2). The model adopts probabilistic functions for some of the buildings' components for which it is difficult to define a deterministic threshold of damage occurrence in relation to hazard parameters. The list of explicit input variables accounted by INSYDE is shown in Table 5.1, with the indication of their respective data sources. Despite the large number of inputs, the model proved to be adaptable to the actual available knowledge of the flood event and building characteristics (Molinari and Scorzini 2017).

Variable	Description	Source	Unit	Name
Hazard features				
Water depth	Maximum depth	Hydro model	m	he
Flow velocity	Maximum velocity	Hydro model	m/s	v
Duration	Hours of inundation	Hydro model	h	d
Exposure and vulnerability of buildings				
Replacement value	Economic value of the building structure	CRESME	EUR/m^2	RV
Area and perimeter	Footprint area and external perimeter	OSM/CTR	m^2, m	FA, EP
Basement	Presence (1) or absence (0) of basement	CRESME	-	B
Number of floors	1, 2, 3 or more than 3 floors	Census/Inspection	-	FN
Building type	Flat (1), semi-detached (2) or detached (3)	Census/Inspection	-	BT
Building structure	Bricks (1) or concrete (2)	Census/Inspection	-	BS
Finishing level	Low (0.8), medium (1) or high (1.2)	Census/Inspection	-	FL
Conservation status	Bad (0.9), normal (1) or good (1.1)	Census/Inspection	-	CS
Observed damage				
Damage claims	Private and shared structural parts	Official survey	EUR	D

Table 5.1. List of variables included in the multivariable analysis.

5.3.2.2 Models developed and trained on the observation dataset

This section provides an overview about the empirical damage model obtained from our events dataset, namely two supervised learning algorithms (Random Forest, Artificial Neural Network) and three uni- and bivariable regression models used to assess the importance of variables as damage predictors. All these models share the same sampling approach for training and validation: the observation dataset is split in three parts, where two thirds are used to train the model and one third for validation.

5.3.2.2.1 Multivariable models: supervised learning algorithms

A probabilistic approach is required in damage estimation in order to control the effects of data variability on the model uncertainty. This is useful to overcome the limitations associated with the choice of a singular model and to increase the statistical value of the analysis (Kreibich et al. 2017). The algorithms we employed to deal with the empirical data share an iterative scrambling and resampling approach (1,000 repetitions) in order to draw the confidence interval of the models independently from source data variability. For the setup of empirically-based MVMs we selected ten variables from those listed in Table 1, excluding those with small

variability (basement, conservation status) or those for which an adequate level of detail is not possible in our case studies (age, heat system). These ten variables serve as input for two machine learning algorithms, namely Random Forest (RF) and Artificial Neural Network (ANN), described in the next paragraphs. Both algorithms produce a distribution of estimates for each record, from which the mean value is calculated.

Random Forest

The RF is a data mining procedure, a tree-building algorithm that can be used for classification and regression of continuous dependent variables (CART method - see Breiman 1984) like the one used by Merz et al. (2013). RF has numerous advantages, such high prediction accuracy, tolerance of outliers and noise, avoidance of overfitting problems, and no need of assumptions about independence, distribution or residual characteristics. Because of this, it has already been employed in the context of natural hazards, including earthquake-induced damage classification (Tesfamariam and Liu 2010), flood hazard assessment (Wang et al. 2015), and flood risk (Merz et al. 2013; Spekkers et al. 2014; Chinh et al. 2015; Kreibich et al. 2017; Carisi et al. 2018).

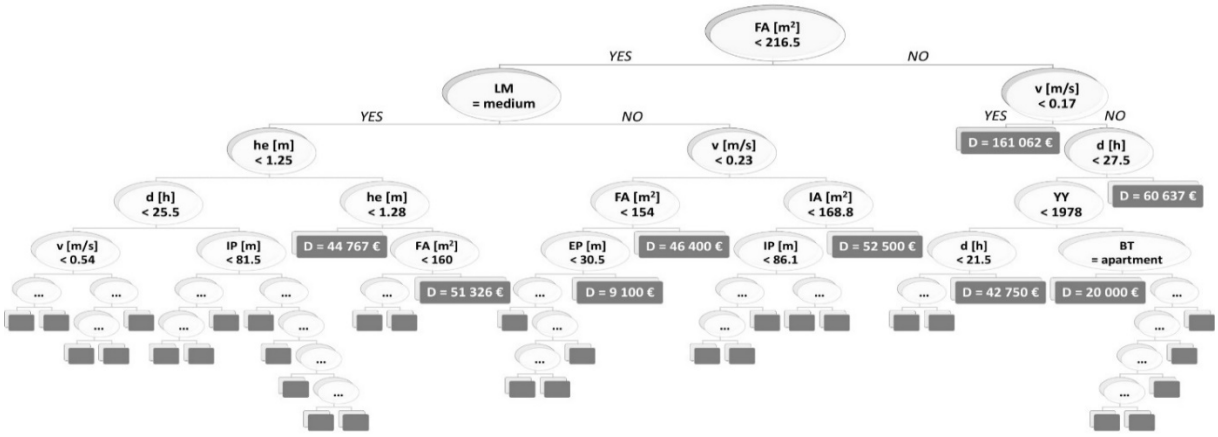


Figure 5.3. Example of one of the regression trees produced by the Random Forest model.

We use the algorithm implemented in the R package *RandomForest* by Liaw and Wiener (2002). The Random Forest algorithm builds and combines many decision trees, where each tree has a non-linear regression structure, recursively splitting the input dataset into smaller parts by identifying the variables and their splitting values which maximize the predictive accuracy of the model. The tree structure has several branches, each one starts from the root node and includes several leaf nodes, until either a threshold for the minimum number of data points in leaf nodes is reached or no further splitting is possible. Each estimated value represented by the resulting terminal node of the tree answers to the partition question asked in the previous interior nodes and depends on the response variable of all the parts of the original

dataset that are needed to reach the terminal node (Merz et al. 2013). In order to reduce the uncertainty associated with the selection of a single tree, the RF algorithm (Breiman 2001) creates several bootstrap replicas of the learning data and grows regression trees for each subsample, considering a limited number of variables at each split. This will result in a great number of regression trees, each based on a different (although similar) set of damage records and each leaving out a different number of variables at each split. The mean value among all prediction of the individual trees is chosen as representative output. An example of a built tree for the present study is shown in Figure 5.3. Another important strength of RF is its capability to evaluate the relative importance of each independent variable in the tree-building procedure, i.e., in our case, in representing the damage process. By randomly simulating the absence of one predictor, the RF algorithm calculates the decreasing of the performance of the model and thus the importance of the variables in the prediction.

Artificial Neural Network

ANNs are mathematical models based on non-linear, parallel data processing (Haykin 2001). They have been applied in several fields of research, such as hydrology, remote sensing, and image classification (Heermann and Khazenie 1992; Giacinto and Roli 2001; Campolo et al. 2003). The model used in this study (Essenfelder 2017) consists of a Multi-Layer Perceptron (MLP) neural network model, using back-propagation as the supervised training technique and the Levenberg-Marquardt as the optimization algorithm (Hagan and Menhaj 1994; Yu and Wilamowski 2011) (see figure 5.4 for the structure of the model).

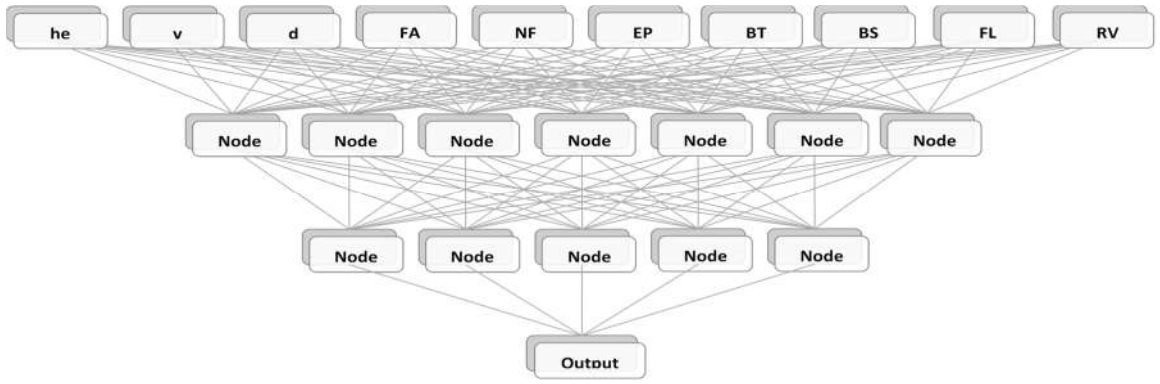


Figure 5.4. Structure of the Artificial Neural Network model applied in this study using two neurons (nodes) layers.

