Check for updates

Journal of Flood Risk Management

WILEY

CIWEM Chartered Institution of Water and Environmental Management | Journal of Flood Risk Management

SPECIAL ISSUE PAPER OPEN ACCESS

Key Drivers of Flash Flood Damage to Private Households

Daniela Rodríguez Castro[1] | Kasra Rafiezadeh Shahi[2,3] | Nivedita Sairam[2] | Melanie Fischer[2] | Guilherme Samprogna Mohor[4] | Annegret Thieken[4] | Benjamin Dewals[1] | Heidi Kreibich[2]

[1]Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, Belgium | [2]Section Hydrology, GFZ German Research Centre for Geosciences, Potsdam, Germany | [3]Planetary Boundaries Science Lab, Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany | [4]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany

Correspondence: Daniela Rodríguez Castro (drodriguez@uliege.be)

Received: 30 July 2024 | **Revised:** 19 March 2025 | **Accepted:** 5 June 2025

Funding: This work was supported by HIDA (Helmholtz Information and Data Science Academy) GRANT, 14297, Interreg NWE project FlashFloodBreaker. Federal Ministry of Education and Research (Bundesministerium f¨ur Bildung und Forschung-BMBF), FKZ 01LN2209A, FKZ 01LR2102I.

Keywords: cross-country data analysis | feature importance | flash floods | flood damage | machine learning

ABSTRACT

Flash floods cause high numbers of casualties and enormous economic damage. Good knowledge of the damage processes is crucial for the implementation of effective flash flood risk management. However, little is known about the damage processes that occur during flash floods, despite their severity. To gain more knowledge, independent data collection initiatives were carried out in the affected areas of Belgium and Germany after the 2021 floods. The resulting datasets include 420 damaged residential buildings in the Vesdre valley in Belgium, 277 in the Ahr valley in Rhineland-Palatinate (Germany) and 332 in North Rhine-Westphalia (Germany). A total of 30 potential damage-influencing variables were harmonized across the regions, providing valuable insights into hazard characteristics, the vulnerability of exposed assets, the coping capacity of inhabitants, and socio-economic factors. Machine learning-based analysis reveals the significant importance of hazard variables, such as water depth and sediment transport, particularly for building damage. In addition to these, exposure (living area) and physical vulnerability factors (building type and wall type) also play a role in determining building damage across the affected regions. For content damage, besides water depth and living area, socio-economic vulnerability (ownership status of the building) and emergency measures were found to be important predictors. These key drivers of building and content damage from flash floods can be utilized to develop more accurate damage models, thereby improving flash flood risk assessments, enhancing risk communication, and supporting better preparedness strategies.

1 | Introduction

Flash floods are among the most devastating natural hazards, causing a high number of fatalities and tremendous economic damage in Europe and worldwide (Ahmadalipour and Moradkhani 2019; Petrucci et al. 2019; Paprotny et al. 2018; Terti et al. 2017). Flash floods are defined as sudden floods in relatively small, often mountainous catchments that occur within a few hours of heavy rainfall (Calianno et al. 2013; Llasat et al. 2016). Due to their rapid onset, high flow velocities, and debris transport, flash floods can cause loss of life and be extremely damaging to property and infrastructure (Laudan et al. 2017; Yin et al. 2016; Fuchs et al. 2012).

In recent decades, much knowledge has been gained about flood damage processes, especially with regard to fluvial floods. However, flash floods show different dynamics, so that the damage processes of flash floods are different from those of river floods (Borga et al. 2014). Damage caused by flash floods is more influenced by flow velocity, and not primarily

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Journal of Flood Risk Management* published by Chartered Institution of Water and Environmental Management and John Wiley & Sons Ltd.

by water level and duration (e.g., Mohor et al. 2020). Lead times are shorter for flash floods, making it more difficult to implement effective emergency measures than for river floods (Thieken et al. 2022).

Little is known about the damage processes associated with flash floods (Laudan et al. 2017), and the quantitative assessment of flash flood damage to buildings remains poorly understood (Zhen et al. 2022; Leal et al. 2021; Alipour et al. 2020). Fuchs et al. (2012) proposed a spatial approach demonstrating that high damage caused by flash floods is not only explained by the severity of the hazard, but also by land use patterns and building characteristics such as building type and year of construction.

Building characteristics and location have also been identified as particularly decisive factors in determining building damage caused by high debris flow (Papathoma-Köhle 2016). The indicator-based method (IBM) is a commonly used approach in flood risk assessment (Zhen et al. 2022), which involves assigning weights to key building characteristics related to flash flood damage and their corresponding categories, resulting in a relative vulnerability index (RVI). However, IBM is not hazard-specific, and despite modifications made to account for hazard indicators that significantly influence flash flood damage (Papathoma-Köhle et al. 2019, 2022), only water depth and sediment indicators have been included. This approach overlooks other critical variables, such as flow velocity and debris impacts, which also play a crucial role in the damage process (Zhen et al. 2022), or additional socio-economic and coping capacity variables that could influence the vulnerability of the population and therefore increase or mitigate the flash flood impacts.

To account for additional hazard variables in describing the event, (Zhen et al. 2022) developed a methodology that quantitatively assesses the impact by combining the flood hazard index (FHI) with the building vulnerability index (VI), resulting in a flood risk index (FRI) for buildings affected by flash floods. These indicator-based approaches are useful for mapping the risk, but it is also necessary to quantitatively estimate the damage to enable cost–benefit analyses and support decision making regarding the implementation of flood prevention, protection, and preparation measures.

As the availability of empirical damage data has increased, machine learning approaches for feature selection have been increasingly used to identify the key drivers of damage. Applied methods include bagging decision trees, conditional inference forests, gradient boosting machines, and linear regression models such as Lasso (Merz et al. 2013; Rözer et al. 2019; Sieg et al. 2017; Rafiezadeh et al. 2024). Laudan et al. (2017) used a random forest model and a random generalized linear model to analyze empirical damage data from a flash flood event in southern Germany in 2016. The study found that water depth, the exposure of a building to the flow direction, and the presence of susceptible building parts (such as shop windows) are the most important damage drivers. Due to the empirical nature of these approaches, the results tend to be limited to local assessments and it is not known to what extent the results are transferable to other regions. Cross-regional datasets that include more variability in the represented flood hazard processes, characteristics of exposed buildings, and socio-economic features help to obtain more generic results (Wagenaar et al. 2018; Kreibich, Di Baldassarre, et al. 2017).

In July 2021, a low-pressure system called 'Bernd' caused severe flooding in the Eifel-Ardennes highlands of Germany, Belgium, and the Netherlands. The floods claimed more than 200 lives and caused around €40 billion in economic damage across Europe (Munich RE 2022; Thieken et al. 2023). In addition to the significant impact on infrastructure and the commercial sector, private households (i.e., residential buildings and content) also experienced severe damage. In the German federal states of North Rhine-Westphalia and Rhineland-Palatinate alone, at least 200,000 buildings were flooded, with insured damages to residential buildings and content reaching €8.4 billion (Gesamtverband der Versicherer 2021). The most affected district of Ahrweiler reported average damage to residential buildings of €255,000 (Gesamtverband der Versicherer 2021). In Belgium, the event caused significant damage to residential buildings, industry, and infrastructure such as railways and roads, with an estimated total cost of over €5 billion (CSR 2022), and 39 fatalities.

Despite the existence of flood damage models in both Belgium and Germany (i.e., FLEMOps, and INSYDE-BE) (Thieken et al. 2008; Scorzini et al. 2022), these models were developed based on data from previous flood events, corresponding primarily to more moderate riverine floods. While these models already incorporate multiple variables, they still fail to adequately capture more extreme damage to buildings. Given this limitation, the aim of this study is to enhance our understanding of the damage processes affecting private households during flash floods, using data collected after the 2021 extreme flood in both countries. This study conducts a cross-regional (cross-country) analysis based on building-level empirical data from households affected by the July 2021 floods in Belgium and the two most affected federal states of Germany. Furthermore, the study aims to identify the key drivers of flash flood damage to private households, which will later serve as input parameters for developing a regional flood damage model specifically for flash floods.

2 | Study Areas and Surveys

This section introduces the selected study areas (Section 2.1), as displayed in Figure 1, and the surveys and datasets used in the analysis (Section 2.2).

2.1 | Study Areas

2.1.1 | Germany

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 Germany, the Ahr Valley in Rhineland-Palatinate (RLP) saw an average of 115 mm rainfall poured in 72 h, and more than 150 mm was registered in 24 h at a few stations in North Rhine-Westphalia

17533 8x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

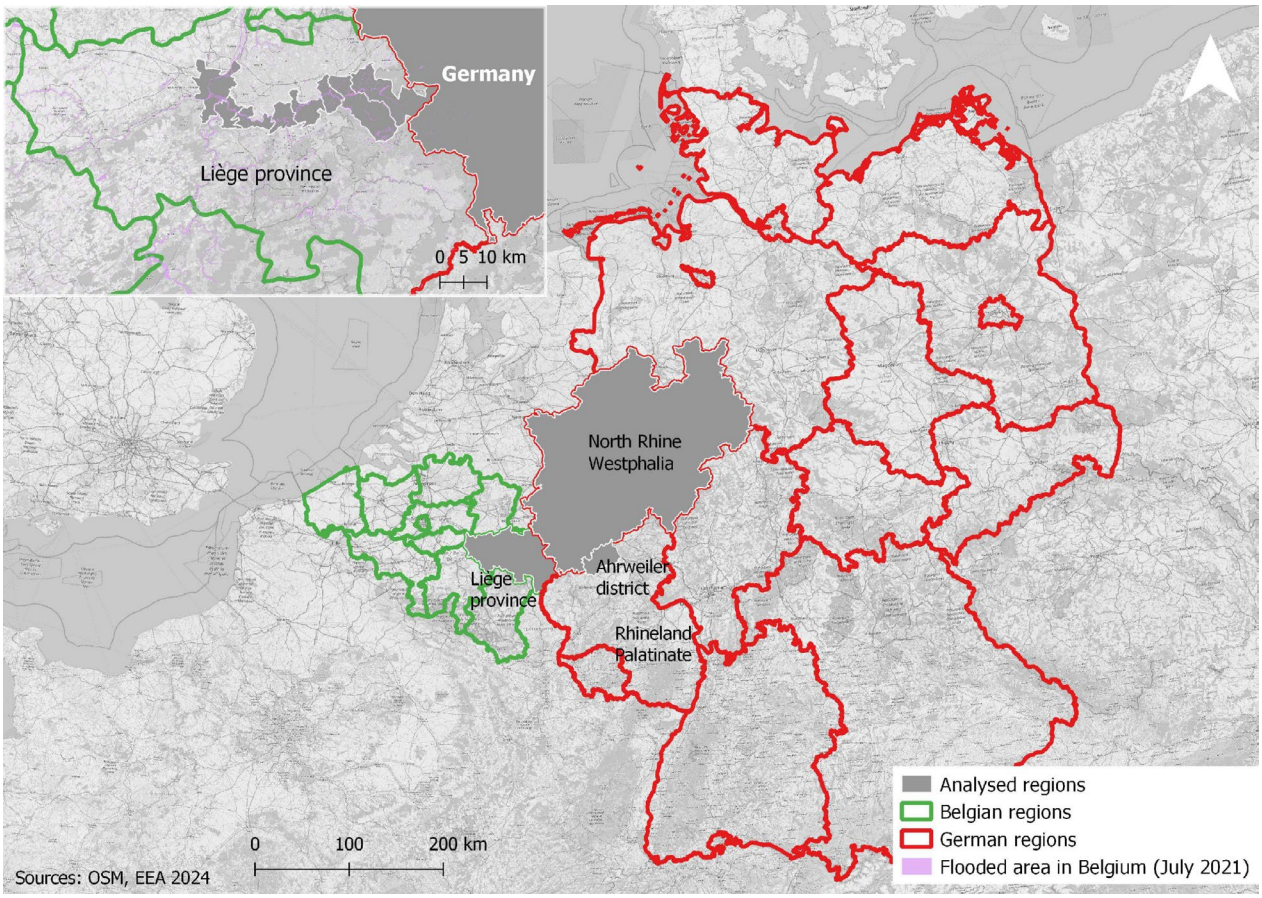

FIGURE 1 | Study areas in Belgium (Vesdre valley in Liège province), and Germany (Ahrweiler district in Rhineland-Palatinate and North Rhine-Westphalia). The flooded area is only displayed for Belgium (see inset).

(NRW) (Junghänel et al. 2021). This weather event was deemed the most extreme precipitation event in Germany (among events included in the CatRaREv2022.01—Lengfeld et al. 2023), when assessed by the cross-scale weather extremity index (xWEI) which considers the area affected besides the intensity and duration. Moreover, soils already saturated by previous rainfall contributed to particularly high runoff quantities (Dietze et al. 2022).

The Rapid Mapping from the Copernicus Emergency Management Services estimated a total of more than 7000 affected buildings in RLP and more than 4700 in NRW. Likewise, more than 400 km of transportation lines were affected in RLP and 140 in NRW. Finally, an estimated 1172 ha of agricultural land, forest, and so forth, were affected in RLP, 2861 in NRW (Copernicus Emergency Management Service 2021). These numbers are, however, lower bounds, as other areas not covered by the rapid mapping have since reported damage.

Twelve districts ('Landkreise') in the two federal states have been strongly affected, leading to hundreds of wounded persons (823, as of 21.07.2021—Schäfer et al. 2021), and 186 fatalities (49 in NRW and 137 in RLP, including one missing person) (Thieken et al. 2023). In NRW, more than 180 municipalities were affected, that is more than 20 thousand people, leading to a current estimate of 12.3 billion euros (Bundesministerium des Innern und für Heimat 2022). In RLP, 65 thousand people have been affected, leading to damages above 18 billion euros. 30 Billion Euros were initially provided for the reconstruction and compensations, although the actual damage sum is still unknown, but has been later estimated to 33 Billion Euros in Germany alone (Munich RE 2022).

