Egusphere-2025-1715: Deciphering the drivers of direct and indirect damages to companies from an unprecedented flood event: A data-driven, multivariate probabilistic approach

Reviewer #1:

This is an interesting and highly relevant topic that can significantly contribute to a deeper understanding of the various factors influencing both direct and indirect damages to businesses caused by flooding events. The research methods employed are notably technical and innovative, offering fresh perspectives and valuable insights into the complexity of flood-related impacts on commercial sectors. However, despite the strengths of the approach, there are certain points that require further attention and refinement. These include the justification of chosen methodologies, the interpretation of the survey results, a more clear interpretation of the results, and the need for a more comprehensive discussion of the limitations and potential implications of the findings. Addressing these aspects would enhance the overall robustness and applicability of the study.

We would like to express our sincere gratitude to the Editor for keeping the discussion open upon request and would like to thank the reviewer for acknowledging the significance of our study and for providing valuable feedback. The comments were extremely helpful in improving the quality of the manuscript and will be acknowledged. We provide detailed responses to each comment below, with reviewer comments shown in blue and our responses in black. P refers to page and L refers to Line number. All references cited in our responses are listed at the end of this letter.

Abstract, 'to date no study has examined the factors influencing company damages during such an extreme event'; is this a correct statement? In the introduction you mention multiple papers that investigated the factors that influenced company damages such as Endendijk et al. (2024), Kreibich et al. (2010). Please clarify or revise this statement.

We thank the reviewer for this important observation. We will revise the original statement in abstract as follows:

"*While the drivers of company damages from riverine flooding are well documented, the drivers of both direct and indirect damages during an extreme flash flood event have not yet been examined.*"

We will also revise the abstract as follows (P1/L13-24):

Floods are among the most destructive natural hazards, causing extensive damage to companies through direct impacts on assets and prolonged business interruptions. The extraordinary July 2021 fast-onset flood in Germany caused unprecedented damage, particularly in North Rhine-Westphalia and Rhineland-Palatinate, affecting companies of all sizes. While the drivers of company damages from riverine flooding are well documented, the drivers of both direct and indirect damages during an extreme flash flood event have not yet been examined. This study addresses this gap using survey data from 431 companies affected by the July 2021 flood. Results show that 62% of companies incurred direct damages exceeding €100,000. Machine learning models and Bayesian network analyses identify water depth and flow velocity as the primary drivers of both direct damage and business interruption. However, company characteristics (e.g., premises size, number of employees) and preparedness also play critical roles. Companies that implemented precautionary measures experienced significantly shorter business interruption durations—up to 58% for water depths below 1 m and 44% for depths above 2 m. These findings offer important insights for policy development and risk-informed decision-making. Incorporation of behavioural indicators into flood risk management strategies and improving early warning systems could significantly enhance business preparedness.

Method

Survey data, it would be good to better define the variables in an appendix for example. It is not clear how business interruption is defined. Does business interruption mean that the business is not operational at all or that there is a reduction in business activity, if so how much is this reduction. This should be better defined.

In the revised manuscript, we will add an overview of all variables in the appendix (Table A1). Business interruption and restriction are defined as follows (P24/L540):

Business interruption duration (bid): The number of days during which business operations were completely suspended as a direct consequence of the flooding event. A value of 0 indicates no interruption, while values up to 540 represent the reported duration of full shutdown. A value of 540 days reflects the survey limit, meaning the business had not yet resumed operations when the survey ended.

Business restriction duration (brd): The number of days it took after the flooding event until the business resumed operations without any restrictions. Restrictions refer to any form of reduced capacity compared to pre-flood conditions. The maximum value is 540 days, meaning the business still had restrictions when the survey ended.

Table A1: Overview of company variables, associated survey questions, response types (continuous, ordinal, nominal), and, for selected variables, the steps applied to develop the index.

Variable		Survey question	Response Type (and Index development)
wd	Water depth	At maximum water level, how high was the water above the Earth's surface on your company premises in cm?	Continuous variable
d	Inundation duration	For how many hours did water remain on the company premises?	Continuous variable
v	Velocity indicator	How strong was the water current in the immediate vicinity of your company?	• 1 – Calm/slowly flowing • 2 • 3 • 4 • 5 • 6 – Wild/turbulent current Recoded categories (used in the analysis): 1. Low flow (original categories 1–2) 2. Moderate flow (original categories 3–4) 3. Torrential flow (original categories 5–6)
con	Contamination	Did contamination from the following substances entered your company during the flood event?	*Categorial response (with multiple options possible):* • Oil/Gasoline • Chemicals

			• Sewage • No contamination Recoded categories (used in the analysis): 0. No contamination 1. Sewage or Chemicals only 2. Oil/Gasoline only 3. Oil/Gasoline + Sewage, or Oil/Gasoline + Chemicals 4. Oil/Gasoline + Chemicals + Sewage
ew	Early warning received	Did your company receive an early warning of the flood event?	0. No Yes
ws	Early warning source	From which source did your company receive the flood warning?	*Response (with multiple options possible):* • Loudspeaker announcements • App or SMS • Telephone call • Radio report • TV report • Newspaper report • Social media • Own research • Own observation • No warning Recoded categories (used in the analysis): 0. No warning 1. Own research 2. Contacts (employees, acquaintances, other companies, phone calls) 3. Media (radio, TV, newspaper, online, social media) 1. Official authorities (direct official warning, apps/SMS, civil protection, loudspeaker announcements, regional services)
wt	Warning lead time	How many hours before the arrival of the flash flood or heavy	Number of hours before the arrival of the flash flood or heavy rainfall that the warning reached the company. Companies that reported