The developed ANN model evaluates the Sum of Squared Errors (SSE) of the model outputs with regards to the targets for each training epoch as a way of assessing the generalization property of a trained ANN model (Hsieh and Tang 1998; Maier and Dandy 2000). The ANN

runs in a multi-core configuration and provides an ensemble of trained ANN models as a result, thus being suitable for probabilistic analysis. The input and target information are normalized by feature scaling before being processed by the model, while the initial number of hidden neurons per hidden layer is approximated as two-thirds of the summation of the number of neurons in the previous and next layers (Han 2002). Regarding the activation functions, a log-sigmoid function is used for the connection with neurons in the first and second hidden layers, while a linear function is used for the connections with neurons in the output layer, allowing values to be either lower or greater than the maximum observed valued in the target dataset. This configuration is interesting as it does not limit the output range of the ANN model to the range of normalized values. The input data is randomly split between three distinct sets, namely training, validation, and test. The training dataset is used to calibrate the ANN model, meaning that the weight connections between neurons are updated with respect to the data available in this dataset. The validation set is utilized to avoid the overtraining or overfitting of the ANN model, being used to stop the training process. The test set is not presented to the model during the training procedure, being used only as a way of verifying the efficiency of a trained ANN when stressed by new data. In order avoid any possible bias coming from the random split of the original dataset into training, validation, and test datasets, about 1,000 training attempts are performed, each with a different initial weight initialization and training dataset composition. The resulting ANN model consists of an ensemble of 4 models, representing the best overall results after the training procedure, that are used to define the confidence interval.

5.3.2.2.2 Univariable and bivariable models

In order to understand if the added complexity of MVMs brings any improvement in the accuracy of damage estimates, we compare them with traditional, deterministic univariable (UVM) and bivariable (BVM) regression models that are empirically derived from the observation dataset. Considering the first (water depth) or the first two variables (water depth and water velocity), we investigate whether a linear, logarithmic or exponential function has the best regression fit to the records. All functions that consider water depth are forced to pass through the origin, because most buildings have no basement and, accordingly, no water means no damage. Similarly to what we did for the MVM training, we uses an iteration of 1,000 scrambling and resampling cycles which is repeated using the two different sampling strategy: first the models are trained on 2/3 of the data and validated on the remaining 1/3 of the records.

5.4 Results and discussion

5.4.1 Recorded hazard and damage

Our combined dataset contains records of 1,158 damaged residential buildings (Table 5.2). More than a half of these were damaged by the Secchia flood, which affected the largest residential area (17.7 ha) and caused the largest total losses. Only verified, spatially-matching records are accounted; economic losses are scaled to EUR2015 inflation value. Note that these losses are related to the structural damage of residential buildings, thus they do not represent the full cost of the events.

Case study [River basin, year]	Affected buildings [n]	Flood extent [ha]	Avg. water depth [m]	Declared damage [M EUR 2015]
Adda, 2002	270	5.5	0.8	4.7
Bacchiglione, 2010	294	7.1	0.5	7.9
Secchia, 2014	594	17.7	1	21.1
Total	1,158	30.3	2.3	33.7

Table 5.2. Summary of residential buildings affected by the three investigated flood events according to hydraulic simulations and damage claims.

Boxplots in Figure 5 show the variance of variables driving the damage. Water depths range from 0.01 to about 2 meters, with most records falling in the interval 0.4 – 1.2 meters. Water velocities range between 0.01 and 1.5 m/s. Footprint areas and observed relative damages have similar average values for all three events, however the Secchia case study presents the longer count of records as well as the largest spread of outliers.

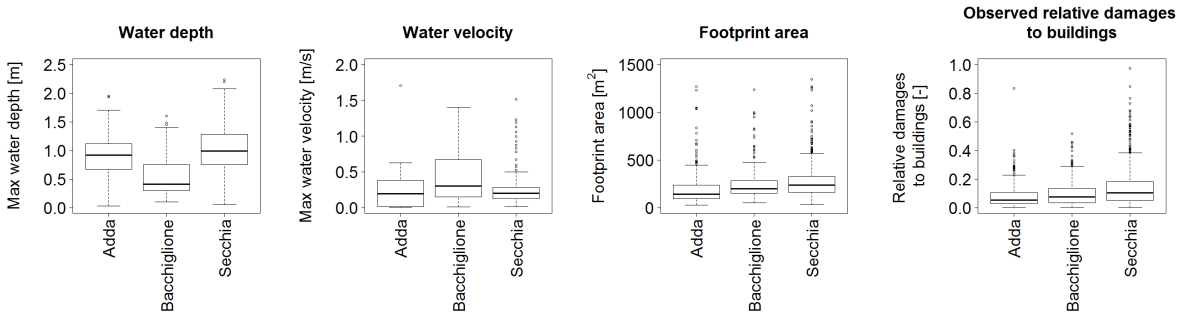


Figure 5.5. Data distribution for four variables from the three sample case studies.

The scatterplot in Figure 6 better shows the density of observed damages records in relation to the maximum water depth. The increase in depth corresponds to a larger range of variability in the economic damage.

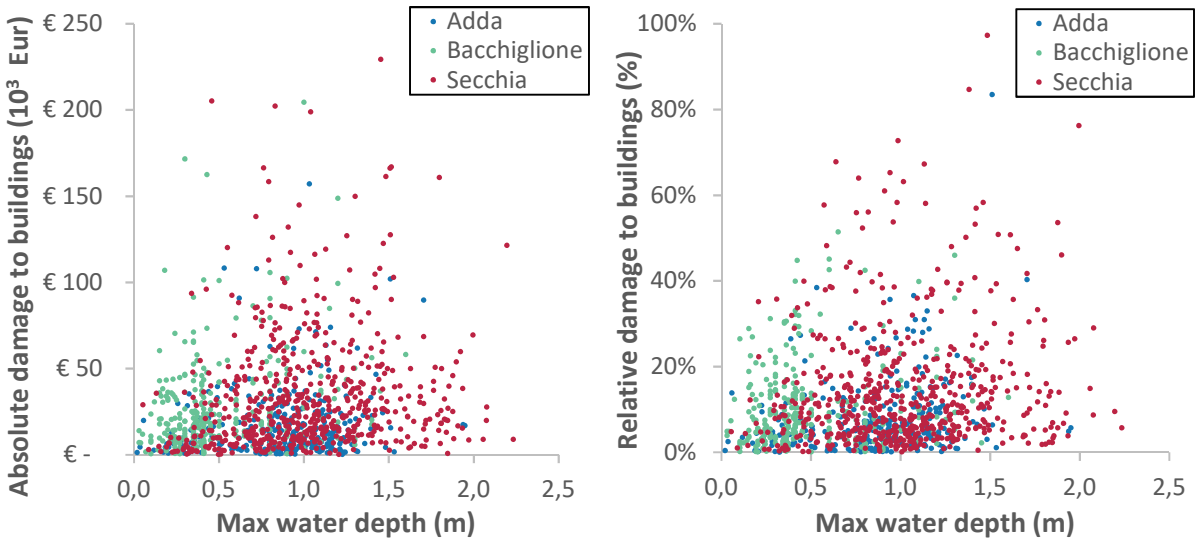


Figure 5.6. Scatterplot of monetary (**left**) and relative (**right**) damage (y-axis) in relation to maximum observed water depth (x-axis). Records from the same event are shown with the same color.

5.4.2 Influence of hazard and exposure variables on predicting flood damage

Water depth (*he*) is identified by RF as the most important predictor of damage (factor 3.4) among the ten examined variables (Figure 6). This confirms previous findings (Wagenaar et al. 2017a) and justifies the use of depth-damage curves for post-disaster need assessment. Flow velocity and geometric characteristics of buildings (area and perimeters) are also important (factor 2.7 to 2.3), followed by other predictors such as building value, flood duration, number of floors, finishing level and type of structure (factor 1 or less). Although water depth is the most influential variable, it is only moderately more important than other predictors. That substantiates the efforts to test the applicability of multivariable approaches to improve the estimation of damage.

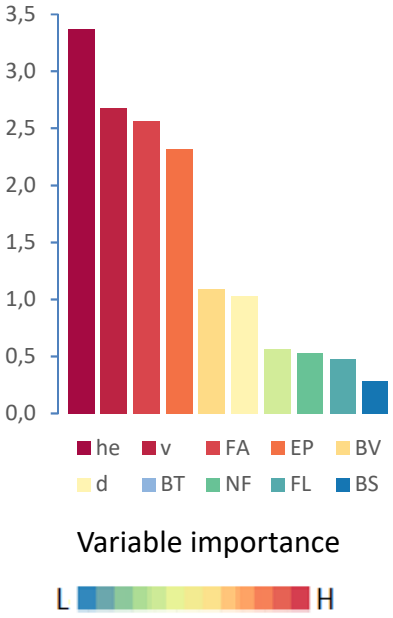


Figure 5.7. Relative importance of variables as predictors of damage according to the RF model.