2.1.2 | Belgium

In Belgium (BE), the precipitation volumes reached up to 300 mm in 48 h, which is comparable to 3 months of typical precipitation experienced in just 2 days (Barcelloni Corte et al. 2022; Journée et al. 2023). Inundation depths above 7 m were observed, resulting in 39 casualties and severe damage to residential buildings, industries, and infrastructure, with more than 300 buildings washed away, around 200 to be demolished, and roughly 3000 partially destroyed, with an estimated total cost above EUR 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 purpose, the regional government classified these municipalities into three categories as a function of the severity of damage. Category 1 gathers the ten most impacted municipalities, Category 2 contains 28 moderately affected municipalities, and the remaining ones are in Category 3. Remarkably, nine out of the ten most impacted municipalities are in a single catchment, namely

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

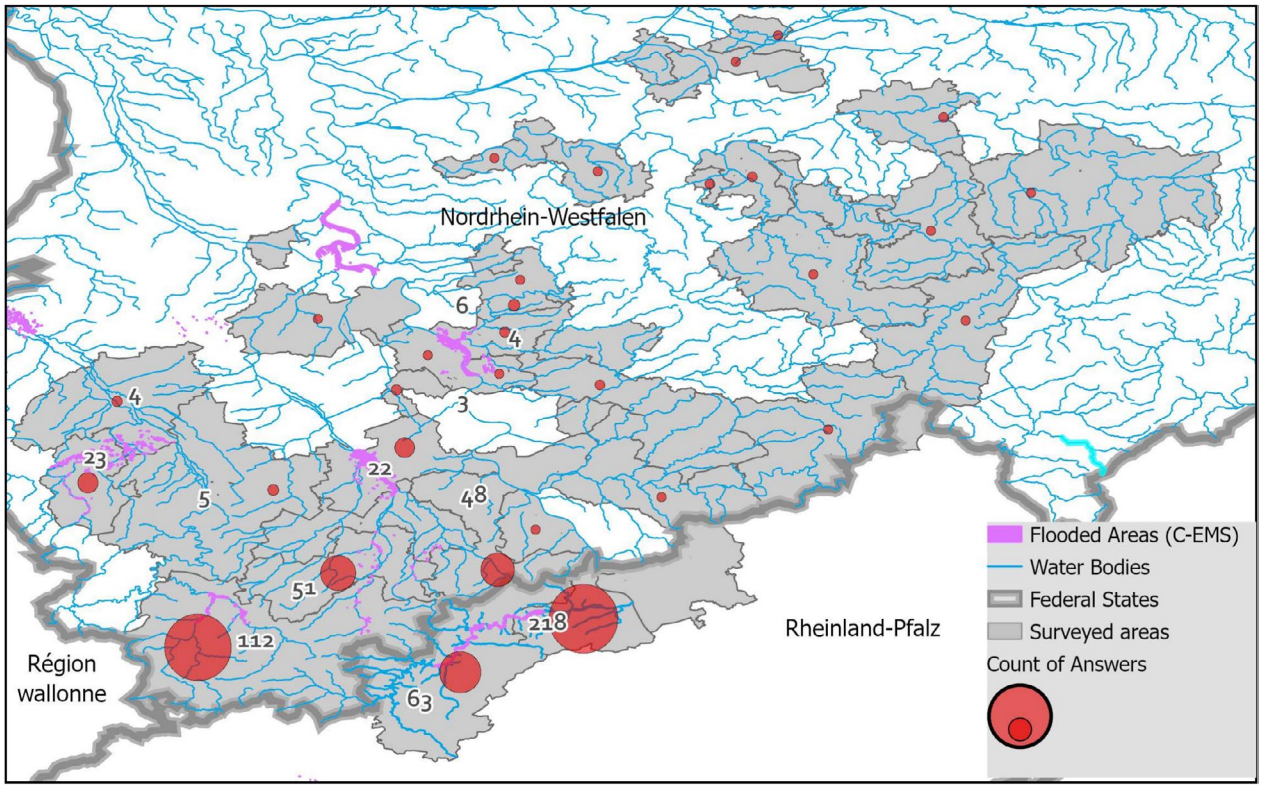

FIGURE 2 | Surveyed areas in Germany. Two hundred and seventy-seven surveys in Rhineland-Palatinate and 332 in North Rhine-Westphalia.

the catchment of river Vesdre (700 km^2). This is the reason why we focus our analyses on this catchment, which is equipped with two large reservoirs (Bruwier et al. 2015; Cuvelier et al. 2018).

2.2 | Surveys

2.2.1 | Germany

After the flood of the July 2021, the University of Potsdam has conducted an online survey with 107 questions addressing the flood characteristics, information on warning, preparedness, socio-economic factors, aspects of mental health, damage incurred, and the state of recovery. The questionnaire is based on previous questionnaires used in Germany after flood events (Thieken et al. 2017). A different strategy was used to recruit respondents in each of the considered federal states (Figure 2). In RLP, the district administration sent letters with a link to the questionnaire to a random sample of households that had applied for emergency aid ('Soforthilfe'; not to be mistaken with the recovery funds-'Wiederaufbaufonds'), as most of the damage occurred in one district. The response rate was of 8%, with more than 5000 letters sent out, but more than 1000 could not be delivered and were returned. In NRW, people from the affected areas were invited to take part in the online survey through advertisements on Facebook and other social media. The survey was conducted between November and December 2022 and a total of 609 completed questionnaires were collected (277 in RLP, 332 in NRW).

The only (optional) information collected on the location of the respondents was a 3-digit postcode (Figure 2), allowing researchers to work with anonymous data, in line with GDPR requirements and the university's ethics committee.

2.2.2 | Belgium

Following the July 2021 flood event, the University of Liège conducted a field survey in residential buildings in the Vesdre valley (Figure 3) to collect data on direct flood damage and explanatory factors. A new questionnaire was designed based on questionnaires previously used to collect flood damage data in other countries (Ballio et al. 2018; Kreibich, Thieken, et al. 2017).

The survey covers all but two municipalities along the main course of the river Vesdre: Chaudfontaine, Trooz, Pepinster, Verviers, Limbourg, and Eupen (Figure 3). The municipality of Liège was left out because the flood hazard there is jointly influenced by several water courses (Ourthe, Meuse) and not only the river Vesdre. The municipality of Baelen was not incorporated in the study because a comparatively small number of buildings were affected (less than 3% of the total number of affected buildings in the valley), with limited damage.

The strategy for deploying the questionnaires was a paper-based field survey. The target areas for the field survey in each municipality were based on water depth observations by the regional river authorities (SPW), which can be viewed and retrieved from the geoportal of the Walloon region (https://geoportail.wallonie.be).

To foster participation in the survey, the population was informed in advance by means of posters displayed in the different municipalities and letters delivered door to door a few weeks before the

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

questionnaire. Both the posters and the letters contained instructions on how the citizens could set an appointment through text message, a phone call, or an online form accessible with a QR code. They could also just wait until the researchers visit their building. Around 38% of the contacted population participated in the survey, which was conducted from March 2022 to May 2023 and gathered a total of 420 completed questionnaires.

Data storage was organized in two parts. Part 1 contained personal information, such as the building address, while Part 2 collected anonymous research data: flood hazard features, building attributes, description of building damage, flood damage estimates, and capacity response variables such aslevel of preparedness, implemented mitigation measures, and received early warning. The two parts are linked by a unique identifier, allowing researchers to work with anonymized data, in line with GDPR requirements.

3 | Data and Methods

Figure 4 illustrates the workflow used to analyze the survey data. After harmonizing the data from the surveys conducted in Belgium and Germany (Section 3.1), the missingness rate of the data was assessed region by region (BE, RLP, and NRW), and variables with a high missingness rate (i.e., greater than 45%) were discarded. Once these variables were removed, an additional dataset (ALL) was created by combining the datasets from the three geographical regions. Following this, imputation was applied for each dataset (BE, RLP, NRW, ALL) to handle missing information in the remaining explanatory variables (Section 3.2). To address the potential dominance of certain features due to their wide-ranging values, normalization was performed on all continuous, ordinal, and integer-type variables, ensuring that the features were on a comparable scale. Next,

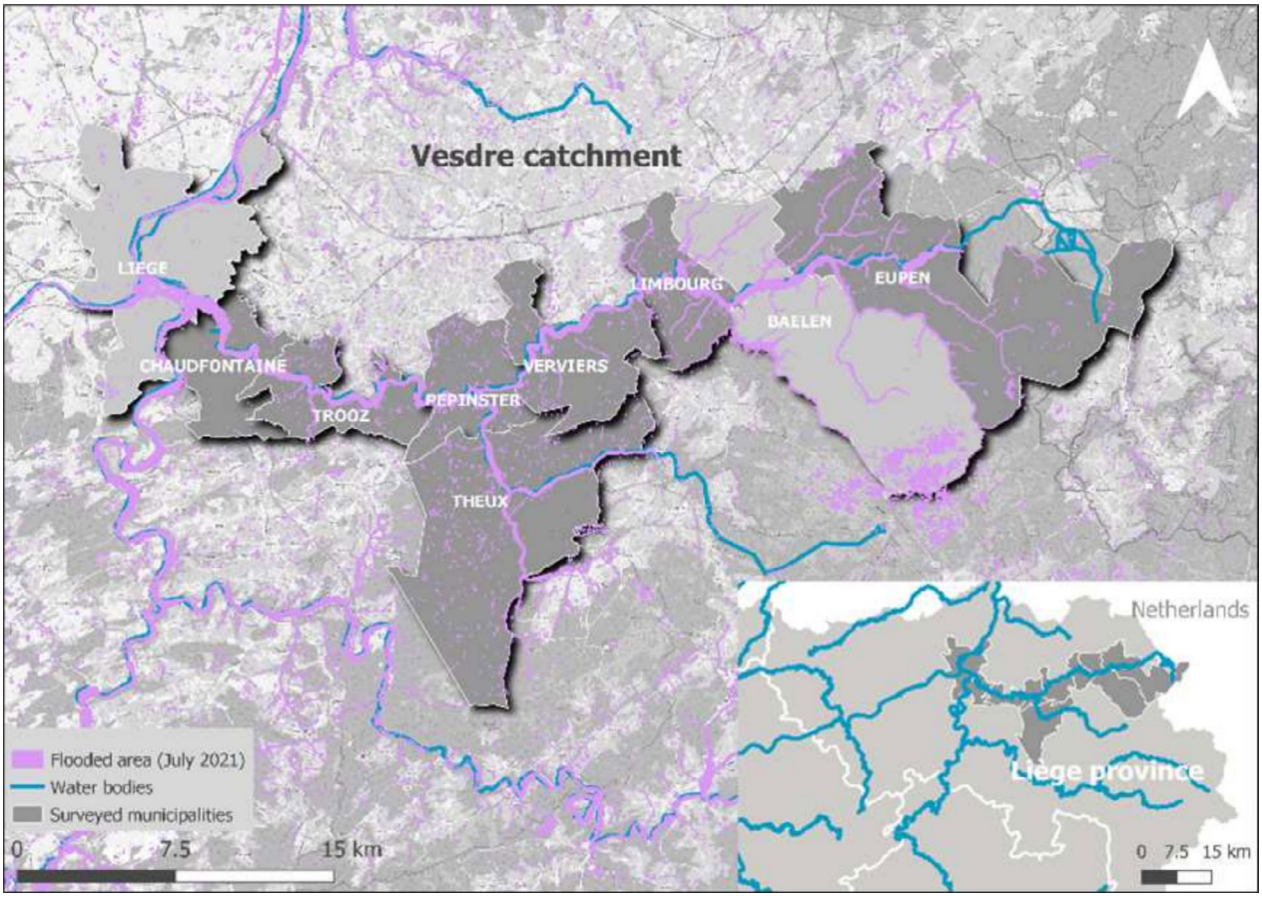

FIGURE 3 | Surveyed municipalities in Belgium—420 valid surveys.

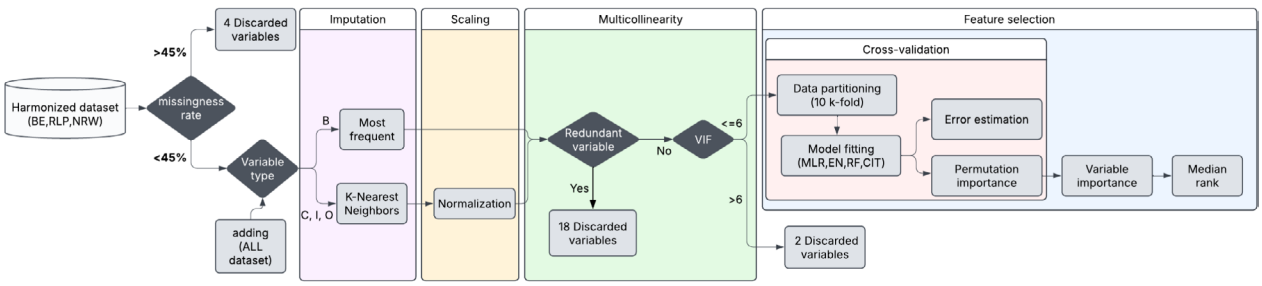

FIGURE 4 | Workflow developed for data processing and feature selection.

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

variables that can be explained by others were discarded, and some ordinal variables were grouped. The final variables were checked for multicollinearity, and highly correlated variables were eliminated (Section 3.3). With this, the datasets were ready (Section 3.4) for calibrating linear and non-linear models and assessing the importance of each individual feature, as detailed in Section 3.5.