		rainfall did the warning reach your company?	"no warning received" were coded as 0 hours, as they were not asked the follow-up question on warning lead time. This approach reduced the proportion of missing values.
ws	Early warning source	From which source did your company receive the flood warning?	*Response (with multiple options possible):* • Loudspeaker announcements • App or SMS • Telephone call • Radio report • TV report • Newspaper report • Social media • Own research • Own observation • No warning Recoded categories (used in the analysis): 4. No warning 5. Own research 6. Contacts (employees, acquaintances, other companies, phone calls) 7. Media (radio, TV, newspaper, online, social media) 8. Official authorities (direct official warning, apps/SMS, civil protection, loudspeaker announcements, regional services)
me	Emergency measures undertaken	Were measures to reduce damage undertaken in your company before or during the flood event?	0. No 1. Yes
ep	Emergency plan	At the time of the flood event, did your company have an emergency or flood protection plan?	0. No 1. Yes
kh	Knowledge about hazard	Had this site already been flooded before? Were you aware that your company is located in a flood-prone area?	This variable was derived from two survey questions. If a site had been flooded before, we coded the company as having knowledge (Yes). If the site had not been flooded before, we then used the follow-up question on awareness of being located in a flood-prone area. Companies that reported awareness were

			coded as Yes, while those that were not aware were coded as No.
ms	Emergency measures success	Were measures to reduce damage undertaken in your company before or during the flood event? How effective were these mitigation measures?	This variable was based on two survey questions. First, respondents were asked whether any measures to reduce damage were undertaken before or during the flood event. If no measures were undertaken, the company was coded as "No measure undertaken." If measures were reported, respondents were then asked to rate their effectiveness. Responses were coded into four categories: Completely ineffective, Partly effective, Mostly effective, Completely effective. Recoded categories (used in the analysis): 0. No measure undertaken 1. Completely ineffective, 2. Partly effective, *3.* Mostly/ completely effective
fe	Flood experience	Q1: Had this company site already been flooded before the event? If yes, how many times?	Number of previous floods: 0. Never 1. Once 2. Twice 3. ≥ Three times
		Q2: When was the company site last affected by a flood prior to the event? (Year)	Time elapsed since the last flood: 1. > 25 years ago 2. 11–25 years ago 3. 2–10 years ago
		Flood experience was calculated from the number of previous floods (Q1) and the time elapsed since the last flood (Q2).	• If only one value (Q1 or Q2) was available, that value was used. • If both values were available, the flood experience score was calculated as the mean of the two.
pr	Precaution indicator	*Measures included* V1. Company insured against flood damages. V2. Heating system adjusted (converted or flood-protected). V3. Emergency plan in place.	Conversion: • Each measure was coded as 1 if implemented prior to the flood, 0 otherwise.

		V4. Frequency of emergency drills conducted before the flood. V5. Tanks, silos, or storage facilities securely anchored. V6. Stationary or mobile water barriers installed. V7. Sensitive equipment relocated to higher floors. V8. Water-hazardous substances relocated to higher floors. V9. Use of flood-prone areas adapted to risk. V10. Air conditioning/ventilation system flood-proofed. V11. Building flood safety improved (e.g., sealing basements, strengthening stability).	• For drills, any positive frequency (≥1 per year) was coded as 1, absence as 0. Weighting scheme: • Low impact / basic preparedness (weight = 1): V1 to V4 • Medium impact / protective but limited scope (weight = 5): V5 to V8 • High impact / comprehensive protection (weight = 10): V9 to V11 Calculation of weighted score (p): $p = v1 + v2 + v3 + v4 + \left(5 \times (v5 + v6 + v7 + v8)\right) + \left(10 \times (v9 + v10 + v11)\right)$ Precaution Indicator (pr): 0. No precautionary measures 1. Medium precaution (p: $1 - 5$) 2. Very good precaution ($p \geq 6$)
in	Insurance	Is the company insured against flood damages before the flood event?	0. No 1. Yes
sp	Size premise	How large is the property on which your company is located?	Continuous variable (m^2)
sec	Sector	Which sector does your company belong to?	Categorial variable: 1. Agriculture 2. Manufacturing 3. Commerce 4. Financial 5. Private and public services
ss	Spatial situation	Which description best fits the spatial situation of this flood-affected company site?	Categorial variable: 1. Business premises with several buildings belonging to the company 2. Entire building fully used by the company 3. One or more floors in a building otherwise used for non-business purposes

			4. Less than one floor in a building otherwise used for non-business purposes
own	Ownership	Are the buildings or rooms owned by the company or rented?	1. Owned 2. Rented 3. Partly owned / partly rented
emp	Number of employees	How many people were employed in the previous month?	Continuous variable
Damage type			
Predictand		**Description**	**Response**
bdam	Relative damage to building	Represents the percentage of costs incurred repairing or replacing elements of the building fabric in relation to its new value.	Degree of damage between 0 and 1
edam	Relative damage to equipment	Represents the percentage of costs incurred repairing or replacing equipment of fixed assets in relation to its new value.	Degree of damage between 0 and 1
gsdam	Relative damage to goods & stock	Represents the percentage of costs incurred repairing or replacing goods, products, and stock in relation to its new values.	Degree of damage between 0 and 1
bid	Business interruption duration	How long, in the aftermath the flooding event, were businesses operations totally interrupted	0 to 540 days (A value of 0 indicates no interruption, while values up to 540 indicate the reported duration of full shutdown. Cases recorded at 540 days reflect the survey limit, meaning that the business had not yet resumed operations at the time of the survey)
brd	Business restriction duration	How long, in the aftermath the flooding event, businesses operations resumed without any restrictions	0 to 540 days (The maximum value is 540 days, meaning the business still had restrictions when the survey ended)

Variable selection, please introduce this section. The variable selection section dives into the three machine learning techniques without introducing why these three techniques are used. In general, the method section needs more structure: it should be better explained why each algorithm/method is used. A clear motivation as to why the three specific techniques are used is needed.

