Can macro- or meso-scale coping capacity variables improve the classification of building flood losses?

Can macro- or meso-scale coping capacity variables improve the classification of building flood losses?

D. Rodriguez Castro[1], M. Cools[2], S. Roucour[1], P. Archambeau[1], D. Molinari[3], A.R. Scorzini[4], C. Dessers[1], S. Erpicum[1], M. Pirotton[1], J. Teller[2], B. Dewals[1].

[1]Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, 4000, Belgium
[2]Local Environment Management and Analysis (LEMA), University of Liège, Liège, 4000, Belgium
[3]Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, 20133, Italy
[4]Department of Civil, Environmental and Architectural Engineering, University of L'Aquila, L'Aquila, 67100, Italy

Abstract

This study proposes a novel approach to improve the classification of severe building losses caused by river floods (i.e., identification of buildings with high flood damages). In addition to traditional variables reflecting flood hazard and building vulnerability, we investigate the impact of coping capacity variables (i.e., variables accounting for the preparedness and disaster response of the population and management authorities). These coping capacity variables are evaluated at three different scales: the building level (micro-scale), the census tract level (meso-scale), and the municipality level (macro-scale). Specifically, at the macro- and meso-scale these include: (*i*) the surprise effect (the ratio of the number of flooded buildings to the number of flooded buildings located in an official flood hazard area), (*ii*) the overwhelming effect (the fraction of flooded buildings compared to the total number of buildings within each census tract or municipalities), and (*iii*) flood rarity (the ratio of the peak discharge of the considered event to the 100-year flood peak). A binomial logistic regression model is used to classify flood losses based on field survey data from the extreme 2021 flood in eastern Belgium. Each variable is assessed for statistical significance, physical relevance, and multicollinearity. The results show that macro- and meso-scale coping capacity variables are insignificant in classifying building losses using the current dataset, suggesting that data on the building level are needed to reliably estimate building losses. Instead, the variables that contribute most to the classification are water depth, building footprint area, building finishing level and the heating system location. The performance of the classifier, measured by the AUC value, achieves an accuracy of 83%.

1 INTRODUCTION

As urbanisation and climate change lead to increasingly frequent and severe river floods (Dottori et al., 2018; Paprotny et al., 2018), the impacts of floods are anticipated to increase dramatically (European Environment Agency., 2024; IPCC, 2015). The need to improve our resilience is evidenced by the growing number of flood events worldwide and the associated losses (CRED, 2024). Traditional flood protection measures organised by authorities, typically through structural interventions, are insufficient to meet the challenges of increasing flood risks, as advocated in the EU Floods Directive (2007/60/EC). Therefore, it has become widely recognised as imperative to broaden the scope from flood protection to flood management, including not only the use of alternative measures (i.e., spatial planning, relocation, insurance, forecasting, early warning and communication) but also the active role of the citizens, for instance by implementing adaptation measures at the level of individual buildings (Begg et al., 2017; Merz et al., 2010).

Numerous studies have emphasised the importance of involving various actors at all stages of flood management (i.e., prevention, protection, preparedness, response, recovery and learning) to enhance the effectiveness of flood risk mitigation and adaptive response strategies (Kaufmann et al., 2016;

Mees et al., 2016, 2017; Pasquier et al., 2019; Renn, 2015; Van Eerd et al., 2017). To highlight the role of citizens and shared responsibility in flood protection, several authors have studied the implementation of personal protection measures, demonstrating the capacity of these measures to reduce flood losses (Endendijk et al., 2023; Kreibich et al., 2015; Thieken et al., 2005). Additionally, research into the factors that motivate citizens to take responsibility has identified risk perception and the perceived effectiveness of these measures as key drivers for action (Begg et al., 2017; O'Neill et al., 2016; Takao et al., 2004; Zhai et al., 2005).

In this study, we investigate whether the classification of flood losses can be improved by considering not only conventional variables representing flood hazard and building features (Carisi et al., 2018; Molinari et al., 2020; Scorzini et al., 2022; Thieken et al., 2008a) but also variables reflecting households' risk awareness and the degree of implementation of individual protection measures. We refer to this set of additional variables as 'coping capacity variables'. In addition, rather than focusing solely on the influence of coping capacity at the level of individual buildings (i.e. at the micro scale), we also examine whether the classification of flood losses can benefit from variables reflecting the coping capacity of the population and emergency services at a larger scale. Two scales are considered: the census tract level (meso-scale) and the municipality level (macro-scale). Three novel variables are introduced to reflect these broader dimensions of coping capacity.

The first variable, 'surprise effect', is defined as the ratio of the number of flooded buildings during the considered event to the number of flooded buildings located in an official flood hazard area. It is a proxy for reduced risk awareness, perception, and preparedness, often occurring in extreme flood events extending beyond the officially mapped flood hazard areas. Reduced risk awareness and preparedness among residents, emergency services or authorities can lead to higher flood losses.

The second variable, 'overwhelming effect', is defined as the ratio of the number of flooded buildings to the total number of buildings in each municipality or census tract. This variable reflects the strain on emergency services and volunteers in the event of widespread flooding relative to the size of the community, where a higher proportion of affected buildings within a community may reduce the availability of emergency response resources, leading to potentially higher losses, for instance, due to a lack of pumps or dehumidifiers in the days following a major flood.

The third variable, 'flood rarity', is computed as the ratio of the peak discharge of the considered flood event to the 100-year flood peak discharge. Similar to the 'surprise effect' variable, 'flood rarity' assesses whether the occurrence of a rare event (e.g., with a flood peak higher than the value commonly used for flood hazard mapping) results in higher building losses due to insufficient preparedness from the community, emergency services and authorities to respond to such an event.

To test the contribution of the 'coping capacity variables', we examine a severe flood event that impacted eastern Belgium in the summer of 2021. The results of the analysis are presented in two parts. First, we highlight significant differences between municipalities in terms of the surprise effect, overwhelming effect, and flood rarity. Second, using various variables collected during a field survey conducted after the July 2021 flood event, a binomial logistic regression model is applied to classify flood losses to residential buildings, incorporating explanatory variables reflecting hazard, vulnerability (such as building characteristics), and coping capacity at the building level, as well as the new coping capacity variables evaluated at macro- or meso-scales (i.e., at the level of municipalities or census tracts).

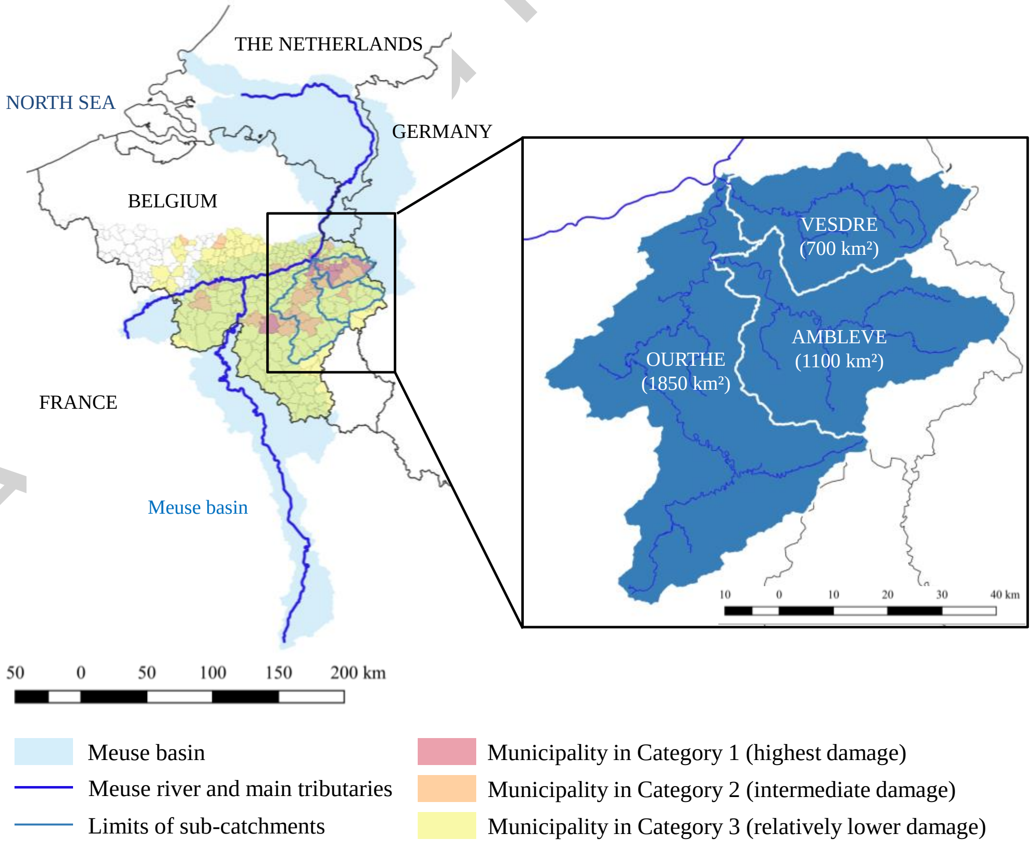
2 DATA AND METHODS

In this section, we introduce the considered case study (Section 2.1), the datasets used in the analyses (Section 0), and the applied statistical methods (Section 0).

2.1 Case study

In July 2021, the Bernd low-pressure system induced extreme floods over parts of Germany, Belgium, and the Netherlands (Mohr et al., 2023; Szönyi et al., 2022). In Belgium, the precipitation volumes reached up to 300 mm in 48 hours, which is comparable to three months of typical precipitation experienced in just two days (Journée et al., 2023). Water depths above 7 meters were observed, resulting in 39 casualties and severe losses to residential buildings, industries, and infrastructure, with more than 300 buildings washed away, around 200 to be demolished and roughly 3,000 partially destroyed, with an estimated total cost above 5 billion € (CSR, 2022).

In the southern part of Belgium (Walloon Region), 209 municipalities out of a total of 262 (i.e., 80 %) were affected by the 2021 flood. For operational purposes, the regional government classified these municipalities into three categories based on the severity of the losses. Category 1 includes the ten most impacted municipalities, Category 2 consists of 28 moderately affected municipalities, and the remaining ones are in Category 3. In the left panel in Figure 1, the municipalities are coloured according to this classification. It reveals that the most severely affected municipalities are mostly concentrated in the eastern part of the country, in the catchment of river Ourthe and some of its tributaries (Figure 1). This spatial distribution aligns with available rainfall data, indicating extreme precipitation volumes in this specific region (Journée et al., 2023). Nine out of the ten most affected municipalities are concentrated in the river Vesdre catchment (700 km^2), leading us to focus our analyses on this area (Bruwier et al., 2015; Cuvelier et al., 2018). In this area, peak discharges were as high as four times the 100-year flood discharge that had been estimated prior to the event, indicating that the 2021 event had return periods in the range of 10^4-10^5 years. (Archambeau et al., 2022). By including the 2021 event in the data used to calibrate the extreme value distributions, Archambeau et al. obtained a return period of 270 years for a station at Chaudfontaine, while return periods of 375 and 465 years were estimated for Verviers and Eupen respectively. This event was the highest on record in the region, which has made its management particularly challenging as it is regularly the case for unprecedented floods (Kreibich et al., 2022; Merz et al., 2021).



3

Figure 1. Map displaying international borders, the Meuse basin, main river courses, as well as municipalities in the south-east of Belgium (Walloon Region). The municipalities are coloured according to an official classification of the severity of the flood impacts.

The analysis presented here covers all municipalities along the main course of river Vesdre, except for Liège and Baelen: Chaudfontaine, Trooz, Pepinster, Verviers, Limbourg and Eupen (Figure 2). The municipality of Liège was excluded due to the combined influence of multiple water courses (Ourthe, Meuse) on flood hazard, rather than solely the River Vesdre. The municipality of Baelen was omitted from both field surveys and statistical analysis due to the limited number of impacted buildings (less than 3 % of the total number of affected buildings in the valley) and limited losses. Nonetheless, the municipality of Baelen has not been discarded in some preliminary analyses (Section 3.1).

