https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

FLEMO$_{flash}$ - Flood Loss Estimation MOdels for companies and households affected by *flash* floods

Apoorva Singh[1,2#], Ravi Kumar Guntu[1#*], Nivedita Sairam[1], Kasra Rafiezadeh Shahi[1], Anna Buch[3], Melanie Fischer[1], Chandrika Thulaseedharan Dhanya[2], and Heidi Kreibich[1]

[1]Section 4.4 Hydrology, GFZ Helmholtz Centre for Geosciences, 14473 Potsdam, Germany

[2]Department of Civil Engineering, Indian Institute of Technology Delhi, 110016 New Delhi, India

[3]Institute of Geography, University of Heidelberg, 69117 Heidelberg, Germany

**Correspondence to*: Ravi Kumar Guntu (ravikumar.guntu@gfz.de)

[†]These authors contributed equally to this work

Key points:

- Applying Machine Learning on multi-event data reveals key drivers of flash flood losses such as flow velocity and emergency response.
- The first probabilistic flash flood loss model provides robust estimates of company and household losses, including uncertainty information.
- High preparedness during extreme flash floods was found to reduce building losses of households by up to 77%.

Abstract. In light of the increasing losses from flash floods intensified by climate change, there is a critical need for improved loss models. Loss assessments predominantly focus on fluvial flood processes, leaving a significant gap in understanding the key drivers of flash floods and the effect of preparedness on losses. To address these gaps, we introduce FLEMO$_{flash}$—a novel multivariate probabilistic Flood Loss Estimation Model compilation for flash floods. The models are developed for companies and households based on survey data collected after flash flood events in 2002, 2016, and 2021 in Germany. FLEMO$_{flash}$ employs a data-driven feature selection approach, combining machine learning techniques (Elastic Net, Random Forest, XGBoost) to identify key drivers influencing flash flood losses and Bayesian networks to model probabilistic loss estimates, including uncertainty. Model-based findings show that in extreme hazard scenarios, high preparedness can reduce building losses by up to 47% for large companies. Households who knew exactly what to do during high water depth were able to reduce their building losses by 77% and contents losses by 55%. Thus, FLEMO$_{flash}$ can support risk communication and management by providing reliable estimation of flash flood losses along with the loss differential considering the level of risk preparedness.

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

1 Introduction

Flash floods characterized by their rapid onset and short duration, rank amongst the most devastating natural disasters, leading to significant loss of life and property (Zain et al., 2021). The flash flood events in western Germany and neighboring countries during July 2021 caused an estimated USD 54 billion in losses (Munich Re), marking it as the costliest natural disaster in the history of Germany to date. Other notable examples include events in the Elbe and Danube rivers in Germany in 2002 resulted in USD 9 billion losses (Kreibich et al., 2007). The 2017 flash floods in the Houston area of Texas during Hurricane Harvey resulted in losses ranging from USD 90 to 160 billion (Rözer et al., 2021). The 2012 event in Beijing caused total losses of USD 1.86 billion (Wang et al., 2013), and the monster flood of Pakistan in 2022 caused losses worth USD 30 billion (Chughtai, 2022). With the ongoing climate change crisis and high population density, the risk of flash floods is anticipated to increase in the future; thus, emphasizing the need for flash flood loss modelling to derive quantitative estimates of expected losses in monetary terms.

While progress has been made related to fluvial flood loss models for households ((Lüdtke et al., 2019; Paprotny et al., 2020; Schoppa et al., 2022; Seifert et al., 2010; Thieken et al., 2008), there remains a limited understanding of the variables and mechanisms influencing flash flood losses. Unlike fluvial floods that have longer lead time and slower rise in water levels, flash floods are characterized by rapid and unprecedented rise in water levels (Laudan et al., 2020). Moreover, due to high flow velocities, sediment transport, and shorter lead times, flash floods often cause more losses than fluvial floods. The sudden nature of flash floods makes them extremely difficult to predict (Dougherty and Rasmussen, 2020), necessitating loss modelling tailored to these events. Unlike floods caused by slowly rising water levels, and dyke breaches, flash floods exhibit heterogeneity in hazard characteristics such as water depth, flow velocity, inundation duration, and contamination indicators (Kreibich and Dimitrova, 2010). Furthermore, earlier studies suggest significant differences in the variables and processes that influence losses in different flood types (Mohor et al., 2020). Thus, understanding the flash flood process is crucial, despite our previous understanding of the losses caused by fluvial floods. A comprehensive understanding of the complex processes behind flash floods is essential to develop sustainable and cost-efficient flash flood risk management strategies.

Traditionally, flood loss estimation has relied on univariate stage-damage functions (SDF) (Middelmann-Fernandes, 2010). Nevertheless, SDF models have faced criticism for ignoring the effects of precaution, emergency, and socioeconomic aspects (Merz et al., 2013; Schröter et al., 2014). To address these limitations, multivariate loss estimation models are developed. The conventional multivariate flood loss estimation models often employ decision tree-based approaches to assess the role of different variables in influencing flood loss. For instance, Merz et al., (2013) utilized regression trees and bagging decision trees to identify the most significant variables based on flood events in Germany during 2002, 2005, and 2006. A similar approach was employed in the Mekong basin (Chinh et al., 2016) and in the Burnett River catchment, Australia (Hasanzadeh Nafari et al., 2016). Additionally, various machine learning models have been developed for flood loss estimation. These include neural networks (Salas et al., 2023), random forest (Ghaedi et al., 2022), Bayesian regression (Mohor et al., 2021),

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

and Bayesian networks (Vogel et al., 2018). Among these, Bayesian networks (BN) are an advanced and preferred approach due to their probabilistic representation of conditional dependencies among multiple variables, their ability to handle missing data, and their transferability (Sairam et al., 2020). For instance, Wagenaar et al., (2018) developed a regional and temporal transferable BN-based flood loss model for microscale residential applications, which was later upscaled to mesoscale by Lüdtke et al., (2019). Despite these advancements, there has been limited work on flood loss models for companies, and almost non-existent for flash floods. We propose a BN-based approach for developing flash flood loss models to address the challenges in assessing the conditional dependencies among multiple variables and handling multievent heterogenous datasets.

The objective of this study is to build a novel Flood Loss Estimation MOdel affected by flash floods ($FLEMO_{flash}$). The model is developed for the company and household sectors in Germany using empirical loss data collected after flash flood events in 2002, 2016, and 2021. The study identifies the important variables and explains the underlying processes that govern the flash flood losses. Next, we examine the predictive performance of $FLEMO_{flash}$ model and compare it with conventional SDF models. Additionally, we used the model to derive the predictive density of losses and illustrate the effect of preparedness in controlling both the extent of loss reduction and the associated uncertainties.

2 Data and Methods

2.1 Multievent empirical data in different regions

$FLEMO_{flash}$ is built based on self-reported flash flood losses along with associated information of the affected companies and households. The data was collected through different surveys using computer-aided telephone interviews with representatives of affected companies and households following three highly damaging flash flood occurrences in Germany (Kellermann et al., 2020; Kreibich et al., 2017). These include the 2002 event in the Elbe catchment in Eastern Germany, the 2016 heavy precipitation event, and the most recent July 2021 event in Western Germany. The variables potentially influencing losses are extracted and homogenized from the datasets of the 2002, 2016, and 2021 flood surveys. The variables are grouped into five categories, as presented in Table 1 for companies and Table 2 for households, respectively. The loss variable represented as relative loss (rloss), is defined as the ratio between the reported loss and replacement cost on the [0,1] interval. A value of 0 indicates no loss, while 1 indicates complete loss (Kreibich and Dimitrova, 2010; Schoppa et al., 2020; Sieg et al., 2017). For companies, losses are estimated for three categories of assets - buildings, equipment, and goods & stock. For households, losses are estimated for buildings and contents. Survey responses with no data on loss were excluded from the dataset.