5.4.3 Performance of the models

For assessing the predictive capacity of the four selected literature models, we compare them with empirically-based, data-trained models structured on the same variables, i.e. the evaluation of the models' performances is carried out by measuring and comparing the error metrics from the aforementioned models (JRC-IT, Luino, OS and INSYDE) to those provided by the

empirical MVMs obtained from supervised learning algorithms, the BVMs and traditional UVMs (depth-damage curves) developed on our dataset. The performances of each model are evaluated by using three metrics, namely Mean Absolute Error, Mean Bias Error and Root Mean Square Error. The MAE indicates the precision of the model in replicating the total recorded damage. The MBE shows the systematic error of the model, which is its mean accuracy. The RMSE measures the average magnitude of the error, enhancing the weight of larger errors. In addition to these error metrics, the total percentage error (E%, difference between observed and simulated damage divided by observed damage) is shown in tables.

5.4.3.1 Literature models

As first step, estimates of empirical and synthetic models from literature are compared with observed damages and the results in terms of total loss and total percentage error are shown in Table 5.3.

Case study	Unit	Obs.	JRC-IT	LUINO	OS	INSYDE
Adda 2002	M EUR 2015	4.7	24.3	13.0	8.1	5.6
	E%		417.0	176.6	72.3	19.1
Bacchiglione 2010	M EUR 2015	7.9	19.2	11.4	6.5	8.3
	E%		143.0	44.3	-17.7	5.1
Secchia 2014	M EUR 2015	21.1	64.5	44.1	19.8	28.8
	E%		205.7	109.0	-6.2	36.5
Full set	M EUR 2015	33.7	108.0	68.5	34.4	42.7
	E%		220.5	103.2	2.0	26.7

Table 5.3. Estimates and error from literature models compared to observed damage. Monetary values are in Million Eur, E% is total percentage error.

JRC-IT is the worst performing model, largely overestimating damage from the three events (E% 143-417), followed by the UV empirical model from Luino which overestimates damage with a percentage error ranging from 44 to 177. These results indicate that meso-scale models are not suitable for application at the micro-scale and that empirical models should be carefully applied for flood events with different characteristics from the ones for which they are developed. In fact, Luino's model was produced for a flash-flood event, with higher velocities and impacts. The two synthetic models, OS and INSYDE, perform much better, yet showing a large variability of the error factor, depending on the considered case. In detail, OS provides better results for the Secchia event (6% underestimation) and worse for the Adda set (72% overestimation), resulting in an estimate that is very close to the observations in terms of percentual error on the total dataset, although this is mainly due to compensation of positive

and negative errors for the different events. Differently, the INSYDE model exhibits a better performance for the Bacchiglione event (5% overestimation) and worse for the Secchia case study (37% overestimation). It is worth noting that, although the accuracy of the OS model is higher than of the INSYDE model for the full set, the latter is more accurate for two out of the three case studies (i.e. Adda 2002 and Bacchiglione 2010). Moreover, the INSYDE model provides more precise results, with a variance in errors 10 times lower than of the OS model and with maximum errors never exceeding an absolute value of 40%. However, INSYDE seems to consistently overestimate the total damages. Figure 5.7 compares the estimated and observed damages for the entire dataset for the two best performing literature models (OS and INSYDE).

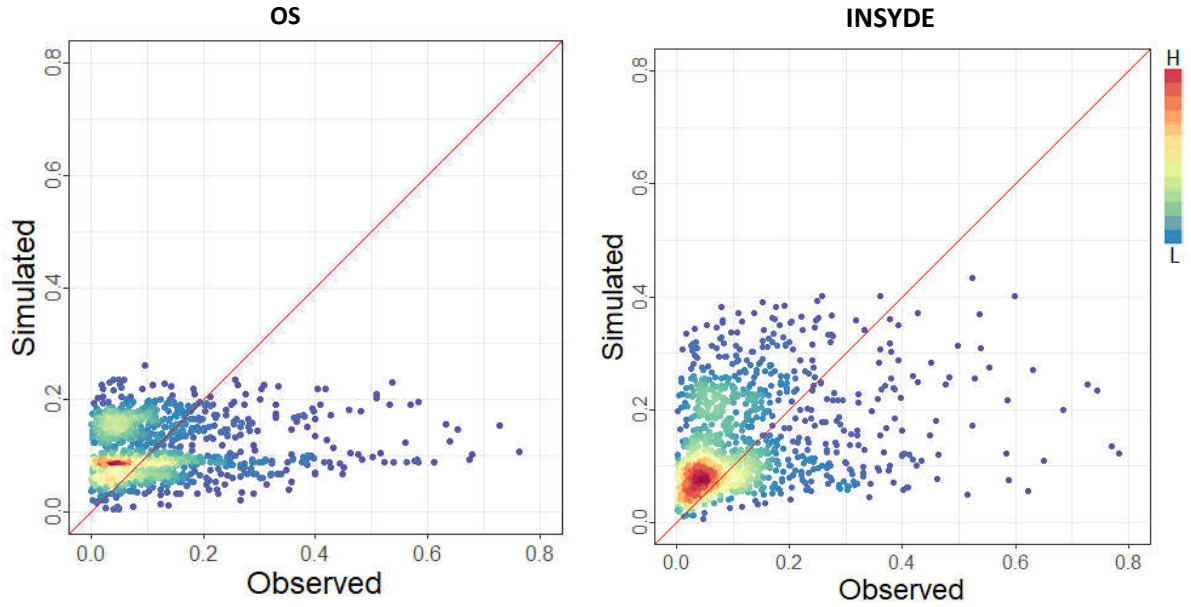


Figure 5.8. Scatterplot comparing relative damage estimates produced by the two best performing literature models, OS **(left)** and INSYDE **(right).** Simulated damage on the y-axis, observed damage on the x-axis. Colors represent records density.

5.4.3.2 Data-trained univariable, bivariable and multivariable models

In this section, damage values estimated by empirical, data-trained UVMs, BVMs and MVMs are compared with observed damage data. The results provided by these empirically-based models are used as a benchmark to understand the capability of tested literature models in predicting damage. The error metrics chosen for comparing the models' performances are presented for relative damage based on official estimates of replacement value, however training and validation were carried out also in terms of monetary damage with similar results, not presented for the sake of brevity.

	UVMs			BVMs		
Function	MBE	MAE	RMSE	MBE	MAE	RMSE
Linear	-0.015	0.087	0.127	-0.012	0.087	0.126
Log	-0.046	0.080	0.131	-0.046	0.080	0.131
Root	-0.003	0.086	0.123	-0.002	0.086	0.123

Table 5.4. Error metrics for the Univariable and Bivariable models.

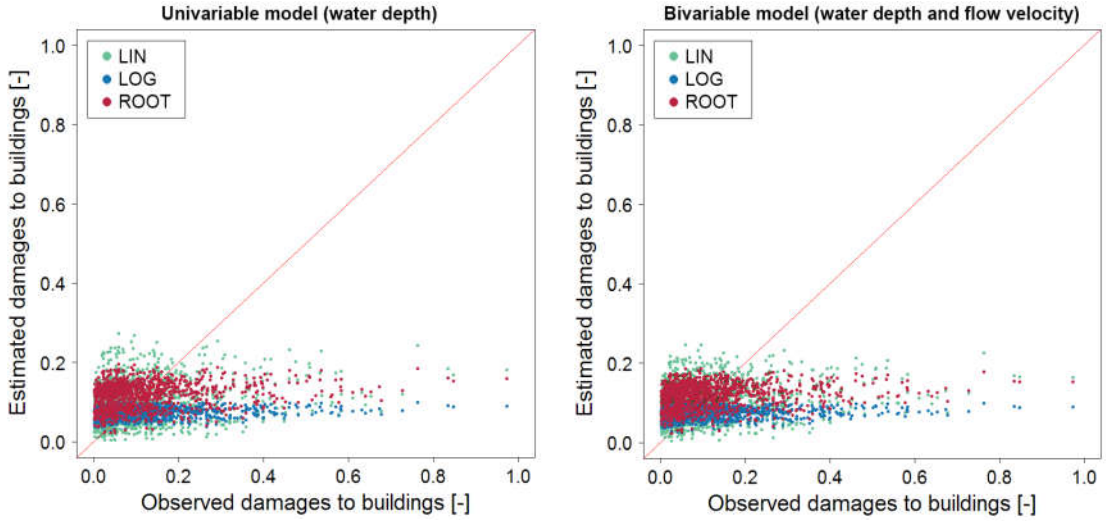


Figure 5.9. Testing the predictive capacity of uni- and bivariable models: estimated relative damage (y-axis) from the UVM **(left)** and BVM **(right)** are plotted against observed relative damage (x-axis) according to the three tested regression functions (LINear, LOGarithmic and ROOT function).