The municipalities along the Vesdre Valley exhibit notable differences in terms of socio-economic context. Figure 3a in Supplement displays the median annual pre-tax income of the citizens, weighted by the number of buildings in each census tract in the entire municipality and on the tracts covered by the field survey detailed in Section 2.2.2). For instance, comparing the income the municipality of Chaudfontaine shows the highest income, and the municipality of Verviers the lowest one (Figure 3a in Supplement).

Analysing the median income of the entire municipality alongside that of the surveyed tracts reveals that the tracts representing the flooded districts tend to have a lower income than the overall municipality. This suggests that these flooded districts are home to less privileged residents. Differences in the income of all the municipality and the income in the surveyed districts varies from 10 to 22%. Additionally, average real estate values differ significantly across some municipalities (Figure 3b in the Supplement). Verviers has the lowest mean property value in the surveyed tracts, with a 26% increase compared to the whole municipality. Chaudfontaine is an exceptional case, where the difference between the flooded areas and the entire municipality is 86%, indicating a substantial disparity between those living close to the river and those farther away.

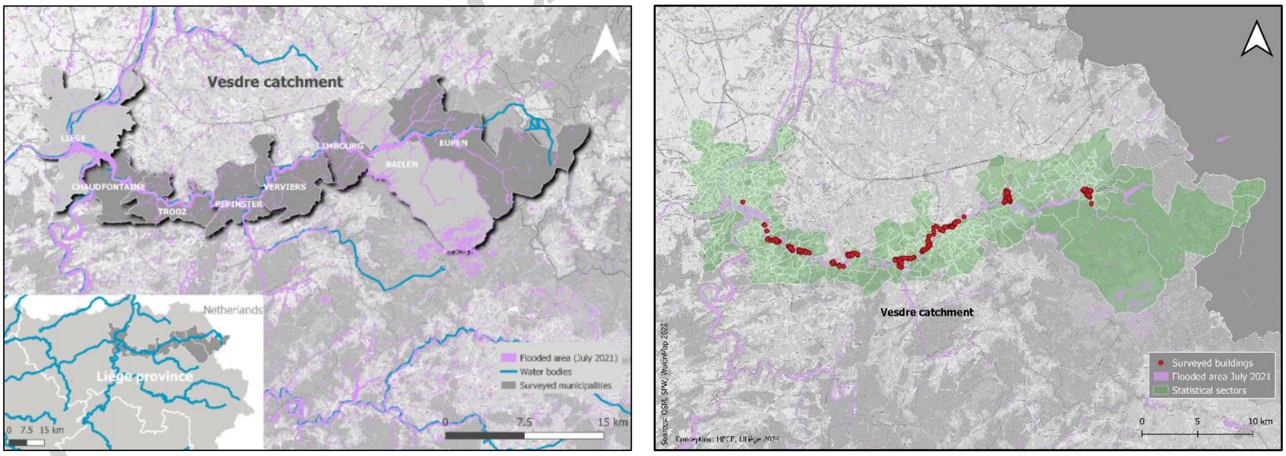


Figure 2. Left: Municipalities in the Vesdre catchment. Surveyed municipalities are represented in dark grey. Right: Census tracts of the surveyed municipalities. The surveyed buildings are displayed in dark red.

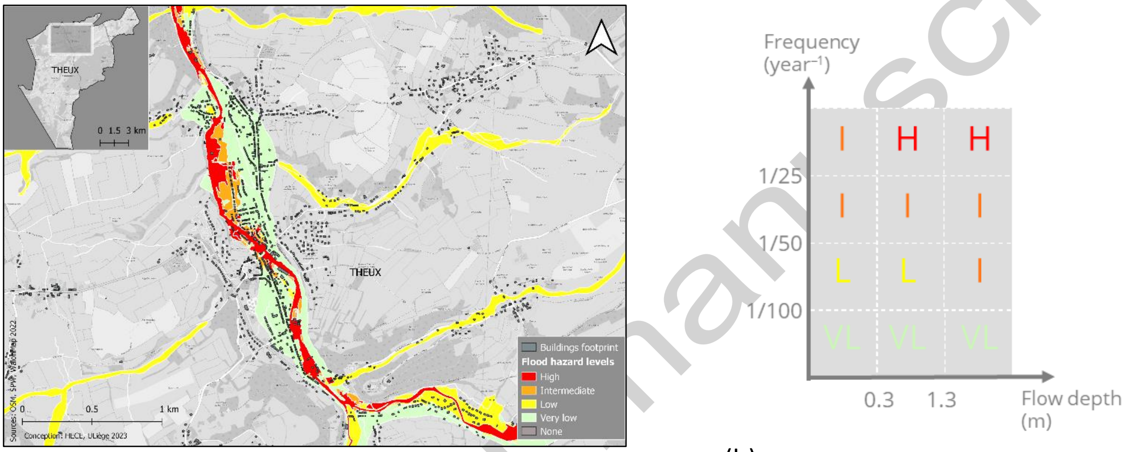
2.2 Data

2.2.1 Existing datasets

Five different existing datasets were used in this study.

Dataset #1: Official flood hazard maps

In the Walloon Region, official flood maps were prepared by combining water depth maps corresponding to flood scenarios of different return periods into one 'hazard map', classifying the hazard into four levels: high, intermediate, low and very low (Figure 3). A hazard level is assigned to each pixel depending on the estimated annual exceedance probability and the maximum water depth expected in the area, as shown in Figure 3b (Mustafa et al., 2018). The considered flood scenarios extend up to the 100-year flood, except for the 'very low hazard category', which does not match a particular annual exceedance probability. It corresponds to an estimate of the 100-year flood under a climate evolution scenario (Kitsikoudis et al., 2020). So, there is no flood frequency officially assigned to this particular scenario.



(a) (b)

Figure 3. (a) Example of an official flood hazard map; (b) Definition of high (H), intermediate (I), low (L), and very low (VL) as a function of water depth and flooding frequency.

Dataset #2: Field survey by the water authorities

Regional water authorities of the Service Public de Wallonie (hereafter referred to as SPW) surveyed outdoor water depths at the building level in the days following the July 2021 flood. Pointwise water depths at more than 15,000 locations were collected throughout the Vesdre Valley (Figure 4). Civil servants and volunteers registered water depth values based on outdoor water marks. The reference level for these measurements is not clearly documented, so it is likely that some inconsistent references were considered, such as either the street or sidewalk levels. The flood extent was estimated from the fusion of the pointwise observed water depths with photogrammetry and satellite imagery (*Zones Inondées IDW - Hauteur d'eau - Juillet 2021*, 2024).

5

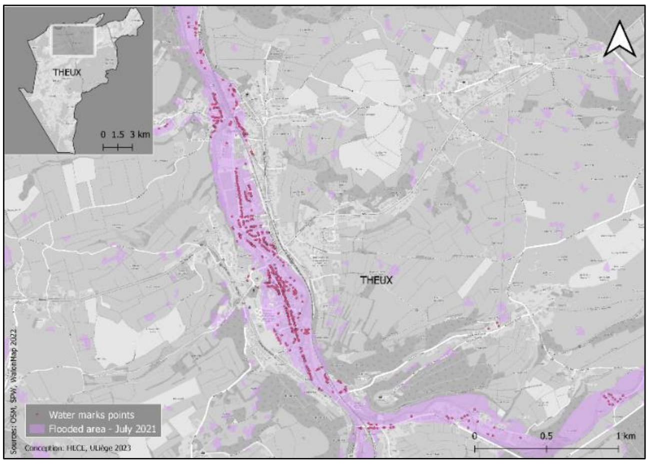


Figure 4. Observed water depths (circular symbols) and estimated flood extent (purple shaded area).

Dataset #3: Hydrological hindcast

Peak discharge estimates were obtained from a hydrological hindcast (Archambeau, 2022). Modelling was necessary since no measurement station captured the peak discharge, as all stations along the river Vesdre were damaged by the flood before the peak was reached. The hydrological computations were conducted based on a 100 m resolution grid, with a time step of 10 minutes. The model was forced by RADCLIM data (Goudenhoofdt, 2023).

The computed peak discharge in the municipality of Eupen (upper part of the valley, Figure 2) reaches about four times the official value for a 100-year flood (Figure 5). In the municipalities of Verviers and Chaudfontaine, the calculated peak is close to three times the previously estimated 100-year flood. As revealed by the hydrographs in Figure 5, all peaks occurred at night or early in the morning, which may have contributed to the surprise of the population.

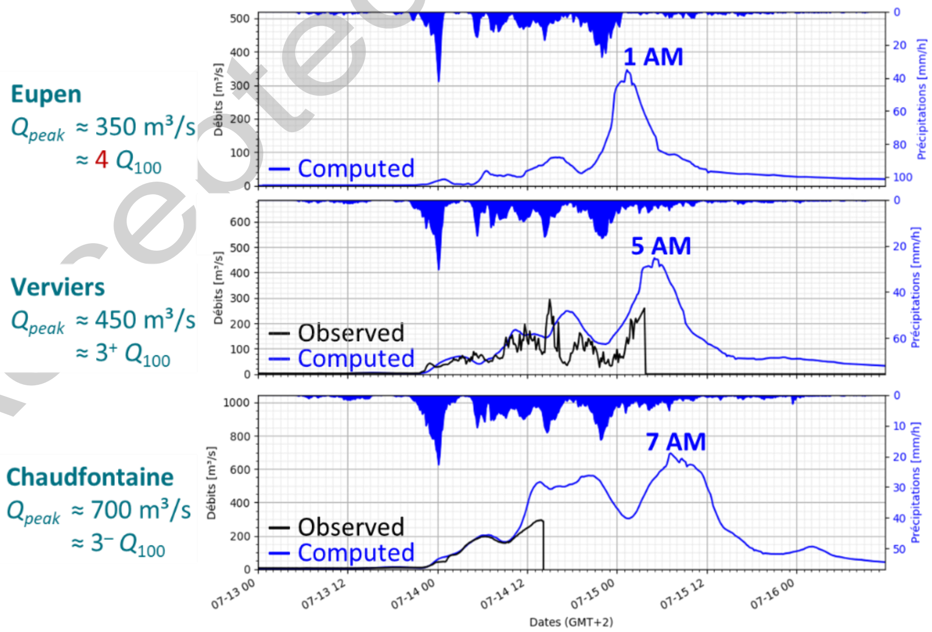


Figure 5. Example of reconstructed hydrographs in some of the surveyed municipalities.

Dataset #4: Index of social disparity

6

An index of social disparity (ISD) was evaluated using principal component analysis applied to 23 variables representative of the following four dimensions related to the population and its socio-economic level: (i) origin and nationality, (ii) taxable income, (iii) unemployment and participation rate, (iv) share of households with social security income (Poussard et al., 2021). As shown in Figure 6, ISD may take five different integer values, ranging from 1 to 5, with 1 corresponding to the most deprived population and 5 to the most affluent one (Poussard et al., 2021). This information is available at the census tract level. Variable ISD is used in the statistical analysis detailed in Section 2.3.1 and Table 1.