2.2 Terrain analysis for the identification of flash flood cases

Since the surveys focussed on regions affected by both flash floods and riverine floods, we conducted terrain analysis using GIS-based topographical selection where, observations related to riverine floods were excluded. We calculated the median

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

slope within a 10 km radius around the location of each observation using the Digital Elevation Model with a 90 meters resolution from the Shuttle Radar Topographic Mission. This process was repeated for 14 reference municipalities known to have experienced flash flood events in the past or described as prone to flash floods (Thieken et al., 2022). The minimum slope from these 14 municipalities was then used as the threshold value. Observations where the slope was equal to or higher than the threshold value were considered to be affected by flash flood dynamics. The final database available for developing loss models consists of 241, 379, 355, 1131, 1448 observations for company building, equipment, goods & stock, household building, and contents, respectively. The percentage of missing data for different variables in each of the asset is summarized in Fig S1 (companies) and Fig S2 (households). To maximise the amount of training data for model building, we employed the nearest neighbour technique to impute the missing data. We tested a range of k-neighbours for our datasets (k =1,3,5,7,9) and selected the value with best performance.

Table 1. List of variables for companies. The variable type stands C for continuous, O for ordinal, and N for nominal.

	Variable		Type and range	
Hazard	wd	Water depth	C:	0 cm to 963 cm above ground
	d	Inundation duration	C:	0 to 1440 h
	v	Velocity indicator	O:	1 = low flow velocity to 3 = torrential flow velocity
	con	Contamination	O:	0 = no contamination to 4 = heavy contamination
Emergency response	wt	Warning lead time	C:	0 to 240 h
	ws	Early warning source	O:	0 = no warning to 4 = official warning through authorities
	ew	Early warning received	N:	0 = no, 1 = yes
	me	Emergency measures undertaken	N:	0 = no, 1 = yes
	ep	Emergency plan	N:	0 = no, 1 = yes
	kh	Knowledge about hazard	N:	0 = no, 1 = yes
	ms	Emergency measures success	O:	0 = no measure undertaken, 1 = completely ineffective to 4 = very effective
Precaution	fe	Flood experience	O:	0 = no experience to 5 = recent flood experience
	pr	Precaution indicator	O:	0 = no/low precautions, 1 = medium precautions, 2 = very good precautions.
	in	Insurance	N:	0 = no, 1 = yes
Company characteristics	sec	Sector	O:	1 = Agriculture, 2 = Manufacturing, 3 = Trade, 4 = Finance, 5 = Services
	ss	Spatial situation	O:	1 = several buildings, 2 = entire building, 3 = one or more floors, 4 = less than one floor
	own	ownership	O:	1 = building owned, 2 = rented, 3= partly owned/ partly rented
	emp	Number of employees	C:	1 to 920
	sp	Size premise	C:	10 to 69000 m^2

2021 data source: Survey "Flooding in Germany in July 2021: Damage of companies", German Research Centre for Geosciences, Deutsche Rückversicherung AG, 2022

2016 data source: Survey "Pluvial Flooding and Flash Floods in May/June 2016: Damage of companies", University of Potsdam, German Research Centre for Geosciences, Deutsche Rückversicherung AG, 2017.

2002 data source: Survey "Flooding in Germany in August 2002: Damage of companies“, German Research Centre for Geosciences, Deutsche Rückversicherung AG, 2003

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

Table 2. List of variables for households. The variable type stands C for continuous, O for ordinal, and N for nominal.

Predictors			Type and range	
Hazard	wd	Water depth	C:	245 cm below ground to 700 cm above ground
	d	Inundation duration	C:	1 to 1440 h
	v	Velocity scaled	O:	0 = no flow velocity to 6 = torrential flow velocity
	hs	Human stability	O:	1 = person can stand effortlessly in calm water to 3 person would have been swept away; 4 = too deep to stand
	con	Contamination	O:	0 = no contamination to 4 = heavy contamination
emergency response	ew	Early warning received	N:	0 = no, 1 = yes
	wt	Warning lead time	C:	0 to 168 h
	ws	Warning source	O:	0 = no warning to 4 = official warning through authorities
	ke	Knowledge about emergency action	O:	0 = no measure undertaken, 1= receiver of warning had no idea what to do to 6= receiver of warning knew exactly what to do
	me	Emergency measures undertaken	N:	0 = no, 1 = yes
	mu	Number of emergency measures undertaken	O:	0 = no measures undertaken to 13 = all measures undertaken
precaution	fe	Flood experience	O:	0 = no experience to 5 = recent flood experience
	pw	Precaution indicator	O:	0 = no/low precautions, 1 = medium precautions, 2 = very good precautions.
building characteristics	fa	Building footprint area	C:	5 to 1000 m^2
	b	Basement	N:	0 = No basement, 1 = Partial basement, 2 = Full basement
socio-economic status	per	Household size, i.e. number of persons	C:	1 to 12 people
	chi	Number of children (< 14 yr)	C:	0 to 9
	eld	Number of elders (> 65 yr)	C:	0 to 4
	inc	Monthly net income in classes	O:	1 = below 500 EUR to 6 = 3000 EUR and more
	$socp$	Socioeconomic status according to Plapp, (2003)	O:	3 = very low socioeconomic status to 13 = very high socioeconomic status

2021 data source: Survey "Flooding in Germany in July 2021: Damage of private households“, University of Potsdam, data collection within the KAHR-project, funded by the German Ministry of Education and Research (BMBF, contract 01LR2102I), approved by the ethical committee of the University of Potsdam (60/2022)
2016 data source: Survey "Pluvial Flooding and Flash Floods in May/June 2016: Damage of private households", University of Potsdam, German Research Centre for Geosciences, Deutsche Rückversicherung AG, 2017.
2002 data source: Survey "Flooding in Germany in August 2002: Damage of private households“, German Research Centre for Geosciences, Deutsche Rückversicherung AG, 2003.

2.3 Machine Learning-based feature selection

Flood damage processes vary by region, flood type, and asset type (Mohor et al., 2020; Sairam et al., 2019; Wagenaar et al., 2018). To derive the drivers of flash flood losses, this study adopts a data-driven feature selection approach to the empirical data. Feature selection involves identifying variables that have the highest influence on the target variable (i.e. relative loss). We train multiple models – nonlinear models: Random Forest (RF), Extreme Gradient Boosting (XGBoost), and linear model: Elastic Net (EN).

RF is an ensemble machine learning method primarily used for classification and regression tasks, developed by Breiman, (2001). RF generates an ensemble of decision trees, each trained on a random subset of the data using bootstrap sampling. At

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

each node within these trees, a random subset of features is considered for splitting. The final prediction for a given input is obtained by averaging the predictions from all individual trees. This approach helps RF reduce overfitting and enhances the model's generalization ability. XGBoost, similarly to RF, is an ensemble learning algorithm that benefits from a decision tree-based structure. However, the key difference compared to RF is that in XGBoost, each tree corrects the errors from the previous ones. The process starts with a simple model and iteratively adds trees that focus on the residuals or errors made by the existing ensemble. With its efficient implementation, XGBoost demonstrates superior performance and handles large-scale data more effectively than RF (Chen and Guestrin, 2016). While RF and XGBoost are non-linear models, EN is a regularization technique used in linear regression, combining both Lasso (L1) and Ridge (L2) regularization penalties. It effectively addresses multicollinearity in datasets by shrinking the less influential predictors toward zero (Lasso) while additionally providing some degree of regularization to prevent overfitting (Ridge). EN's ability to handle correlated features and select relevant predictors makes it a valuable tool in regression tasks (Zou and Hastie, 2005).

During training, we employed a nested cross-validation framework with 10 splits and 10 repeats, resulting in a total of 100 evaluations. We selected the best set of hyperparameters, which obtained the least mean absolute error, which was then applied to the final feature selection. From each resulting final model, we derived the feature importance. Next, we calculated each variable's weighted feature importance and overall rank. The final selection of the variables (Fig 1) is elaborated upon in the results section.