We thank the reviewer for the feedback. In the revised manuscript, we will introduce the variable selection section with a clear rationale for choosing the three machine learning techniques, as follows (P6/L156-165):

"2.2 Variable Selection

Flood damage processes vary by region, flood type, and asset type (Mohor et al., 2020; Sairam et al., 2019; Wagenaar et al., 2018). Since our analysis focuses on flash floods and covers direct and indirect damages, we use a data-driven approach to identify which variables strongly influence these diverse outcomes. We adopt three feature selection approaches that are robust to multicollinearity and capable of capturing nonlinear relationships and interactions. To this end, we employ three complementary machine learning techniques: Elastic Net (EN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). EN efficiently handles multicollinearity and performs variable selection through regularization; RF captures nonlinear relationships and complex interactions via ensemble decision trees; and XGBoost, a gradient boosting algorithm, provides high predictive accuracy and models intricate dependencies. By combining the strengths of these methods, we assume to ensure a comprehensive assessment of variable importance. To mitigate potential biases from relying on a single model, we aggregate the variable importance scores across all three methods to derive a final ranking.''

Minimizing J(β) should be called Obj(β) or it should be made more clear that J stands for the objective function as in equation 1 it is defined as Obj(β) and not J(β).

P7/L174: $J(\beta)$ will be replaced with $Obj(\beta)$ to match the notation used in Equation 1.

Variable importance: Please better explain/introduce why Bayesian Networks are used in this case.

We thank the reviewer for the feedback. In this study, BNs are employed to complement the machine learning models by providing a probabilistic framework for analyzing multivariate dependencies and scenario-based inference. Whereas EN, RF, and XGBoost primarily emphasize predictive accuracy and variable ranking, BNs explicitly capture conditional dependencies among variables. We will revise the section to better explain this motivation as follows (P9/L226-233):

"Bayesian networks (BNs) are probabilistic graphical models that represent dependencies among multiple variables and enable multivariate predictive density estimation (Sucar, 2021). In this study, BNs are employed to complement the machine learning models by providing a probabilistic framework for analyzing multivariate dependencies and scenario-based inference. Whereas EN, RF, and XGBoost primarily emphasize predictive accuracy and variable ranking, BNs explicitly capture conditional dependencies among the variables. This is particularly valuable in flood damage analysis, where damage outcomes result from complex interactions between hazard intensity, company characteristics, and preparedness measures. Moreover, BNs are able to estimate posterior probabilities of damages given partial evidence (e.g., observed water depth or company preparedness), thereby offering a transparent and interpretable tool for risk assessment under uncertainty".

Results and discussion

Overview of affected companies: It is implied that sales figures would be a better metric of company size although number of employees is more often used to classify whether the company is an SME or a large company. Therefore, this sentence is unnecessary in my opinion.

We agree with the reviewer and have removed this sentence in the revised manuscript.

'These disruptions can result in partial or complete business interruptions, triggering consequences ranging from loss of sales to bankruptcy'. This sentence is unclear, loss of sales is a form of business interruption.

We will revise the sentence as follows (P11/L276-278):

"*Floods not only cause damage to tangible assets such as buildings or machinery but also lead to significant disruptions in supply chains and transportation. Such disruptions can result in partial or complete business interruptions and, in extreme cases, bankruptcy (Thieken et al., 2016).*"

It would also be interesting to show the differences in vulnerability and exposure levels between sectors instead of only between company sizes. This should be added or otherwise be explained why it is left out.

We thank the reviewer for this insightful suggestion. We agree that differences among sectors can provide additional perspectives on vulnerability and exposure. In our analysis, company size was found to be the dominant factor in explaining variations in damages. To maintain focus and clarity, our main emphasis in this study is on company size. However, sectoral difference is shown in the Supplementary Information (Fig. S4) and will be referred in the revised manuscript as follows (P11/L278-282):

"*Figure 2 illustrates the percentage of companies affected by various types of impacts, categorized by company size, while Figure S4 presents the same results by sectors. Since company size emerged as the dominant factor explaining variations in damages or revealed differences in vulnerability levels, our main emphasis in this study is on company size.*"