3.1 | Data Harmonization

A subset of 57 variables, common to both the Belgian and German surveys, was selected. These variables include features that, according to expert knowledge, are essential for estimating flood damage. They encompass factors related to the hazard characteristics, building features, as well as the level of awareness, mitigation, and prevention measures implemented by residents, and the cost of damage to buildings and household contents induced by the flood.

The format of each of these variables in the datasets from Belgium and Germany was compared as regards the variable type (continuous, categorical, binary, or integer), the units, and their discretization into categories (for categorical variables). The format of each of the variables was harmonized between the two datasets. The type of the harmonized variables, their units, as well as their discretization are detailed in Table S1.

As much as possible, the initial variables were preserved without alteration. However, some adjustments were necessary.

- The most typical implemented adjustments consist in regrouping categories due to differences in discretization. For example, sediment classifications varied. In Belgium, sediments were categorized as mud, sand, gravel, stones, and boulders. To be comparable to data from Germany, these categories were consolidated into 'mud and sand' and 'stones and boulders'.
- In the Belgian questionnaire, water depth was recorded in ranges rather than as a continuous variable. To enable comparison with the German database and for modeling, a numerical variable was needed. Thus, the reported water depth, whether outside the building or on the ground floor, was represented by the midpoint of the ranges recorded in the Belgian survey.
- In Belgium, respondents indicated whether they had used any of the precautionary measures listed in Table S1 (part 'Preparedness'). Three options were given: during the flood event, after the flood event, or not possible/not planned. To align with German data, only measures reported as implemented during the flood event were included in the harmonized database.

Due to substantial differences in the hazard and building characteristics between the two surveyed areas in Germany (NRW and RLP), and given the context-specific nature of flood damage processes, the German dataset was divided into two subsets, each one corresponding to one region: Rhineland-Palatinate (RLP) and North Rhine-Westphalia (NRW). The Belgian dataset was left unchanged since all data collected in Belgium originate from the same valley, hence from a similar context (e.g., climate, morphological, or macro- economic characteristics). In the following, the term regions' refers both to the individual geographical regions (BE, RLP, NRW) and to their combination (ALL).

3.2 | Data Imputation and Scaling

Each dataset exhibits missing values. This missingness is characterized by the so-called missing at random (MAR) phenomenon (Gravesteijn et al. 2021), which means that the probability of a particular value to be missing depends on the values of other observed features. For instance, the duration of the inundation could not be reported for buildings where the inhabitants were evacuated.

According to the MAR phenomenon, the likelihood that missing data take a particular value is linked to the values of other observed variables. Therefore, the K-nearest neighbours (K-NN) method is employed for the imputation of continuous, ordinal, and integer explanatory variables. Based on a systematic evaluation of the performance of various imputation methods, with missingness rates between 10% and 50%, Jadhav et al. (2019) showed that K-NN consistently performed best, regardless of the dataset and the missingness rate. Using imputation has some drawbacks, such as 'weakening' or 'artificially strengthening' the relationship between variables and distorting statistical inference (Ramosaj et al. 2020; Shadbahr et al. 2023). However, this method could not be avoided here because, in each dataset, each entry had at least one missing information. In other words, simply discarding entries with missing data would have resulted in empty datasets.

The K-NN method assumes that comparable data points are characterized by similar values for the missing data. Hence, the missing values were estimated as the unweighted mean of the values at a number of nearest neighbors, determined by Euclidean distance. Here, this number of neighbors was taken equal to 4. In case of missing values at some of the nearest neighbors, the calculation excluded those data points. For Boolean variables, missing data were filled with the most frequently observed value for this variable. This is similar to defining a default value, a method also used in existing flood damage models (Dottori et al. 2016).

Before performing imputation, the degree of missingness of each variable was determined in each dataset (BE, RLP, NRW). Variables with a missingness rate exceeding 45% are reported in Table 1.

- A high rate of missing data was observed in two of the three considered regions (BE and RLP) for the time taken by residents to implement emergency measures. This can be attributed, on one hand, to the absence of flood warnings and, on the other hand, to insufficient time available to implement the necessary measures.
- In Belgium, only 20% of the surveyed residents responded to questions about their previous experience with flooding.
- Respondents in Belgium found it challenging to recall the precise moment when the flooding ended. This difficulty

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Variables with high missingness rate.

Variable	Label	Dataset	Missingness rate
Time to implement measures (h)	time_ measures	BE	0.82
Time since last flood event experienced (year)	fe_last_ flood_time	BE	0.80
Inundation duration (h)	f_duration	BE	0.80
Warning lead time (h)	warning_ time	BE	0.62
Time to implement measures (h)	time_ measures	RLP	0.55

arose for various reasons: in some instances, residents evacuated and were not present to witness the water receding from their homes, while in others, the summer period meant that occupants were away from home during the disaster.

- In several Belgian municipalities, authorities failed to issue warnings to the population. Hence, warning lead time was reported by less than 40% of the respondents.

These variables were excluded from the imputation process and the subsequent analysis (feature selection) due to insufficient field information to enable a reliable estimation of their values.

After discarding some variables due to their high missingness rate, the missing values of the remaining variables were imputed. Once the imputation of the explanatory variables was completed in each of the three datasets, the continuous, integer, and ordinal variables were normalized, meaning they were re-scaled to a range between 0 and 1. This transformation facilitates a meaningful comparison between variables and prevents certain features from overshadowing others based solely on their magnitude. The normalization was conducted using the min-max scaling method:

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where $x_{normalized}$ denotes the normalized variable, x the value of the variable before normalization, x_{min} the minimum value taken by the variable in the considered dataset (BE, RLP, NRW or ALL), and x_{max} the maximum value taken by the variable in the considered dataset.

No imputation was applied to the target variables (building and content damage); but these variables were also normalized according to Equation (1).

3.3 | Multicollinearity Assessment

To limit issues arising from multicollinearity between explanatory variables, we proceeded in three steps.

First, we discarded explanatory variables which are either strongly correlated with others or can be derived from others, as detailed in Table S2. For instance, the variable reflecting the presence of a heating system in the basement (heater_basement) was preserved, while the variable basement_size was discarded due to a high value of the Spearman rank correlation coefficient between these two variables. Similarly, the variable footprint_ area was kept but not the variables living_area, number_storey, and basement_ex, because footprint_area is calculated based on the three other variables, as detailed in Table 2. The same applies to variables describing the location of the heating system, the warning source, a socio-economic indicator, the ownership status (tenant vs. owner), and the size of the household.

Second, when a feature was represented by several boolean variables (i.e., presence of sediments or contaminants, building type, wall type, precautionary measures), two different approaches were used. For features 'Building type' and 'Wall type', the number of categories was simply reduced by grouping several possible answers, as can be seen by comparing Table S15 to Table 2. For the feature 'Precautionary measures', an integer indicator was introduced. It is calculated as the sum of individual scores assigned to each of the possible responses, similarly to the indicator for 'Emergency measures' already present in the initial dataset (Table S1). The individual scores depend on the assumed effectiveness of the corresponding measures. The indicator for 'Precautionary measures' accounts also for the insurance coverage of the respondents and their knowledge of hazard, while these aspects were previously coded in separate variables (insurance_binary and knowledge_hazard). An integer indicator was also introduced for features 'Sediments' and 'Contamination'. These two approaches result in a reduced number of variables, with more balanced distributions.

The feature reflecting previous flood experience of the respondents was initially coded in the form of an ordinal variable; but the distribution of the answers was particularly skewed (Figure S14). It was therefore simplified into a binary variable by grouping categories (Table 2).

After this processing of the data, multicollinearity was checked for each dataset (BE, RLP, NRW, ALL) using the Variance Inflation Factor (VIF). Though no single threshold value of VIF has gained universal acceptance, VIF values lower than 5–10 are generally considered suitable (Jeng 2023). There is general consensus that a VIF greater than 10 indicates serious multicollinearity. In this analysis, a threshold of 6 was chosen.

Two variables were excluded from the dataset to ensure low enough values of VIF for all variables: water_depth_gf and velocity_scaled (Tables S3 and S4). Despite a low value of VIF was obtained for variable velocity_scale, discarding it enabled a substantial decrease in the value of VIF for water_depth, as both variables are correlated. Although the variable water_depth lead to a

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Analysed variables for feature selection.

Common variables (C: continuous, B: Binary, I: Integer)				
Category	**Feature**	**Variables**	**Range/description**	**Type**
Absolute damage	Building damage	loss_b	Reported absolute building damage (€)	C
	Content damage	loss_c	Reported absolute content damage (€)	C
Hazard	Water depth	water_depth	Water depth outside the building (m)	C
	Flood intensity	flood_intensity	1: Would be swept away	B
			0: Else	
	Sediments	sediment_ind	1: Mud and sand	I
			2: Gravels, and possibly finer material	
			3: Stones and boulders, and possibly finer material	
	Contamination	contaminant_ind	1: Chemicals/sewage water/faeces	I
			2: Hydrocarbons, and possibly others	
	Debris impact	debris_impact	1: Yes	B
			0: No	
Building characteristics	Living area	living_area	$\text{footprint} \times (\text{number_storey} - \text{basement_ex})$ (m^2)	C
	Building type	building_type	1: Detached/semi-detached	B
			0: Else	
	Wall type	wall_type	1: Masonry/prefabricated/timber	B
			0: Else	
	Heating location	heat_basement	Location of the heating system before the flood	B
Socio-economic	Number of children	children	Number of children (< 14 years)	I
	Number of elderly	elderly	Number of elderly (> 65 years)	I
	Ownership	tenant	1: Yes	B
			0: No	
Awareness	Flood protection education	protection_ knowledge	Look up in advance how to flood-proof the home	B
	Flood experience	flood_experience	1: At least once	B
			0: Never	
Warning	Warning received	warning_binary	1: Yes	B
			0: No	
Preparedness	Precautionary measures indicator	precaution_ind	– Flood insurance (1 point) – Informed about flood hazard (1 point) – Moving heating system (5 points) – Non return valves (5 points) – Oil tank floating protection (5 points) – Adapted use of building (10 points)	I
	Emergency measures indicator	emergency_ind	– Oil tank floating protection (2 points) – Flood gates (5 points) – Pumping (5 points) – Moving furniture (2 points) – Turn off gas/electricity (2 points)	I

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

higher value of VIF than water_depth_gf, retaining water_depth was preferred because this quantity is more accessible in practice based on outcomes from hydrodynamic simulations. Similarly, keeping the variable flood_intensity was preferred compared to velocity_scaled because the latter was originally treated as an ordinal variable, where 1 represents a slow flow and 6 represents a torrential one. This has resulted in more scattered responses, whereas the information contained in variable flood_intensity was defined in an easily understandable manner for the respondents (e.g., 'water was strong enough to sweep away an adult person').

To ensure consistency in the full dataset (ALL), any independent variable discarded due to high multicollinearity in at least one of the regions was also discarded for the other regions.

3.4 | Analysed Variables

Table 2 shows the final set of variables, along with their range and units (before normalization), as used for the subsequent analysis.

- The target variables are the absolute damage to the residential buildings (tloss_b) and the absolute damage to their content (tloss_c), both originally expressed in euro.
- Five features describe the hazard characteristics likely to influence damage. These include the following variables: reported outdoor water depth (water_depth), as well as the perceived ability of a typical adult to withstand the flow (i.e., being able to stand in the flow vs. being swept away), coded as a binary variable (flood_intensity). The presence of various types of sediments, such as mud, sand, gravel, stones, and boulders, is also reported (sediment_ind), as well as the presence of contaminants in the floodwater (contaminant_ind). Two groups of contaminants are considered: (i) sewage water and/or feces and/or other chemical products, (ii) hydrocarbons, such as oil or gasoline. This information is encapsulated in a single integer variable. Additionally, the occurrence of impacts of floating objects (e.g., tree logs, debris, washed-away vehicles) is recorded in the binary variable debris_impact.
- Four features describe the characteristics of the exposed buildings, that is, the vulnerability of the structure in the engineering sense. These features include the living area (living_area), which is calculated by multiplying the building footprint area by the number of floors in the building, without considering the basement, due to its primary usage in the analyzed regions as a storage room or technical room and not as living space. Additionally, the building type is described by whether the building is either detached or semi-detached (i.e., with three or four facades), or attached (building_type). The wall material for the building structure is represented as a binary variable, indicating whether the building was constructed either with masonry, timber, or prefabricated materials, or with other materials (wall_type). These categories were selected because they have been recognized as reflecting the structural resistance of a given building to flood damage (Merz et al. 2013; Thieken et al. 2005). Whether the heating system was located in the basement is recorded through a boolean variable (heat_basement). This aspect was shown to be a significant damage-influencing feature in Belgium (Rodríguez Castro et al. 2025).
- The socio-economic characteristics of the flood-affected households are described through three features, including the number of children and elderly people living in the building and the ownership status (tenant or owner).
- The dataset also includes two variables reflecting the risk awareness of household members. The questionnaires contained yes-or-no questions about whether the interviewees were informed prior to the flood event about possible measures for flood-proofing their homes, and whether they had any prior flood experience. The latter was recorded using a binary variable: 0 for no prior flood experience, and 1 for at least one flood experience before the 2021 event.
- To assess whether warning of the household members in the immediate run-up to the flood event had an influence on the resulting damage, the variable (warning_binary) is used to indicate whether a warning was received by the inhabitants. The considered warnings could originate from any source (e.g., contacts, media, authorities, or official early-warning systems) and at any time, irrespective of the lead time before the event.
- The surveys differentiate between flood 'precautionary measures' implemented well in advance of the 2021 flood and 'emergency measures' taken shortly before, at the onset or even during the flood event. Interviewees were asked whether they had implemented precautionary measures as well as emergency measures. We aggregated this information into two indicators: one for precautionary measures and one for emergency measures. In these indicators, the points assigned to each implemented measure are summed. The number of points attributed to a specific measure is directly related to its potential damage-reducing effect, as described by Thieken et al. (2005). For example, emergency measures that protect the building from incoming water (e.g., sealing doors) or the installation of water pumps are given the highest score (5 points). Measures such as moving furniture or securing oil tanks are ranked lower (2 points), while switching off electricity and gas is scored even lower (1 point).