2.4 Probabilistic FLEMO$_{flash}$ development

Based on the identified features, a multivariate probabilistic Flood Loss Estimation MOdel (Bayesian Network – BN) is calibrated for predicting *flash* flood losses (Jensen and Nielsen, 2007; Kitson et al., 2023; Scutari and Denis, 2021). BNs are probabilistic graphical models where the predictor and target variables are connected through a directed acyclic graph (DAG). Each variable is depicted as a node, and related nodes are connected through arcs. These connections allow for the estimation of conditional probability distributions, facilitating an understanding of the underlying processes and probabilistic estimation of rloss (Vogel et al., 2018). The continuous variables are discretized using an equal frequency discretization approach (Wagenaar et al., 2017), and a discrete BN is formulated. The number of bins is determined based on the model performance against different numbers of bins – the number of bins resulting in the best model performance is chosen. The specification of a discrete BN involves defining a set of variables $(X_1, \dots X_n)$, constructing a DAG representing the probabilistic dependencies among variables, and obtaining the conditional probability distribution $P\left(\frac{X_i}{parents(X_i)}\right)$ for each variable (X_i) in the DAG, where $parents(X_i)$ denotes the parents of node X_i in the DAG. The final joint probability distribution for the set of variables connected in a discrete BN is formulated as (Pearl, 1988):

$$P(X_1, \dots, X_n) = \prod_{i=1}^{n} P\left(\frac{X_i}{parents(X_i)}\right) \tag{1}$$

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

Within the predicted bins of the discrete BN (rloss bins), we fit a distribution based on weighted sampling of the empirical loss data. This results in a continuous distribution of the losses (Schoppa et al., 2020). For further details on the BN structure learning we refer to S1.

2.4.1 Comparison to stage damage function

We compared FLEMO$_{flash}$ to a univariate stage-damage function (SDF), a conventional model in flood loss estimation (Gerl et al., 2016). We implemented the linear functional form of deterministic SDF (SDF-D) and probabilistic SDF (SDF-P) to assess the added value of the multivariate and probabilistic model. SDF-D is formulated as:

$$rloss_i = \alpha + \beta(wd)_i + \varepsilon_i \tag{2}$$

where $rloss_i$ is the relative loss for a given water depth (wd_i). α, β and ε_i are the intercept, regression coefficient, and error of observation i, respectively. Further, to implement the SDF-P, we assume that the relative loss follows a zero-and-one-inflated Beta distribution (Schoppa et al., 2020):

$$Y_i \; BEINF(\lambda, \gamma, \mu_i, \phi) \tag{3}$$

$$logit(\mu_i) = \alpha + \beta(wd)_i \tag{4}$$

In the above equation 3, we only predict the μ, whereas other distribution parameters are assumed constant across the observations. SI (Table. S1) contains further information on the prior choice for model parameters as well as specification for Markov chain Monte Carlo sampling.

2.4.2 Model validation and sensitivity

We evaluate the performance of the FLEMO$_{flash}$, SDF-D, and SDF-P models individually for the different types of assets in both, companies and private households. We employed 10-fold cross-validation, which was repeated (n=100) with independent random seeds to obtain robust estimates. To address parameter sensitivity of FLEMO$_{flash}$ model performances due to Bayesian network structure learning, we systematically evaluated three critical factors:

1. Number of predictors (f1-f5): Section 2.3 identifies the ensemble-based important predictors, and the top five were used to develop the BN model. We demonstrated the model performance with varying number of predictors.
2. Number of discretization bins (b3-b8): Continuous variables were binned using quantile-based stratification.
3. Number of neighbours (k1-k9): Missing data were imputed with k-nearest neighbours.

For each combination, we validated the model for each cross-validation fold using three performance metrics: mean absolute error (MAE), continuous ranked probability score (CRPS), and mean bias error (MBE) (Gneiting and Katzfuss, 2014; Jensen and Nielsen, 2007; Krüger et al., 2021; Schoppa et al., 2020). Detailed information on the validation procedure and three scores used to compare the models are provided in SI.

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

3 Results and Discussion

3.1 Drivers of flash flood losses

An ensemble of linear and non-linear machine learning models ensures that both linear and non-linear relationships between the predictors and flood loss are captured. Water depth emerges as the most important predictor of damage across all asset types (Fig 1), which is consistent with previous loss models (Kreibich et al., 2010; Merz et al., 2013; Schoppa et al., 2020; Sieg et al., 2017; Thieken et al., 2008). For companies, emergency measures success and number of employees are also significant factors influencing the flood loss estimation. Among other flood characteristics, duration is ranked fourth for building (Fig. 1a), velocity is ranked third for equipment (Fig. 1b), and contamination is the fifth most significant driver for goods & stock (Fig. 1c), respectively. In case of households, human stability and contamination are most important hazard variables after water depth.

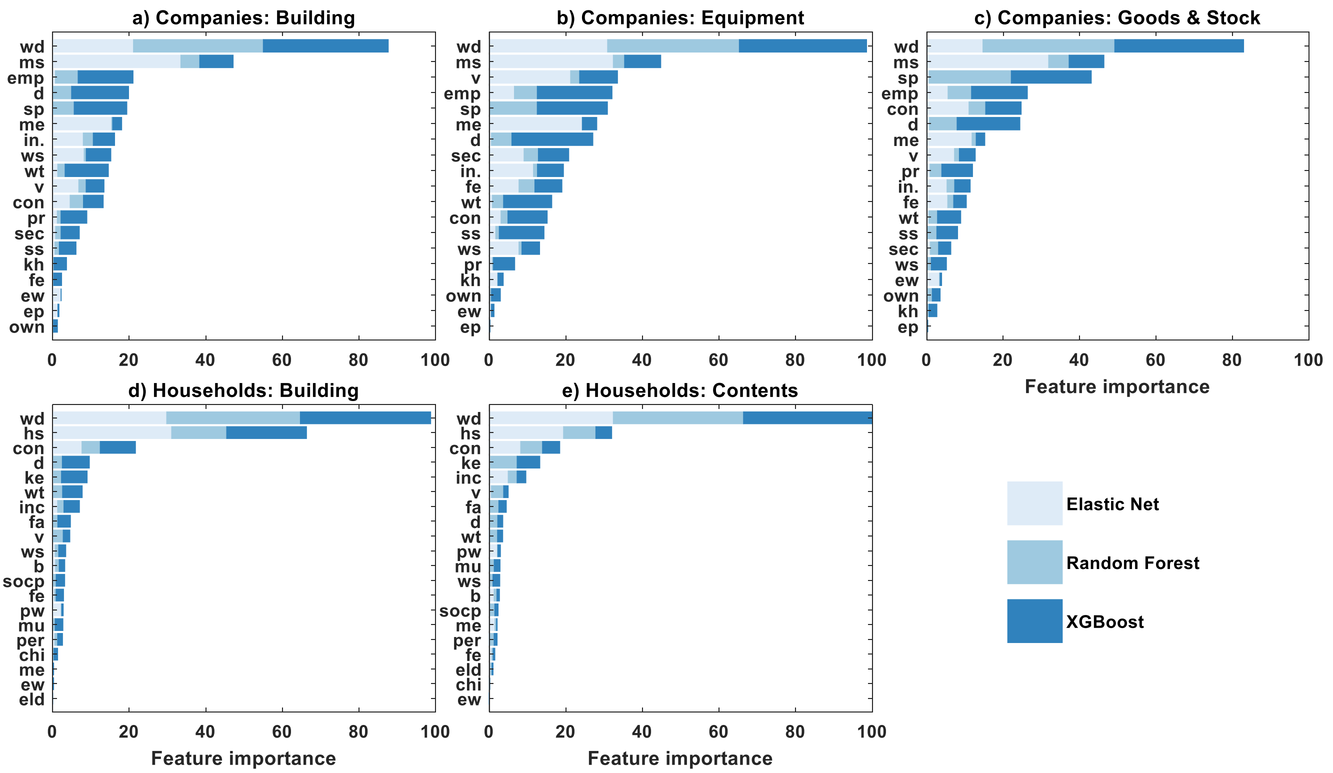

Figure 1. Illustration of the feature importance in predicting the relative loss of companies categorized into (a) Building, (b) Equipment, (c) Goods and Stock. Similarly, private household is categorized into (d) Building and (e) Contents. X-axis denotes the weighted importance derived from an ensemble approach, combining two non-linear models (Random Forest and Extreme Gradient Boosting) and one linear model (Elastic Net).

The significance of water depth and emergency measures has also been emphasized by Hasanzadeh Nafari et al., (2016) in case of fluvial flood losses. Exposure variables such as number of employees significantly influence company losses.

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

Additionally in case of households, losses are more influenced by variables representing flood vulnerability such as knowledge about emergency action. (Fig. 1d-e). These findings are in line with the previous studies (Kreibich et al., 2005; Sairam et al., 2019; Zander et al., 2023) that highlight the potential for adaptation measures to reduce flood losses. By identifying the varied drivers of flash flood losses, these results also emphasize the importance of multivariable loss estimation models that capture the interplay across these drivers and their influence on losses.