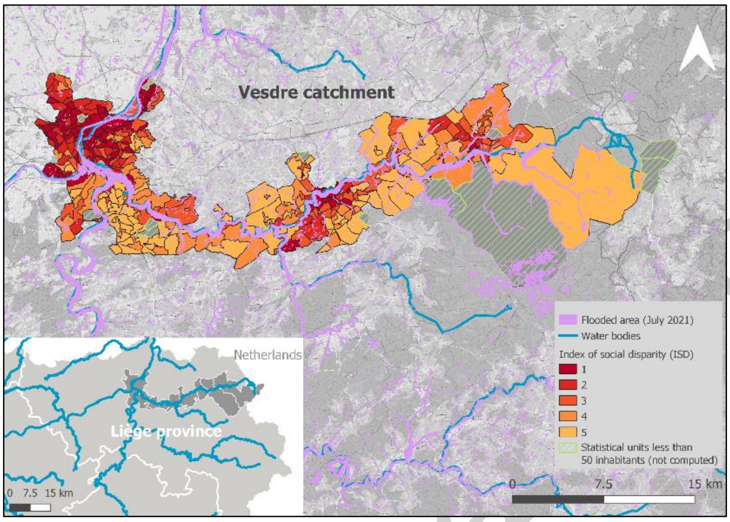


Figure 6.Index of social disparity (ISD) at the census tract level along the Vesdre valley.

Dataset #5: Buildings footprint

Polygons representing the shape, location and footprint area of the buildings in the flooded area were retrieved from the geoportal WalOnMap (https://geoportail.wallonie.be/), in which the vector layer 'Projet Informatique de Cartographie Continue' (PICC) was used. Figure 7 displays the footprint area of residential buildings surveyed by the University of Liège (see Sub-section 2.2.2) in each municipality of the Vesdre Valley. Surveyed buildings in the municipalities of Chaudfontaine and Eupen are characterised by comparatively larger footprint areas, with median values of 114 and 91 m^2, respectively. In contrast, the surveyed buildings in the municipality of Pepinster tend to have a lower footprint area, with a median value of 59 m^2.

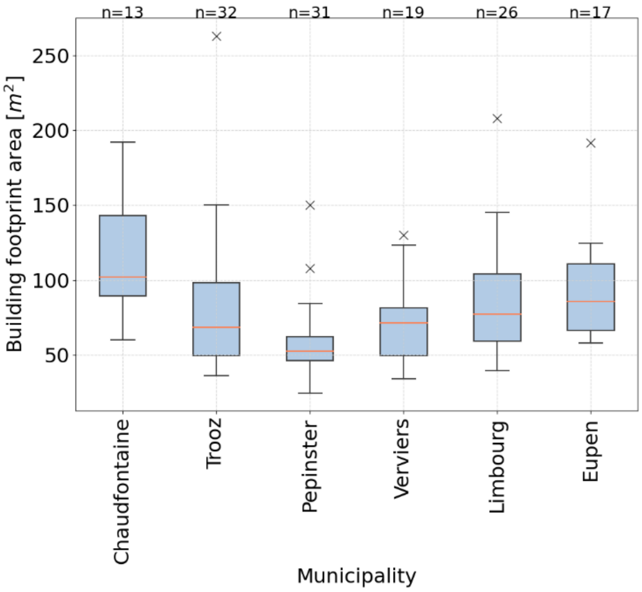


Figure 7. Footprint area of the surveyed buildings in the different municipalities. Notation *n* refers to the number of entries.

2.2.2 Flood loss surveys

In the aftermath of the 2021 flood, the University of Liège (ULiège) undertook a field survey of residential buildings in six municipalities of the Vesdre Valley (Section 2.1). This survey consisted of a questionnaire with households to collect data on direct flood losses and explanatory variables. The surveys were based on a questionnaire similar to those used in previous research (Ballio et al., 2018; Kreibich et al., 2017).

On-site surveys were preferred over online or phone surveys to maximise data quality and reliability. Some variables, such as water marks or construction type, could be verified on-site by the surveyors. In each municipality, the target area was delineated from the observed water depths and the flooded area contained in Dataset #2 (Sub-section 2.2.1).

To foster participation in the surveys, the population was informed in advance through posters displayed in each municipality and letters delivered door-to-door. Both the posters and the letters contained instructions on how the citizens could set an appointment through a text message, a phone call, or an online form accessible with a QR code. They could also wait for the researchers to visit their building. From the population with which contact was established, 39 % accepted to participate, leading to a total of 286 surveys. However, out of these 286 surveys, only 142 enabled the collection of monetary values of flood losses. In this study, we focus on this subset of surveys for which flood loss data is available.

Data from each survey was encoded in an online platform. To ensure data quality, this encoding was repeated twice by two different researchers. The two entries were automatically compared using a Python script, and the detected differences were corrected. Research data are anonymous, since personal data, such as building addresses, were recorded in a separate database, with a unique identifier linking corresponding entries in the two datasets.

As listed in Table 1, collected data contain building-level information on flood hazard, building features, coping capacity, and monetary losses. Descriptive statistics of the data are provided in Supplement (Section 5).

Flood hazard

Information on maximum water depth was collected in the form of ranges based either on water marks directly observed by the surveyors or on values reported by the respondents. This variable was converted into a continuous one by selecting the central value of each range (notation *h* in Table 1). The resulting distribution of water depth values reported by the respondents was compared to the distribution of water depths surveyed by SPW (Dataset #2 in Sub-section 2.2.1), which cannot be individually attributed to specific buildings but constitutes a considerably larger sample. The two distributions agree well, particularly in the higher range of water depths (Figure 8). The median value obtained from the surveys is 1.45 m (with the building's ground floor as a reference), while the median in SPW observations is 1.20m (no clearly documented reference). Notably, the SPW surveys include some values up to 5 m, while the highest reported value in ULiège surveys is 3.5 m. Some deviations are visible for lower values, which may be attributed to a priority given to flooded areas with higher registered water depths in the surveyors' choice of routes.

Since estimating the flow velocity is not straightforward for the population living in floodplains, two semi-quantitative indicators of flow velocity were collected during the surveys. First, the respondents were asked to give a score reflecting the velocity magnitude (notation v_1 in Table 1). It ranges from 1 (low velocity) to 6 (high velocity). Second, the respondents were asked to choose between the following textual descriptions: (i) I could easily stand in the flow; (ii) efforts were necessary to stand in the flow; (iii) the flow would have swept me away; or (iv) water was too deep to stand (Thieken et al., 2005). The responses were stored in a nominal variable, noted v_2 in Table 1. Two Boolean variables were used to record the presence or absence of sediments and contaminants. In Table 1, they are noted as *sed* and *c*, respectively.

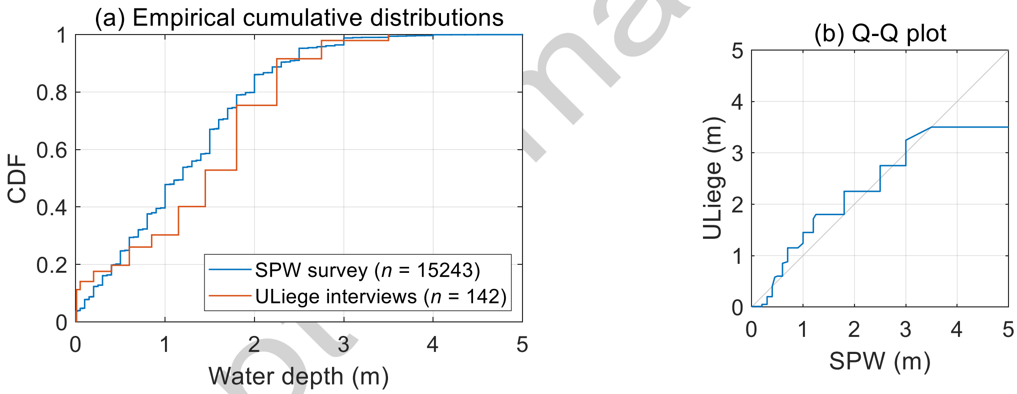


Figure 8. (a) Cumulative distribution functions of the water depths in the SPW and ULiège surveys for each municipality. (b) Quantile-quantile plots (Q-Q plots) of the same data.

Building features

Additional variables describing the building characteristics in the Vesdre Valley were collected in the field surveys.

A typical feature characterising many houses in the case study area is an elevation of the ground floor level compared to the street level, with a few steps to climb to reach the main door. This suggests a form of adaptation of the buildings situated in flood-prone areas. Information on the ground floor level was collected in the surveys and stored in a continuous variable noted as *gl* (Table 1). The median of the surveyed values of *gl* is 0.42 m.

Most buildings were single-family houses distributed on four levels: basement, ground floor, first floor and attic. This information was stored in two variables: building type (i.e., attached, detached or semi-detached building) registered as a nominal feature noted b_t, and the number of floors stored in the variable n_f (Table 1). Attached buildings were the most frequent in the considered area, with 49 % of the surveyed houses.

These residential buildings were typically old, low-rise masonry structures with basements primarily used for storage and housing the building's technical systems, such as heating and electrical controls. The type of construction material was stored in the nominal variable b_s. 36% of the surveyed buildings were constructed with clay bricks and 40% with concrete blocks. The height of the basement was recorded in the continuous variable b_h. The period for construction was recorded in the continuous variable y_y. In the survey data, 86% of the buildings were constructed before 1970.

During the field surveys, it was observed that, in most cases, the systems controls (i.e., electric panel, boiler for the heating system) were located in the basement, making them particularly prone to damage in case of flooding. Therefore, two additional variables were recorded to document the type of heating system p_d and p_t (Table 1).

As the building quality could influence the flood building losses (Dottori et al., 2016; Thieken et al., 2008), two additional variables were included to describe how luxurious and well-maintained the buildings were. This information was stored in ordinal variables, noted as f_l (standing for finishing level) and l_m (standing for level of maintenance), respectively. Each of them is represented by a score assigned by the surveyors, distinguished into three levels: 'high', 'medium', or 'low'. Such variables are not standard in flood loss modelling but were used in the model INSYDE-BE, the only flood loss model adapted for the Walloon Region in Belgium (Scorzini et al., 2022). Despite the drawback of being affected by some subjectivity, they constitute an attempt to capture two building vulnerability features which certainly influence the amount of flood losses.

Coping capacity

Four types of data were collected during the surveys to estimate the population coping capacity at the building level:

- a binary variable, *fe*, indicating whether the respondent had already experienced a flood before the 2021 event;
- another binary variable, *w*, recording whether the households reported having received an official warning before their house was flooded (considered official warnings included those issued by local authorities, such as fire brigade or police, by a federal warning system, BE-Alert, or through an online platform informing in real-time about the hydrological situation);
- the variable *pm,* representing the number of preparedness measures (i.e., consulting existing hazard maps, registration to national/regional warning system, searching for information on individual house flood protection, taking out insurance) implemented by the respondents;
- similarly, two additional variables were considered to account for the number of implemented mitigation measures: *smm* (counting the number of short-term actions, such as using a pump during the flood) and *lmm* (counting the number of long-term adaptations, such as dry-proofing the building).

Monetary losses

During the field surveys, the estimates of the total monetary losses (which consider flood losses in the building and its content) were collected (continuous variable *td* in Table 1). More than 80% of the reported losses were either based on invoices for the repair works carried out or on assessments made by an expert from the insurance company.

Figure 9 represents the dataset in the form of one boxplot per municipality. The data are also shown in the form of cumulative distribution functions in Supplement (Section 2). Although all municipalities are in the same region and were affected by the same hydrometeorological event, substantial differences can be observed between them. The lowest value of the median loss at the municipality level is 55,000 € (Eupen). This value is about 2.5 times lower than the highest median loss, i.e., 136,182 € (Trooz). This pattern in flood losses across municipalities does not match the pattern of flood magnitude, expressed in terms of peak discharge normalised by the 100-year flood peak (Figure 5). This pattern between municipalities is a motivation for exploring other building characteristics, as well as unconventional explanatory variables at the macro- and meso-scales, as introduced in the next section.