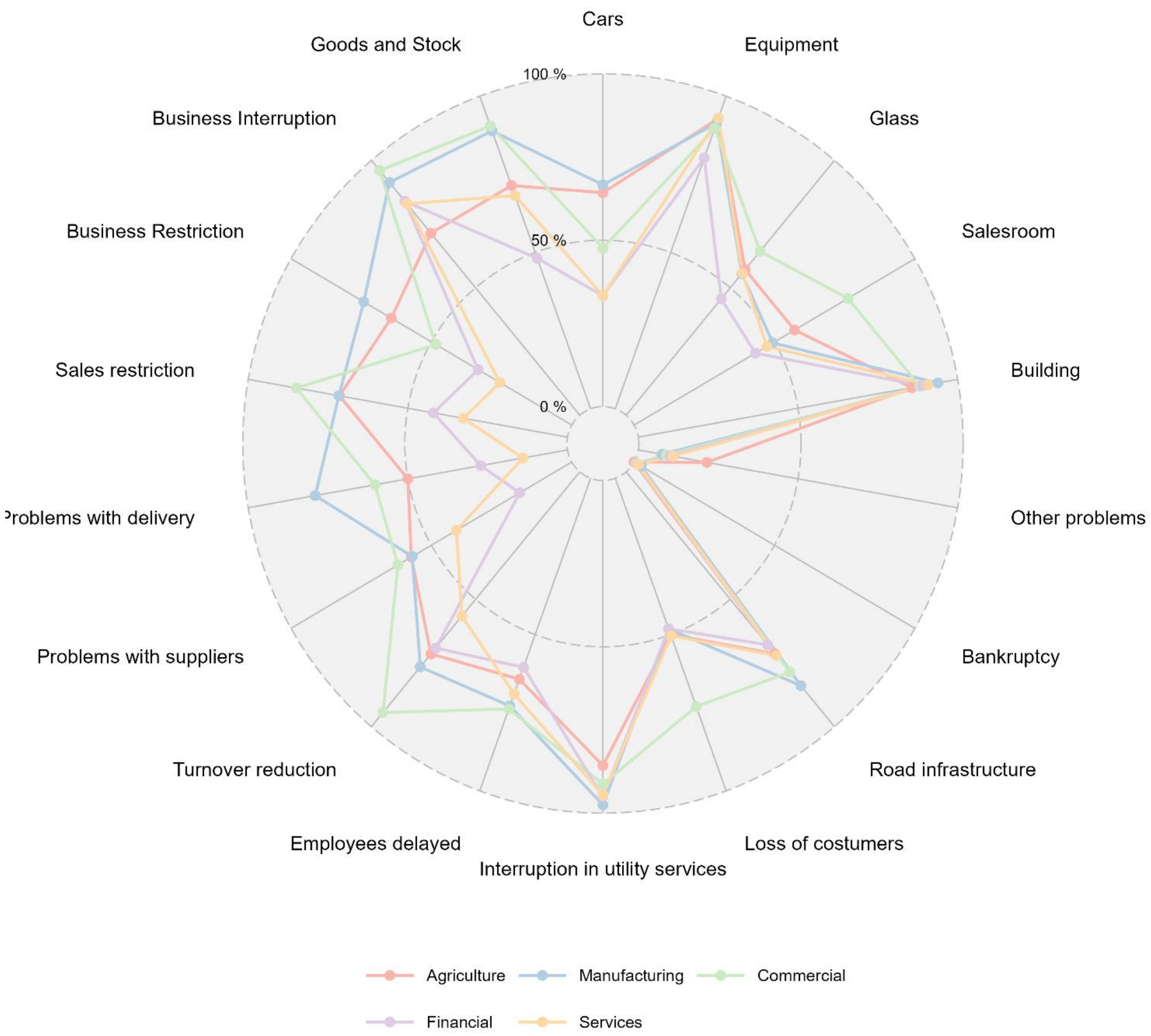

Figure S4: Spider chart illustrating the percentage of companies experiencing different types of flood impacts, categorized by the sector.

‘Bankruptcy risks remain generally low across all company sizes’. How is bankruptcy risk defined? Isn’t this a very biased variable given that bankrupt companies are probably not surveyed? Please clarify or leave this out.

We have tried contacting companies that went bankrupt, however it is challenging, as they are less likely to participate in survey and probably have already moved out of the affected area at the time of the survey. To avoid potential misinterpretation, we will remove this statement in the revised manuscript and also highlight it is one of the limitations.

‘They highlight the need for tailored risk management (...)’. Please clarify to what it should be tailored, to company size or also to company sector?

We will revise the sentence as follows (P13/L308-311):

“*Overall, the results illustrate the complex and diverse impacts of flooding on companies, varying by size. Micro and small companies are more susceptible to supply chain disruptions and sales restrictions, while larger companies face higher asset-related risks. Accordingly, risk management and resilience strategies should be tailored to company size*”.

‘Tend to recover more quickly, likely benefiting from greater resilience’. This sentence sounds tautological -> recovering more quickly is part of the definition of resilience.

We will revise the text as follows (P14/L337-338):

“*In contrast, medium and large companies tend to recover more quickly, likely because they benefit from diversified operations, and access to more substantial resources*.”

Figure 3: Please explain why the outlier levels differ between the business sizes. It seems weird to leave out observations for one class and leave them in for another. This does not look correct. Also, why are there no outliers removed for business restriction duration?

We thank the reviewer for this important observation. Figure 3 (provided below) presents the distribution of business interruption duration and business restriction duration across companies of varying sizes. Because quartiles (Q1, Q3) and interquartile ranges (IQR) differ between company size groups, the thresholds for detecting outliers also vary.

For business interruption duration, some smaller and medium-sized companies have values exceeding these thresholds, while (the few) larger companies do not. For business restriction duration, all observed values fall within the calculated bounds, so no outliers were detected. All data points are retained in the analysis, and the table below (shown only for response) provides the detailed statistics and identified outliers.

We will add the following text in the revised manuscript (P13/L323-326):

“The number of outliers differs across company sizes because thresholds were determined using the standard 1.5 × IQR rule. For business restriction duration, no outliers were detected, as the upper thresholds were consistently high (e.g., >650 days for micro and small companies) and all observations fell within these ranges.”

Company size	Q1	Q3	IQR	Outliers
Business Interruption Duration				

1–9	10.00	180.00	170.00	540.00, 500.00, 510.00, 499.00, 450.00, 540.00, 462.00, 540.00, 540.00
10–49	8.50	75.00	66.50	270.00, 270.00, 300.00, 180.00, 210.00, 180.00, 250.00, 365.00, 180.00, 420.00, 450.00, 420.00, 340.00, 270.00
50–249	6.50	60.00	53.50	180.00, 540.00
250–480	5.00	112.50	107.50	–
Business Restriction Duration				
1–9	36.25	290.00	253.75	–
10–49	28.00	280.00	252.00	–
50–249	12.00	150.00	138.00	–
250–480	120.00	120.00	0.00	–