3.2 Probabilistic multivariate flash flood loss model

3.2.1 FLEMO*flash* Bayesian network structure

The FLEMO*flash* models are developed using a score-based structure learning algorithm and BN models are developed to capture the multivariate dependencies among variables. Data-driven BNs with the best performance were evaluated by domain experts to ensure consistency with the existing understanding about the underlying dynamics of flood loss processes (Fig. 2). The direction of the arrow represents an association between two variables but doesn't necessarily represent causality (Lüdtke et al., 2019; Sairam et al., 2020). Water depth emerged as the most important predictor for loss estimation across all asset types and is directly connected to the rloss node in all BN structures (Fig. 2). In case of companies and households, the losses are influenced predominantly by water depth. Company characteristics (number of employees and size premise) significantly impact the losses for company assets (Fig. 2a-c). The measure success (ms) predictor is directly connected to rloss in case of building and equipment (Fig 2a-b) and indirectly connected to rloss through water depth for goods & stock (Fig. 2c).

In private households, building loss is directly connected to water depth, contamination, and knowledge about emergency action (Fig 2d). For contents, loss is directly connected to water depth and knowledge about emergency action, with contamination, human stability, and income also playing important roles (Fig. 2e). Additionally, human stability, which is a factor of both depth and velocity, influences household losses through water depth. The direct connection of rloss with measures success for companies' assets (Fig. 2a-c), and with knowledge about emergency action for households (Fig. 2d-e) highlights the significance of risk preparedness in mitigating flood losses. The measure success (ms) indicates the respondent's perception of the efficiency of emergency measures undertaken. The direct connection of measures success to company (buildings and equipment), and knowledge about emergency to private household (buildings and contents) indicates that given similar hazard and exposure situation, better preparedness could lead to reduced losses. We derived joint probability distribution for all asset types and the findings reveal trade-offs between preparation strategies for different target losses.

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

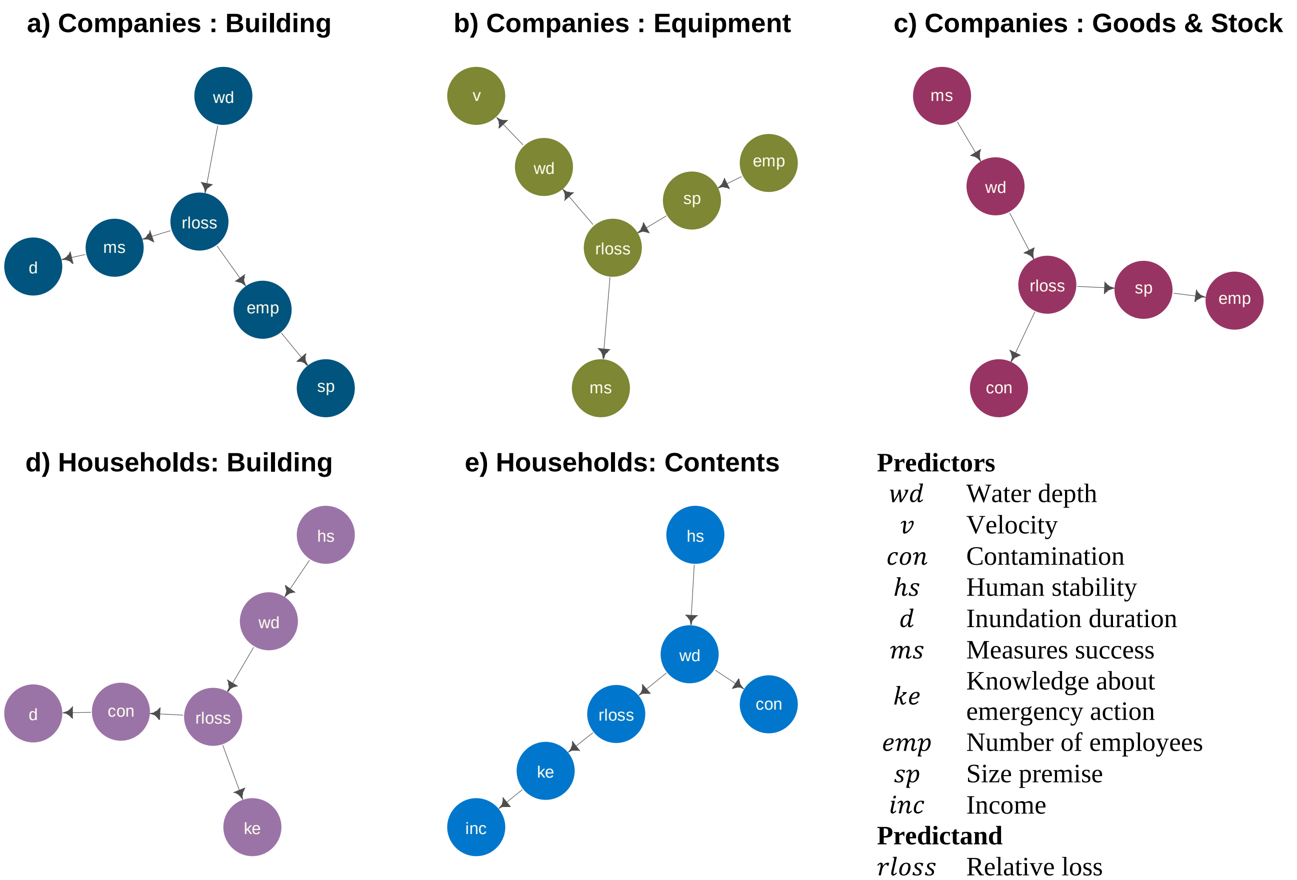

Figure 2. Bayesian Network structures for (a) Companies: Building, (b) Companies: Equipment, (c) Companies: Goods and Stock, (d) Households: Building, (d) Households: Contents, obtained from score-based structure learning algorithm.

3.2.2 Performance and comparison

The performance of the FLEMO$_{flash}$ BN structure was tested considering varying numbers of predictors, bins, and k-nearest neighbors. In the BN structure, null value was assigned to one predictor at a time (from lowest-5 rank to highest-1 rank based on feature importance – Fig 1) to examine the performance using the remaining variables (Fig 3). The performance metrics (MAE, CRPS) vary with the number of predictors. The predictive performance improved with the number of predictors. For instance, the MAE values for all asset types decreased as the number of predictors increased. Based on the performance metrics, the optimum number of predictors was found to be five. However, the Markov Blanket of loss consists of two predictors (water depth and knowledge of emergency measures) for households: contents loss, and three predictors otherwise. Similarly, CRPS values also showed a declining pattern with increasing number of predictors, indicating better probabilistic predictions. The MBE was relatively stable, suggesting that bias in the prediction did not significantly change with the number of predictors.

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

Examining the performance with optimal predictors while modifying the number of bins, we found significant differences for companies but not for households. This could be attributed to the fact that the number of data points for companies is relatively limited and more heterogenous compared to households (Schoppa et al., 2020). Company buildings model (C:BUI) with too few bins tend to lose information, resulting in higher MAE and CRPS values. Conversely, the goods & stocks model (C:GNS) performed better with fewer bins. For household models, MAE was not sensitive to the choice of bins, but model performance was slightly better with 8 bins. The MBE for company models showed fluctuations around zero, indicating some sensitivity to the choice of bins and potential bias.