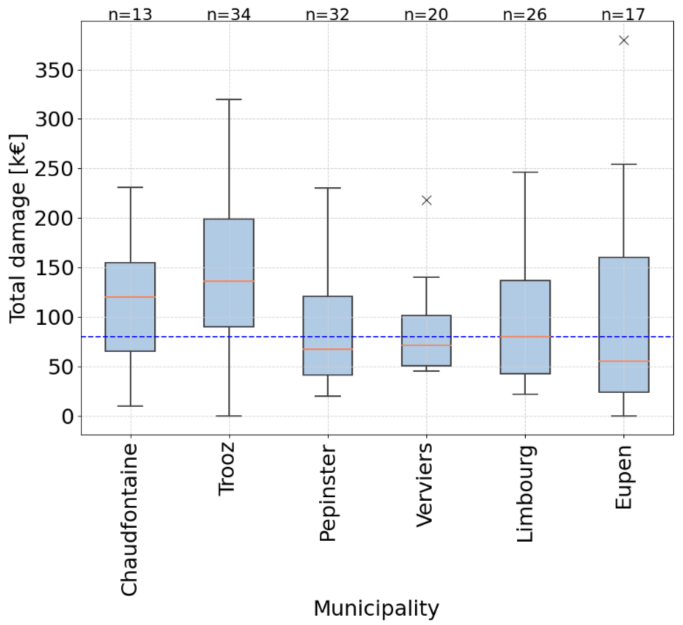


Figure 9. Reported total loss in surveyed residential buildings after the 2021 flood event. The horizontal dashed blue line represents the threshold of 80,000 € selected in the statistical analysis.

2.2.3 Macro- and meso-scale variables

Besides the building variables collected through field surveys (Section 2.2.2), we defined three macro- and meso-scale copying capacity variables (i.e., variables evaluated at the municipality and census tract level, respectively).

Surprise effect

For each municipality and census tract, we defined the variable *s* as the ratio of the number of flooded buildings in the considered event to the number of buildings located in official flood hazard zones:

$$s = \frac{\text{number of flooded buildings}}{\text{number of buildings in official flood hazard zones}} \tag{1}$$

We referred to this variable as a quantification of a 'surprise effect' as it relates to the number of affected buildings which were unlikely to be prepared for floods, being them not located within pre-event official flood hazard maps.

Overwhelming effect

We also introduced the variable *o* representing the ratio of the number of flooded buildings in the considered event to the total number of buildings in the municipality or census tract:

$$o = \frac{\text{number of flooded buildings}}{\text{total number of buildings (municipality or census tract)}} \tag{2}$$

In the following, this variable is referred to as a quantification of an 'overwhelming effect'. Indeed, local emergency services have a lower chance to effectively intervene if the number of flooded buildings is high compared to the size of the community. Similarly, the support provided by a community in the aftermath of a flood event is likely to become less comprehensive as the value of *o* rises (e.g., some of the political decision-makers, police officers, health experts, and construction workers were personally affected by the flood and not available for reacting and repairing).

Flood rarity

Several studies showed that direct flood losses correlate with the frequency of loss-inducing floods (Ghaedi et al., 2022; Merz et al., 2013; Schröter et al., 2014). In the considered case study, the peak discharge in the analysed municipalities ranged between about three and four times the corresponding 100-year flood peak (as estimated before the event). This suggests that the frequency of the 2021 flood varied across the municipalities. Due to the extreme nature of this flood, we considered that assigning a clear-cut flood frequency to the event was too uncertain. Therefore, we introduce the variable *r*, defined as the ratio of the peak discharge Q_p of the considered event to the 100-year flood peak Q_{100}:

$$r = \frac{Q_p}{Q_{100}} \tag{3}$$

This variable is referred to as 'flood rarity' and is only assessed at the macro-scale (i.e., at the municipality level). Estimating this variable at the meso-scale is not feasible because the locations of the gauging stations do not align with the census tracts, with fewer gauging stations compared to the number of census tracts.

2.3 Statistical analysis

This section introduces the statistical tool used to classify flood losses in the considered case study, along with the corresponding procedure for the statistical analysis.

2.3.1 Logistic regression

Binary logistic regression was utilised to classify data on total flood losses. This technique estimates the probability that a given input point belongs to a certain class, assuming that the input space can be separated into two regions (Murphy, 2022; Veloso et al., 2022). In the present case, the two regions correspond to lower or higher flood losses than a predefined threshold. The selected threshold is 80,000 €, which is close to the median of the total flood losses of the surveyed residential buildings. The calibration of the logistic regression model is detailed in the next sub-section.

The performance of a classification model can be evaluated by means of metrics such as sensitivity and specificity, which measure the true positive and true negative rates, respectively:

$$\text{sensitivity} = \frac{TP}{TP + FN} \tag{4}$$

12

$$\text{specificity} = \frac{TN}{TN + FP}$$

(5)

where TP refers to the number of true positives (i.e., the number of inputs for which the classifier predicts a total loss above the threshold, while the observed loss is indeed above the threshold), FN the number of false negatives, TN the number of true negatives, and FP the number of false positives.

The performance of a classification model varies with the value selected as a threshold in the estimated probability for an input to be classified as true (Murphy, 2022). By systematically varying this threshold between 0 and 1, a so-called ROC curve can be plotted. It displays the value of sensitivity as a function of the specificity for each possible value of the threshold applied to the probability estimated by the classification model. In turn, the ROC curve is often summarised in a single number (AUC), which is the area under the curve (Murphy, 2022). The closer the AUC to 1, the better the classifier performance.

2.3.2 Three-step strategy

The data analysis followed four steps. It started with a baseline classification model (Step 1). Next, it assessed the value of adding extra building features as well as macro- and meso-scale variables (Step 2). Finally, a classification model was considered based on all variables deemed significant in the first two steps (Step 3). Throughout the three steps, the absolute total monetary loss in residential buildings was considered as the dependent variable.

Step 1

In the initial step, all the variables were standardised, and the classification model was trained with two usual variables in flood loss modelling: water depth (h) and building footprint area (A). The building footprint area was included since greater absolute losses are expected in the case of a larger flooded building. The square of h was also included to capture non-linear effects (see Section 4 in Supplement). This model, including water depth (h), building footprint area (A) and the squared of h (h^2), served as a baseline for comparing the more advanced classifiers tested in the next steps.

Step 2

The goal of Step 2 was to identify which variables at the building level enable improvement of the baseline classification model (Step 1). Two criteria were used to include an individual additional variable in the classification model:

- the influence of the tested variable on the classification model must be statistically significant (i.e., $p < 0.05$);
- the sign of the estimate corresponding to the variable must align with physical sense (e.g., the sign of the estimate corresponding to an ordinal variable reflecting velocity magnitude must be positive, in line with a higher flow velocity leading to a higher probability for the flood loss to be above the considered threshold).

The significance of each nominal variable was assessed using the likelihood ratio test.

The logistic regression model was recalibrated multiple times, each time using the variables from Step 1 (A, h, h^2) along with one additional set of variables (Table 1). These additional sets included variables representing either hazard, vulnerability (building characteristics), coping capacity at micro-scale, or coping capacity at the meso- or macro-scale, as introduced in Section 2.2.3. Each set of variables was tested on a separate subsample. Indeed, the full dataset contained 142 observations of 31 variables. However, some values were missing in the dataset. Therefore, if all variables were tested simultaneously and observations containing at least one missing value of any of the variables were excluded, the remaining sample size would end up being high-dimensional (i.e., space data, in which the number of variables is too high compared to the number of observations), leading to substantial variability or instability in the estimation of the parameters of the classification model.

Once the variables from each group were individually tested, those appearing as statistically significant and having an estimate sign consistent with physics were pre-selected for incorporation in an improved classifier.

Step 3

Step 3 was based on the outcomes of Step 2: an improved classification model was calibrated by considering all variables found individually significant in the first two steps. The statistical significance of each variable in this improved model was reassessed taking into account the relationships between the pre-selected variables. If the significance of any pre-selected variable was below the defined threshold, it was discarded. Next, some of the variables were filtered out based on an assessment of multicollinearity using the variance inflation factor (VIF), with a threshold value of 5 as recommended by Schroeder et al., 1990. Only the remaining variables were retained to end up with a recommended classification model based on the available dataset.

3 RESULTS

In this section, the macro- and meso-scale variables introduced in Section 2.2.3 are first evaluated for each municipality in the case study (Section 3.1). Second, the classifier described in Section 0 is calibrated based on building features as well as macro- and meso-scale variables (Section 3.2).

3.1 Macro- and meso-scale variables

3.1.1 Surprise effect

Figure 10a illustrates the number of residential buildings located within the extend of the observed flooded area in each municipality. For each bar, the colours reflect the number of flooded residential buildings in the respective flood hazard zones (Dataset #1) or outside the officially mapped flood hazard zones.

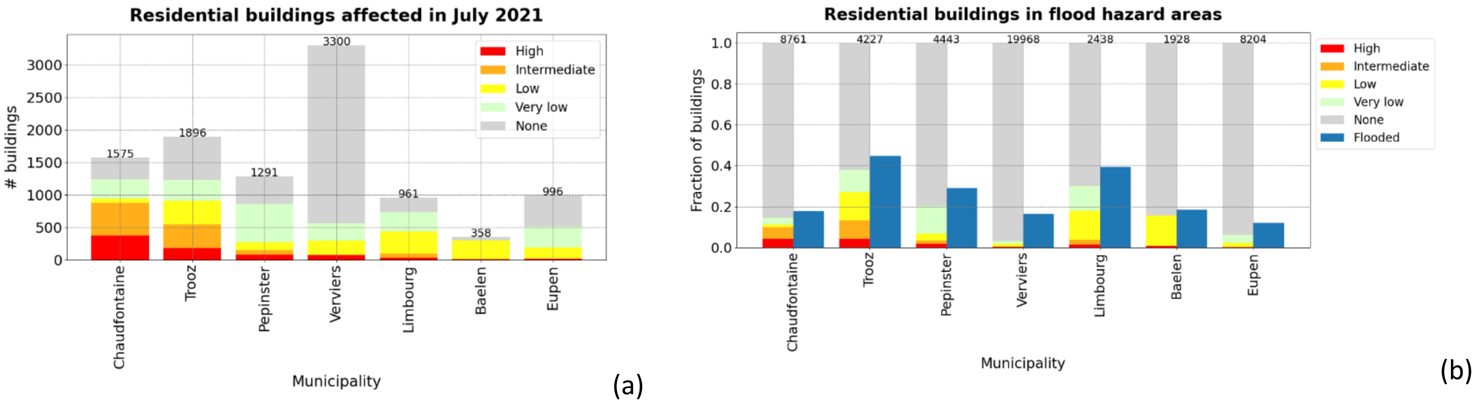
In most municipalities, the share of flooded residential buildings which are not in an officially mapped flood hazard zone (i.e., variable s at macro-scale) varies between 15 % and 35 %. In the uppermost municipality, Eupen, this share rises to almost 50 %. This may be related to the more extreme nature of the flood peak ratio (Q_p / Q_{100}) in this part of the valley: approximately four times the previously estimated 100-year flood, compared to about three times this value in the middle and lower parts of the valley (Dataset #3, Figure 5). Another exception is the municipality of Verviers, where the portion of flooded residential buildings located outside the official flood hazard zones exceeds 80%.