The results shown in Table 5.4 and figure 5.9 indicate no significant differences between UVMs and BVMs. We can affirm that the inclusion of water flow velocity as complementary explanatory variable does not improve the performance of simple regression models in our case study. For this reason, BVMs are dropped from further discussion from now on, to focus on a direct comparison between UVMs and MVMs.

Taking into consideration only UVMs, MAE and RMSE are very similar for the three tested regression functions. However, the root function described by the general formula $y = b(\sqrt[a]{x})$ has a slightly better fit (correlation is higher, MBE is lower) compared to linear and log functions. We select the function described by the equation $y = 0.13(\sqrt{x})$ as the best performing UVM to be included in the comparison with MVMs. Our findings confirm previous results indicating that the root curve as the most adequate to describe the flood damage process (Buck and Merkel 1999; Sluijs et al. 2000; Penning-Rowsell et al. 2005; Scawthorn et al. 2006; Kreibich and Thieken 2008; Thieken et al. 2008; Elmer et al. 2010; Cammerer et al. 2013; Wagenaar et al. 2017a). Figure 8 shows a direct comparison between the damage estimated by the empirically-based models against observed damage. The upper panel shows the output from

the UVM described by the root function. The lower panels show the output of the RF (left) and ANN (right) algorithms. The two machine learning algorithms produce comparable results, with both RF and ANN models tending to slightly overestimate the average damage (higher density of points, in red) and to significantly underestimate extreme values (lower density of values, in blue). This is a common result of data-driven models, where better results are biased to high-frequency values in comparison to low-frequency values due to the larger sample of those data in the calibration dataset. In Figure 5.10, the range of estimates, shown as min-max, describes the confidence of the model for individual records. In the case of RF, it shows the min-max range over all the 1,000 iterations of the model, while in the case of ANN only an ensemble of the four best-fit models is shown (see Section 5.3.2.2.1).

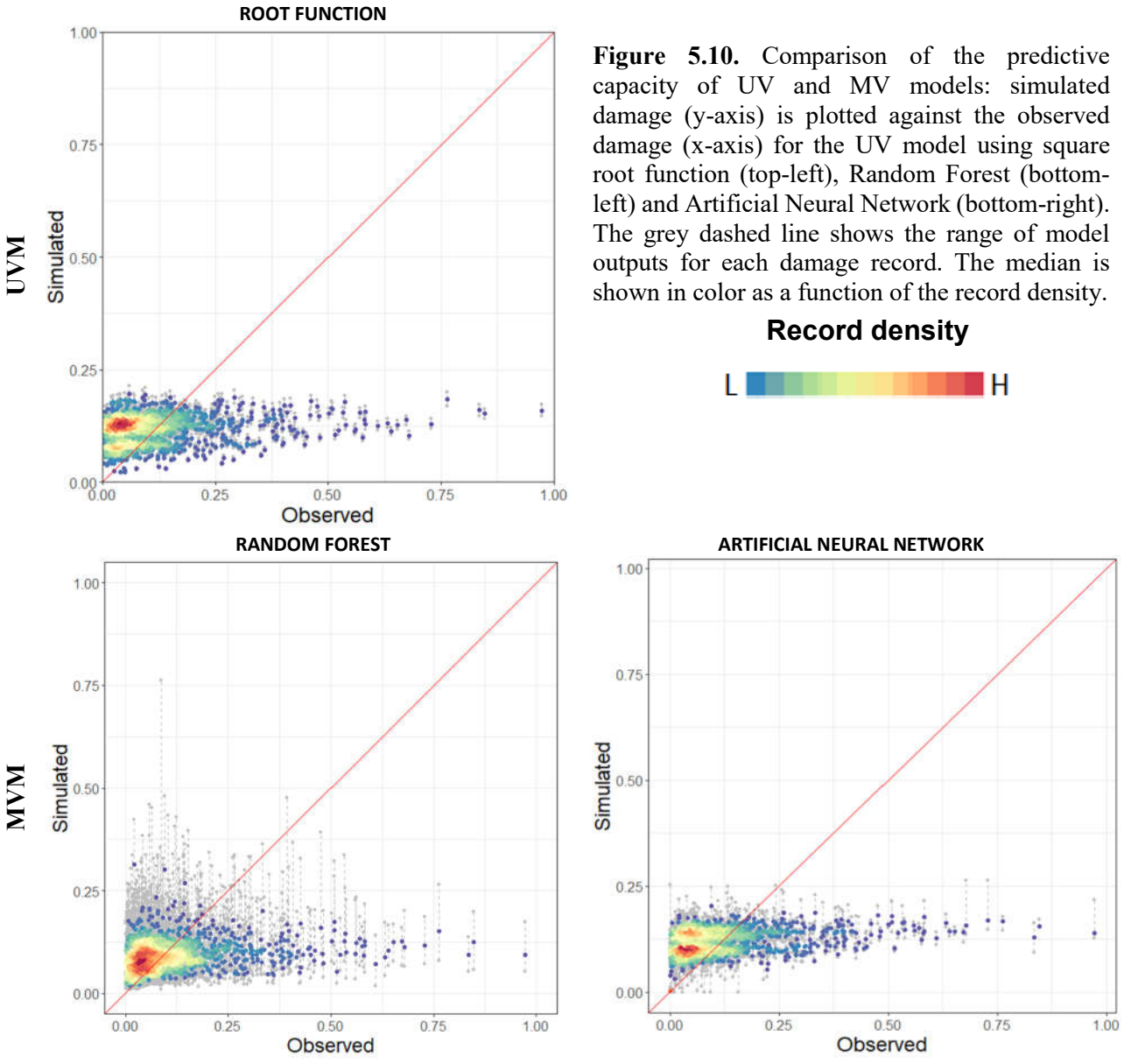


Figure 5.10. Comparison of the predictive capacity of UV and MV models: simulated damage (y-axis) is plotted against the observed damage (x-axis) for the UV model using square root function (top-left), Random Forest (bottom-left) and Artificial Neural Network (bottom-right). The grey dashed line shows the range of model outputs for each damage record. The median is shown in color as a function of the record density.

Theoretically, MVMs should simulate the complexity of the flooding mechanism better than UVMs. In our test, the ANN model has the best fit to the data, but UVMs (depth-damage curves) appear to perform similarly: the MVMs describe recorded damage with a percentage error between 0.2 and 10, while UVMs' error is around 12 (see table 5 in the next paragraph). Accordingly, when extensive descriptive data are not available, UVMs appear to be a reasonable alternative to describe the flood damage process. These empirically data-driven models are useful to understand the capability of multivariable approaches in predicting damage, i.e. which is the range of uncertainty that can be expected when assessing the flood damage process, comparing to simpler models like UVMs.

5.4.4 Comparing models' performances

First, we evaluate how selected literature UVMs (JRC-IT, Luino and OS) compare to the root function trained on the empirical dataset. Figure 5.9 shows the distribution and the density of observed relative damage as a function of water depth for the full dataset, together with the UV curves selected for testing. This figure explains the results presented in Section 5.4.3.1, with the JRC-IT and Luino models growing too fast for shallow water depths, as opposed to OS (shown as two separate curves for different number of floors of the building), which has a good mean fit to the data.

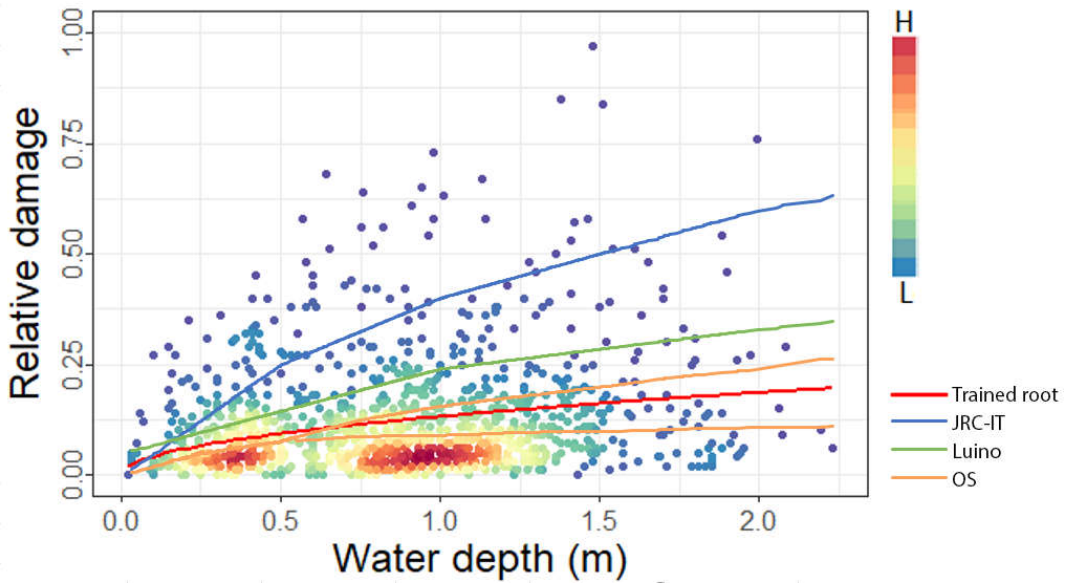


Figure 5.11. Scatterplot of relative damage records (y-axis) and water depth (x-axis). Points color represents record density. The red line shows the empirical root function ($y = 0.13(\sqrt{x})$, selected as best fit. The other lines represent the three UV literature models (JRC-IT, Luino, and OS) selected for the test. OS model is made of two curves, in relation to the number of floors of the building.