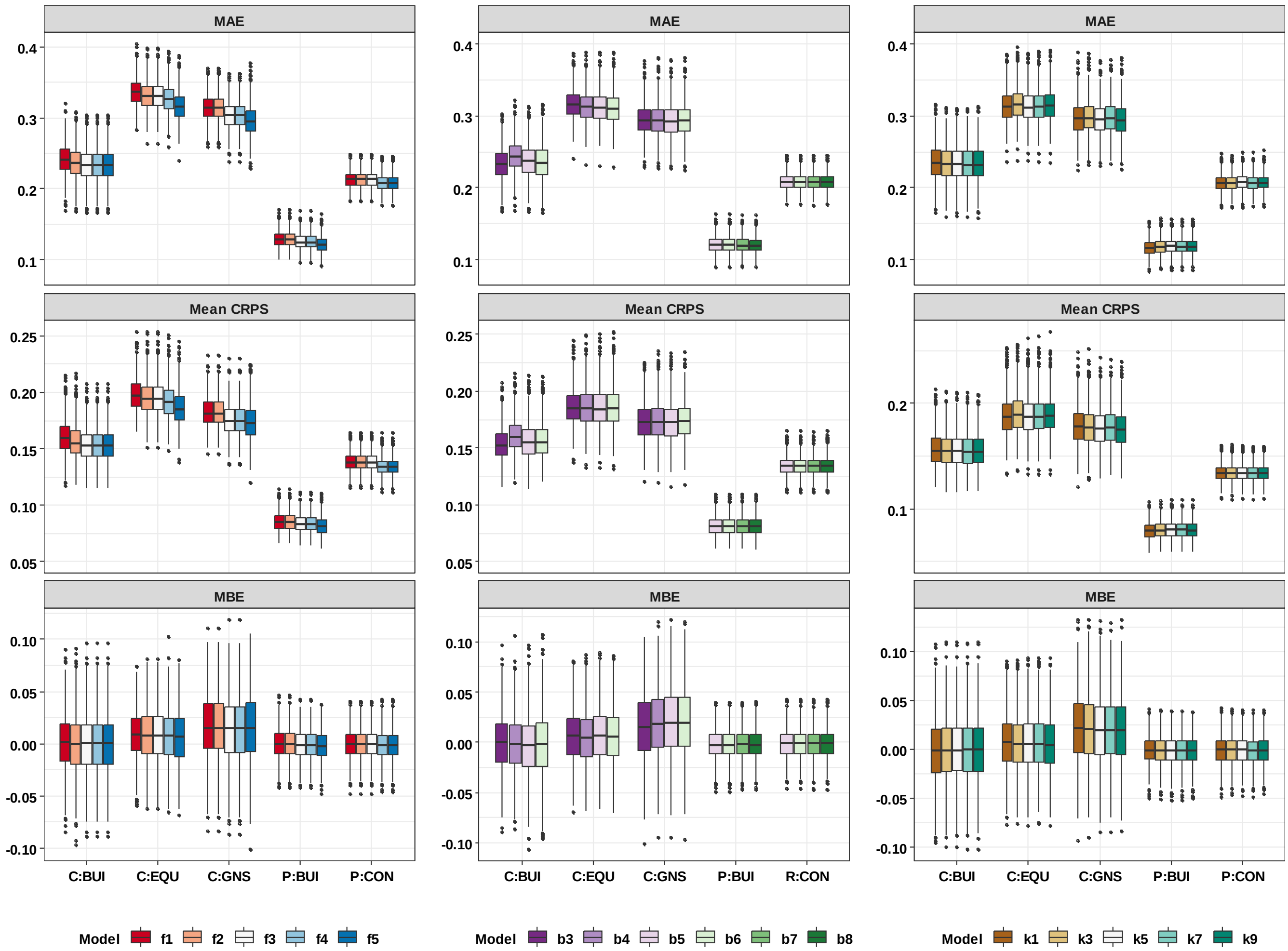

Figure 3. Model sensitivity of FLEMO$_{flash}$ to number of predictors (f1-f5), bins (b3-b8), number of neighbours used for data imputation (k1-k9) evaluated using mean average error (MAE), continuous ranked probability score (CRPS), and mean bias error (MBE) for the five asset types (x-axis). Each bocplot summarizes 100 repetitions of a fivefold cross-validation (companies) and tenfold cross-validation (households) with randomized data partitioning.

The number of neighbours used for data imputation does not show any significant difference in predictive performance. The MBE showed minor variations, remaining close to zero across different numbers of neighbours, suggesting that the bias in

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

the predictions was not significantly affected by the imputation. Overall, building assets in both company and household sectors performed better compared to other asset types. This superior performance is due to the effective inference of the relationship between water depth and building rloss (Gerl et al., 2016; Lüdtke et al., 2019; Merz et al., 2013; Vogel et al., 2018). Additionally, the emergency response of mitigation measures for companies and knowledge about emergency action for households also boosted performance.

The FLEMO*flash* model with the best performance, identified in Fig 3 was compared to SDF models (Fig 4). The FLEMO*flash* model consistently outperformed both the SDF-probabilistic and SDF-deterministic models across all five asset types due to its comprehensive representation of loss processes (Schröter et al., 2014). However, the FLEMO*flash* model overestimated losses for C:GUI compared to the SDF-D and SDF-P models. For households (P:BUI and P:CON), the losses were are significantly underestimated by the SDF-P. Study by Schoppa et al., (2020) also noted a similar observation reporting that probabilistic multivariate models performed better than univariate models for fluvial flood loss estimation. These findings highlight the superior predictive performance of FLEMO*flash* in estimating losses while also indicating the need for improvement to address biases in the predictions.

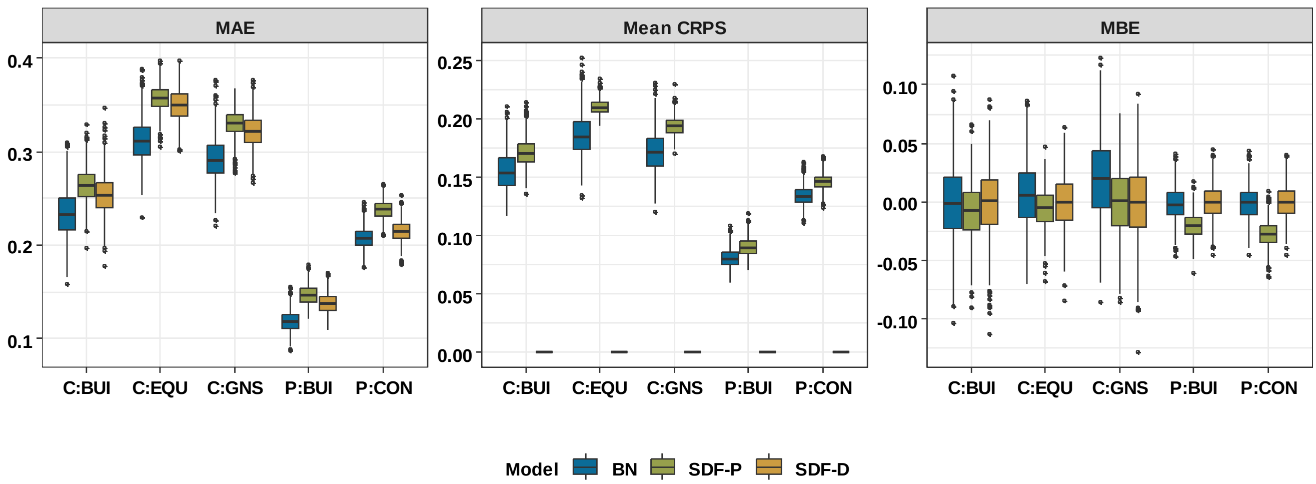

Figure 4. Comparison of FLEMO*flash* predictive performance against SDF-probabilistic and SDF-deterministic using MAE, CRPS, and MBE for the five assets (x-axis). Each boxplot summarizes 100 repetitions of a fivefold cross-validation (companies) and tenfold cross-validation (households) with randomized data partitioning.

3.3 Description of loss processes by FLEMO*flash*

The loss processes described by FLEMO*flash* is illustrated using the predictive density of predicted losses under scenarios of hazard, exposure and vulnerability. For brevity, this section focusses on the FLEMO*flash* model for household buildings (Fig 5), with a similar interpretation extended to other asset types (Fig S3-S6). The nodes of the model comprise of water depth, human stability, inundation duration, contamination, knowledge about emergency action, and relative losses, each with 7, 4, 7, 5, 6, and 8 classes, respectively, indicating different conditions. The Conditional Probability Table (CPT) was populated

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

with joint probabilities to find the predictive density of loss given the condition of other nodes in the network. The predictive density of loss based only on water depth reveals that, low water depth is associated with low losses and vice versa. This pattern is illustrated in the CPT, where the highest probabilities are concentrated along the diagonal, indicating a monotonic increasing relationship between water depth and losses (Fig 5). The CPT suggests that low water depths (< 0.28 m) are most likely associated with low loss (< 0.05), while high water depths (> 0.15m) with high loss (> 0.24). Probabilities decrease as we move away from this diagonal, indicating lower likelihood of extreme discrepancies between water depth and loss.