3.5 | Feature Selection

In order to identify the most influential variables for predicting flood damages in residential buildings, linear (Multi-linear regression and Elastic Net) and non-linear approaches (Random Forest and Conditional Inference Trees) were used (Sections 3.5.1 and 3.5.2). Taking into account the models' performance based on several evaluation metrics (Section 3.5.3), the most important variables to describe the flood damages were assessed (Section 3.5.4) and selected (Section 3.5.5).

3.5.1 | Linear Models

Two linear methods were utilized: multi-linear regression and elastic net.

We used a multi-linear regression (MLR) as described by Hastie et al. (2009) to predict the target variable y, from a vector of p

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

independent variables $\boldsymbol{x} = (x_1, x_2, \ldots, x_p)^T$, using the following equation:

$$\hat{y} = \beta_0 + \sum_{j=1}^{p} x_j \beta_j \tag{2}$$

where the term β_0 is the intercept, β_j the coefficients associated to each independent variable, and j the index assigned to each independent variable ($1 \leq j \leq p$).

The MLR model was calibrated using the ordinary least square method, which consists in finding the coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^T$ to minimize the residual sum of squares (RSS) over N observations:

$$\mathrm{RSS}(\boldsymbol{\beta}) = \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 \tag{3}$$

with $1 \leq p \leq N$, and i the index assigned to each observation ($1 \leq i \leq N$).

Elastic net (EN) is a method derived from MLR, in which an additional term is considered in $\mathrm{RSS}(\boldsymbol{\beta})$ to penalize high values of the β coefficients (i.e., regularization), which may lead to model overfitting (Altelbany 2021). The regularization used is based on a combination of the Lasso and Ridge penalties, which respectively involve the square and the absolute value of the coefficients (Zou and Hastie 2005):

$$\mathrm{RSS}(\beta) = \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^{p} \left(\alpha \beta_j^2 + (1-\alpha) |\beta_j| \right) \tag{4}$$

with λ a so-called complexity parameter which controls the degree of regularization/amount of shrinkage ($\lambda \geq 0$), and α a parameter determining the relative weight of the two types of penalties ($0 \leq \alpha \leq 1$).

3.5.2 | Non-Linear Models

In this section, we describe the non-linear methods used to identify important damage-influencing variables: Random Forest and Conditional Inference Trees. Both methods use a collection of trees (i.e., tree-like structured algorithms with internal nodes that split the data based on specific features, branches connecting the nodes representing possible outcomes, and leaf nodes at the end of the branches providing the final prediction). These trees are created from new datasets formed by randomly sampling the original dataset with replacement (i.e., bootstrapping). The methods differ in how they select and split variables to build the trees.

In random forest (RF), each tree from the collection is trained on one of the bootstrapped datasets, using randomly selected features. The result of the prediction of the RF is the average of the results of all the trees. This bagging method is used to decrease the high dependency on the training data which occurs when a single decision tree is used (Lamberti 2023). To build each individual tree, the classification and regression tree (CART) algorithm is used. The splitting criterion is based on the impurity of a sample, which is defined as the variance of the output y in the considered sample (Louppe 2014). The split which minimizes the variance is considered as the best split (Geurts and Wehenkel 2022).

The Conditional Inference Tree (CIT) algorithm (Hothorn et al. 2006) was also considered here. It aims at overcoming limitations of the CART algorithm in terms of variable selection biased toward categorical predictors (i.e., variables with many levels). These predictors allow numerous partitions and, therefore, more splitting possibilities, which can lead to model overfitting (Sieg et al. 2017; Hastie et al. 2009).

The main differences between CART and CIT are the methods used to select and split variables (splitting criterion) as well as to identify leaf nodes (stop criterion). CART identifies the variable with the best split based on a measure of node impurity. The splitting process stops either if a certain threshold of node impurity is reached or if no further splitting is possible. CIT makes use of hypothesis tests to identify the splitting variable at each node. At each node, a randomly chosen set of variables is used as candidate variables for splitting. The variable with the strongest association, measured by the p-value of the hypothesis test, with the response variable is selected as the splitting variable. If no association between the response variable and the predictor variable in the current node can be established, this node is defined as a leaf node (Sieg et al. 2017; Levshina 2021).

The RF method, as well as the linear models (ML and EN), were coded in the 'Scikit-learn' library for Python. The CIT model is included in the 'party' package of R, which can be connected to Python using the 'rpy2' library.

3.5.3 | Models Performance

The linear (MLR, EN) and non-linear models (RF and CIT) were calibrated on the datasets of the different regions (BE, RLP, NRW, ALL) adjusted to fix multicollinearity issues (Section 3.3). To avoid model overfitting (i.e., when the model memorizes noise and fits too closely to the training data, resulting in poor performance on unseen data), a cross-validation process was employed to train and evaluate the models. The k-fold cross-validation procedure involves splitting the data into K smaller sets. The model is trained on $K-1$ folds and validated on the remaining fold. The performance measure reported by k-fold cross-validation is the median of the validation results across all K folds. This method is chosen over a simple training, test, and validation split because it allows a larger training sample, which is crucial for improving model accuracy due to the relatively small sample sizes we have here.

The two linear (MLR, EN) and the two non-linear (RF, CIT) models underwent individual training for each target variable (content and building damage), using data from each region (BE, RLP, NRW and ALL). The models' performance were evaluated based on a 10-fold cross-validation approach (i.e., $K = 10$). The outcomes were assessed using the median of the mean squared error (MSE) of the 10 folds.

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.5.4 | Variable Importance

In order to assess the most important variables in the prediction of direct flood damage, we used the permutation importance (PI) technique. PI is a model inspection technique used to obtain the importance of features based on their impact on a trained model's predictions. It consists in randomly reshuffling the values of a single feature and evaluating the resulting degradation of the model's performance. By breaking the relationship between a particular feature and the target, the method enables determining how much the model accuracy relies on that feature. This random reshuffling was repeated L times in each fold. Here, L was set to 10.

For each model (ML, EN, RF, CIT), the importance $i_{j,k}$ of feature j in fold k ($1 \leq k \leq K$) was estimated as:

$$i_{j,k} = \frac{1}{L} \sum_{l=1}^{L} MSE_{j,l,k} - MSE_k \tag{5}$$

where MSE_k represents the reference value (i.e., without reshuffling) of MSE for the considered model in fold k, $MSE_{j,l,k}$ is the MSE value of the model trained on the l^{th} set ($1 \leq l \leq L$) of corrupted data (i.e., with reshuffling of the values of feature j) in fold k.

The final importance of each feature i_j was computed as the median of the feature importance values $i_{j,k}$ for all the K folds:

$$i_j = \text{median}\{i_{j,1}, \ \dots \ , i_{j,K}\} \tag{6}$$

3.5.5 | Feature Selection

Although water depth is recognized as the most important factor for predicting flood damages (Thieken et al. 2005; Gabriels et al. 2022; Huizinga et al. 2017; Endendijk et al. 2023; Ghamrawi et al. 2023), studies suggest that more accurate predictions are obtained by considering additional damage-influencing variables, that is, by opting for multi-variable damage models (Schröter et al. 2014; Molinari et al. 2020; Paulik et al. 2022; Carisi et al. 2018). These extra variables may provide further information on the hazard but also on other aspects such as the physical vulnerability of assets at risk, the socio-economic vulnerability of exposed households, their level of preparedness, and mitigation efforts.

Conversely, including numerous variables in a model also leads to drawbacks. The uptake (let alone the mainstreaming) of a sophisticated multi-variable flood damage model by operational agencies may be lower due to the (actual or perceived) burden of collecting the input data and keeping them up to date in the long run. Paucity in input data (typically on vulnerability aspects) may in turn hinder the accuracy of multivariable models when applied in practice. Besides, the explainability of the model outcomes may be reduced in the case of complex multivariable damage models compared to simple monotonous depth-damage functions. This may reduce the acceptability of the damage model by various stakeholders. It is therefore of utmost importance to point to the most critical features for achieving valid estimations of flood damage.

Here, the outcomes of the permutation importance method described in Section 3.5.4 were used to rank the 17 considered independent variables from the most important one to the less important one for estimating flood damage. This is performed independently for each model (MLR, EN, RF and CIT), in each region (BE, RLP, NRW, ALL), and for each target variable (building and damage content). A normalized rank was introduced, ranging from 0 (least important variable) to 1 (most important, variable). The median normalized rank across the four models was calculated to determine the final score for each feature within each region.

Besides ranking the variables as a function of their importance, the number of variables to be included in a model needs to be chosen. This number typically results from a trade-off between the accuracy of predictions and the operationality of the model. The number of variables included in existing damage models for riverine floods (by far the most common) (Thieken et al. 2008; Scawthorn et al. 2006; Bubeck and de Moel 2010; Huizinga et al. 2017; Carisi et al. 2018), flash floods (Arrighi et al. 2020; Fuchs et al. 2012) and pluvial floods (Rözer et al. 2019) ranges from two (e.g., water depth and building or land value; Huizinga et al. 2017) to ten or more (Dottori et al. 2016; Vogel et al. 2012; Schröter et al. 2014). Hence, there is currently no consensus on the optimal number of variables to be considered in a damage model. The results discussed hereafter focus primarily on the six most important variables. This was considered a reasonable trade-off between the amount of considered information on damage-influencing factors and the effort for data collection, as well as the chances of model uptake in practice.

4 | Results and Discussion

The performance of the four tested models is reported in Section 4.1. The most influential variables are presented in Section 4.2. The selection of the most important features is discussed in Section 4.3 for both building and content damage.

4.1 | Models Performance

This section presents the performance of the four considered models (MLR, EN, RF, CIT) in terms of MSE. Here, the models are not utilized for the purpose of predicting flood damage, but they serve as a basis for identifying the most significant damage drivers (Section 4.3) and understanding the challenges faced when developing cross-country transferable models, particularly in the step of feature selection.

Figure 5 and S30 in Supplement compare the median (over the K folds used for cross-validation) of the MSE for the four implemented models (MLR, EN, RF, CIT) and for each of the four datasets (ALL, BE, RLP, NRW). For building damage, the dataset covering all three regions (ALL) leads to a lower MSE than when the regions are considered individually. This is likely due to the larger sample size, covering a broader range of combinations of variable values, which is known to lead to an improved model performance (Sieg et al. 2017).

For a given region, the differences in performance between the three models EN, RF, and CIT are generally insignificant (Figure 5 and S30). When regions are analyzed individually (BE, RLP,

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

NRW), only MLR shows a higher MSE value than the other three models in some cases. Nonetheless, when the dataset combining all regions is used (ALL), the value of MSE becomes independent of the chosen model (see left panel in Figure S31). This hints at the value of combining data from several regions to achieve better accuracy. Including cases from different regions helps to better capture the diversity of damage-influencing processes in residential buildings. Such a comprehensive dataset (ALL) unlocks a more robust understanding of flood damage mechanisms, potentially improving the predictive capacity of the models.

The model performance analysis presented here indicates that none of the four models consistently outperforms the others. Consequently, it is deemed appropriate to continue with the remainder of the analysis using all four models. In addition, the relatively low error values across the models suggest that they are reliable in capturing the underlying damage mechanisms, allowing us to use them with confidence to identify the most important variables.

4.2 | Variables Importance

For each considered region and for the four tested models, Figure 6 shows the ten most influential explanatory variables (i.e., those leading to the highest values of variable importance i_j) for building damage, while the most influential variables for content damage are displayed in Figure S23.

Overall, Figure 6 reveals a remarkable agreement in the most influential predictors identified for the different regions and using different models. This is shown by the similar ranking of several explanatory variables:

- In terms of importance, the variable water depth is either first or second in all cases, regardless of the model used and of the region considered. This result is consistent with most previous studies and existing models (e.g., Sieg et al. 2017; Dottori et al. 2016; Fuchs et al. 2012; Rözer et al. 2019; Vogel et al. 2012; Nafari et al. 2016; Kreibich, Botto, et al. 2017).
- Similarly, the living area is ranked first or second irrespective of the model and in all regions except one (NRW). This is explained by the fact that, in the dataset NRW, living area and water depth are significantly correlated (Figure S22). It is not the case for the other two regions (Figures S20 and S21).
- For the datasets ALL and BE, the building type ranks third or fourth, whatever the model used. Exactly the same is found for the dataset RLP, except when RF is used. In this case, the building type ranks seventh (instead of third or fourth). The dataset NRW leads to building type being ranked fourth to tenth depending on the model. Hence, a greater difference is found between the two German regions than between Belgium and Germany (RLP). This suggests a limited influence of the data collection methods, which differ between the two countries.
- The presence of sediments is in all cases among the ten most influential features. When the transnational dataset ALL is considered, the corresponding variable ranks third, irrespective of the model. For each model, the importance of this variable remains very similar between the two regions BE and RLP, which both experienced extreme flooding. In contrast, the results appear different for NRW. In this case, the sediment indicator is systematically ranked first or second for all four models. Hence, once again, more difference is found between NRW and RLP than between BE and RLP.

Remaining differences between regions are expected, as they reflect differences in hazard, vulnerability, and coping capacity across regions, as highlighted by Scorzini et al. (2022).

As can be seen in Figures 6 and S23, the features ranked from one to six generally have a much greater influence on flood damage than all the following ones. This is highlighted by a drop in the value of the importance observed between the variables ranked in the first six positions and the others. In most cases, the importance of variables ranked from seven to ten is almost insignificant. Therefore, only the six most influential variables in each case are kept for the following analyses.