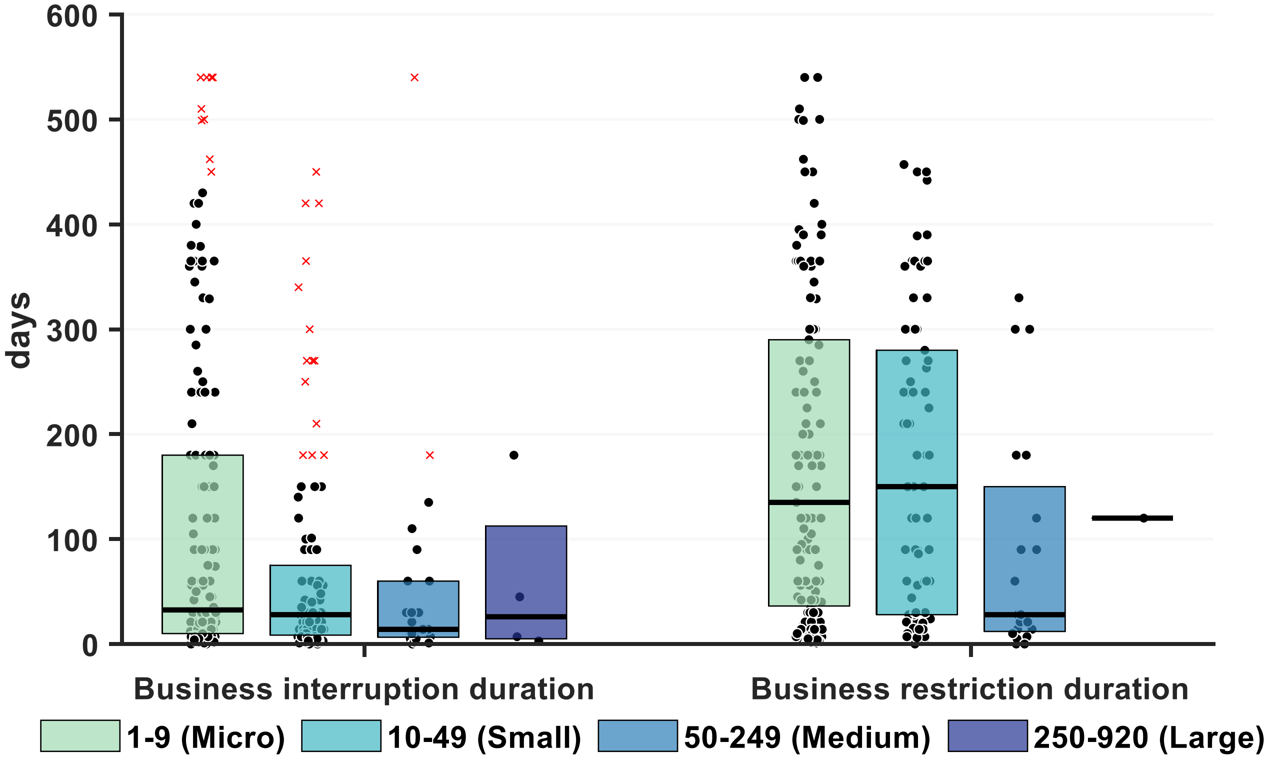

Figure 1: Boxplot of (a) Business interruption duration (days) and (b) Business restriction duration (days) for companies categorized by the number of employees. Black circular markers represent individual data points, and red crosses indicate outliers identified using the 1.5 × IQR rule.

Having an n=3 for large companies is too low for any inference. Please make this more clear. 'should be interpreted with caution' does not cover it fully in my opinion.

We agree that a sample size of three for large companies is extremely small and does not allow for reliable statistical inference. We have revised the manuscript to explicitly highlight this limitation and to clarify that these results are illustrative rather than generalizable. We will add the following text in the revised manuscript (P16/L376-378):

"Due to the extremely limited number of large companies surveyed, these results cannot be generalized. These values are presented for illustration purposes only and cannot be considered representative of large companies in general."

‘However, substantial variance within each category highlights the influence of extreme cases’. Maybe it is better to infer about the median values instead of the averages then. Please do this or clarify why not.

In the manuscript, we have clearly presented both median and mean values (Table 2) to provide a comprehensive picture of the financial losses. Both the medians and averages are included to highlight the impact of extreme cases. To improve clarity, we will revise the text (P14-16/L345-378) to explicitly note that while the median values reflect more frequent losses, the mean is nonetheless informative, reflecting skewness of the data.

“The average costs (in euros) for each company size are presented in Table 2, alongside medians and the number of companies (n) contributing to each calculation. Building damages accounted for the highest average costs across all company sizes, particularly impacting medium and large companies. Micro companies reported building damages of €711,459 on average, with a median of €250,000. This wide gap between the mean and median suggests that while many small firms experienced moderate losses, a few outliers faced severe damages. For small companies, the mean building damage increased to €908,482 (median €500,000). Medium companies faced substantial building-related losses, averaging €2,838,103 with a median of €1,350,000. Large companies, though represented by a very small sample (n = 4), reported the highest mean building damages of €7,350,000, reflecting the scale of structures at risk within large industrial facilities.

In terms of equipment damages, micro companies incurred a mean loss of €297,854 (median €50,000), while small companies experienced mean loss of €541,898 (median: €150,000). Medium companies reported the highest mean losses at €3,630,652 with a median of €600,000 likely driven by the presence of high-value machinery. Interestingly, large companies recorded a comparatively lower mean loss of €160,000 (median €200,000), though this is based on a very small sample size (n = 3). Lower median values across all groups suggest the presence of extreme cases skewing the mean, particularly among medium-sized companies. Goods and stock damages were generally lower across all company sizes (Table 2). Micro companies faced mean losses of €159,422 (median: €30,000), while small companies reported similar mean damages of €134,470 (median: €31,500). Medium companies experienced higher mean losses of €1,503,250 (median: €150,000), indicating greater inventory exposure. Large companies reported much smaller mean losses of €55,000 (median: €10,000), but are not representative due to the small sample. Lower median values that most companies incurred relatively less damages in this category, with a few outliers.

Business interruption losses also varied by company size. Micro companies faced interruption costs of €139,931 on average (median: €30,000), while small companies reported higher mean losses of €311,173 (median: €100,000). Medium companies were the most affected, with mean losses of €703,250 (median: €200,000). Large companies, despite the small sample size (n = 3), recorded an average business interruption cost of €400,000, with the median even higher at €500,000, reflecting significant operational disruptions. Overall, the financial costs associated with building, equipment, goods & stock, and business interruption showed that larger companies typically incurred more significant costs. Due to the extremely limited number of large companies surveyed, these results cannot be generalized. These values are presented for illustration purposes only and cannot be considered representative of large companies in general.”