Table 1. Tested variables (C: continuous, O: ordinal, N: nominal, B: Binary)

		Variable		Type and range	N° of observations *
Hazard	1	h	Ground floor water depth	C: from 0.01 m to 3.50 m	142
	2	h^2	Water depth squared	C: squared value of h	142
	3	v_1	Flow velocity perception 1	O: 1=low velocity to 6=high velocity	135
	4	v_2	Flow velocity perception 2	N: 1=could easily stand. 2=efforts necessary to stand; 3=would be swept away; 4=deep water	134

	5	*sed*	Presence of sediments	B: 0 = no sediment; 1 = presence of sediments	142
	6	*c*	Presence of contaminants	B: 0 = no contaminant; 1=presence of contaminant	142
Vulnerability (building features)	7	*a*	Building footprint area	C: from 24 to 263 m^2	138
	8	g_l	Ground floor level	C: from 0.0 m to 1.7 m	107
	9	b_t	Building type	N: 1=Attached; 2=semi-detached, 3=detached	142
	10	n_f	Number of floors	C: from 1 to 5 floors	141
	11	b_s	Building structure	N: Reinforced concrete; concrete masonry; clay masonry; timber frame; prefabricated; natural stone; half-timbered [in French: colombage]	133
	12	b_h	Basement height	C: from 1.4 m to 3.0 m	99
	13	y_y	Year of construction	N: before 1875; 1875-1918; 1919-1945; 1946-1970; 1971-1990; 1991-2000; 2001-2010; after 2011	133
	14	p_d	Heating distribution system	B: 0 = distributed heating; 1 = central heating	141
	15	p_t	Heating system type	N: coal; gas; oil; electric; pellet	112
	16	l_m	Level of maintenance	N: high, medium, low	142
	17	f_l	Finishing level	N: high, medium, low	135
Micro-scale coping capacity	18	*hz*	Hazard zone	N: high; intermediate; low; very low; none	142
	19	*fe*	Previous flood experience	B: 0 = no flood experience; 1 = previous flood experience	141
	20	*pm*	Preparedness measures	C: from 1 to 4 preparedness measures were implemented	141
	21	*smm*	Short mitigation measures	C: 0 to 3 short mitigation measures implemented	141
	22	*lmm*	Long mitigation measures	C: 0 to 1 long mitigation measures implemented	141
	23	*isd*	Index of social disparity (Dataset #2)	O: from 1 to 4, being 1 = most deprived population to 5 = most affluent population	118
	24	*w*	Official warning	B: 0 = no official warning before the event; 1 = official warning before the event	125
Macro-scale coping capacity	25	*s*	Surprise effect	C: from 1.27 to 5.87	142
	26	*o*	Overwhelming effect	C: from 0.12 to 0.45	142
	27	*r*	Flood rarity	C: from 2.73 to 4.11	142
	28	*m*	Municipality	N: Chaudfontaine, Trooz, Pepinster, Verviers, Limbourg, Eupen	142
Meso-scale coping capacity	29	*s*	Surprise effect	C: from 1.0 to 169	142
	30	*o*	Overwhelming effect	C: from 0.017 to 0.78	142
Direct flood loss	31	*td*	Total losses	C: from 0 to 380,000 €	142

*Since all the people were not willing to answer all the questions, not all information is available for each survey



15

Figure 10. (a) Flooded buildings within the different flood hazard zones. (b) For each municipality, fraction of buildings in flood hazard zones compared to the fraction of flooded buildings.

These contrasting results between municipalities may be interpreted by taking a closer inspection at maps comparing the location of the flooded area and the flood hazard zones. In Figure 11a, the flooded area (represented in purple) and the flood hazard zones are compared for the municipality of Trooz. The map reveals that, in the western part of the municipality, the flooded area extends slightly beyond the flood hazard zones. This seems reasonable, as the estimated flood peak exceeds the 100-year flood discharge by a factor of three. In contrast, the highest flood scenario considered in the preparation of the maps is approximately 1.3 times the 100-year flood (Dataset #1).

In contrast, in the municipality of Verviers (Figure 11b), a whole district appears to have been flooded while being situated outside the official flood hazard zones. This situation is like a levee effect, as described by Di Baldassare et al. (2015). In Verviers, the Vesdre river is channelised over almost its whole course through the town. The design standard of this channelised river section is particularly high since none of the flood scenarios considered for hazard mapping led to river overflow. Hence, besides the main riverbed, virtually no part of the town of Verviers is mapped as a flood hazard zone, while the observed flooded area extends over a considerable width. This is a typical situation in which surprise effects are likely to have been substantial for both the inhabitant and the authorities.

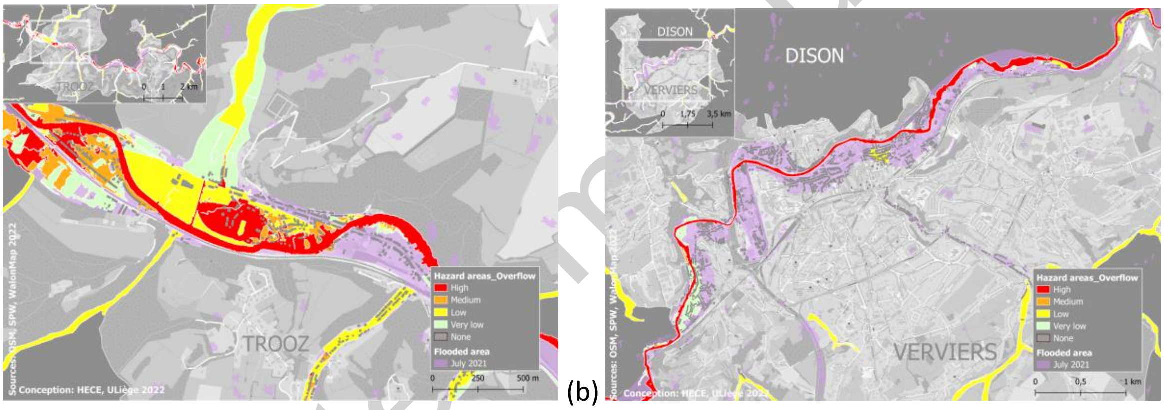
(a) (b)


Figure 11. Map of the municipality of Trooz (a) and Verviers (b) showing the observed flood extent (in purple) and the coverage of the flood hazard maps (red, orange, yellow and light green).

Figure 12 shows the number of sampled flood depth values *h* for each municipality (using Dataset #2), which fall into five classes. They are labelled as follows: up to ankle ($0 \text{ m} < h < 0.15 \text{ m}$), between ankle and knee ($0.15 \text{ m} \leq h < 0.45 \text{ m}$), between knee and waist ($0.45 \text{ m} < h < 1.0 \text{ m}$), between waist and shoulders ($1.0 \text{ m} < h < 1.5 \text{ m}$), and above shoulders ($h \geq 1.5 \text{ m}$). Like in Figure 10, each bar is coloured according to the official flood hazard zone in which the survey point is located. Figure 12 reveals that all classes of flood depths were observed in each municipality, and the class corresponding to the highest water depth (above shoulders) was the most populated in all municipalities but one. This reflects again the extreme nature of the flood event throughout the valley. In Verviers, which was pointed out in Figure 10a, about 800 points correspond to flood depths above the waist or shoulders while being surveyed outside any official flood hazard zone. Since the numbers in Figure 12 come from a sample of all flooded residential buildings, it indicates that a minimum (probably many more) of 800 residential buildings experienced flood depths above the waist while being located outside of any official flood hazard zone. This suggests that inhabitants in those flooded buildings were most probably unaware of being at risk of flooding and, therefore, unprepared. This reminds us of the need to manage residual risk in the case of channelised rivers or rivers equipped with high-standard flood defences.

The same analysis has been repeated at the meso-scale, i.e., by evaluating variable *s* for each census

16

tract (instead of for each municipality). The results are displayed in Figure 12 in Supplement. They reveal a considerably larger variability in the values of s when evaluated at the meso-scale ($1.05 \leq s \leq 110.5$) instead of the macro-scale ($1.27 \leq s \leq 5.87$).

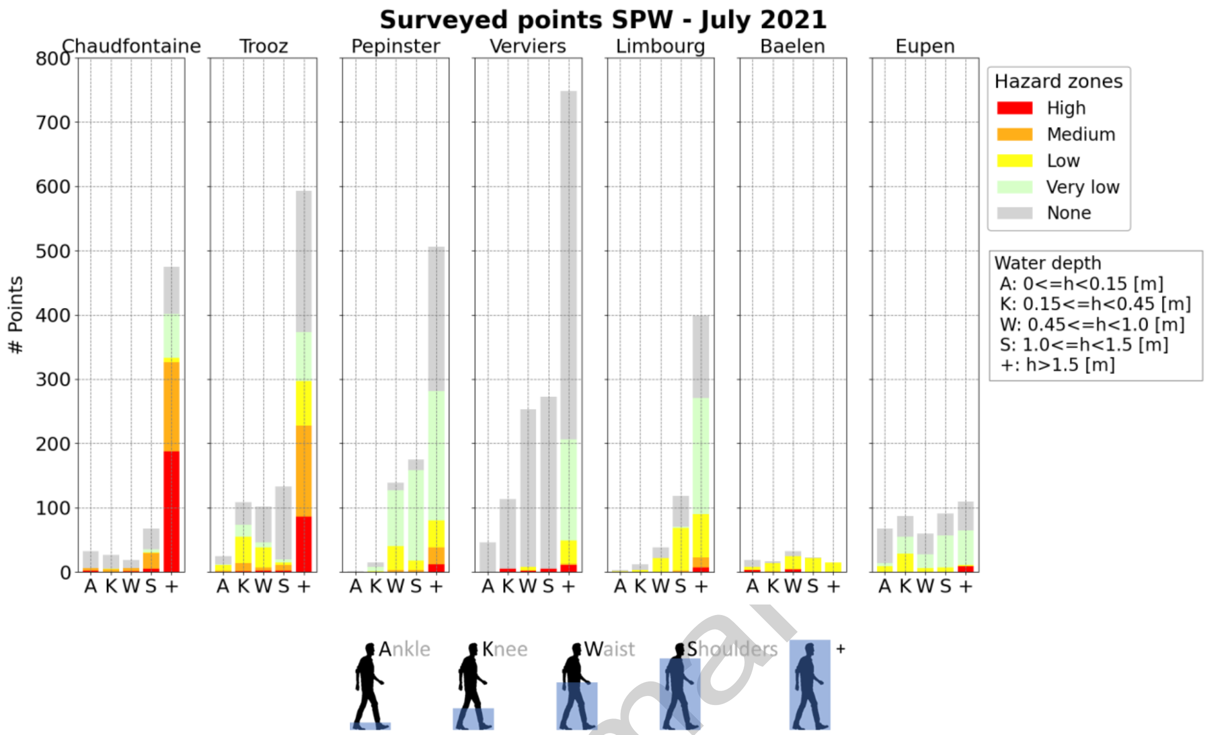


Figure 12. Water depth distribution per municipality and per flood hazard zone (Dataset #2).

3.1.2 Overwhelming effect

Figure 10b displays the share of buildings situated in the observed flooded area (in blue) compared to the total number of buildings in each municipality, classified according to the official flood hazard zones. Figure 10b hints that the municipalities were affected by the flood at different degrees in relation to their size, and flood risk may be at different levels in their agendas due to contrasting shares of the municipality buildings being located in official flood hazard zones.

On one hand, some municipalities have a relatively high share of all their residential buildings situated in a flood hazard zone. This share may be as high as 30 % to 38 % in municipalities such as Limburg and Trooz. It is anticipated that, in such municipalities, flood risk management is at a relatively high level in the priorities of local authorities. In contrast, in the town of Verviers, the number of residential buildings in a flood hazard zone does not exceed 3%, suggesting that flood risk management may not be so high a priority in the local agenda.

On the other hand, the share of flooded residential buildings, compared to the total number of residential buildings in the municipality, suggests that some municipalities have been far more affected than others in comparison to their size. In the municipality of Eupen, approximately 12 % of the residential buildings experienced flooding, while in the municipality of Trooz, this number is as high as 45 %. In that municipality, where nearly half of the population was affected by the flood event, it is expected that both the emergency and recovery phases will have more challenges due to limited available resources.