Table 5.5 summarises the main results from all the models in terms of error metrics. Specifically, among all models, MVMs RF and ANN are those with the lowest MAE and RMSE, followed by UVM ROOT with a MAE of 0.086 and a RMSE of 0.123. In terms of percentage error, the ranking is the same, with the only exception of OS, whose result in terms of this metric lies between the two empirical data-trained MVMs. Overall, the two expert-based literature models OS and INSYDE, are the best performing ones when compared to empirically-trained models, as shown by MAE, MBE and RMSE. As mentioned before, the performance of the UVM OS is very close to those of the MVM INSYDE, although this result may depend on the fact that the large share of records come from the Secchia event, for which OS outperforms INSYDE.

	Model	MBE	MAE	RMSE	Est. dmg [M EUR 2015]	Abs. error [M EUR 2015]	Percent error [%]
Trained models	**UVM (ROOT)**	-0.003	0.086	0.123	37.8	+4.1	+12.3
	MVM (RF)	-0.024	0.081	0.126	30.4	-3.3	-9.8
	MVM (ANN)	+0.009	0.091	0.115	33.8	-0.1	-0.2
Literature models	**UVM (JRC_IT)**	+0.217	0.239	0.27	108	+74.3	+220.5
	UVM (Luino)	+0.082	0.13	0.154	68.5	+34.8	+103.2
	UVM (OS)	-0.009	0.088	0.127	34.4	+0.8	+2.0
	MVM (INSYDE)	+0.019	0.093	0.132	42.7	+9.0	+26.7

Table 5.5. Comparing error metrics between empirically-base models and INSYDE.

Based on these results, the synthetic models INSYDE and OS currently represent very good alternatives for flood risk assessment in Italy, in cases where no empirical loss data are available to develop specific damage models. Indeed, our analysis has shown that particular care should be taken when transferring models derived from specific events (Luino curve) or from different scales (JRC-IT), while synthetic models can be considered more robust tools, relying on a physically-based description of flood damage mechanisms. Overall, for the investigated dataset, the synthetic MVM INSYDE has not been found to provide an improvement in the accuracy of damage estimates compared to those of the UV OS. However, the results of INSYDE are more precise if considering the different flood events, with a general, although limited, damage overestimation in all the cases, as opposed to OS which exhibited more accurate performance only for the Secchia flood and larger variability for the other two events, consequently being less precise. Further validation exercises, combined with the application of standardised and detailed procedures for damage data collection (e.g. Molinari et al. 2014) could improve INSYDE's predictive accuracy; being an open-source model, it is possible to modify the

damage functions for the different building components; for example, the availability of datasets with building losses subdivided into different categories (e.g. structural/non-structural elements, finishing, systems, etc.) could help to identify which damage components are responsible for the larger share of the error. The same cannot be said for OS, which is presented as a simple stage-damage curve, without a detailed explanation of the modelling assumptions on the considered flood-damage mechanisms.

As a final consideration, the accuracy and precision of damage observations are key aspects for the correct development of an MVM. This makes synthetic and empirical MVMs better fit for applications at the micro-scale (up to the census block scale (Molinari and Scorzini 2017)), where explanatory variables can be spatially disaggregated. Indeed, the aggregation scale is of primary importance in the application of MVMs: if we can compare our results to those reported in other studies applying similar multivariable approaches on an extensive damage dataset (bagging of regression trees), as in Wagenaar et al. (2017a) and in Kreibich et al (2017), we observe that our range of uncertainty is drastically smaller. This difference is likely imputable to the fact that, in the referred studies, information is aggregated at the municipality level, as opposed to our case, where each variable is precisely linked to buildings' location.

5.5 Conclusions

Risk management requires a reliable assessment tool to identify priorities in risk mitigation and adaptation. Such tool should be able to describe potential damage based on the available data related to hazard features and exposure characterisation. Recent studies suggest that multivariable flood damage modelling can outperform customary univariable models (depth-damage functions). In this study we collected a large empirical dataset which includes multiple hazard and exposure variables for three riverine flood events in Northern Italy, including the declared economic damage to residential buildings. On this basis, we produced three univariable, three bivariable and two multivariable models that are compared in terms of predictive accuracy and precision. We found that water depth is the most important predictor of flood damage, followed by secondary variables related to hazard (flow velocity, duration) and exposure features (area, perimeter and replacement value of the building). However, our results suggest that the inclusion of one additional variable (flow velocity) does not improve the estimates produced by simple regression models in a bivariable setup. On the other side, the analysis confirms the literature notion that the root function is the best fitting curve to describe damage in relation to water depth. Two MVMs were trained using two different machine

learning algorithms, namely Random Forest and Artificial Neural Network. These empirically-trained MVMs performed well (with an error ranging from 1 to 10%) in reproducing the damage output from the three events and thus were set as a reference for assessments in the same geographic context. In this perspective, other case studies are needed to confirm their robustness. Moreover, our results corroborate previous findings about the advantages of supervised machine learning approaches for developing or improving flood damage models. Still, their application remains limited by the availability of the data required for the MVM setup. In case of scarce number of variables, however, simple univariable models trained on the specific contexts seem to be a good alternative to MVMs.

We then considered four literature models of different nature and complexity to be tested on our extended case study dataset. We compared their error metrics with those of the empirically-trained UVMs and MVMs in order to evaluate their performance as predictive tool for flood risk assessment. The results have shown that both UV (Oliveri and Santoro 2000) and MV (INSYDE, Dottori et al. 2016) synthetic models can provide similar (although obviously larger) errors to those observed from empirical models. On the contrary, we found important errors when transferring models derived from other specific events (Luino curve) or different scales (JRC-IT). Therefore, the tested synthetic models can be currently considered as the best option for damage prediction purposes in the Italian context, in cases where no extensive loss data are available to derive a location-specific flood damage model. Overall, we found that errors produced by synthetic models were smaller than 30% of observed damage, with INSYDE providing more precise results over the different, single case study events (with a percentage overestimation of 19, 5 and 37% for Adda, Bacchiglione and Secchia, respectively) and is more accurate for two out of the three case studies (i.e. Adda and Bacchiglione), while the OS model is generally less precise but more accurate for the Secchia flood event only (2% error, as opposed to a 72% overestimation for the Adda and 18% underestimation for the Bacchiglione event).

Observed errors depend in part on the inherent larger variability found in the dataset related to that particular event. Nevertheless, the collection of additional independent flood records from different geographic contexts in Italy would help to further evaluate the adaptability of the models, especially of the open-source INSYDE, to estimate their uncertainty, and to increase their predictive accuracy. Finally, the work presented here has assembled a dataset that is currently one of the most extended and advanced for Italy; on this track, we aim to promote a shared effort towards an updated catalogue of floods that includes hazard, exposure and damage

information at the micro-scale. To this purpose, the adoption of a standardised and detailed procedure for damage data collection is a mandatory step.

5.6 Data availability

The INSYDE model is available as R open source code from https://github.com/ruipcfig/insyde

The hazard simulation of the Secchia flood event was kindly provided by Ing. Vacondio (University of Parma), whom we sincerely thank.

5.7 Acknowledgments

The research leading to this paper has received funding through the CLARA project from the EU's Horizon 2020 research and innovation programme under the Grant Agreement No 730482.

Authors acknowledge with gratitude Daniela Molinari, who provided the data for the Adda case study, within the framework of the Flood-Impat+ project, funded by Fondazione Cariplo.

5.8 References

AdBPo: Caratteristiche del bacino del fiume Po e primo esame dell'impatto ambientale delle attività umane sulle risorse idriche, Autorità di Bacino del Fiume Po ., 2006.

Agenzia delle Entrate: Osservatorio del Mercato Immobiliare - Quotazioni zone OMI, [online] Available from: http://www.agenziaentrate.gov.it/wps/content/Nsilib/Nsi/Documentazione/omi/Banche+dati/Quotazioni+immobiliari/ (Accessed 1 July 2018), 2018.