Table 1: Average financial costs (in euros) incurred for building, equipment, goods and stock, and business interruption categorized by the number of employees (values in brackets represent

medians, and n denotes the number of companies included in the calculation of the means and medians)

Number of employees (Company size)	Building	Equipment	Goods & stock	Business interruption
1-9 (Micro)	711,459 (250,000) n = 167	297,854 (50,000) n = 203	159,422 (30,000) n = 154	139,931 (30,000) n = 143
10-49 (Small)	908,482 (500,000) n = 83	541,898 (150,000) n = 96	134,470 (31,500) n = 82	311,173 (100,000) n = 74
50-249 (Medium)	2,838,103 (1,350,000) n = 29	3,630,652 (600,000) n = 23	1,503,250 (150,000) n = 20	703,250 (200,000) n =16
249-920 (Large)	7,350,000 (1,700,000) n = 4	160,000 (200,000) n = 3	55,000 (10,000) n = 3	400,000 (500,000) n = 3
Total	1,080,999 (350,000) n = 283	604,528 (100,000) n = 325	254,083 (30,000) n = 259	215,910 (50,000) n=236

18 out of 19 variables had less than 7% missing data which was imputed. How much missing data did the other variable have and was this imputed too? Be more clear here.

We thank the reviewer for this observation. We will revise the text in the manuscript for clarity as follows (P16/L385-387):

"The dataset exhibited less than 7% missing data for 18 out of 19 variables (Fig. S1), which were imputed using the kNN technique with $k = 5$ *neighbors (Askr et al., 2024). The remaining variable, warning lead time (wt), had approximately 12% missing data, which was also imputed using the same approach."*

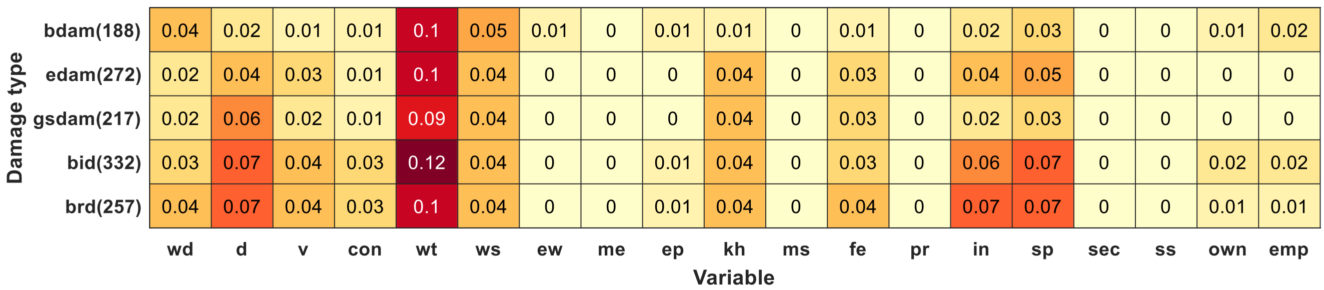

Figure S1: Percentage of missing values per factor (x-axis) for each damage type (y-axis). The values shown in the heatmap are the percentages of missing data, where 0.1 corresponds to 10%. The value in parentheses for each damage type indicates the number of responses available out of 431. For warning time (wt), cases where no warning was received are treated as zero.

Figure 4 and Figure 5: these abbreviations are unclear, write them out or find another way of making them more informative. A figure should be understandable on its own.

Figures 4, and 5 will be revised as follows, and all abbreviations are now written out in full to ensure that the figures are self-explanatory.

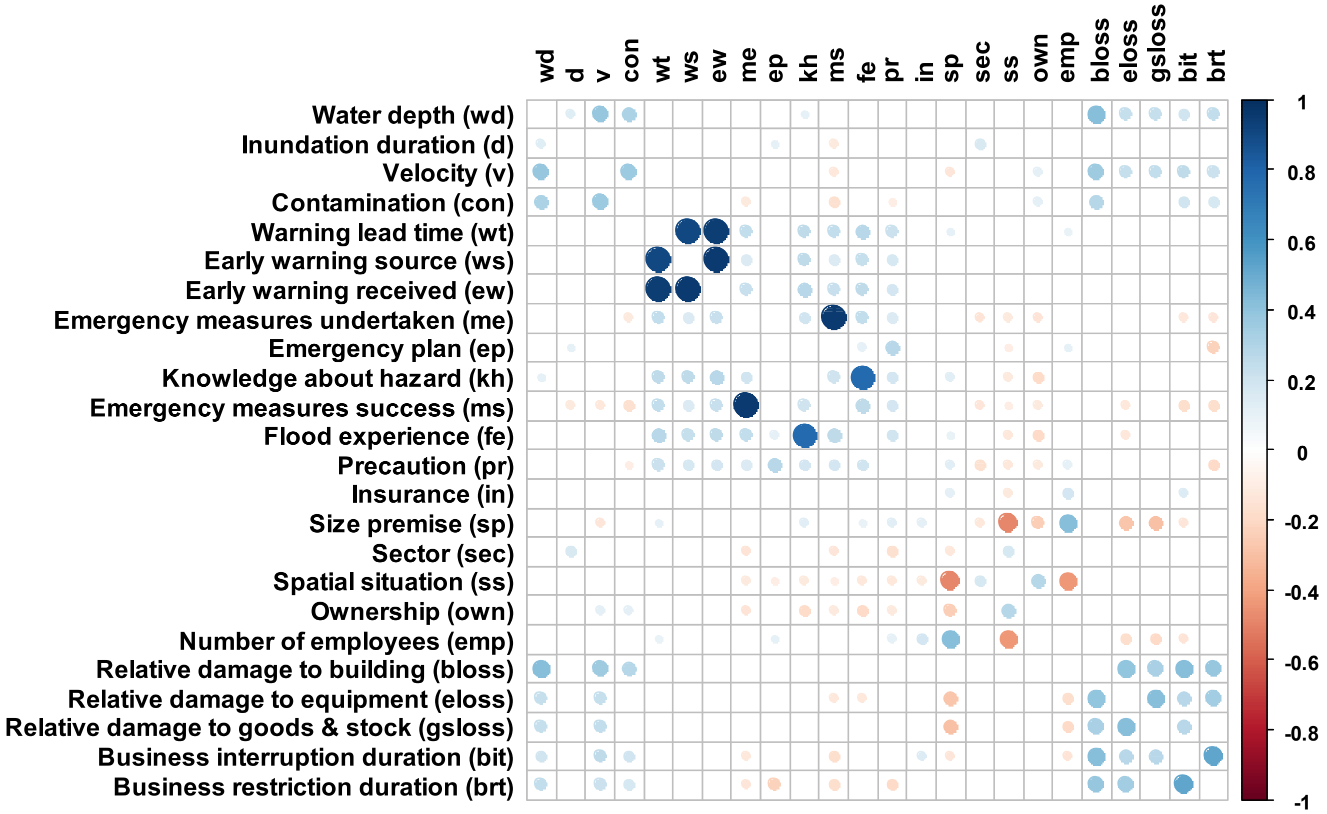

Figure 4: Spearman rank correlation coefficients between 19 influencing factors and five damage types. Full names with abbreviations in brackets are shown in the rows, and abbreviations only in the columns. Only statistically significant correlations ($p < 0.05$) are displayed, highlighting key relationships between influencing factors and damage outcomes.