Figure 7 shows the Spearman's rank correlation coefficients between the variables in the complete dataset (ALL). The left panel shows the correlation of the variables with the building damage, while the right panel shows the correlation with the content damage. Water depth has the highest positive correlation with both building and content damages. Other variables describing the hazard, such as the presence of sediments and contaminants, show significant positive correlations with the damage. Two building characteristics (wall type and location of the heating system in the basement) also show a significant positive correlation with both types of damage. These variables have

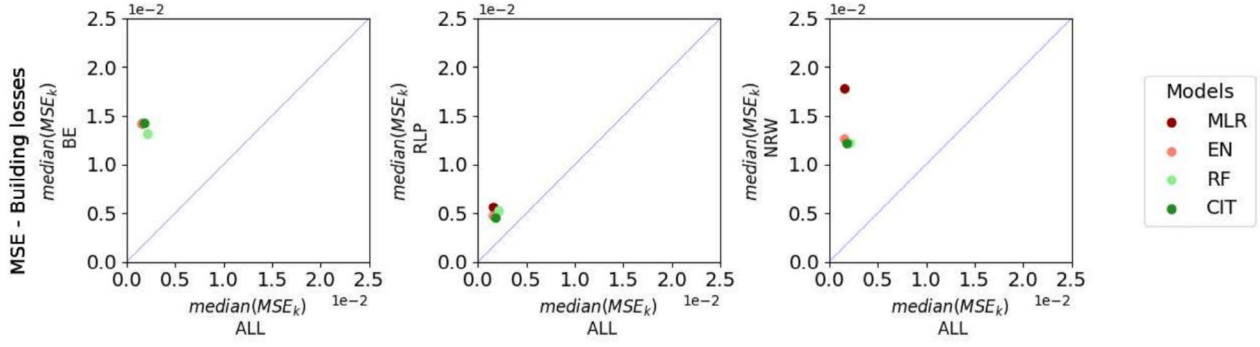

FIGURE 5 | Median over the K folds of the mean squared error (MSE) for building damage estimation across all regions.

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

previously been identified as important predictors of building damage. The building type is primarily associated with structural damage (Maiwald et al. 2022; Suppasri et al. 2013; Laudan et al. 2017). In a recent study focusing on Belgium, the location of a costly item such as a central heating boiler in the basement was also reported to have a significant impact on flood damage in residential buildings (Rodríguez Castro et al. 2025).

The variables indicating the type of building and the ownership status of the household (tenant vs. home owner) are also significantly correlated with the content damage. The corresponding variables show a significant negative correlation between them. It is therefore plausible that both being a tenant and living in an attached building (i.e., building_type = 0) are associated with a generally lower socio-economic level, hence less valuable possessions being exposed and a lower content damage.

Not all the most important features identified by the different models (Figure 6) were found to be significantly (nor monotonically) correlated with damage. This can be attributed to the nature of non-linear methods, which can capture non-monotonic and multivariable relationships, as well as take into account interactions between explanatory variables (Sieg et al. 2017). These aspects are not captured by a simple Spearman's rank correlation analysis.

Figures S20–S22 show the correlation matrix for the individual regions. The differences between the most important variables in the studied areas, along with their correlations, are analyzed and discussed in the following section.

4.3 | Feature Selection

A multi-model approach was chosen for selecting the main damage drivers, as this strategy allows us to balance the strengths and weaknesses of each model, minimizing the bias that could arise from relying on a single model. We hypothesize that the robustness of the feature selection will be enhanced if it builds upon the outcomes of the four considered models (MLR, EN, RF, CIT), rather than on a single one. Therefore, we propose to ultimately select the important features in each region based on the median of the normalized ranks provided by the four tested models (see Section 3.5.5).

For both building and content damage, water depth stands out as a key influential variable, irrespective of the considered region and the utilized model (Tables 3 and 4). Other variables which consistently appear as highly influential across regions differ between building and content damage.

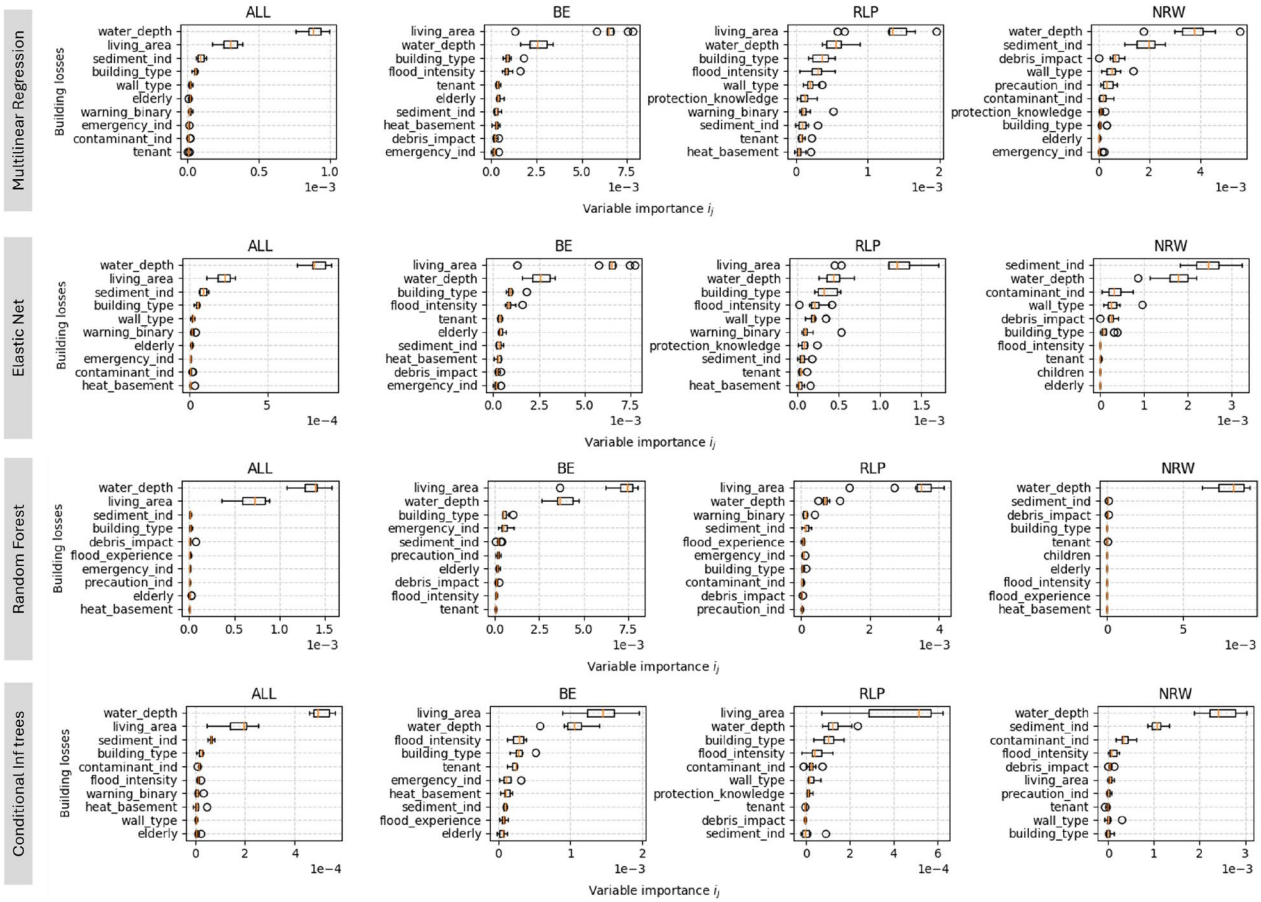

FIGURE 6 | Variable importance $i_{j,k}$ for the prediction of building damage. The boxplots reflect the variability of $i_{j,k}$ between the K folds. The box extends from the first to the third quartiles, with the orange mark indicating the median value (i_j). The whiskers extend to the farthest data point lying within 1.5 times the inter-quartile range from the box. The data situated beyond the whiskers are plotted individually using the marker 'o'.

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

- For building damage, three building characteristics are identified as key factors in at least three of the four datasets: living area (reflecting the exposure), building type (i.e., the number of facades potentially exposed to water) and wall type (i.e., the construction material).
- For content damage, two variables are identified as highly influential in all or most of the datasets: ownership status (i.e., whether the inhabitant is a tenant or an owner) and emergency indicator (actions taken in response to a flood event).

In the following, we provide a more detailed account of the obtained results, and we discuss them in light of differences in the variables distributions across the studied regions.

4.3.1 | Building Damage

Table 3 shows that water depth and living area are the most significant variables for predicting building damage. These two variables are commonly used in existing flood damage models. Additional hazard variables, building characteristics, and socio-economic factors also emerge as important.

In terms of building characteristics, two variables appear to be important for damage estimation: building type and wall type. Building type (like water depth) is one of the most important factors in all the regions. This can be attributed to the fact that a detached or semi-detached building has a greater area of floodwater in direct contact with the building than an attached building. In the latter case, a shielding effect by neighboring buildings can help to limit damage to the building, whereas this is less the case for (semi-)detached buildings.

For all regions except Belgium, the type of walls is found to be a highly relevant feature. Figure S8 shows that masonry, timber, or prefabricated construction are predominant in all regions, and these types of materials are generally considered less resistant to floods compared to reinforced concrete or natural stone. The absence of this factor among the most influential ones in the Belgian dataset may be explained by a low variability of the data, as over 80% of the buildings surveyed in Belgium are constructed with masonry.

In two of the four datasets (ALL and BE), the socio-economic variable 'elderly', which counts the number of people over 65 years of age living in the building, emerges as important for estimating building damage. This may be related to the reduced ability of the elderly to implement emergency measures (as confirmed in Figure 7) or precautionary measures. Older adults may have mobility limitations, health problems, or cognitive impairments that make it more difficult for them to prepare effectively for a flood. These factors can significantly increase their vulnerability, not only in terms of physical safety, but also in terms of their ability to protect their property. As a result, the presence of elderly residents in a building can influence the damage and the overall flood risk faced by the household.

The hazard variable 'flood intensity' (i.e., how strong the floodwater momentum was perceived to be) emerges as a key driver for the regions BE and RLP. This is consistent with previous studies, which have found the intensity of floodwaters to be relevant for predicting structural damage and associated losses (Kreibich et al. 2009). This is consistent with the extreme nature of the flood in these two regions. In the other two datasets (ALL and NRW), the sediment indicator stands out as an important predictor, reflecting the additional costs, such as cleaning, associated with the presence of mud and other sediments in floodwaters. Compared to the datasets BE and RLP, the dataset NRW contains more evenly distributed values of the variable sediment_ind, including a substantial number of cases where clear water was reported. This may explain the higher importance of this variable for the dataset NRW. In contrast, in Belgium, about 95% of the respondents reported the presence of mud (Figure S3). In this case, the lack

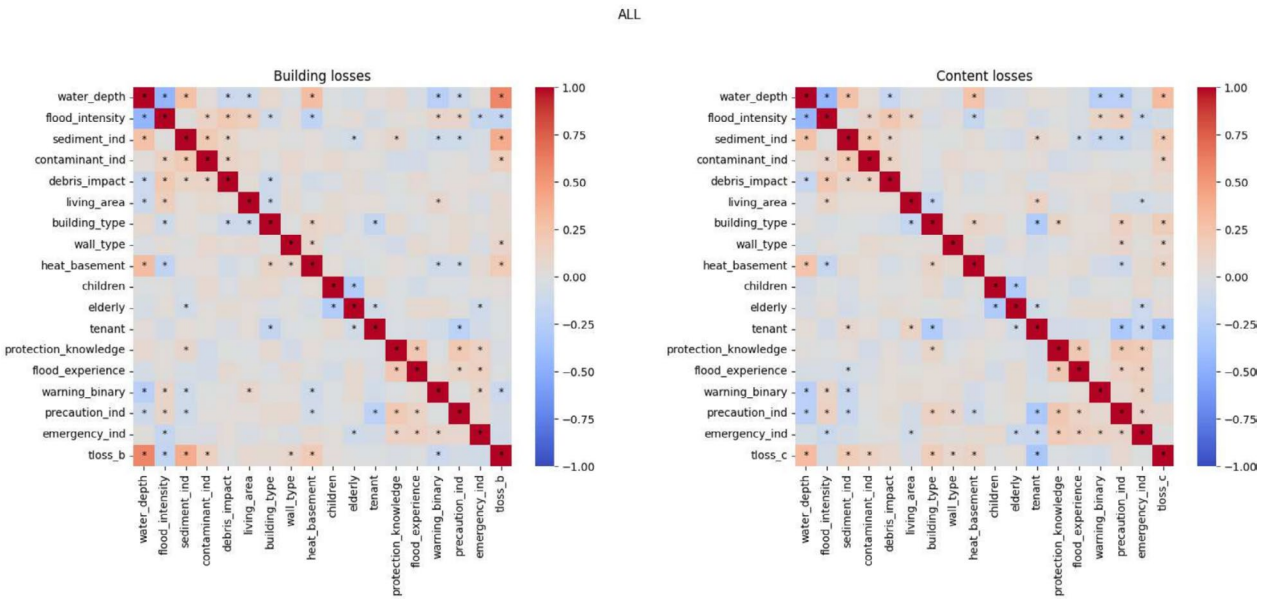

FIGURE 7 | Spearman correlation for the selected features in dataset ALL. Stars show a significant correlation at the 5% significance level.

17533 8x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 | Normalized feature importance ranks for feature selection in building damage prediction using linear and non-linear methods.