Alfieri, L., Feyen, L., Salamon, P., Thielen, J., Bianchi, A., Dottori, F. and Burek, P.: Modelling the socio-economic impact of river floods in Europe, Nat. Hazards Earth Syst. Sci., 16(6), 1401–1411, doi:10.5194/nhess-16-1401-2016, 2016.

Amadio, M., Mysiak, J., Carrera, L. and Koks, E.: Improving flood damage assessment models in Italy, Nat. Hazards, 82(3), 2075–2088, doi:10.1007/s11069-016-2286-0, 2016.

Apel, H., Thieken, a. H., Merz, B. and Blöschl, G.: Flood risk assessment and associated uncertainty, Nat. Hazards Earth Syst. Sci., 4(2), 295–308, doi:10.5194/nhess-4-295-2004, 2004.

Apel, H., Aronica, G. T., Kreibich, H. and Thieken, a. H.: Flood risk analyses—how detailed do we need to be?, Nat. Hazards, 49(1), 79–98, doi:10.1007/s11069-008-9277-8, 2009.

ARPAV: Scheda Evento "Idro" 31 Ottobre - 5 Novembre 2010., 2010.

Associazione Nazionale fra le Imprese Assicuratrici: Le alluvioni e la protezione delle abitazioni, , 22, 2015.

Barton, C., Viney, E., Heinrich, L. and Turnley, M.: The Reality of Determining Urban Flood Damages, in NSW Floodplain Management Authorities Annual Conference, Sydney., 2003.

Belcaro, P., Gasparini, D. and Baldessari, M.: 31 ottobre - 2 novembre 2010: l'alluvione dei Santi., 2011.

Beta Studio: Interventi per la sicurezza idraulica dell'area metropolitana di Vicenza: bacino di laminazione lungo il Torrente Timonchio in comune di Caldogno – Progetto definitivo.

Relazione idrologica e idraulica., 2012.

Breiman, L.: Classification and regression trees, Chapman & Hall. [online] Available from: https://books.google.it/books/about/Classification_and_Regression_Trees.html?id=JwQx-WOmSyQC&redir_esc=y (Accessed 30 July 2018), 1984.

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

Buck, W. and Merkel, U.: Auswertung der HOWASSchadendatenbank, edited by H. Institut für Wasserwirtschaft und Kulturtechnik der Universität Karlsruhe., 1999.

Cammerer, H., Thieken, a. H. and Lammel, J.: Adaptability and transferability of flood loss functions in residential areas, Nat. Hazards Earth Syst. Sci., 13(11), 3063–3081, doi:10.5194/nhess-13-3063-2013, 2013.

Campolo, M., Soldati, A. and Andreussi, P.: Artificial neural network approach to flood forecasting in the River Arno, Hydrol. Sci. J., 48(June), 381–398, doi:10.1623/hysj.48.3.381.45286, 2003.

Carisi, F., Schröter, K., Domeneghetti, A., Kreibich, H. and Castellarin, A.: Development and assessment of uni-and multi-variable flood loss models for Emilia-Romagna (Italy), , 2015(October), doi:10.5194/nhess-2017-342, 2018.

Carrera, L., Standardi, G., Bosello, F. and Mysiak, J.: Assessing direct and indirect economic impacts of a flood event through the integration of spatial and computable general equilibrium modelling, Environ. Model. Softw., 63, 109–122, doi:10.1016/j.envsoft.2014.09.016, 2015.

Chinh, D., Gain, A., Dung, N., Haase, D. and Kreibich, H.: Multi-Variate Analyses of Flood Loss in Can Tho City, Mekong Delta, Water, 8(1), 6, doi:10.3390/w8010006, 2015.

Cipolla, F., Guzzetti, F., Lolli, O., Pagliacci, S., Sebastiani, C. and Siccardi, F.: Catalogo delle località colpite da frane e da inondazioni: verso un utilizzo più maturo dell'informazione, , 1–9, 1998.

CRESME/CINEAS/ANIA: Definizione dei costi di (ri)costruzione nell'edilizia, edited by CINEAS., 2014.

D'Alpaos, L., Brath, A. and Fioravante, V.: Relazione tecnico-scientifica sulle cause del collasso dell' argine del fiume Secchia avvenuto il giorno 19 gennaio 2014 presso la frazione San Matteo., 2014.

Domeneghetti, A., Carisi, F., Castellarin, A. and Brath, A.: Evolution of flood risk over large areas: Quantitative assessment for the Po river, J. Hydrol., doi:10.1016/j.jhydrol.2015.05.043, 2015.

Dottori, F., Figueiredo, R., Martina, M. L. V., Molinari, D. and Scorzini, A. R.: INSYDE: A synthetic, probabilistic flood damage model based on explicit cost analysis, Nat. Hazards Earth Syst. Sci., 16(12), 2577–2591, doi:10.5194/nhess-16-2577-2016, 2016.

EASAC: Extreme weather events in Europe. Preparing for climate change adaptation: an update on EASAC's 2013 study., 2018.

EEA: Flood risks and environmental vulnerability - Exploring the synergies between floodplain restoration, water policies and thematic policies., 2016.

Elmer, F., Thieken, a. H., Pech, I. and Kreibich, H.: Influence of flood frequency on residential building losses, Nat. Hazards Earth Syst. Sci., 10(10), 2145–2159, doi:10.5194/nhess-10-2145-2010, 2010.

Essenfelder, A. H.: Climate Change and Watershed Planning: Understanding the Related Impacts and Risks, Universita' Ca' Foscari Venezia., 2017.

European Environment Agency: Mapping the impacts of recent natural disasters and technological accidents in Europe - An overview of the last decade., 2010.

Feranec, J. Ot'ahel', J.: Final version of the 4th level CORINE land cover classes at scale 1:50000, Bratislava., 1998.

Figueiredo, R., Schröter, K., Weiss-Motz, A., Martina, M. L. V. and Kreibich, H.: Multi-model ensembles for assessment of flood losses and associated uncertainty, Nat. Hazards Earth Syst. Sci., 18(5), 1297–1314, doi:10.5194/nhess-18-1297-2018, 2018.

Geofabrik GmbH: OpenStreetMap data extracts, [online] Available from: http://download.geofabrik.de/ (Accessed 30 March 2017), 2018.

Giacinto, G. and Roli, F.: Design of effective neural network ensembles for image classification purposes, Image Vis. Comput., 19(9–10), 699–707, doi:10.1016/S0262-8856(01)00045-2, 2001.

Gissing, A. and Blong, R.: Accounting for variability in commercial flood damage estimation, Aust. Geogr., 35(2), 209–222, doi:10.1080/0004918042000249511, 2004.

Govi, M. and Turitto, O.: Casistica storica sui processi d'interazione delle correnti di piena del Po con arginature e con elementi morfotopografici del territorio adiacente, Sci. e vita nel momento attuale, V, 105–160, 2000.

Hagan, M. T. and Menhaj, M. B.: Training Feedforward Networks with the Marquardt Algorithm, IEEE Trans. Neural Networks, 5(6), 989–993, 1994.

Hallegatte, S.: An adaptive regional input-output model and its application to the assessment of the economic cost of Katrina, Risk Anal., 28(3), 779–99, doi:10.1111/j.1539-6924.2008.01046.x, 2008.

Han, J.: Application of Artificial Neural Networks for Flood Warning Systems, North Carolina State University., 2002.

Hasanzadeh Nafari, R., Amadio, M., Ngo, T. and Mysiak, J.: Flood loss modelling with FLF-IT: a new flood loss function for Italian residential structures, Nat. Hazards Earth Syst. Sci., 17(7), 1047–1059, doi:10.5194/nhess-17-1047-2017, 2017.

Haykin, S.: Neural Networks: A Comprehensive Foundation, 2nd ed., Prentice Hall, Inc., Upper Saddle River, NJ, USA., 2001.

Heermann, P. D. and Khazenie, N.: Classification of multispectral remote sensing data using a back-propagation neural network, IEEE Trans. Geosci. Remote Sens., 30(1), 81–88, 1992.

Hsieh, W. W. and Tang, B.: Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography, Bull. Am. Meteorol. Soc., 79(9), 1855–1870, doi:10.1175/1520-0477(1998)079<1855:ANNMTP>2.0.CO;2, 1998.

Huizinga, H. J.: Flood damage functions for EU member states, Technical report implemented in the framework of the contract # 382441-F1SC awarded by the European Commission - Joint Research Centre., 2007.

Huizinga, J., De Moel, H. and Szewczyk, W.: Methodology and the database with guidelines Global flood depth-damage functions., 2017.

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

Jonkman, S. N., Bočkarjova, M., Kok, M. and Bernardini, P.: Integrated hydrodynamic and economic modelling of flood damage in the Netherlands, Ecol. Econ., 66(1), 77–90, doi:10.1016/j.ecolecon.2007.12.022, 2008.