Similarly, as for the surprise effect, the overwhelming effect was also assessed at the meso-scale, i.e.,

17

per census tract instead of municipality (Figure 13 in Supplement). Again, the value of the meso-scale variable o varies much more substantially ($0.13 \leq o \leq 1.0$) than its counterpart evaluated at the macro-scale ($0.12 \leq o \leq 0.45$).Overall, the results above suggest that the municipalities probably faced different levels of surprise, were differently prepared, and faced different percentages of buildings flooded in relation to their size. In Section 3.2.2, we investigate whether an influence of these contrasting situations can be detected in the monetary loss data collected through our field survey.

3.2 Statistical analysis

This section sequentially presents the results of the three-step procedure described in Section 2.3.2.

3.2.1 Step 1: Baseline classification model

The performance of the baseline classifier is summarised in Table 2. The AUC reaches a value of 0.79, and all three variables are significant ($p < 0.05$). A classification characterised by such a value of AUC is generally considered 'good' (De Hond et al., 2022). Water depth is by far the most significant variable ($p < 10^{-5}$), consistent with the findings of numerous previous studies (Laudan et al., 2017; Sieg et al., 2017). The signs of the estimates also appear logical. On one hand, a higher building footprint area increases the probability that the flood loss exceeds the pre-selected threshold. The same applies to the water depth, with some non-linear effects as reflected by the estimate for h^2. Note that if the square of h is not included, the classification shows a slightly lower performance, with an AUC of 0.76.

Table 2. Significance of the variables in the baseline classifier and hazard variables when added to this classifier (*** $p \leq 10^{-4}$, ** $10^{-4} < p \leq 10^{-3}$, * $10^{-3} < p \leq 10^{-2}$, ▪ $10^{-2} < p \leq 10^{-1}$, not significant $p > 10^{-1}$)

	Baseline classifier			*Hazard variables*	
	h	h^2	A	v_1	v_2
p	$< 10^{-5}$ ***	$9\ 10^{-3}$ **	$1.2\ 10^{-2}$ *	0.18 ns	0.71 ns
n	142			129	

3.2.2 Step 2: Significance of individual additional variables

In addition to the variables included in the baseline model, we tested three groups of building variables, corresponding respectively to additional hazard, vulnerability and coping capacity features, as well as the newly introduced coping capacity variables at macro- and meso-scales. To improve their statistical significance, some variables were reclassified, as detailed in Supplement (Section6).

Hazard variables

As shown in Supplement (Section 4.1), the binary variables *sed* and *c* bring no information at all since all respondents reported the presence of both sediments and contaminants in the flood water. This may result from the binary nature of these variables but also from the extreme characteristics of the considered event.

Variables v_1 and v_2 related to velocity perception take a wider range of values in the survey responses, but the score corresponding to the highest range of flow velocity is considerably overrepresented in the responses (Section 5.1 in Supplement). This is probably the reason why these two variables are also found to be not significant when included in the calibration of the classification model (Table 2). Alternate discretisations of variables v_1 and v_2 were tested by lumping several classes, as shown in Supplement (Section 6), but this did not improve the significance of these variables. Therefore, based on the criterion of significance, no additional building hazard variable was retained for the next steps of the analysis.

Building features

As shown in Table 3, the only vulnerability variables having a statistically significant influence ($p < 0.1$) are the type of building structure (*bs*), the distribution type of the heating system (*pd*), the level of maintenance (*lm*), and the finishing level (*fl*). The strong statistical significance of variables *lm* and *fl* is a notable result because it suggests that those variables play an important part in the prediction of the losses, while they are rarely included in flood loss modelling (Scorzini et al., 2022; Thieken et al., 2008).

Moreover, the signs of the estimates corresponding to variables *bs*, *pd*, *lm* and *fl* are all in accordance with physics. Indeed, a detailed inspection of the outcomes of the classifier calibration indicates that the class 'concrete masonry' leads to a lower probability that flood losses exceed the threshold compared to clay masonry and other building structures, which is deemed reasonable (Suppasri et al., 2015). Similarly, a heating system with central distribution ($pd = 1$) is associated with a higher probability of higher losses, which is plausible since, in the case study area, critical components of the heating system are mostly located in the basement, i.e., typically the most severely flooded part of the building. For variables *lm* and *fl*, which both involve three classes (low, medium, high), class 'high' is found to correspond to a higher probability of high flood loss than the two others, and class 'medium' leads to a greater chance of high flood loss than class 'low'. This reflects the likely higher cost of repairing more luxurious assets.

Coping capacity variables

Among the building-level coping capacity variables, only the hazard zones (*hz*) and the number of implemented short-term mitigation measures (*smm*) were found statistically significant when tested individually (Table 4).

A negative sign of the estimate corresponding to variable *smm* was found, which plausibly reflects a certain level of effectiveness of mitigation measures to limit flood losses. The considered short-term mitigation measures include the installation of flood gates or flood barriers at the building level, relocating furniture to higher floors, operating pumps, and proactively shut-off of electricity, gas, and water supply.

In contrast, a positive estimate sign was obtained for variable *hz*, suggesting that buildings in flood hazard zones face a higher probability of high flood losses. This trend is certainly true, but it fails to reflect the intended purpose of introducing variable *hz* as a proxy for the level of preparation of the population (i.e., 'being situated in a flood hazard zone,' meaning 'being better informed and prepared'). Therefore, this variable was not retained.

Variable *w* (official warning) was handled together with the macro-scale variables because flood warning in the case study was mostly managed at the municipality level.

The macro- and meso-scale variables introduced in Section 2.2.3 were individually tested, along with variables *w* and *m* (the municipality where the building is located). As shown in Table 4, only the variable representing the municipality was found to be significant. This may hint at some differences in flood losses by municipalities, as presented in Figure 9. The fact that 'overwhelming effect' was found not significant may be related to the fact that, though some municipalities were overwhelmed, solidarity developed fast at the regional, national and even international levels in the case of the 2021 flood. This may have compensated to some degree the consequences of overwhelming of authorities and first responders at the municipality level.

Significance of the analysed variables

At the end of Step 2, the following six additional variables were retained to improve the baseline classification model (Step 1): type of building structure (*bs*), type of heating system (*pd*), level of maintenance (*lm*), finishing level (*fl*), number of implemented short-term mitigation measures (*smm*), and the municipality (*m*).

Table 3. Significance of the building variables when individually added to the baseline classification model (*n*=63).

	Building features								
	gl	*bt*	*nf*	*bs*	*yy*	*pd*	*pt*	*lm*	*fl*
p	0.21 ns	0.88 ns	0.81 ns	**0.025** *	**0.25** ns	0.041 *	**0.37** ns	**9 10^{-4}** ***	**4 10^{-4}** ***
n	63								

NB: The numbers in bold refer to the results obtained using the Likelihood Ratio Test due to the nominal character of the variables.

Table 4. Significance of the building and macro-scale coping capacity variables when individually added to the baseline classification model (*n* = 117).

	Micro-scale (building level)						*Macro-scale (municipality level)*					*Meso-scale (statistical tract level)*	
	hz	*fe*	*prep*	*smm*	*lmm*	*isd*	*s*	*o*	*r*	*w*	*m*	*s*	*o*
p	0.04 *	0.18 ns	0.75 ns	0.057 *	0.42 ns	0.61 ns	0.77 ns	0.56 ns	0.11 ns	0.94 ns	0.08 .	0.39 ns	0.22 ns
n	117						121					121	

3.2.3 Step 3: Improved classification model

An improved classification model was calibrated by considering eight variables, i.e., the three variables already included in the baseline classification model (Step 1), as well as the five additional variables selected in Step 2. Two variables, 'level of maintenance' (*lm*) and 'finishing level' (*fl*), were found to be multicollinear (VIF above 10). Therefore, only one of them was eventually retained. Variable *fl* was selected due to its higher level of statistical significance compared to *lm*. Moreover, the statistical significance of variables *bs*, *smm* and *m* was found to be weaker in such a classifier involving more variables ($p \approx 0.2$) than when combined with only the three variables h, h^2, and A. Therefore, variables *bs*, *smm* and *m* were also discarded due to their lack of statistical significance here.

Table 5 shows the estimates obtained by recalibrating the classification model on the same set of variables except for *lm*, *bs, smm,* and *m*, which were removed. For the remaining variables, the signs of the estimates are not altered compared to those found in Steps 1 and 2. Hence, they also remained consistent with the physical interpretation of the variables.

As expected, there is an increase in the losses with an increase in water depth, along with an expansion in the footprint area of the building. This aligns with the positive sign reported for those variables in Table 5. the sign of the estimate for the square of water depth suggests that losses increase less than linearly with the water depth. This can be explained by the fact that, in single-family homes—which are common in the region of the case study—higher floors typically contain less valuable items. For instance, control systems are often located in the basement or on the ground floor, while the attic or the top floor usually holds items of lesser value. Additionally, it is possible that high water levels are associated with areas that experience lower flow velocities compared to regions with shallower water and higher flow velocity, hence leading to different damage mechanisms.

20

Additionally, a higher loss is expected when the building is characterised by central heating due to the common practice in the studied region of locating the boiler in the basement, which is the first part of the building to be flooded. In contrast, estimates of the finishing level variable have a negative sign, meaning a decrease in the losses for lower building conditions. Table 5 shows two variables for the 'fl' variable due to its nominal type. The graph displays two out of the three possible options (low and medium finishing levels), comparing them with the third one (high finishing level), which is not shown (because it is considered as a reference). In this context, it indicates that buildings with low and medium finishing levels experience lower losses compared to those with higher finishing levels.

No multicollinearity issue was detected within the final set of variables (h, h^2, A, pd and fl), which are also all statistically significant (Table 5). As shown by the ROC curve displayed in Figure 13, the final classification model moderately increases the performance metrics AUC by just 4 points (AUC = 0.83) compared to the baseline classifier (AUC = 0.79).

In summary, the improved classification model based on logistic regression includes the following independent variables: water depth and water depth squared, building footprint area, type of heating system (pd), and building finishing level (fl). Interestingly, besides standard variables considered in flood loss modelling (h, h^2 and A), two less conventional building-level variables stand out. The influence of pd is probably closely linked to the ubiquitous practice of positioning the central heating boiler in the basement in the studied region. This can be partly explained by the high proportion of relatively old buildings typical of the working class, with limited floor area, hence making it more challenging to allocate space for technical systems on the higher floors. The significant statistical influence of variable fl, representing the building finishing level, stresses the importance of progressing towards better documentation of the exposed assets to enhance the accuracy of flood loss prediction. However, incorporating this variable into flood risk management could unintentionally lead to a disproportionate allocation of resources aimed at protecting socio-economically privileged populations. This could create an imbalance in the distribution of flood protection efforts, favouring wealthier communities while potentially neglecting more vulnerable groups. Therefore, flood risk management should take a broader, more holistic approach, going beyond simply assessing the damage. In addition to damage assessment, it is crucial to consider other key factors such as the number of people at risk, the presence of critical infrastructure, and the value of important assets that could be impacted by flooding, providing a more balanced and realistic approach to preventing and mitigating flood risks.

Table 5. Results of logistic analysis in the improved classifier.

Parameter	*Coefficient*	*Standard Error*	*Significance level*	*95 per cent CI*	
				Below	*Upper*
Intercept	0.815	0.305	<0.001	0.231	1.437
h	1.067	0.238	<0.001	0.622	1.562
h^2	-0.445	0.189	1.84e-2	-0.839	-0.084
A	0.555	0.271	4.10e-2	0.061	1.123
pd	0.663	0.299	2.67e-2	0.158	1.421
fl_low	-1.994	0.872	**3.5e-2**	-3.999	-0.425
fl_medium	-0.575	0.549		-1.669	0.503

NB: The numbers in bold refer to the results obtained using the Likelihood Ratio Test due to the nominal character of the variables.