Variable	MLR	EN	RF	CIT	Median rank	Dispersion	Correlation
All regions							
Water depth	1.00	1.00	1.00	1.00	1.00	0.00	(+)
Living area	0.94	0.94	0.94	0.94	0.94	0.00	(+)
Sediment indicator	0.88	0.88	0.82	0.88	0.88	0.03	(+)
Building type	0.82	0.82	0.88	0.82	0.88	0.03	(+)
Wall type	0.76	0.76	0.29	0.53	0.65	0.23	(+)
Elderly	0.71	0.65	0.65	0.47	0.65	0.10	(−)
Belgium							
Living area	1.00	1.00	1.00	1.00	1.00	0.00	(+)
Water depth	0.94	0.94	0.94	0.94	0.94	0.00	(+)
Building type	0.88	0.88	0.88	0.82	0.88	0.03	(+)
Flood intensity	0.82	0.82	0.53	0.88	0.82	0.16	(−)
Tenant	0.76	0.76	0.47	0.76	0.76	0.15	(−)
Elderly	0.71	0.71	0.65	0.47	0.68	0.11	(+)
Rhineland-Palatinate							
Living area	1.00	1.00	1.00	1.00	1.00	0.00	(+)
Water depth	0.94	0.94	0.94	0.94	0.94	0.00	(+)
Building type	0.88	0.88	0.65	0.88	0.88	0.12	(−)
Flood intensity	0.82	0.82	0.35	0.82	0.82	0.24	(+)
Wall type	0.76	0.76	0.24	0.71	0.74	0.26	(+)
Warning	0.65	0.71	0.82	0.35	0.68	0.20	(−)
North Rhine-Westphalia							
Water depth	1.00	0.94	1.00	1.00	1.00	0.03	(+)
Sediment indicator	0.94	1.00	0.94	0.94	0.94	0.03	(+)
Debris impact	0.88	0.76	0.88	0.78	0.82	0.07	(+)
Contaminant indicator	0.71	0.88	0.18	0.88	0.79	0.33	(+)
Wall type	0.82	0.82	0.41	0.53	0.68	0.21	(+)
Building type	0.59	0.71	0.82	0.47	0.65	0.15	(+)

of variability in the values taken by this variable hampers its predictive capacity.

In NRW, additional hazard variables such as the contaminant indicator and debris impact emerge as important predictors. Although this region was less affected than BE and RLP, the prominence of these variables may again be due to a more even distribution of the data compared to the other two regions. While hydrocarbons were the predominant contaminant in BE and RLP (with around 50% of respondents reporting their presence), only 30% of respondents in NRW reported the presence of this type of contaminant.

The findings described above align with existing research showing that more variables than water depth alone are critical for accurate prediction of flood damage. Several building characteristics consistently emerge as significant in all regions, as they allow for a better assessment of the physical vulnerability of exposed assets. In some cases, socio-economic variables, such as ownership status or age of respondents, also stand out, providing insight into the building's economic value or potential mitigation measures. Similarly, additional hazard variables (presence of sediments or flood intensity) are found to be important in some, but not all, regions. The distribution of the data collected appears to play a role in the relative importance of the variables, highlighting the need for a representative sampling when conducting post-flood surveys.

When comparing the variables identified as relevant for flash floods with those typically used in riverine flood damage models, we find

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | Normalized feature importance ranks for feature selection in household damage prediction using linear and non-linear methods.

Variable	MLR	EN	RF	CIT	Median rank	Dispersion	Correlation
All regions							
Water depth	1.00	1.00	1.00	1.00	1.00	0.00	(+)
Tenant	0.94	0.94	0.88	0.94	0.94	0.03	(−)
Living area	0.88	0.65	0.94	0.88	0.88	0.13	(−)
Emergency indicator	0.71	0.53	0.82	0.82	0.76	0.14	(−)
Sediment indicator	0.47	0.76	0.71	0.76	0.74	0.14	(+)
Flood intensity	0.82	0.88	0.59	0.59	0.71	0.15	(−)
Belgium							
Living area	1.00	0.88	1.00	1.00	1.00	0.06	(+)
Emergency indicator	0.94	1.00	0.88	0.94	0.94	0.05	(+)
Water depth	0.88	0.94	0.94	0.88	0.91	0.03	(+)
Debris impact	0.82	0.82	0.76	0.82	0.82	0.03	(+)
Building type	0.76	0.76	0.53	0.65	0.71	0.11	(+)
Tenant	0.53	0.71	0.76	0.76	0.71	0.10	(−)
Rhineland-Palatinate							
Tenant	1.00	1.00	1.00	1.00	1.00	0.00	(−)
Water depth	0.94	0.94	0.94	0.88	0.94	0.03	(+)
Precaution indicator	0.88	0.82	0.82	0.94	0.85	0.06	(+)
Contamination indicator	0.71	0.88	0.65	0.71	0.71	0.10	(+)
Protection knowledge	0.82	0.76	0.24	0.53	0.65	0.27	(−)
Emergency indicator	0.53	0.53	0.76	0.76	0.65	0.14	(+)
North Rhine-Westphalia							
Water depth	1.00	0.82	1.00	1.00	1.00	0.09	(+)
Tenant	0.94	1.00	0.88	0.94	0.94	0.05	(−)
Sediment indicator	0.82	0.94	0.71	0.88	0.85	0.10	(+)
Contaminant indicator	0.65	0.88	0.76	0.82	0.79	0.10	(+)
Wall type	0.88	0.76	0.47	0.76	0.76	0.18	(+)
Warning	0.71	0.71	0.35	0.71	0.71	0.18	(−)

both similarities and differences. Riverine flood damage models generally rely on a small set of key variables—most commonly, water depth, the economic value of the building, and living area (Huizinga et al. 2017; Vanneuville et al. 2006; de Bruijn et al. 2015; Dutta et al. 2003; Arrighi et al. 2020). In our study of flash floods, water depth and the living area are also found to be significant. It is now generally accepted that considering additional hazard and vulnerability factors is crucial to accurately estimate flood damage. More sophisticated models have emerged that account for hazard-related variables, such as return period, flow velocity, water contamination, as well as building-specific characteristics such as building type, basement area, and structural features (Thieken et al. 2008; Carisi et al. 2018). Even more comprehensive models like Vogel et al. (2012); Dottori et al. (2016); FEMA (2009) integrate numerous hazard and building variables to provide a more detailed and robust damage assessment. Our findings align with this broader perspective, indicating that a more nuanced understanding of hazard and building characteristics is essential for better predicting flash flood impacts. Specifically, variables such as sediment transport and flood intensity emerge as crucial in flash flood damage assessment, where high flow velocities and the debris impact can have severe consequences. These variables highlight the unique characteristics of flash floods, which are often associated with rapid onset and intense local impacts. Furthermore, building characteristics, such as building type, which are considered in some riverine models (Thieken et al. 2008; Scorzini et al. 2022), are also relevant in the context of flash floods. Likewise, wall type, which plays a significant role in some riverine models (Carisi et al. 2018), proves to be important in flash flood damage estimation, whereas it is not always as critical in the context of riverine floods.

17533 8x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In conclusion, although the variables found to be relevant for flash flood damage prediction are indeed included in some previously developed riverine flood damage models, their consideration in these models is not systematic. The present study highlights the need to integrate these more diverse hazard and building-related variables into flood damage models, particularly for flash floods, where traditional riverine models may fail to capture the full range of impacts.

4.3.2 | Content Damage

As with building damage, water depth is consistently identified as particularly important for content damage estimation (Table 4). In addition, the ownership status of the respondent appears to be an important variable for content damage estimation in all datasets. All four datasets show that being a tenant is negatively correlated ($p < 0.05$) with the amount of content damage (Figure 7 and Figures S20–S22). This result could be attributed to a generally lower purchasing power of tenants. Also, since tenants often live in a property for shorter periods of time, they may be less likely to invest in expensive or bulky items that would be at risk in the event of a flood, unlike homeowners who may accumulate more belongings over time. These two factors could explain that less valuable possessions are exposed in the case of tenants compared to homeowners, and therefore reported damage to content is lower in dwellings occupied by tenants.

From the combined dataset, as well as from the datasets BE and RLP, the indicator of implemented emergency measures stands out as one of the most important drivers of content damage (Table 4). In many cases, moving content items from basements or lower floors to upper levels could prevent them from coming into contact with floodwaters, thereby reducing the content damage. This differs from the results obtained for building damage (Table 3). Nonetheless, the right-hand panel of Figure S20 shows that, for Belgium, the indicator of implemented emergency measures is positively correlated with content damage, suggesting that, in this region, the damage was not reduced by the implemented emergency measures. This result may be attributed to the extreme nature of the event, with a median water depth of approximately 1.80 m (Figure S1), suggesting that in some cases the protective measures were insufficient to prevent damage to household content. Moreover, Figure S17 suggests that the adoption of emergency measures remains limited in all analysed regions, particularly in Belgium.

In the German region of NRW, the indicator for implemented emergency measures shows relatively low importance (median normalized rank of 0.47). However, it is significantly negatively correlated with content damage ($p < 0.05$). This could be explained by the fact that the hazard in NRW was less intense, with lower water levels than those reported in BE and RLP. As a result, the emergency measures may have helped reduce content damage. This is in contrast to the other regions, where no such negative correlation is observed.

Similar to the building damage, when analyzing all the datasets (ALL), additional hazard characteristics emerge as key drivers of content damage, namely the sediment indicator and flood intensity (e.g., high flow velocity causing appliances and other belongings to be knocked against walls or even swept away). Regarding the sediment indicator, it shows a significant positive correlation in the German region of NRW. In the other regions, however, the very low variability in the responses (as shown in Figure S3) is probably the reason why this variable is not identified as one of the most influential variables. Although flood intensity does not appear in the top six most important features in the individual regions, it is still important in Belgium, where it ranks seventh with a median importance score of 0.68. As shown in Figure S2, more than 90% of respondents in Belgium reported that an adult would have been swept away by the floodwaters, while in the German regions, respondents predominantly chose 'deep water' as the most appropriate description. When all datasets are combined, there is a more even distribution of responses between 'swept away' and 'deep water', which helps to better differentiate the influence of each factor in the models. This may explain why flood intensity is ultimately identified as an important driver of content damage.

In addition to these hazard variables, the building characteristic 'living area' appears to be a key driver for predicting content damage, as the absolute damage is influenced by the area exposed to the flood.

Although early warning does not appear in the top six most important drivers in most of the studied regions, it remains a significant feature in NRW, where the variable associated with early warning ranks sixth in terms of importance, with a median normalized rank of 0.71. As shown in Figure S22, this variable is significantly negatively correlated with content damage, suggesting that when an early warning is issued, people are more likely to take protective measures, such as moving their belongings to higher floors, that reduce content damage.

Figure S15 in reveals that in NRW, about 50% of respondents reported receiving an early warning, regardless of the lead time for implementing emergency measures. In contrast, in Belgium, about 80% of respondents reported receiving an early warning, although most of these warnings (about 42%) came from personal observations rather than official sources. This highlights a significant disparity in the effectiveness and sources of early warning across regions, which is likely to have had an impact on the extent of content damage.

Another important feature identified when analyzing the regions individually is the occurrence of debris impact in the Belgian dataset, where it was found to be significantly positively correlated with content damage (Figure S20). As detailed by Wüthrich et al. (2025), the event was characterized not only in the Ahr valley (Germany) but also in the Vesdre valley (Belgium) by a large amount of floating debris that caused significant damage to infrastructure. However, this variable was not found to be relevant in the analysis of damage to buildings, where it might have been expected to be more significant.

This discrepancy could be explained by the way the question was phrased in the questionnaires and subsequently recorded and analysed. Residents were asked, 'What do you think was the main cause of the damage to the building?', with the impact of floating debris being one of the response options. Given that many other factors were relevant during the event—such as

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

water level, flow velocity, sediment and debris accumulation—the impact of debris was reported in only a few cases. In most regions, less than 20% of respondents mentioned debris impact, with Belgium reporting the highest proportion at 30% (Figure S5).

In Rhineland-Palatinate, in addition to emergency measures, which were also identified as key drivers in other regions, precautionary measures and the variable 'protection knowledge' were found to be important. Over 80% of the respondents in all regions stated that they has not received any information on building-level flood protection. This may be due to the extreme nature of the event, which severely affected areas not identified as flood prone in the official regional flood hazard maps. However, this finding emphasizes the importance of knowledge on building-level protection in reducing content damage, especially in less extreme floods. It underscores the need for better public education and awareness on building-level flood protection measures, especially as floods are expected to become more frequent and affect larger areas.

In summary, similar to the prediction of building damage, predicting household damage also requires the inclusion of additional variables to provide a more accurate estimation. Despite differences in key drivers between regions, it is essential to incorporate not only explanatory variables related to the hazard but also factors that address awareness, preparedness, and the emergency measures implemented by residents. These elements are particularly important for mitigating damage to household content.

5 | Limitations

A limitation of this study is the relatively small sample size, which may affect the generalisability of the findings, particularly in the context of training and testing the machine learning models. As previous studies have shown, the performance of these models improves with a larger training dataset, as increased sample size enhances statistical power. With more observations, the models can generate more accurate and reliable results, leading to better performance and stronger predictions (Rajput et al. 2023; Chen et al. 2024). However, collecting such data is inherently challenging due to the complexity and resource-intensive nature of data collection in different regions and the variability of local conditions. To address this, we have increased the sample size by merging datasets from three regions, which helps to mitigate the limitations of a single-region dataset. Moreover, the data collection processes were originally conducted separately, each serving different research purposes and addressing distinct questions, with varying collection strategies. However, the results show at several instances that the differences are greater between RLP and NRW (very similar data collection method) than between RLP and BE (different data collection methods). This suggests that the differences in hazard and contextual factors are more significant than the impact of the differing data collection methods.