Koks, E. E., Carrera, L., Jonkeren, O., Aerts, J. C. J. H., Husby, T. G., Thissen, M., Standardi, G. and Mysiak, J.: Regional disaster impact analysis: comparing Input-Output and Computable General Equilibrium models, Nat. Hazards Earth Syst. Sci. Discuss., 3(11), 7053–7088, doi:10.5194/nhessd-3-7053-2015, 2015.

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

Kreibich, H., Thieken, A. H., Petrow, T., Müller, M. and Merz, B.: Flood loss reduction of private households due to building precautionary measures – lessons learned from the Elbe flood in August 2002, Nat. Hazards Earth Syst. Sci., 5(1), 117–126, doi:10.5194/nhess-5-117-2005, 2005.

Kreibich, H., Piroth, K., Seifert, I., Maiwald, H., Kunert, U., Schwarz, J., Merz, B. and Thieken, a. H.: Is flow velocity a significant parameter in flood damage modelling?, Nat. Hazards Earth Syst. Sci., 9(5), 1679–1692, doi:10.5194/nhess-9-1679-2009, 2009.

Kreibich, H., Botto, A., Merz, B. and Schröter, K.: Probabilistic, Multivariable Flood Loss Modeling on the Mesoscale with BT-FLEMO, Risk Anal., 37(4), 774–787, doi:10.1111/risa.12650, 2017.

Lastoria, B., Simonetti, M. R., Casaioli, M., Mariani, S. and Monacelli, G.: Socio-economic impacts of major floods in Italy from 1951 to 2003, Adv. Geosci., 7, 223–229, doi:10.5194/adgeo-7-223-2006, 2006.

Liaw, A. and Wiener, M.: Classification and Regression by randomForest, R News, 2(3), 2002.

Luino, F., Cirio, C. G., Biddoccu, M., Agangi, a., Giulietto, W., Godone, F. and Nigrelli, G.: Application of a model to the evaluation of flood damage, Geoinformatica, 13(3), 339–353, doi:10.1007/s10707-008-0070-3, 2009.

Maier, H. R. and Dandy, G. C.: Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications, Environ. Model. Softw., 15(1), 101–124, doi:10.1016/S1364-8152(99)00007-9, 2000.

Masoero, A., Claps, P., Asselman, N. E. M., Mosselman, E. and Di Baldassarre, G.: Reconstruction and analysis of the Po River inundation of 1951, Hydrol. Process., 27(9), 1341–1348, doi:10.1002/hyp.9558, 2013.

McBean, E., Fortin, M. and Gorrie, J.: A critical analysis of residential flood damage estimation curves, Can. J. Civ. Eng., 13(1), 86–94, doi:10.1139/l86-012, 1986.

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

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

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

Messner, F., Penning-rowsell, E., Green, C., Tunstall, S., Veen, A. Van Der, Tapsell, S., Wilson, T., Krywkow, J., Logtmeijer, C., Fernández-bilbao, A., Geurts, P. and Haase, D.: Evaluating flood damages: guidance and recommendations on principles and methods, Flood Risk Manag. Hazards, Vulnerability Mitig. Meas., 189, 2007.

Meyer, V. and Messner, F.: National flood damage evaluation methods : A review of applied methods in England, the Netherlands, the Czech republic and Germany, , 49, 2005.

de Moel, H. and Aerts, J. C. J. H.: Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates, Nat. Hazards, 58(1), 407–425, doi:10.1007/s11069-010-9675-6, 2011.

de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E. and Ward, P. J.: Flood risk assessments at different spatial scales, Mitig. Adapt. Strateg. Glob. Chang., 20(6), 865–890, doi:10.1007/s11027-015-9654-z, 2015.

Molinari, D. and Scorzini, A. R.: On the influence of Input data quality to flood damage estimation: the performance of the INSYDE model, Water, 9(9), 688, doi:10.3390/w9090688, 2017.

Molinari, D., Menoni, S., Aronica, G. T., Ballio, F., Berni, N., Pandolfo, C., Stelluti, M. and Minucci, G.: Ex post damage assessment: an Italian experience, Nat. Hazards Earth Syst. Sci., 14(4), 901–916, doi:10.5194/nhess-14-901-2014, 2014.

Molinari, D., De Bruijn, K., Castillo, J., Aronica, G. T. and Bouwer, L. M.: Review Article: Validation of flood risk models: current practice and innovations, , (August), 1–18, doi:10.5194/nhess-2017-303, 2017.

Molinari, D., Radice, A., Agosti, A., Crippa, J., Scorzini, A. R., Dazzi, S., Bertuzzi, F. and Vacondio, R.: Modellazione idraulica dell'evento alluvionale che ha interessato la città di Lodi nel Novembre 2002, in Atti del XXXVI Convegno di Idraulica e Costruzioni Idrauliche. ISBN: 9788894379907, Ancona., 2018.

Oliveri, E. and Santoro, M.: Estimation of urban structural flood damages: the case study of Palermo, Urban Water, 2(3), 223–234, doi:10.1016/S1462-0758(00)00062-5, 2000.

Van Ootegem, L., Verhofstadt, E., Van Herck, K. and Creten, T.: Multivariate pluvial flood damage models, Environ. Impact Assess. Rev., 54, 91–100, doi:10.1016/j.eiar.2015.05.005, 2015.

Paprotny, D., Sebastian, A., Morales-Nápoles, O. and Jonkman, S. N.: Trends in flood losses in Europe over the past 150 years, Nat. Commun., 9(1), 1985, doi:10.1038/s41467-018-04253-1, 2018.

Penning-Rowsell, E., Johnson, C., Tunstall, S., Morris, J., Chatterton, J., Green, C., Koussela, K. and Fernandez-bilbao, A.: The Benefits of Flood and Coastal Risk Management: a Handbook of Assessment Techniques, Middlesex Univ. Press, 89, doi:10.1596/978-0-8213-8050-5, 2005.

Pistrika, A. K. and Jonkman, S. N.: Damage to residential buildings due to flooding of New Orleans after hurricane Katrina, Nat. Hazards, 54(2), 413–434, doi:10.1007/s11069-009-9476-y, 2010.

Rossetti, S., Cella, O. W. and Lodigiani, V.: Studio idrologico-idraulico del tratto di F. Adda inserito nel territorio comunale di Lodi, Milano., 2010.

Scawthorn, C., Flores, P., Blais, N., Seligson, H., Tate, E., Chang, S., Mifflin, E., Thomas, W., Murphy, J., Jones, C. and Lawrence, M.: HAZUS-MH Flood Loss Estimation Methodology. II. Damage and Loss Assessment, Nat. Hazards Rev., 7(2), 72–81, doi:10.1061/(ASCE)1527-6988(2006)7:2(72), 2006.

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

Scorzini, A. R. and Frank, E.: Flood damage curves: new insights from the 2010 flood in Veneto, Italy, J. Flood Risk Manag., n/a-n/a, doi:10.1111/jfr3.12163, 2015.

Scorzini, A. R. and Leopardi, M.: River basin planning: from qualitative to quantitative flood risk assessment: the case of Abruzzo Region (central Italy), Nat. Hazards, 88(1), 71–93,

doi:10.1007/s11069-017-2857-8, 2017.

Sluijs, L., Snuverink, M., van den Berg, K. and Wiertz, A.: Schadecurves industrie ten gevolge van overstroming, Tebodin Consultant, RWS DWW, Den Haag., 2000.

Smith, D.: Flood damage estimation. A review of urban stage-damage curves and loss function, Water SA, 20(3), 231–238, 1994.

Spekkers, M. H., Kok, M., Clemens, F. H. L. R. and ten Veldhuis, J. A. E.: Decision-tree analysis of factors influencing rainfall-related building structure and content damage, Nat. Hazards Earth Syst. Sci., 14(9), 2531–2547, doi:10.5194/nhess-14-2531-2014, 2014.

Steffler, P. and Blackburn, J.: River2D – Two-dimensional depth averaged model of river hydrodynamics and fish habitat. [online] Available from: www.river2d.ca, 2002.

Tesfamariam, S. and Liu, Z.: Earthquake induced damage classification for reinforced concrete buildings, Struct. Saf., 32(2), 154–164, doi:10.1016/j.strusafe.2009.10.002, 2010.

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

Thieken, A. H., Olschewski, A., Kreibich, H., Kobsch, S. and Merz, B.: Development and evaluation of FLEMOps – a new Flood Loss Estimation MOdel for the private sector, Flood Recover. Innov. Response, WIT Press, 315–324, 2008.

Thieken, a. H., Ackermann, V., Elmer, F., Kreibich, H., Kuhlmann, B., Kunert, U., Maiwald, H., Merz, B., Müller, M., Piroth, K., Schwarz, J., Schwarze, R., Seifert, I. and Seifert, J.: Methods for the evaluation of direct and indirect flood losses, RIMAX Contrib. 4th Int. Symp. Flood Def., 1–10, 2009.