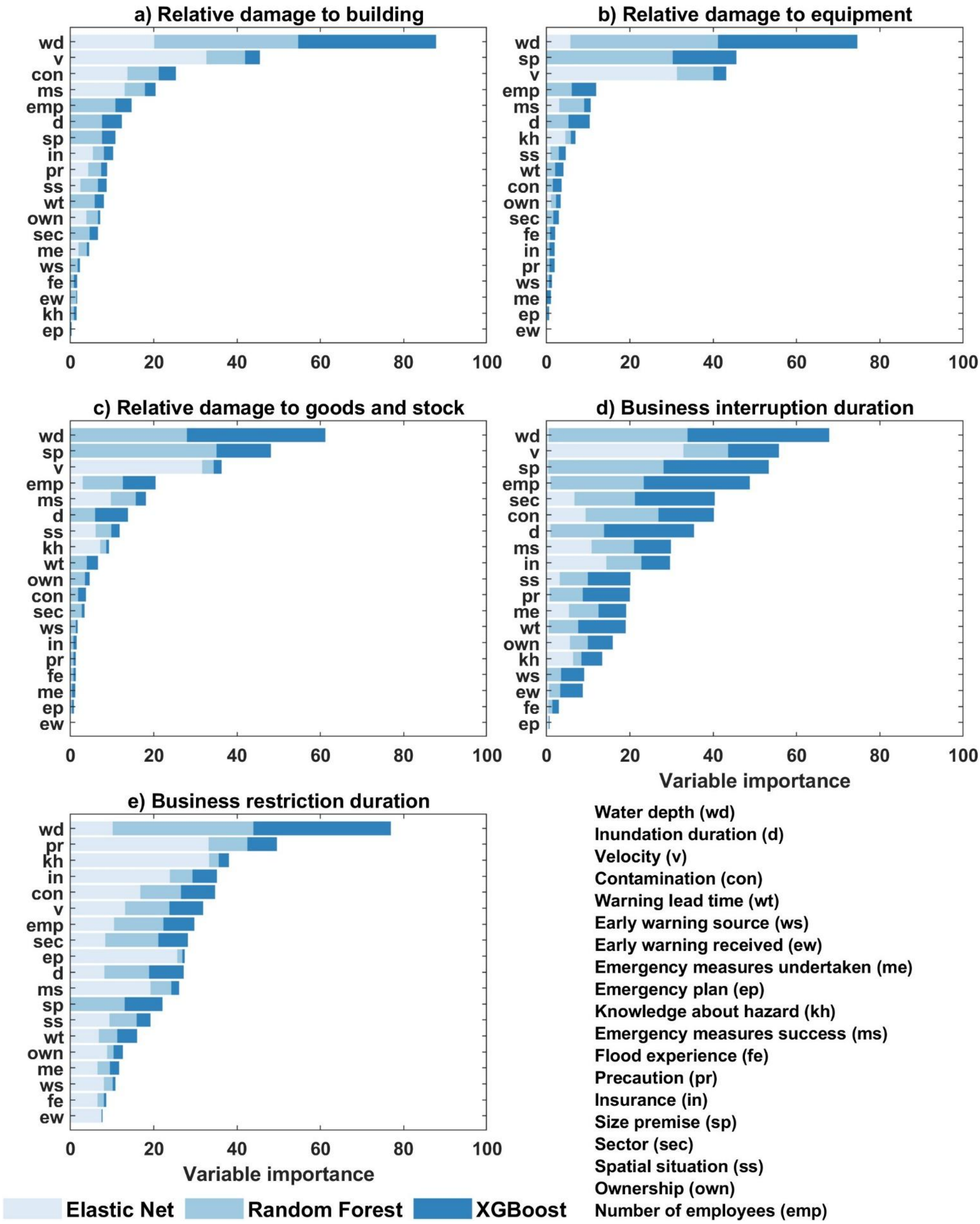

Figure 5: Importance of influencing variables for damage types: (a) buildings, (b) equipment, (c) goods & stock, d) business interruption duration, and e) business restriction duration. The x-axis shows the weighted importance of each variable, obtained from the three models (Random Forest, Elastic Net, and XGBoost).

‘This finding underscores (...) even during unprecedented events like the 2021 flood.’ The analysis was carried out for the unprecedented 2021 flood so the word ‘even’ feels misplaced.

Thank you. We will remove the word ‘even’.

Figure 6: same comment as for Figure 4 and Figure 5. In addition, the resolution of this figure should be higher.

We will revise the figure 6 as follows and all abbreviations are now written out in full form to ensure that it is self-explanatory. We will provide the high-resolution of the figure as a separate file.