To address the issue of differences in data collection approaches between regions, which require assumptions in the data harmonization process, a coordinated transnational data collection approach is recommended. In this approach, key variables would be collected consistently across regions to ensure comparability. This would facilitate the analysis of data from future flood events, create longitudinal, transnational datasets that can be updated as new flood events occur, and ultimately enhance the robustness of the models over time.

Much of the data used here relies on people's testimonies, which provide valuable insights but can be subjective. To obtain more reliable estimates of hazard (e.g., water depth, flow velocity, flow duration), it is crucial to consider validated hydrodynamic models to provide accurate inputs for damage modelling. By combining these hazard variables from hydrodynamic models with physical and socio-economic vulnerability data of buildings and their residents collected in the field, multivariable damage models could provide more accurate predictions of flood risk. The 'reference' values of the target variables (i.e., damage) are also subject to high uncertainties. Although most of the collected damage data are not primarily based on the respondent's perception, but are supported by more factual material (compensation claims, expert reports, quotes from construction companies...), it would be valuable to perform a systematic comparison between data collected through in-person interviews or online questionnaires (as utilized here) and data surveyed by researchers and/or experts in the field. Along the same line, a promising future research direction consists in combining lessons learned from survey-based analyses as carried out here with structural engineering investigations focused on the failure mechanisms induced by hydrodynamic loads on structural components of buildings, possibly using laboratory experiments (Jansen et al. 2020).

In this study, we used four models to identify the key variables for flood damage estimation, but there are other promising methods that could be explored in future research. For instance, Support Vector Machines (SVM) (Mosavi et al. 2018) have shown potential in various predictive tasks and could provide valuable insights for flood damage modeling. Additionally, XGBoost, a highly effective gradient boosting algorithm, has gained popularity for its ability to handle large datasets and complex relationships, making it a strong candidate for improving prediction accuracy (Dey et al. 2024; Ma et al. 2021). Testing these and other advanced techniques could further enhance the robustness and accuracy of flood damage models in the future.

6 | Conclusion

Data from independent flood damage surveys conducted in Germany and Belgium were jointly analyzed to identify the most important explanatory features. All data refer to the European flood of 2021, which affected Germany and Belgium and showed typical flash flood characteristics in many places. A multi-model approach (MLR, EN, RF and CIT), combined with a permutation importance technique, allowed the following conclusions to be drawn regarding the main drivers of flood damage to buildings and contents.

In line with previous research, water depth emerges as the most influential variable for predicting both building and content damage, regardless of the region and the model used. When combining the data from Belgium and Germany, factors other

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

than water depth were also identified as important for improving the assessment of flood damage.

Three physical characteristics of the buildings stand out as important variables in the assessment of building damage. These factors are living area, building type (i.e., the number of facades potentially exposed to floodwaters), and wall type (i.e., the construction material used). Previous studies (Laudan et al. 2017; Suppasri et al. 2013; Maiwald et al. 2022) have already identified the wall type as crucial for understanding structural damage. Unlike riverine floods, flash floods, such as those observed during the 2021 flood event, can cause different patterns of damage, highlighting the significance of these building physical characteristics in flood damage assessment.

To a lesser extent, sediments in the floodwater and the presence of elderly people in the household also emerge as important features for predicting building damage. The former reflects the additional cleaning and waste disposal costs associated with the presence of sediments, which are often overlooked in existing flood damage assessment methods. The latter is likely to be due to the challenges faced by the elderly in implementing precautionary or emergency measures. In the two most severely affected regions (Belgium and Rhineland-Palatinate), the propensity of the flood to sweep away an adult (referred to as 'flood intensity') also appears to be among the most influential variables for predicting damage to buildings, but this importance was found to be lower in the transnational dataset (ALL).

For the prediction of content damage, the ownership status (tenant vs. homeowner) is an important variable, along with water depth and the living area. We hypothesize that tenants tend to accumulate fewer (and/or less valuable) personal belongings compared to homeowners, leading therefore to a generally lower exposure in dwellings occupied by tenants. The implementation of emergency measures is another important variable for predicting content damage. Unlike building damage, where it was often difficult to implement effective measures—especially for buildings that experienced high water depths—there were more opportunities to mitigate content damage. This emphasizes the key role that emergency measures can play in reducing flood impacts. Two additional hazard variables are also found to be important: the presence of sediments and flood intensity, which both tend to induce more severe content damage.

Combining the data from all three regions leads to an enhanced performance (i.e., a lower MSE) for all the tested models compared to their performance obtained based on a training on data from a single region. This showcases the value of gathering transnational flood damage datasets. This trend was found to apply to the building damage as well as to the content damage, except for Belgium, for which all models perform surprisingly well when trained on damage to content data. When the dataset combining all regions is used, the model performance becomes not only better but also virtually independent of the choice of a particular model, suggesting that utilizing a comprehensive transnational flood damage dataset leads to more robust (model-independent) results.

Overall, this study confirms that damage to buildings and content depends on multiple hazard, vulnerability, and coping capacity factors, beyond water depth alone. This is particularly important for predicting flash flood damage due to the unique characteristics of this flood type, such as rapid onset, high flow velocities, and debris transport. As such, the outcomes of the present study are highly valuable in guiding the development of cross-country transferable models, particularly in the model feature selection step.

Acknowledgements

This research project was partly funded and conceived in the frame of the HIDA HELMHOLTZ VISITING RESEARCH GRANT. It was also partly supported by the Interreg NWE project FlashFloodBreaker. The data collection from Germany was conducted by the University of Potsdam in the frame of the project "KAHR-Climate Adaptation, Floods and Resilience", funded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung—BMBF) grant number (FKZ) 01LR2102I, approved by the ethical committee of the University of Potsdam (60/2022). Nivedita Sairam is funded by BMBF grant number FKZ 01LN2209A. The Authors gratefully acknowledge all individuals who have given their time to respond to the damage surveys and shared valuable information, as well as the students and colleagues who helped with the field data collection in Belgium.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

Abid, S., N. Sulaiman, C. Wei, and U. Nazir. 2021. "Flood Vulnerability and Resilience: Exploring the Factors That Influence Flooding in Sarawak." *IOP Conference Series: Earth and Environmental Science* 802: 012059.

Ahmadalipour, A., and H. Moradkhani. 2019. "A Data-Driven Analysis of Flash Flood Hazard, Fatalities, and Damages Over the Conus During 1996–2017." *Journal of Hydrology* 578: 124106.

Alipour, A., A. Ahmadalipour, and H. Moradkhani. 2020. "Assessing Flash Flood Hazard and Damages in the Southeast United States." *Journal of Flood Risk Management* 13: e12605.

Altelbany, S. 2021. "Evaluation of Ridge, Elastic Net and Lasso Regression Methods in Precedence of Multicollinearity Problem: A Simulation Study." *Journal of Applied Economics and Business Studies* 5, no. 1: 131–142.

Arrighi, C., B. Mazzanti, F. Pistone, and F. Castelli. 2020. "Empirical Flash Flood Vulnerability Functions for Residential Buildings." *SN Applied Sciences* 2: 1–12.

Ballio, F., D. Molinari, G. Minucci, et al. 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.

Barcelloni Corte, M., B. Bianchet, J. Privot, C. Schelings, and J. Teller. 2022. "Schéma Stratégique Multidisciplinaire Du Bassin Versant de La Vesdre. Diagnostic Approfondi Et Multithématique." Contributions de La Team Vesdre–Uliège.

Borga, M., M. Stoffel, L. Marchi, F. Marra, and M. Jakob. 2014. "Hydrogeomorphic Response to Extreme Rainfall in Headwater

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Systems: Flash Floods and Debris Flows." *Journal of Hydrology* 518: 194–205.

Bruwier, M., S. Erpicum, M. Pirotton, P. Archambeau, and B. Dewals. 2015. "Assessing the Operation Rules of a Reservoir System Based on a Detailed Modelling Chain." *Natural Hazards and Earth System Sciences* 15, no. 3: 365–379.

Bubeck, P., and H. de Moel. 2010. "Sensitivity Analysis of Flood Damage Calculations for the River Rhine."

Bundesministerium des Innern und für Heimat. 2022. *Bericht Zur Hochwasserkatastrophe 2021: Katastrophenhilfe*. Wiederaufbau und Evaluierungsprozesse.

Calianno, M., I. Ruin, and J. J. Gourley. 2013. "Supplementing Flash Flood Reports With Impact Classifications." *Journal of Hydrology* 477: 1–16.

Carisi, F., K. Schröter, A. Domeneghetti, H. Kreibich, and A. Castellarin. 2018. "Development and Assessment of Uni-and Multivariable Flood Loss Models for Emilia-Romagna (Italy)." *Natural Hazards and Earth System Sciences* 18, no. 7: 2057–2079.

Chen, S., J. Zheng, and J. Li. 2024. "The Impact of Sample Size After Sampling on the Accuracy of Machine Learning Models." In *2024 International Conference on Computers, Information Processing and Advanced Education (CIPAE)*, 61–66. IEEE.

Copernicus Emergency Management Service. 2021. "(2021 European Union), EMSR517." Accessed November 20, 2023.

CSR. 2022. "1 an Après Les Inondations ... Bilan de La Gestion Post-Inondations Et Continuité de La Reconstruction." Commissariat Spécial à la Reconstruction. Technical Report.

Cuvelier, T., P. Archambeau, B. Dewals, and Q. Louveaux. 2018. "Comparison Between Robust and Stochastic Optimisation for Long-Term Reservoir Management Under Uncertainty." *Water Resources Management* 32: 1599–1614.

de Bruijn, K., D. Wagenaar, K. Slager, M. De Bel, and A. Burzel. 2015. "Updated and Improved Method for Flood Damage Assessment: SSM2015. 2:1220043-003."

Dey, H., M. M. Haque, W. Shao, M. VanDyke, and F. Hao. 2024. "Simulating Flood Risk in Tampa Bay Using a Machine Learning Driven Approach." *Npj Natural Hazards* 1, no. 1: 1–16.

Dietze, M., R. Bell, U. Ozturk, et al. 2022. "More Than Heavy Rain Turning Into Fast-Flowing Water – A Landscape Perspective on the 2021 Eifel Floods." *Natural Hazards and Earth System Sciences* 22, no. 6: 1845–1856.

Dottori, F., R. Figueiredo, M. L. Martina, D. Molinari, and A. R. Scorzini. 2016. "Insyde: A Synthetic, Probabilistic Flood Damage Model Based on Explicit Cost Analysis." *Natural Hazards and Earth System Sciences* 16, no. 12: 2577–2591.

Dutta, D., S. Herath, and K. Musiake. 2003. "A Mathematical Model for Flood Loss Estimation." *Journal of Hydrology* 277, no. 1–2: 24–49.

Elmer, F., A. H. Thieken, I. Pech, and H. Kreibich. 2010. "Influence of Flood Frequency on Residential Building Losses." *Natural Hazards and Earth System Sciences* 10, no. 10: 2145–2159.

Endendijk, T., W. W. Botzen, H. de Moel, J. C. Aerts, K. Slager, and M. Kok. 2023. "Flood Vulnerability Models and Household Flood Damage Mitigation Measures: An Econometric Analysis of Survey Data." *Water Resources Research* 59, no. 8: e2022WR034192.

FEMA, M. 2009. *Multi-Hazard Loss Estimation Methodology Flood Model: Hazus-Mh Mr4 Technical Manual*.

Fuchs, S., C. Ornetsmüller, and R. Totschnig. 2012. "Spatial Scan Statistics in Vulnerability Assessment: An Application to Mountain Hazards." *Natural Hazards* 64: 2129–2151.

Gabriels, K., P. Willems, and J. Van Orshoven. 2022. "A Comparative Flood Damage and Risk Impact Assessment of Land Use Changes." *Natural Hazards and Earth System Sciences* 22, no. 2: 395–410.

Gesamtverband der Versicherer, G. 2021. "Versicherungsschäden Durch Flutkatastrophe Bei Rund Sieben Milliarden Euro."

Geurts, P., and L. Wehenkel. 2022. "Classification and Regression Trees."

Ghamrawi, M. F., B. M. Rehan, J. Miller, et al. 2023. "Comparisons of Flood Depth-Damage Relations for Community-Scale Aggregated Building-Level Damage Assessment." Available at SSRN 4408291.

Gravesteijn, B. Y., C. A. Sewalt, E. Venema, D. Nieboer, E. W. Steyerberg, and Collaborators C.-T. 2021. "Missing Data in Prediction Research: A Five-Step Approach for Multiple Imputation, Illustrated in the Center-Tbi Study." *Journal of Neurotrauma* 38, no. 13: 1842–1857.

Hastie, T., R. Tibshirani, J. H. Friedman, and J. H. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Volume 2*. Springer.

Hothorn, T., K. Hornik, and A. Zeileis. 2006. "Unbiased Recursive Partitioning: A Conditional Inference Framework." *Journal of Computational and Graphical Statistics* 15, no. 3: 651–674.

Huizinga, J., H. De Moel, and W. Szewczyk. 2017. "Global Flood Depth-Damage Functions: Methodology and the Database With Guidelines." Technical Report, Joint Research Centre.

Jadhav, A., D. Pramod, and K. Ramanathan. 2019. "Comparison of Performance of Data Imputation Methods for Numeric Dataset." *Applied Artificial Intelligence* 33, no. 10: 913–933.

Jansen, L., P. Korswagen, J. Bricker, S. Pasterkamp, K. De Bruijn, and S. Jonkman. 2020. "Experimental Determination of Pressure Coefficients for Flood Loading of Walls of Dutch Terraced Houses." *Engineering Structures* 216: 110647.

Jeng, C.-C. 2023. "Why a Variance Inflation Factor of 10 Is Not an Ideal Cutoff for Multicollinearity Diagnostics." *Journal of Education Studies* 57, no. 2: 67–93.