Tobin, G. A.: The levee love affair: a stormy relationship?, J. Am. Water Resour. Assoc., 31(3), 359–367, doi:10.1111/j.1752-1688.1995.tb04025.x, 1995.

Vacondio, R., Dal Palù, A. and Mignosa, P.: GPU-enhanced finite volume shallow water solver for fast flood simulations, Environ. Model. Softw., 57, 60–75, doi:10.1016/j.envsoft.2014.02.003, 2014.

Vacondio, R., Aureli, F., Ferrari, A., Mignosa, P. and Dal Palù, A.: Simulation of the January 2014 flood on the Secchia River using a fast and high-resolution 2D parallel shallow-water numerical scheme, Nat. Hazards, 80(1), 103–125, doi:10.1007/s11069-015-1959-4, 2016.

Vacondio, R., Dal Palù, A., Ferrari, A., Mignosa, P., Aureli, F. and Dazzi, S.: A non-uniform efficient grid type for GPU-parallel Shallow Water Equations models, Environ. Model. Softw., 88, 119–137, doi:10.1016/J.ENVSOFT.2016.11.012, 2017.

Vogel, K., Riggelsen, C., Scherbaum, F., Schroeter, K., Kreibich, H. and Merz, B.: Challenges for Bayesian Network Learning in a Flood Damage Assessment Application, in 11th International Conference on Structural Safety & Reliability, pp. 3123–3130., 2013.

Wagenaar, D., de Jong, J. and Bouwer, L. M.: Data-mining for multi-variable flood damage modelling with limited data, Nat. Hazards Earth Syst. Sci. Discuss., 1–23, doi:10.5194/nhess-2017-7, 2017a.

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

Wagenaar, D. J., De Bruijn, K. M., Bouwer, L. M. and De Moel, H.: Uncertainty in flood damage estimates and its potential effect on investment decisions, Nat. Hazards Earth Syst. Sci., 16(1), 1–14, doi:10.5194/nhess-16-1-2016, 2016.

Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S. and Bai, X.: Flood hazard risk assessment

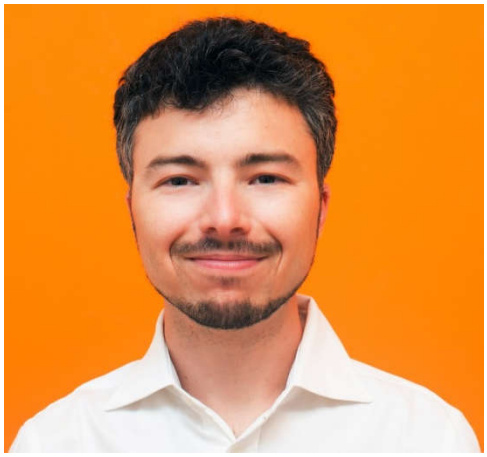
model based on random forest, J. Hydrol., 527, 1130–1141, doi:10.1016/j.jhydrol.2015.06.008, 2015.

Yu, H. and Wilamowski, B. M.: Levenberg-Marquardt Training, Intell. Syst., 5, 12–18, doi:10.1201/b10604-15, 2011.

Zampetti, G., Ottaviani, F. and Minutolo, A.: I costi del rischio idrogeologico, Dossier Legambiente, Roma., 2012.

Zhou, Q., Panduro, T. E., Thorsen, B. J. and Arnbjerg-Nielsen, K.: Verification of flood damage modelling using insurance data, Water Sci. Technol., 68(2), 425, doi:10.2166/wst.2013.268, 2013.

ABOUT THE AUTHOR



Mattia Amadio was born in Venice, Italy, on December 1st, 1984. He obtained a BSc degree in Environmental Sciences in 2009 and a MSc degree in Sustainable Development in 2012, both from the Ca' Foscari University of Venice. He developed his master thesis on flood risk during an internship at the Euro-Mediterranean Centre on Climate Change (CMCC), and since then he has worked as researcher within the Risk Assessment and Adaptation Strategies group led by Dr. Mysiak.

During these years his research revolved around the development and optimisation of a general flood risk modelling framework that could be effectively applied for estimating the expected cost of flood hazard according to current and future climatic scenarios.
The development of updated, validated damage models has been at the core of his work, starting from the spatial representation of economic value, to the characterisation of hazard magnitude and the elaboration of advanced damage models. In addition to that, part of his effort has been dedicated to link direct damage and indirect losses at interregional level by combining GIS impact analysis with macroeconomic modelling. These instruments can ultimately provide a measure which is critical to inform risk management when evaluating climate adaptation strategies. His future research will include the investigation of impacts from multiple climate-related hazards.

List of publications

Amadio, M., Scorzini A.R., Carisi F., Hessenfelder A.H., Domeneghetti A., Mysiak, J., Molinari, D., Castellarin, A. (2018). Validation of a multivariable flood damage model using machine learning. Submitted to Special Issue "Flood risk assessment and management" in NATURAL HAZARDS AND EARTH SYSTEM SCIENCES - ISSN:1684-9981

Amadio, M., Mysiak, J. Marzi, S. (2018). Mapping Socioeconomic Exposure for Flood Risk Assessment in Italy. In RISK ANALYSIS. doi:10.1111/risa.13212

Mysiak, J., Torresan, S., Bosello, F., Mistry, M., **Amadio, M.,** Marzi, S., Furlan, Sperotto, A. (2018). Climate risk index for Italy. Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 376 (2121), 20170305. http://doi.org/10.1098/rsta.2017.0305

Hasanzadeh Nafari, R., **Amadio, M.**, Ngo, T., Mysiak, J. (2017). Flood loss modelling with FLF-IT: a new flood loss function for Italian residential structures. DOI:10.5194/nhess-17-1047-2017. pp.1047-1059. In NATURAL HAZARDS AND EARTH SYSTEM SCIENCES - ISSN:1684-9981 vol. 17 (7)

Amadio, M., Mysiak, J., Carrera, L., Koks, E. (2016). Improving flood damage assessment models in Italy. In NATURAL HAZARDS 82, 2075–2088. doi:10.1007/s11069-016-2286-0

Amadio, M., Dell'Aquila, V. Mysiak, J. (2014). Multi-hazard assessment of water security in the context of the Caribbean Small Island Developing States. In European Water 47:45-54, EWRA 2014.

Amadio, M., Mysiak, J., Pecora, S., Agnetti, A. (2013). Looking Forward from the Past: Assessing the Potential Flood Hazard and Damage in Polesine Region by Revisiting the 1950 Flood Event. FEEM Working Paper No. 99.2013. http://dx.doi.org/10.2139/ssrn.2362900

ESTRATTO PER RIASSUNTO DELLA TESI DI DOTTORATO

Studente: Mattia Amadio

Matricola: 825260

Dottorato: Science and Management of Climate Change

Ciclo: 31°

Titolo della tesi: Integrated framework for flood risk assessment in Italy

Abstract:

This thesis is related to climate-risk management and adaptation. Specifically, it focuses on the assessment of flood risk to preventively understand the costs of hazard scenarios in a changing climate in order to timely elaborate adequate adaptation policies. Italy is a flood-prone country that suffers the highest economic impact in all of the EU. Despite this, there is no established framework or method at national level to estimate flood risk, and existing country-wide assessments are very broad-brushed. This is critical considering that impacts from extreme meteorological events are expected to increase by 2050 in Europe.

The objective of the thesis is to develop an improved methodology for the assessment of flood risk combining the most updated and reliable data available by means of advanced spatial and statistical approaches. By reducing the uncertainty typical of simple customary methods, the improved flood risk model can be used to translate any change in flood hazard probability and magnitude into variation of Expected Annual Damage. Key for the improvement of damage modelling is the collection and analysis of empirical data from observed flood events; starting from a large dataset collected after the 2014 flood on the Secchia basin (Emilia-Romagna), the thesis shows different approaches, more and more sophisticated, to elaborate the available information into a prognostic tool that can be reliably employed for risk management. Starting from a general investigation on the performances of transferred univariable models (depth-damage curves) over different damage categories (Ch.2), my research proceeded with the development of an empirical-base univariable model by using a statistical calibration procedure (Ch.3). In my third study (Ch.4) I collected heterogeneous country-wide data (e.g. land use, soil sealing, population and building census, cadastral information, production value) and combined them by using a dasymetric approach in order to draw a detailed and homogeneous representation of exposure in terms of asset, population, GDP, and social vulnerability. Exposure to different hazard scenarios is then estimated in relation to JRC flood hazard modelling. My last study (Ch.5) takes a far more advanced step in the identification of a tool that can be practically employed for country-wide risk assessment by validating an innovative multivariable, synthetic damage model for residential structures by means of machine learning.

Firma dello studente