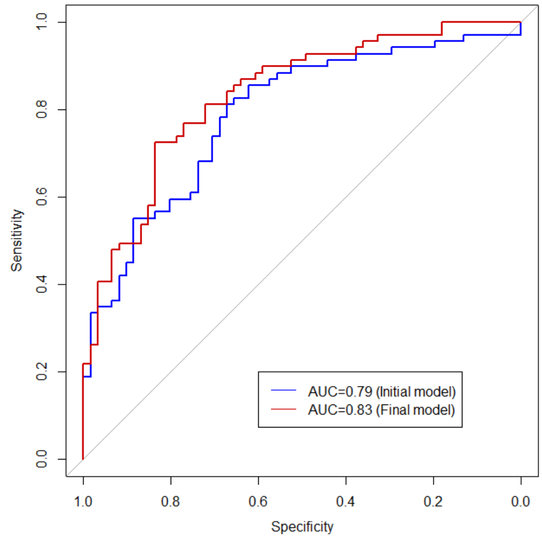


Figure 13. Model performance. In blue is the baseline model, and in red is the final model.

4 DISCUSSION

4.1 Influence of macro- and meso-scale coping capacity variables

New macro- and meso-scale variables have been introduced. They relate respectively to the spatial distribution of risk perception (surprise effect) and of response effectiveness after a flood event (overwhelming effect). Despite considerable differences in these variables between the municipalities or census tracts considered, they appear to be insignificant in improving the classification of building losses in the residential tract. Three hypotheses can be put forward to explain the lack of additional insight provided by these new variables.

First, the variable 'surprise effect' is evaluated as the ratio of the number of flooded buildings to the number of buildings in flood hazard zones. The fact that this quantity reflects the degree of surprise of the community assumes that residents in flood hazard zones have a certain level of flood risk awareness. However, this may not be the case, particularly in the case of an extreme flood event, where the residents are unlikely to be aware of the possibility of experiencing flood depths as deep as those experienced in 2021. The magnitude of the actual flood exceeded the risk perceived by residents.

Second, a certain level of risk perception does not necessarily imply the actual implementation of risk reduction measures by residents or emergency services, while the actual implementation of such measures is a prerequisite for risk perception and awareness to reduce flood losses. Both risk perception and response effectiveness are known to increase the likelihood that people will take action and implement measures (Begg et al., 2017; O'Neill et al., 2016; Takao et al., 2004; Zhai et al., 2005). However, there is no guarantee that the measures were actually implemented. As highlighted by (Scolobig et al. (2012), the link between risk awareness and preparedness is not straightforward, and there is a gap between the risk evaluation of individuals and that of the community. Therefore, it is also possible that even risk-aware residents do not take responsibility and do not implement individual loss-reduction measures. This aligns with the status quo effect, where communities fail to learn and adapt to prevent damage, even when exposed to frequent floods (Mendoza Leal et al., 2024)

A third possible explanation is the ineffectiveness of the implemented risk-reduction measures due to the extreme nature of the considered flood event. A large proportion of the buildings experienced flood depths greater than 1.5 meters (Figure 12). In such cases, most building-level mitigation measures are ineffective in preventing or reducing losses. This suggests that even well-informed individuals and respondents may not have been able to reduce losses through the measures they were able to implement. This explains why lower levels of surprise and/or overwhelming do not correlate with reduced flood losses.

Overall, these results emphasise that the risk associated with extreme floods and more frequent floods should be managed through very different strategies. Communicating about flood risk, sharing responsibility for protection, and implementing property-level mitigation measures can help reduce the impacts of frequent to moderate floods. For extreme floods, the effectiveness of property-level mitigation measures is more limited. This makes other flood management strategies more effective, such as strict spatial planning prohibiting new development in floodplains and the gradual relocation of people already living in high-risk flood areas.

4.2 Influence of the threshold in the binomial logistic regression

The binomial logistic regression model utilises a threshold to classify flood losses as either below or above this particular value. The value of the threshold is an input parameter for the model. Here, an approximation of the median value of the total building losses was used (EUR 80,000). However, if the model is applied for predictive purposes, this value is unknown. It is, therefore, important to assess the extent to which the model performance is sensitive to changes in this threshold. Systematic tests were conducted with threshold values between half the initial value (EUR 40,000) and twice this value (EUR 160,000). Figure 14 in Supplement shows that the obtained AUC remains relatively stable even when the threshold varies over a broad range. The median value of AUC is found equal to 0.81, with a maximum value of 0.86 and a minimum value of 0.75.

In addition to evaluating the model's performance as the threshold changes, the significance of the final selected variables for classification was also examined. Figure 15 in Supplement shows that the significance of these variables changes when the considered threshold is strongly varied. This can be attributed to a rising imbalance in the amount of data in the two classification groups as the threshold deviates from the median value of flood losses, as illustrated in Figure 16 in Supplement. However, regardless of the model's threshold, the variable 'water depth' remains consistently very significant, confirming its overwhelming importance in estimating residential flood losses. It can be noted that, in addition to water depth, there is always another significant variable that enhances prediction accuracy, irrespective of the considered threshold. This underscores the value of multivariate flood damage models in achieving better flood damage assessment.

4.3 Model dependency on data

The final proposed model was constructed using a relatively small dataset comprising 130 entries without missing values. Therefore, it was impossible to separate the data in a training and a test sets to evaluate the model performance. Consequently, the model performance was evaluated by creating slightly different datasets using bootstrapping sampling. A sample of the same size as the original dataset was generated by randomly drawing with replacement.

The median value of the model AUC using 500 bootstrapped datasets is equal to 0.85, with a standard deviation of 0.03. Therefore, it can be concluded that the constructed model performs well for classifying the losses above or below a specific threshold, independent of small variations in the dataset.

Even though the model classifies the flood losses with good accuracy, it still has some limitations. The first limitation is the small database with which the model is constructed, as the accuracy of the model increases with the size of the learning sample. However, despite the relatively small sample size, it contains high-quality data guaranteed from the whole data collection process: on-site surveys of around one-hour, double encoding conducted by two different researchers, comparison, and

23

correction of the two encodings in case differences are found, leading to a final trustworthy database.

One limitation of the study is that it is based on a single flood event. It is therefore necessary to investigate whether the statistical significance of some variables would be altered if the dataset is expanded with data from other countries and other flood events. This hints at the value of following up on this study with transnational analyses of flood loss databases gathered in various countries and for a broader range of flood events. The current analyses should also be extended to setting up a model able to predict flood losses instead of merely classifying them. This also requires a larger dataset to ensure the robustness of the model calibration. Another possible bias in the conducted surveys may result from the difficulty faced in approaching former inhabitants of washed-away, demolished or abandoned buildings. This aspect deserves dedicated research focusing on buildings having encountered severe structural damage as a complement to the surveys conducted here.

5 CONCLUSIONS

Coping capacity variables were tested as additional predictor variables to estimate the classification of severe flood building losses. Along with building-level variables, three coping capacity variables were defined and tested at different scales (i.e., meso and macro-scale) to describe the preparedness and disaster response of the population and management authorities. (*i*) the surprise effect (the ratio of the number of flooded buildings to the number of flooded buildings located in an official flood hazard area), (*ii*) the overwhelming effect (the fraction of flooded buildings compared to the total number of buildings within each census tract or municipalities), and (*iii*) flood rarity (the ratio of the peak discharge of the considered event to the 100-year flood peak).

A binomial logistic regression was trained and tested using 142 observations from six municipalities in the Vesdre Valley in Belgium, following the July 2021 flood event. The inclusion of meso- or macro-scale coping capacity variables was found to be not significant for classifying severe flood losses at the building level. Several plausible reasons can explain this result. First, the extreme nature of the examined flood event may have caught residents by surprise. While they were aware of being in a flood hazard zone, they did not expect flood depths as severe as those actually experienced in 2021. Second, while risk awareness is known to increase the likelihood that individuals will take action, it does not guarantee that measures will be implemented. Finally, even when measures were implemented, they may have been ineffective due to the extreme flood depths encountered. These results suggest that managing the risks of both frequent and extreme floods requires different strategies. While property-level mitigation measures may be effective for frequent floods, they do not seem to be sufficient for extreme floods. Regarding extreme floods, property-level mitigation measures are more limited, requiring other strategies such as strict spatial planning prohibiting new development in floodplains and the gradual relocation of people already living in high-risk flood areas.

Consequently, flood risk management strategies should incorporate a range of measures that address low, moderate, and extreme flood events. These strategies should focus not only on preventing and protecting against flood damage but also on mitigating risks and enabling rapid recovery from unprotected residual risk. By doing so, communities will be better equipped to cope with flood hazards in a comprehensive manner, ensuring resilience across all stages of the disaster cycle.

Limited to the studied dataset, the final model classifies severe flood building damages with an accuracy of 83%. This classifier includes four predictor variables: water depth, building footprint area, building finishing level, and heating system location, all of which have been found to be statistically significant. The latter two variables describe the vulnerability of the assets at risk. The 'finishing level' reflects the higher cost of repairing more luxurious dwellings, while the 'heating system location' refers to the placement of key components (e.g., boiler, control panel), which are typically located in the basement of buildings in the region of interest. This result highlights the importance of explicitly considering the characteristics of building systems, an often-overlooked factor in flood loss models.

24

The focus was set here on the residential sector, which has been one of the most affected by the 2021 flood event. However, similar studies are necessary for other sectors tracts that were also heavily impacted, including with indirect consequences. Therefore, it is of utmost importance to pursue systematic collection of the flood loss data that allows updating, improving, and validating flood loss models.

Acknowledgements

This research was partly supported by the Interreg NWE project FlashFloodBreaker. The Authors gratefully acknowledge the insightful comments received from two anonymous Reviewers. The Authors are also grateful to all individuals who have given their time to respond to the flood loss survey and have shared valuable information.

Declaration of interest statement

The authors declare no conflict of interest.

6 REFERENCES

Archambeau, P. (2022). *Reconstruction des valeurs hydrologiques associées aux évènements de juillet 2021 dans la vallée de la Vesdre*. https://orbi.uliege.be/handle/2268/264719

Archambeau, P., Dewals, B., Erpicum, S., Pirotton, M., Teller, J., & Schellings, C. (2022). *Zone d'aléas et périodes de retour. In Schéma stratégique multidisciplinaire du bassin versant de la Vesdre: Diagnostic. Studio Paola Viganò and Team Vesdre-ULiège.* (pp. 45–46).

Ballio, F., Molinari, D., Minucci, G., Mazuran, M., Arias Munoz, C., Menoni, S., Atun, F., Ardagna, D., Berni, N., & Pandolfo, C. (2018). The RISPOSTA procedure for the collection, storage and analysis of high quality, consistent and reliable damage data in the aftermath of floods: The RISPOSTA procedure for the management of damage data. *Journal of Flood Risk Management*, *11*, S604–S615. https://doi.org/10.1111/jfr3.12216

Barcelloni Corte, M., Bianchet, B., Privot, J., Schelings, C., & Teller, J. (2022). *Schéma stratégique multidisciplinaire du bassin versant de la Vesdre. Diagnostic approfondi et multithématique. Contributions de la TEAM Vesdre–ULiège.*

Begg, C., Ueberham, M., Masson, T., & Kuhlicke, C. (2017). Interactions between citizen responsibilization, flood experience and household resilience: Insights from the 2013 flood in Germany. *International Journal of Water Resources Development*, *33*(4), 591–608. https://doi.org/10.1080/07900627.2016.1200961

Bruwier, M., Erpicum, S., Pirotton, M., Archambeau, P., & Dewals, B. (2015). Assessing the operation rules of a reservoir system based on a detailed modelling chain. *Natural Hazards and Earth System Sciences*, *15*(3), 365–379.

CRED. (2024). *EM-DAT - The international disaster database*. https://www.emdat.be/

CSR. (2022). *1 an après les inondations … Bilan de la gestion post-inondations et continuité de la reconstruction. Commissariat Spécial à la Reconstruction*. https://www.wallonie.be/sites/default/files/2022-07/Bilan%20complet%20CSR%2026%20juillet%202022.pdf

Cuvelier, T., Archambeau, P., Dewals, B., & Louveaux, Q. (2018). Comparison between robust and stochastic optimisation for long-term reservoir management under uncertainty. *Water Resources Management*, *32*, 1599–1614.