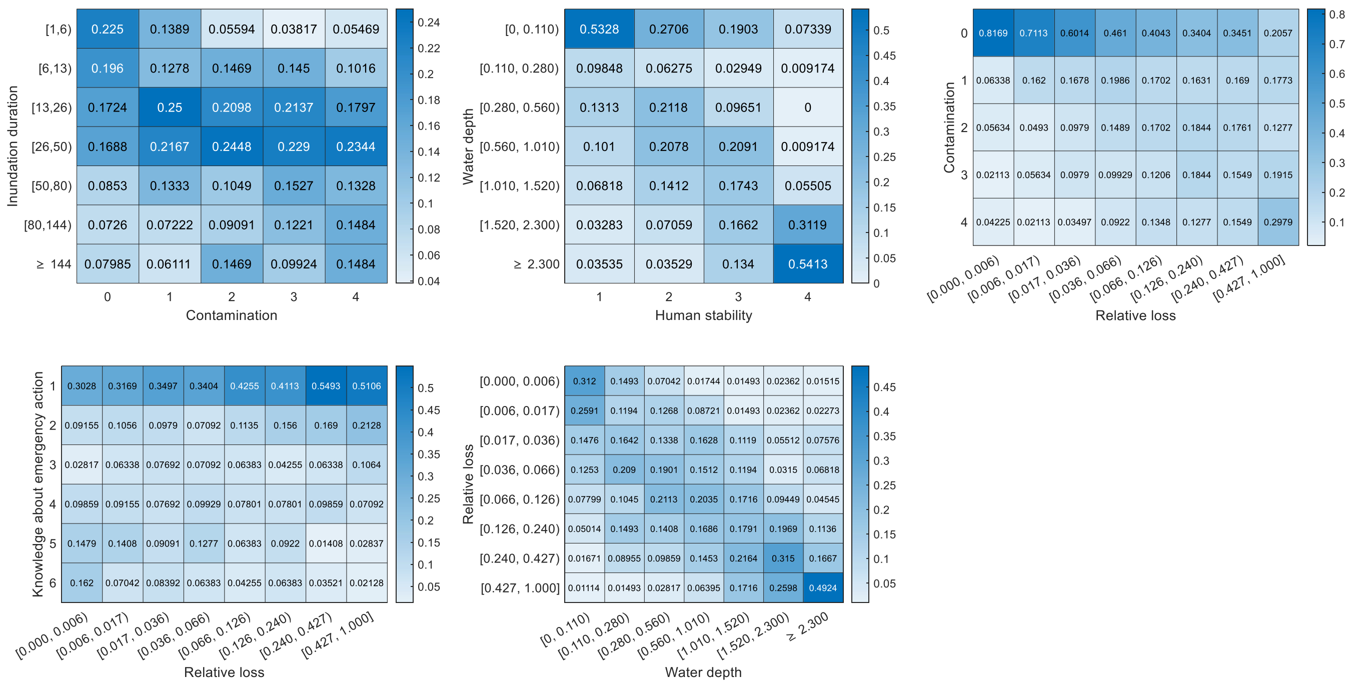

Figure 5. Conditional probability tables of the Residential: building Bayesian network. Each heatmap illustrates the conditional probabilities between a child and its parent node, with increasing intensity of blue indicating higher probability values. Numerical values are displayed in each cell, and a colorbar in each subplot shows the corresponding probability scale.

The integration of contamination into the FLEMO$_{flash}$ model, in addition to water depth, provides a more comprehensive understanding of the factors influencing flash flood damage. Contamination is classified into five levels, ranging from zero (no contamination) to four (high contamination). The CPT clearly indicates that contamination significantly amplifies the likelihood of experiencing higher loss (Fig 5). Specifically, when there is no contamination (class 0), the probability of experiencing loss is low (< 0.01). Conversely, if there is high contamination (class 4), the probability of experiencing loss is high (> 0.24), reflecting the impact of oils, chemicals, and sewage entering the building (Kreibich et al., 2005; Laudan et al., 2020).

The integration of knowledge about emergency action into the FLEMO$_{flash}$ model alongside water depth and contamination provides a comprehensive understanding of how preparedness can mitigate loss during flash floods. Knowledge about

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

emergency action is categorized into six classes, ranging from 1 (low knowledge) to 6 (high knowledge). The CPT clearly illustrates that a high level of emergency action knowledge can significantly reduce loss (Fig 5e). Specifically, when households doesn't knew what to do (1), there is a high likelihood of incurring higher loss. Conversely, when households with good preparedness (> 4), the incurred loss significantly decreases. Residents with high levels of preparedness are more likely to take effective emergency measures, thereby reducing the severity of flood loss (Kreibich et al., 2005; Sairam et al., 2019).

3.3.1 Effect of preparedness

Preparedness is crucial in mitigating potential losses (Barendrecht et al., 2020; Berghäuser et al., 2023; Bubeck et al., 2013; Sairam et al., 2019; Surminski and Thieken, 2017). In the past decade, significant progress has been made in quantifying the role of preparedness in mitigating the losses of fluvial or riverine floods (Lüdtke et al., 2019; Schoppa et al., 2020; Wagenaar et al., 2018). However, the effect of preparedness in controlling the losses during flash floods has not yet been quantified. We used the FLEMO$_{flash}$ to estimate loss distributions under different hazard & exposure conditions, while incorporating varying levels of preparedness (Fig 6).

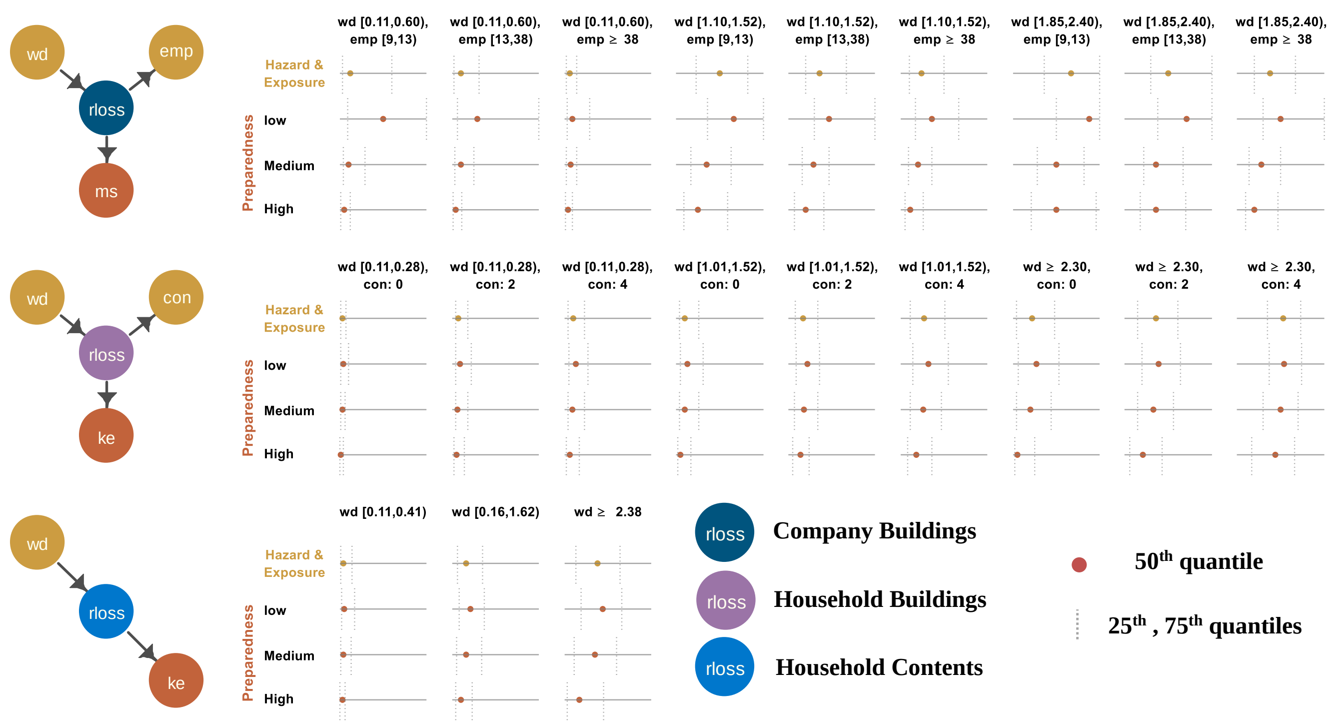

Figure 6. FLEMO$_{flash}$ application for company buildings, private household buildings and contents, considering relative loss Markov blankets. The first row in each panel shows the probabilistic predictive density of relative loss on the interval [0,1] based on the specific scenarios of hazard and exposure combination. The second to fourth rows in each panel illustrate the changes in relative loss with different levels of preparedness for the given hazard and exposure combinations.