Journée, M., E. Goudenhoofdt, S. Vannitsem, and L. Delobbe. 2023. "Quantitative Rainfall Analysis of the 2021 Mid-July Flood Event in Belgium." *Hydrology and Earth System Sciences* 27, no. 17: 3169–3189.

Junghänel, T., P. Bissolli, J. Daßler, et al. 2021. "Hydro-Klimatologische Einordnung Der Stark- und Dauerniederschläge in Teilen Deutschlands im Zusammenhang mit dem Tiefdruckgebiet "Bernd" Vom 12. Bis 19. Juli 202." Technical Report, DWD [German Weather Bureau].

Kreibich, H., A. Botto, B. Merz, and K. Schröter. 2017. "Probabilistic, Multivariable Flood Loss Modeling on the Mesoscale With Bt-Flemo." *Risk Analysis* 37, no. 4: 774–787.

Kreibich, H., G. Di Baldassarre, S. Vorogushyn, et al. 2017. "Adaptation to Flood Risk: Results of International Paired Flood Event Studies." *Earths Future* 5, no. 10: 953–965.

Kreibich, H., K. Piroth, I. Seifert, et al. 2009. "Is Flow Velocity a Significant Parameter in Flood Damage Modelling?" *Natural Hazards and Earth System Sciences* 9, no. 5: 1679–1692.

Kreibich, H., A. Thieken, S.-N. Haubrock, and K. Schröter. 2017. "Howas21, the German Flood Damage Database." In *Flood Damage Survey and Assessment: New Insights From Research and Practice*, 65–75.

Lamberti, W. F. 2023. "An Overview of Explainable and Interpretable Ai." In *AI Assurance*, 55–123. Elsevier.

Laudan, J., V. Rözer, T. Sieg, K. Vogel, and A. H. Thieken. 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, no. 12: 2163–2179.

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Leal, M., E. Reis, S. Pereira, and P. P. Santos. 2021. "Physical Vulnerability Assessment to Flash Floods Using an Indicator-Based Methodology Based on Building Properties and Flow Parameters." *Journal of Flood Risk Management* 14, no. 3: e12712.

Lengfeld, K., P. Voit, F. Kaspar, and M. Heistermann. 2023. "Brief Communication: On the Extremeness of the July 2021 Precipitation Event in Western Germany." *Natural Hazards and Earth System Sciences* 23, no. 3: 1227–1232.

Levshina, N. 2021. "Conditional Inference Trees and Random Forests." In *A Practical Handbook of Corpus Linguistics*, 611–643. Springer.

Llasat, M. C., R. Marcos, M. Turco, J. Gilabert, and M. Llasat-Botija. 2016. "Trends in Flash Flood Events Versus Convective Precipitation in the Mediterranean Region: The Case of Catalonia." *Journal of Hydrology* 541: 24–37.

Louppe, G. 2014. "Understanding Random Forests: From Theory to Practice." Preprint, arXiv:1407.7502.

Ma, M., G. Zhao, B. He, et al. 2021. "Xgboost-Based Method for Flash Flood Risk Assessment." *Journal of Hydrology* 598: 126382.

Maiwald, H., J. Schwarz, C. Kaufmann, T. Langhammer, S. Golz, and T. Wehner. 2022. "Innovative Vulnerability and Risk Assessment of Urban Areas Against Flood Events: Prognosis of Structural Damage With a New Approach Considering Flow Velocity." *Water* 14, no. 18: 2793.

Merz, B., H. Kreibich, and U. Lall. 2013. "Multi-Variate Flood Damage Assessment: A Tree-Based Data-Mining Approach." *Natural Hazards and Earth System Sciences* 13, no. 1: 53–64.

Mohor, G., P. Hudson, and A. Thieken. 2020. "A Comparison of Factors Driving Flood Losses in Households Affected by Different Flood Types." *Water Resources Research* 56, no. 4: e2019WR025943.

Mohr, S., U. Ehret, M. Kunz, et al. 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, no. 2: 525–551.

Molinari, D., A. R. Scorzini, C. Arrighi, et al. 2020. "Are Flood Damage Models Converging to "Reality"? Lessons Learnt From a Blind Test." *Natural Hazards and Earth System Sciences* 20, no. 11: 2997–3017.

Mosavi, A., P. Ozturk, and K.-w. Chau. 2018. "Flood Prediction Using Machine Learning Models: Literature Review." *Water* 10, no. 11: 1536.

Munich RE. 2022. "Hurrikane, Kältewellen, Tornados: Wetterkatastrophen in USA Dominieren Naturkatastrophen-Schadenstatistik 2021." https://www.munichre.com/de/unternehmen/media-relations/medieninformationen-und-unternehmensnachrichten/medieninformationen/2022/bilanz-naturkatastrophen-2021.html.

Nafari, R. H., T. Ngo, and W. Lehman. 2016. "Development and Evaluation of Flfacs–A New Flood Loss Function for Australian Commercial Structures." *International Journal of Disaster Risk Reduction* 17: 13–23.

Papathoma-Köhle, M. 2016. "Vulnerability Curves vs. Vulnerability Indicators: Application of an Indicator-Based Methodology for Debris-Flow Hazards." *Natural Hazards and Earth System Sciences* 16, no. 8: 1771–1790.

Papathoma-Köhle, M., M. Schlögl, L. Dosser, et al. 2022. "Physical Vulnerability to Dynamic Flooding: Vulnerability Curves and Vulnerability Indices." *Journal of Hydrology* 607: 127501.

Papathoma-Köhle, M., M. Schlögl, and S. Fuchs. 2019. "Vulnerability Indicators for Natural Hazards: An Innovative Selection and Weighting Approach." *Scientific Reports* 9, no. 1: 15026.

Paprotny, D., A. Sebastian, O. Morales-Nápoles, and S. N. Jonkman. 2018. "Trends in Flood Losses in Europe Over the Past 150 Years." *Nature Communications* 9, no. 1: 1985.

Paulik, R., A. Wild, C. Zorn, and L. Wotherspoon. 2022. "Residential Building Flood Damage: Insights on Processes and Implications for Risk Assessments." *Journal of Flood Risk Management* 15, no. 4: e12832.

Petrucci, O., L. Aceto, C. Bianchi, et al. 2019. "Flood Fatalities in Europe, 1980–2018: Variability, Features, and Lessons to Learn." *Water* 11, no. 8: 1682.

Poussin, J. K., W. W. Botzen, and J. C. Aerts. 2015. "Effectiveness of Flood Damage Mitigation Measures: Empirical Evidence From French Flood Disasters." *Global Environmental Change* 31: 74–84.

Rafiezadeh, S. K., N. Sairam, L. Schoppa, L. T. Sang, D. L. H. Tan, and H. Kreibich. 2024. *BN-FLEMOΔ: A Bayesian Network-Based Flood Loss Estimation Model for Adaptation Planning in Ho Chi Minh City*. Vietnam.

Rajput, D., W.-J. Wang, and C.-C. Chen. 2023. "Evaluation of a Decided Sample Size in Machine Learning Applications." *BMC Bioinformatics* 24, no. 1: 48.

Ramosaj, B., L. Amro, and M. Pauly. 2020. "A Cautionary Tale on Using Imputation Methods for Inference in Matched-Pairs Design." *Bioinformatics* 36, no. 10: 3099–3106.

Rodríguez Castro, D., M. Cools, S. Roucour, et al. 2025. "Can Macro-Or Meso-Scale Coping Capacity Variables Improve the Classification of Building Flood Losses?" *Natural Hazards* 121: 1–31.

Rözer, V., H. Kreibich, K. Schröter, et al. 2019. "Probabilistic Models Significantly Reduce Uncertainty in Hurricane Harvey Pluvial Flood Loss Estimates." *Earth's Future* 7, no. 4: 384–394.

Sairam, N., K. Schröter, S. Lüdtke, B. Merz, and H. Kreibich. 2019. "Quantifying Flood Vulnerability Reduction via Private Precaution." *Earth's Future* 7, no. 3: 235–249.

Scawthorn, C., P. Flores, N. Blais, et al. 2006. "Hazus-Mh Flood Loss Estimation Methodology. Ii. Damage and Loss Assessment." *Natural Hazards Review* 7, no. 2: 72–81.

Schäfer, A., B. Mühr, J. Daniell, et al. 2021. "Hochwasser Mitteleuropa, Juli 2021 (Deutschland): 21. Juli 2021 – Bericht Nr. 1 "Nordrhein-Westfalen & Rheinland-Pfalz"." Technical Report, CEDIM Forensic Disaster Analysis (FDA) Group-Karlsruher Institut Für Technologie (KIT).

Schröter, K., H. Kreibich, K. Vogel, C. Riggelsen, F. Scherbaum, and B. Merz. 2014. "How Useful Are Complex Flood Damage Models?" *Water Resources Research* 50, no. 4: 3378–3395.

Scorzini, A. R., B. Dewals, D. Rodriguez Castro, P. Archambeau, and D. Molinari. 2022. "Insyde-Be: Adaptation of the Insyde Model to the Walloon Region (Belgium)." *Natural Hazards and Earth System Sciences* 22, no. 5: 1743–1761.

Shadbahr, T., M. Roberts, J. Stanczuk, et al. 2023. "The Impact of Imputation Quality on Machine Learning Classifiers for Datasets With Missing Values." *Communications Medicine* 3, no. 1: 139.

Sieg, T., K. Vogel, B. Merz, and H. Kreibich. 2017. "Tree-Based Flood Damage Modeling of Companies: Damage Processes and Model Performance." *Water Resources Research* 53, no. 7: 6050–6068.

Suppasri, A., E. Mas, I. Charvet, et al. 2013. "Building Damage Characteristics Based on Surveyed Data and Fragility Curves of the 2011 Great East Japan Tsunami." *Natural Hazards* 66: 319–341.

Szönyi, M., V. Roezer, T. Deubelli, et al. 2022. *PERC Floods Following "Bernd"*. Publisher.

Terti, G., I. Ruin, S. Anquetin, and J. J. Gourley. 2017. "A Situation-Based Analysis of Flash Flood Fatalities in the United States." *Bulletin of the American Meteorological Society* 98, no. 2: 333–345.

Thieken, A. H., P. Bubeck, A. Heidenreich, J. Von Keyserlingk, L. Dillenardt, and A. Otto. 2023. "Performance of the Flood Warning System in Germany in July 2021 – Insights From Affected Residents." *Natural Hazards and Earth System Sciences* 23, no. 2: 973–990.

Thieken, A. H., H. Kreibich, M. Müller, and J. Lamond. 2017. "Data Collection for a Better Understanding of What Causes Flood Damage-Experiences With Telephone Surveys: Chapter 7." In *Flood Damage*

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Survey and Assessment, Geophysical Monograph, edited by D. Molinari, F. Ballio, and S. Menoni, 95–106. Wiley and American Geophysical Union.

Thieken, A. H., M. Müller, H. Kreibich, and B. Merz. 2005. "Flood Damage and Influencing Factors: New Insights From the August 2002 Flood in Germany." *Water Resources Research* 41, no. 12.

Thieken, A. H., A. Olschewski, H. Kreibich, S. Kobsch, and B. Merz. 2008. "Development and Evaluation of Flemops–a New Flood Loss Estimation Model for the Private Sector." *WIT Transactions on Ecology and the Environment* 118: 315–324.

Thieken, A. H., T. Petrow, H. Kreibich, and B. Merz. 2006. "Insurability and Mitigation of Flood Losses in Private Households in Germany." *Risk Analysis* 26, no. 2: 383–395.

Thieken, A. H., G. Samprogna Mohor, H. Kreibich, and M. Müller. 2022. "Compound Inland Flood Events: Different Pathways, Different Impacts and Different Coping Options." *Natural Hazards and Earth System Sciences* 22, no. 1: 165–185.

Vanneuville, W., R. Maddens, C. Collard, P. Bogaert, P. De Maeyer, and M. Antrop. 2006. Impact Op Mens En Economie Tgv Overstromingen Bekeken in Het Licht Van Wijzigende Hydraulische Condities, Omgevingsfactoren En Klimatologische Omstandigheden.

Vogel, K., C. Riggelsen, B. Merz, H. Kreibich, and F. Scherbaum. 2012. "Flood Damage and Influencing Factors: A Bayesian Network Perspective." In *6th European Workshop on Probabilistic Graphical Models (PGM 2012), Volume 625*. University of Granada.

Wagenaar, D., S. Lüdtke, K. Schröter, L. M. Bouwer, and H. Kreibich. 2018. "Regional and Temporal Transferability of Multivariable Flood Damage Models." *Water Resources Research* 54, no. 5: 3688–3703.

Wüthrich, D., P. A. Korswagen, H. Selvam, J. Oetjen, J. Bricker, and H. Schüttrumpf. 2025. "Field Survey Assessment of Flood Loads and Related Building Damage From the July 2021 Event in the Ahr Valley (Germany)." *Journal of Flood Risk Management* 18, no. 1: e13024.

Yin, J., D. Yu, Z. Yin, M. Liu, and Q. He. 2016. "Evaluating the Impact and Risk of Pluvial Flash Flood on Intra-Urban Road Network: A Case Study in the City Center of Shanghai, China." *Journal of Hydrology* 537: 138–145.

Zhen, Y., S. Liu, G. Zhong, et al. 2022. "Risk Assessment of Flash Flood to Buildings Using an Indicator-Based Methodology: A Case Study of Mountainous Rural Settlements in Southwest China." *Frontiers in Environmental Science* 10: 931029.

Zou, H., and T. Hastie. 2005. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 67, no. 2: 301–320.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1.** Supporting Information.

1753318x, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.70088 by Helmholtz-Zentrum Potsdam GFZ, Wiley Online Library on [07/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License