De Hond, A. A. H., Steyerberg, E. W., & Van Calster, B. (2022). Interpreting area under the receiver operating characteristic curve. *The Lancet Digital Health*, *4*(12), e853–e855.

https://doi.org/10.1016/S2589-7500(22)00188-1
Dottori, F., Figueiredo, R., Martina, M. L. V., Molinari, D., & Scorzini, A. R. (2016). INSYDE: A synthetic, probabilistic flood damage model based on explicit cost analysis. *Natural Hazards and Earth System Sciences*, *16*(12), 2577–2591. https://doi.org/10.5194/nhess-16-2577-2016
Dottori, F., Szewczyk, W., Ciscar, J.-C., Zhao, F., Alfieri, L., Hirabayashi, Y., Bianchi, A., Mongelli, I., Frieler, K., Betts, R. A., & others. (2018). Increased human and economic losses from river flooding with anthropogenic warming. *Nature Climate Change*, *8*(9), 781–786.
Endendijk, T., Botzen, W. J. W., de Moel, H., & Kok, M. (2023). *Experience From the 2021 Floods in the Netherlands: Household Survey Results on Impacts and Responses*. *2*.
European Environment Agency. (2024). *European climate risk assessment: Executive summary.* Publications Office. https://data.europa.eu/doi/10.2800/204249
Ghaedi, H., Reilly, A. C., Baroud, H., Perrucci, D. V., & Ferreira, C. M. (2022). Predicting flood damage using the flood peak ratio and Giovanni Flooded Fraction. *PLOS ONE*, *17*(8), e0271230. https://doi.org/10.1371/journal.pone.0271230
Goudenhoofdt E. (2023). *RADCLIM User Guide. Royal Meteorological Institute*. https://opendata.meteo.be/cases/202107/radclim_userguide.pdf
IPCC. (n.d.). *3.5.2 How will climate change affect flood damages? - AR4 WGII Chapter 3: Fresh Water Resources and their Management*. Retrieved September 18, 2024, from https://archive.ipcc.ch/publications_and_data/ar4/wg2/en/ch3s3-5-2.html
Journée, M., Goudenhoofdt, E., Vannitsem, S., & Delobbe, L. (2023). Quantitative rainfall analysis of the 2021 mid-July flood event in Belgium. *Hydrology and Earth System Sciences*, *27*(17), 3169–3189. https://doi.org/10.5194/hess-27-3169-2023
Kaufmann, M., Mees, H., Liefferink, D., & Crabbé, A. (2016). A game of give and take: The introduction of multi-layer (water) safety in the Netherlands and Flanders. *Land Use Policy*, *57*, 277–286. https://doi.org/10.1016/j.landusepol.2016.05.033
Kitsikoudis, V., Becker, B. P., Huismans, Y., Archambeau, P., Erpicum, S., Pirotton, M., & Dewals, B. (2020). Discrepancies in Flood Modelling Approaches in Transboundary River Systems: Legacy of the Past or Well-grounded Choices? *Water Resources Management*, *34*, 3465–3478.
Kreibich, H., Bubeck, P., Van Vliet, M., & De Moel, H. (2015). A review of damage-reducing measures to manage fluvial flood risks in a changing climate. *Mitigation and Adaptation Strategies for Global Change*, *20*(6), 967–989. https://doi.org/10.1007/s11027-014-9629-5
Kreibich, H., Thieken, A., Haubrock, S.-N., & Schröter, K. (2017). HOWAS21, the German flood damage database. *Flood Damage Survey and Assessment: New Insights from Research and Practice*, 65–75.
Kreibich, H., Van Loon, A. F., Schröter, K., Ward, P. J., Mazzoleni, M., Sairam, N., Abeshu, G. W., Agafonova, S., AghaKouchak, A., Aksoy, H., & others. (2022). The challenge of unprecedented floods and droughts in risk management. *Nature*, *608*(7921), 80–86.
Laudan, J., Rözer, V., Sieg, T., Vogel, K., & Thieken, A. H. (2017). Damage assessment in Braunsbach 2016: Data collection and analysis for an improved understanding of damaging processes during flash floods. *Natural Hazards and Earth System Sciences*, *17*(12), 2163–2179. https://doi.org/10.5194/nhess-17-2163-2017
Mees, H., Crabbé, A., Alexander, M., Kaufmann, M., Bruzzone, S., Lévy, L., & Lewandowski, J. (2016). Coproducing flood risk management through citizen involvement: Insights from cross-country comparison in Europe. *Ecology and Society*, *21*(3), art7. https://doi.org/10.5751/ES-08500-210307
Mees, H., Crabbé, A., & Driessen, P. P. J. (2017). Conditions for citizen co-production in a resilient, efficient and legitimate flood risk governance arrangement. A tentative framework. *Journal of Environmental Policy & Planning*, *19*(6), 827–842. https://doi.org/10.1080/1523908X.2017.1299623
Mendoza Leal, C., Coloma, R., Ponce, D., Alarcón, B., Guerra, M., Stehr, A., Carrasco, J. A., Alcayaga, H., Rojas, O., Link, F., & Link, O. (2024). The *status quo* effect in the sociohydrology of floods.

Hydrological Sciences Journal, *69*(16), 2423–2435. https://doi.org/10.1080/02626667.2024.2412734

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C. J. H., Bates, P., Bertola, M., Kemter, M., Kreibich, H., Lall, U., & Macdonald, E. (2021). Causes, impacts and patterns of disastrous river floods. *Nature Reviews Earth & Environment*, *2*(9), 592–609. https://doi.org/10.1038/s43017-021-00195-3

Merz, B., Hall, J., Disse, M., & Schumann, A. (2010). Fluvial flood risk management in a changing world. *Natural Hazards and Earth System Sciences*, *10*(3), 509–527. https://doi.org/10.5194/nhess-10-509-2010

Merz, B., Kreibich, H., & Lall, U. (2013). Multi-variate flood damage assessment: A tree-based data-mining approach. *Natural Hazards and Earth System Sciences*, *13*(1), 53–64. https://doi.org/10.5194/nhess-13-53-2013

Mohr, S., Ehret, U., Kunz, M., Ludwig, P., Caldas-Alvarez, A., Daniell, J. E., Ehmele, F., Feldmann, H., Franca, M. J., Gattke, C., Hundhausen, M., Knippertz, P., Küpfer, K., Mühr, B., Pinto, J. G., Quinting, J., Schäfer, A. M., Scheibel, M., Seidel, F., & Wisotzky, C. (2023). A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe – Part 1: Event description and analysis. *Natural Hazards and Earth System Sciences*, *23*(2), 525–551. https://doi.org/10.5194/nhess-23-525-2023

Murphy, K. P. (2022). *Probabilistic machine learning: An introduction*. The MIT Press.

Mustafa, A., Bruwier, M., Archambeau, P., Erpicum, S., Pirotton, M., Dewals, B., & Teller, J. (2018). Effects of spatial planning on future flood risks in urban environments. *Journal of Environmental Management*, *225*, 193–204.

O'Neill, E., Brereton, F., Shahumyan, H., & Clinch, J. P. (2016). The Impact of Perceived Flood Exposure on Flood-Risk Perception: The Role of Distance. *Risk Analysis*, *36*(11), 2158–2186. https://doi.org/10.1111/risa.12597

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

Pasquier, U., Few, R., Goulden, M., Hooton, S., He, Y., & Hiscock, K. (2019). Integrating stakeholder and scientific knowledge of future flood risk to inform climate change adaptation planning in a coastal region. *Geophysical Research Abstracts*, *21*.

Poussard, C., Dewals, B., Archambeau, P., & Teller, J. (2021). Environmental Inequalities in Flood Exposure: A Matter of Scale. *Frontiers in Water*, *3*, 633046. https://doi.org/10.3389/frwa.2021.633046

Renn, O. (2015). Stakeholder and Public Involvement in Risk Governance. *International Journal of Disaster Risk Science*, *6*(1), 8–20. https://doi.org/10.1007/s13753-015-0037-6

Schroeder, M. A., Lander, J., & Levine-Silverman, S. (1990). Diagnosing and Dealing with Multicollinearity. *Western Journal of Nursing Research*, *12*(2), 175–187.

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

Scolobig, A., De Marchi, B., & Borga, M. (2012). The missing link between flood risk awareness and preparedness: Findings from case studies in an Alpine Region. *Natural Hazards*, *63*(2), 499–520. https://doi.org/10.1007/s11069-012-0161-1

Scorzini, A. R., Dewals, B., Rodriguez Castro, D., Archambeau, P., & Molinari, D. (2022). INSYDE-BE: adaptation of the INSYDE model to the Walloon region (Belgium). *Natural Hazards and Earth System Sciences*, *22*(5), 1743–1761. https://doi.org/10.5194/nhess-22-1743-2022

Sieg, T., Vogel, K., Merz, B., & Kreibich, H. (2017). Tree-based flood damage modeling of companies: Damage processes and model performance. *Water Resources Research*, *53*(7), 6050–6068.

Suppasri, A., Charvet, I., Imai, K., & Imamura, F. (2015). Fragility Curves Based on Data from the 2011 Tohoku-Oki Tsunami in Ishinomaki City, with Discussion of Parameters Influencing Building

Damage. *Earthquake Spectra*, *31*(2), 841–868. https://doi.org/10.1193/053013EQS138M
Szönyi, M., Roezer, V., Deubelli, T., Ulrich, J., MacClune, K., Laurien, F., & Norton, R. (2022). *PERC floods following "Bernd."*
Takao, K., Motoyoshi, T., Sato, T., Fukuzondo, T., Seo, K., & Ikeda, S. (2004). Factors determining residents' preparedness for floods in modern megalopolises: The case of the Tokai flood disaster in Japan. *Journal of Risk Research*, *7*(7–8), 775–787. https://doi.org/10.1080/1366987031000075996
Thieken, A. H., Müller, M., Kreibich, H., & Merz, B. (2005). Flood damage and influencing factors: New insights from the August 2002 flood in Germany. *Water Resources Research*, *41*(12), 2005WR004177. https://doi.org/10.1029/2005WR004177
Thieken, A. H., Olschewski, A., Kreibich, H., Kobsch, S., & Merz, B. (2008). Development and evaluation of FLEMOps – a new *F* lood *L* oss *E* stimation *MO* del for the *p* rivate *s* ector. *Flood Recovery, Innovation and Response I*, *I*, 315–324. https://doi.org/10.2495/FRIAR080301
Van Eerd, M. C. J., Wiering, M. A., & Dieperink, C. (2017). Solidarity in transboundary flood risk management: A view from the Dutch North Rhine–Westphalian catchment area. *Climate Policy*, *17*(3), 261–279. https://doi.org/10.1080/14693062.2015.1075376
Veloso, C., Flores, E., Noguera, I., Faúndez, R., Arriagada, P., Rojas, O., Carrasco, J. A., & Link, O. (2022). Preparedness against floods in nearly pristine socio-hydrological systems. *Hydrological Sciences Journal*, *67*(3), 319–327. https://doi.org/10.1080/02626667.2021.2023156
Zhai, G., Fukuzono, T., & Ikeda, S. (2005). MODELING FLOOD DAMAGE: CASE OF TOKAI FLOOD 2000. *Journal of the American Water Resources Association*, *41*(1), 77–92. https://doi.org/10.1111/j.1752-1688.2005.tb03719.x
Zones inondées IDW - Hauteur d'eau—Juillet 2021. (2024, October 8). http://geoportail.wallonie.be/catalogue/59f97c7e-c501-4e92-8c47-047d000f4c13.html

28