For company building, the incurred loss increases with increasing water depth. Considering the company characteristics, the relative loss experienced by companies with 9 to 12 employees is higher compared to those with 38 or more employees. Taking preparedness into account, we found that the effectiveness of these measures significantly influences loss. For example, when considering only hazard & exposure, companies with 38 or more employees, inundated by water depths ranging from 1.85 m to 2.40 m above the ground, experience a relative loss of 0.38. However, with high preparedness, this relative loss decreases to 0.20, representing a substantial 47% reduction. On the other hand, with low preparedness, the relative loss increases to 0.50, marking a 32% increase in estimated losses.

For household buildings, considering hazard and exposure with a water depth of ≥ 2.3 m above ground and no contamination, the predicted relative loss is 0.22. However, with high preparedness with good knowledge in emergency actions, this loss decreases to 0.05, reflecting a 77% reduction. Conversely, with low preparedness, the relative loss rises to 0.27, a 23% increase. Preparedness also plays a crucial role in mitigating losses to household contents. For instance, when considering only hazard, the relative loss for water depths of ≥ 2.3 m above ground is 0.38. With high preparedness, this loss drops to 0.17, a 55% reduction. In contrast, with low preparedness, the relative loss increases to 0.44, representing a 16% rise. In both companies and households, losses are consistently higher with low preparedness and lower with high preparedness. The extent of this difference varies depending on hazard and exposure characteristics. These findings highlight the effect of preparedness in reducing flash flood risks.

4 Conclusions

This study introduces FLEMO$_{flash}$ - new multivariate flash flood loss models for estimating relative losses for companies and households. The machine learning feature selection identified key loss drivers comprising hazard, company characteristics, and emergency response variables. It enhances our understanding of damage processes during flash floods. Model-based findings show that preparedness of companies and households can significantly reduce losses due to flash floods. In extreme hazard scenarios, such as high water depth, high preparedness can reduce building losses by up to 47% for companies with 38 or more employees. Households that were well-prepared and knew precisely what actions to take during high water depth were able to reduce building losses by 77% and contents losses by 55%. The knowledge gained about flash flood damage processes and FLEMO$_{flash}$ can support risk analyses, impact-based predictions of flash floods and emergency communication to improve flash flood risk management and reduce losses.

Acknowledgements

This research has been supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the AVOSS project (grant no. FKZ 02WEE1629C). Apoorva Singh acknowledges Helmholtz Information & Data Science Academy (HIDA) for providing financial support for the research stay at the Section Hydrology, GFZ. Nivedita Sairam is funded by the BMBF project “HI-CliF”, FKZ: 01LN2209A. Collection of the 2021 private household data was undertaken by the Geography and Disaster Risk Research Lab, University of Potsdam within the KAHR-project, funded by BMBF (contract 01LR2102I). Collection of the 2021 company data was undertaken by Section Hydrology, GFZ and Deutsche Rückversicherung AG, funded by the GFZ-HART initiative and Deutsche Rückversicherung AG. Collection of the 2016 data was undertaken within the DFG Research Training Group "NatRiskChange" (GRK 2043/1), funding is gratefully acknowledged, as are additional funds provided by GFZ and the Deutsche Rückversicherung AG. Collection of the 2002 data was undertaken within the German Research Network Natural Disasters (DFNK), in cooperation between GFZ and the Deutsche Rückversicherung AG, we thank the Deutsche Rückversicherung AG and BMBF (01SFR9969/5) for financial support.

Data Availability Statement

The survey data is partly accessible at the German flood damage database, HOWAS21 (http://dx.doi.org/10.1594/GFZ.SDDB.HOWAS21)

Author contribution

Conceptualization: AS, RKG, KRS, NS, HK; **Methodology**: AS, RKG, NS, KRS, AB, MF, HK; **Data curation**: AS, RKG, MF; **Writing- original draft**: AS, RKG; **Writing- review & editing**: RKG, NS, KRS, HK; **Visualization**: AS, RKG; **Supervision**: CTD, HK

Competing interests

The author Heidi Kreibich is a member of the editorial board of Natural Hazards and Earth System Sciences.

References

Barendrecht, M. H., Sairam, N., Cumiskey, L., Metin, A. D., Holz, F., Priest, S. J., and Kreibich, H.: Needed: A systems approach to improve flood risk mitigation through private precautionary measures, Water Security, 11, 100080, https://doi.org/10.1016/j.wasec.2020.100080, 2020.

Berghäuser, L., Bubeck, P., Hudson, P., and Thieken, A. H.: Identifying and characterising individual flood precautionary behaviour dynamics from panel data, International Journal of Disaster Risk Reduction, 94, 103835, https://doi.org/10.1016/j.ijdrr.2023.103835, 2023.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Bubeck, P., Botzen, W. J. W., Kreibich, H., and Aerts, J. C. J. H.: Detailed insights into the influence of flood-coping appraisals on mitigation behaviour, Global Environmental Change, 23, 1327–1338, https://doi.org/10.1016/j.gloenvcha.2013.05.009, 2013.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 785–794, https://doi.org/10.1145/2939672.2939785, 2016.

Chinh, D. T., Gain, A. K., Dung, N. V., Haase, D., and Kreibich, H.: Multi-Variate Analyses of Flood Loss in Can Tho City, Mekong Delta, Water, 8, 6, https://doi.org/10.3390/w8010006, 2016.

Chughtai, A.: Mapping the scale of damage by the catastrophic Pakistan floods | Infographic News | Al Jazeera, https://www.aljazeera.com/news/longform/2022/9/16/mapping-the-scale-of-destruction-of-the-pakistan-floods, 2022.

Dougherty, E. and Rasmussen, K. L.: Changes in Future Flash Flood–Producing Storms in the United States, https://doi.org/10.1175/JHM-D-20-0014.1, 2020.

Gerl, T., Kreibich, H., Franco, G., Marechal, D., and Schröter, K.: A Review of Flood Loss Models as Basis for Harmonization and Benchmarking, PLOS ONE, 11, e0159791, https://doi.org/10.1371/journal.pone.0159791, 2016.

Ghaedi, H., Reilly, A. C., Baroud, H., Perrucci, D. V., and Ferreira, C. M.: Predicting flood damage using the flood peak ratio and Giovanni Flooded Fraction, PLoS ONE, 17, e0271230, https://doi.org/10.1371/journal.pone.0271230, 2022.

Gneiting, T. and Katzfuss, M.: Probabilistic Forecasting, Annu. Rev. Stat. Appl., 1, 125–151, https://doi.org/10.1146/annurev-statistics-062713-085831, 2014.

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

Jensen, F. V. and Nielsen, T. D.: Bayesian Networks and Decision Graphs, Springer New York, New York, NY, https://doi.org/10.1007/978-0-387-68282-2, 2007.

Kellermann, P., Schröter, K., Thieken, A. H., Haubrock, S.-N., and Kreibich, H.: The object-specific flood damage database HOWAS 21, Natural Hazards and Earth System Sciences, 20, 2503–2519, https://doi.org/10.5194/nhess-20-2503-2020, 2020.

Kitson, N. K., Constantinou, A. C., Guo, Z., Liu, Y., and Chobtham, K.: A survey of Bayesian Network structure learning, Artif Intell Rev, 56, 8721–8814, https://doi.org/10.1007/s10462-022-10351-w, 2023.

Kreibich, H. and Dimitrova, B.: Assessment of damages caused by different flood types, FRIAR 2010, Milan, Italy, 3–11, https://doi.org/10.2495/FRIAR100011, 2010.

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

Kreibich, H., Müller, M., Thieken, A. H., and Merz, B.: Flood precaution of companies and their ability to cope with the flood in August 2002 in Saxony, Germany, Water Resources Research, 43, https://doi.org/10.1029/2005WR004691, 2007.

Kreibich, H., Seifert ,Isabel, Merz ,Bruno, and and Thieken, A. H.: Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector, Hydrological Sciences Journal, 55, 1302–1314, https://doi.org/10.1080/02626667.2010.529815, 2010.

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

Kreibich, H., Thieken, A., Haubrock, S.-N., and Schröter, K.: HOWAS21, the German Flood Damage Database, in: Flood Damage Survey and Assessment, American Geophysical Union (AGU), 65–75, https://doi.org/10.1002/9781119217930.ch5, 2017.

Krüger, F., Lerch, S., Thorarinsdottir, T., and Gneiting, T.: Predictive Inference Based on Markov Chain Monte Carlo Output, International Statistical Review, 89, 274–301, https://doi.org/10.1111/insr.12405, 2021.

Laudan, J., Zöller, G., and Thieken, A. H.: Flash floods versus river floods – a comparison of psychological impacts and implications for precautionary behaviour, Natural Hazards and Earth System Sciences, 20, 999–1023, https://doi.org/10.5194/nhess-20-999-2020, 2020.

Lüdtke, S., Schröter, K., Steinhausen, M., Weise, L., Figueiredo, R., and Kreibich, H.: A Consistent Approach for Probabilistic Residential Flood Loss Modeling in Europe, Water Resources Research, 55, 10616–10635, https://doi.org/10.1029/2019WR026213, 2019.

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

Middelmann-Fernandes, M. h.: Flood damage estimation beyond stage–damage functions: an Australian example, Journal of Flood Risk Management, 3, 88–96, https://doi.org/10.1111/j.1753-318X.2009.01058.x, 2010.

Mohor, G. S., Hudson, P., and Thieken, A. H.: A Comparison of Factors Driving Flood Losses in Households Affected by Different Flood Types, Water Resources Research, 56, e2019WR025943, https://doi.org/10.1029/2019WR025943, 2020.

Mohor, G. S., Thieken, A. H., and Korup, O.: Residential flood loss estimated from Bayesian multilevel models, Nat. Hazards Earth Syst. Sci., 21, 1599–1614, https://doi.org/10.5194/nhess-21-1599-2021, 2021.

Munich Re: Flood risks on the rise - Greater loss prevention is needed, https://www.munichre.com/en/risks/natural-disasters/floods.html, n.d.

Paprotny, D., Kreibich, H., Morales-Nápoles, O., Castellarin, A., Carisi, F., and Schröter, K.: Exposure and vulnerability estimation for modelling flood losses to commercial assets in Europe, Science of The Total Environment, 737, 140011, https://doi.org/10.1016/j.scitotenv.2020.140011, 2020.

Pearl, J.: Probabilistic Reasoning in Intelligent Systems, Elsevier, https://doi.org/10.1016/C2009-0-27609-4, 1988.

Plapp, S.: Risk perception of natural catastrophes–an empirical investigation in six endangers areas in south and west Germany, Karlsruher Reihe II–Risikoforschung und Versicherungsmanagement, 2, 2003.

Rözer, V., Peche, A., Berkhahn, S., Feng, Y., Fuchs, L., Graf, T., Haberlandt, U., Kreibich, H., Sämann, R., Sester, M., Shehu, B., Wahl, J., and Neuweiler, I.: Impact-Based Forecasting for Pluvial Floods, Earth’s Future, 9, 2020EF001851, https://doi.org/10.1029/2020EF001851, 2021.

Sairam, N., Schröter, K., Lüdtke, S., Merz, B., and Kreibich, H.: Quantifying Flood Vulnerability Reduction via Private Precaution, Earth’s Future, 7, 235–249, https://doi.org/10.1029/2018EF000994, 2019.

Sairam, N., Schröter, K., Carisi, F., Wagenaar, D., Domeneghetti, A., Molinari, D., Brill, F., Priest, S., Viavattene, C., Merz, B., and Kreibich, H.: Bayesian Data-Driven approach enhances synthetic flood loss models, Environmental Modelling & Software, 132, 104798, https://doi.org/10.1016/j.envsoft.2020.104798, 2020.

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

Salas, J., Saha, A., and Ravela, S.: Learning inter-annual flood loss risk models from historical flood insurance claims, Journal of Environmental Management, 347, 118862, https://doi.org/10.1016/j.jenvman.2023.118862, 2023.

Schoppa, L., Sieg, T., Vogel, K., Zöller, G., and Kreibich, H.: Probabilistic Flood Loss Models for Companies, Water Resources Research, 56, e2020WR027649, https://doi.org/10.1029/2020WR027649, 2020.

Schoppa, L., Barendrecht, M., Sieg, T., Sairam, N., and Kreibich, H.: Augmenting a socio-hydrological flood risk model for companies with process-oriented loss estimation, Hydrological Sciences Journal, 67, 1623–1639, https://doi.org/10.1080/02626667.2022.2095207, 2022.

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

Scutari, M. and Denis, J.-B.: Bayesian Networks: With Examples in R, 2nd ed., Chapman and Hall/CRC, Boca Raton, https://doi.org/10.1201/9780429347436, 2021.

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

Sieg, T., Vogel, K., Merz, B., and Kreibich, H.: Tree-based flood damage modeling of companies: Damage processes and model performance, Water Resources Research, 53, 6050–6068, https://doi.org/10.1002/2017WR020784, 2017.

Surminski, S. and Thieken, A. H.: Promoting flood risk reduction: The role of insurance in Germany and England, Earth's Future, 5, 979–1001, https://doi.org/10.1002/2017EF000587, 2017.

Thieken, A. H., Olschewski, A., Kreibich, H., Kobsch, S., and Merz, B.: Development and evaluation of FLEMOps – a new *F* lood *L* oss *E* stimation *MO* del for the *p* rivate *s* ector, in: Flood Recovery, Innovation and Response I, FLOOD RECOVERY, INNOVATION AND RESPONSE 2008, London, England, 315–324, https://doi.org/10.2495/FRIAR080301, 2008.

Thieken, A. H., Samprogna Mohor, G., Kreibich, H., and Müller, M.: Compound inland flood events: different pathways, different impacts and different coping options, Nat. Hazards Earth Syst. Sci., 22, 165–185, https://doi.org/10.5194/nhess-22-165-2022, 2022.

Vogel, K., Weise, L., Schröter, K., and Thieken, A. H.: Identifying Driving Factors in Flood-Damaging Processes Using Graphical Models, Water Resources Research, 54, 8864–8889, https://doi.org/10.1029/2018WR022858, 2018.

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

Wagenaar, D., Lüdtke, S., Schröter, K., Bouwer, L. M., and Kreibich, H.: Regional and Temporal Transferability of Multivariable Flood Damage Models, Water Resources Research, 54, 3688–3703, https://doi.org/10.1029/2017WR022233, 2018.

Wang, K., Wang, L., Wei, Y.-M., and Ye, M.: Beijing storm of July 21, 2012: observations and reflections, Nat Hazards, 67, 969–974, https://doi.org/10.1007/s11069-013-0601-6, 2013.

https://doi.org/10.5194/egusphere-2025-1512
Preprint. Discussion started: 22 April 2025

Zain, A., Legono, D., Rahardjo, A. P., and Jayadi, R.: Review on Co-factors Triggering Flash Flood Occurrences in Indonesian Small Catchments, IOP Conf. Ser.: Earth Environ. Sci., 930, 012087, https://doi.org/10.1088/1755-1315/930/1/012087, 2021.

Zander, K. K., Nguyen, D., Mirbabaie, M., and Garnett, S. T.: Aware but not prepared: understanding situational awareness during the century flood in Germany in 2021, International Journal of Disaster Risk Reduction, 96, 103936, https://doi.org/10.1016/j.ijdrr.2023.103936, 2023.

Zou, H. and Hastie, T.: Regularization and Variable Selection via the Elastic Net, Journal of the Royal Statistical Society. Series B (Statistical Methodology), 67, 301–320, 2005.