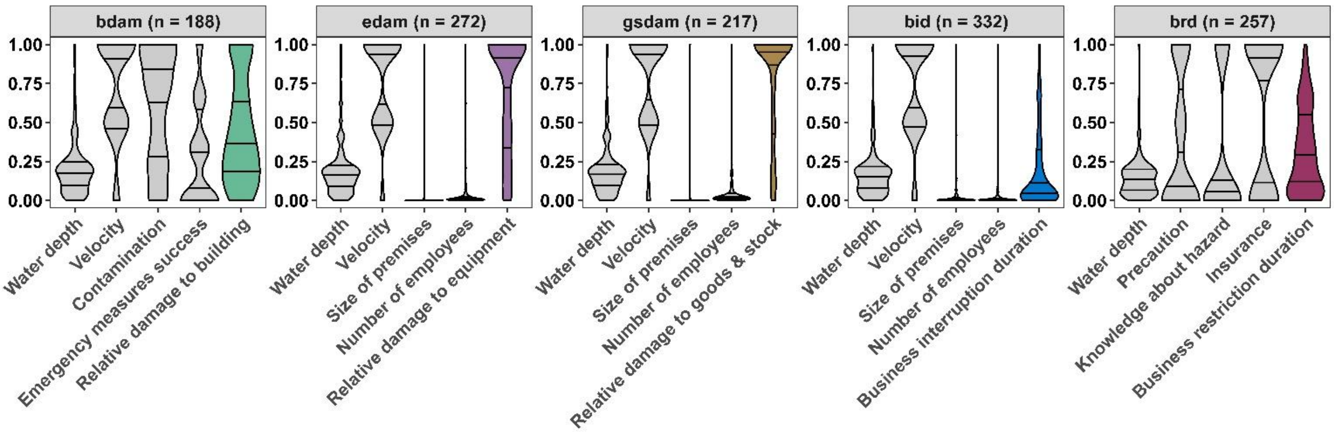

Figure 6: Kernel density estimations of influencing factors and damage types, with all variables scaled between 0 and 1. The lines in the violin plots indicate the quartiles.

The fact that the observed damage and business interruption/restriction durations are scaled from 0 to 1 make interpretation difficult. Saying that the 75th percentile decreases from 0.68 to 0.61 for example is hard to interpret. It would be better to make the results a bit more tangible, this way the results will also appeal more to policymakers and it makes the conclusion easier.

We thank the reviewer for this helpful suggestion. In the revised manuscript, business interruption and restriction durations will be presented in actual days rather than scaled values as follows (P23/L503-518):

“The duration of business interruption varies with velocity and company size. Micro-companies (1–9 employees) show a consistent pattern under low and moderate flow conditions, with a median interruption duration of around 22 days. Under torrential flows, the median interruption duration rises sharply to nearly 60 days. Small companies (10–49 employees) exhibit a similar pattern, though their interruption duration under torrential flow is slightly lower. Medium and large companies (>49 employees) demonstrate greater resilience, with median interruption durations ranging from about 11 to 33 days across all flow conditions. The results indicate that small companies, especially micro-companies, have been disproportionately affected by the 2021 unprecedented flood event.

The analysis of business restriction duration, based on the Markov blanket, further shows that companies without precautionary measures experience the longest restrictions. For instance, the median restriction duration for companies without precautions increases from roughly 102 days when water depth < 1 meter to about 210 days for water depth > 2 meters. Implementing medium precautionary measures (see variable precaution in Table A1) results in a modest reduction in restriction duration, particularly for deeper water, where the 75th percentile decreases from 368 days *to 330 days. A more substantial reduction is observed in companies with strong precautionary measures, where the median restriction durations remain below 150 days. For shallow water depth (< 1 meter), effective precautionary measures reduce the 75^th^ percentile to 178 days, compared to 238 days in companies without precautions. These results highlight the effectiveness of precautionary measures in reducing business restriction durations".*

"In addition, for smaller premises (75–500 m^2) the uncertainty is very less", remove the "very" or replace with "much".

We will revise the sentence as follows (P23/L499):

"In addition, for smaller premises (75–500 m^2) the uncertainty is less".

Please also add a discussion that elaborates on any shortcomings such as low sample size for some company sizes/sectors and outliers, potential selection bias etc. Directions for future research.

We thank the reviewer for this valuable suggestion. We will add the limitations and future scope in the conclusions section, as shown below (P24/L530-539):

"*While the study has deciphered the drivers of company damages during the 2021 flood event, it does have some limitations. First, the sample size for some company categories, particularly large companies, was small, which limits the generalization of findings. Second, survey participation was voluntary, which may have introduced self-selection bias. Although 431 responses are a notable sample size given the challenges of post-disaster data collection, future studies should aim for more diverse representation across different company sizes and sectors. This would further strengthen the generalizability of the findings. Moreover, comparative analyses across multiple extreme flood events in different geographical regions and socio-economic contexts, for instance, in Belgium and the Netherlands in the case of the 2021-event would allow for broader generalization of findings.*"

Conclusion:

The conclusion should be more extensive, this conclusion seems a bit too short and concise for an academic paper. There should be more links with the results section.

We thank the reviewer for the feedback. In the revised manuscript, we will add the limitations and future scope in the conclusions section as mentioned in the previous comment. The revised conclusion reads as follows (P23-24/L520-539):

"The July 2021 flood in Germany highlighted the significant vulnerability of companies to extreme floods, with both direct and indirect damages resulting in substantial financial costs. A central question of this study was whether the influencing factors behind flood damage during the extreme July 2021 event differ from those in earlier floods from 2002 to 2016. Our findings indicate that core hazard related variables, including water depth, flow velocity, and contamination, remain consistent predictors of damage across different events. Similarly, company characteristics such as size of the premises and number of employees continue to play an important role. What sets the 2021 flood apart is the elevated importance of emergency preparedness and behavioral responses, particularly in shaping indirect damages such as business restriction duration, while the sector was not that important. A novel insight from this study is the demonstrated link between knowledge about flood hazard and amount of precaution taken, highlighting its relevance in reducing business restriction duration. Small and micro-companies that implemented very good precaution measures experienced notably shorter restriction durations.

While the study has deciphered the drivers of company damages during the 2021 flood event, it does have some limitations. First, the sample size for some company categories, particularly large companies, was small, which limits the generalization of findings. Second, survey participation was *voluntary, which may have introduced self-selection bias. Although 431 responses are a notable sample size given the challenges of post-disaster data collection, future studies should aim for more diverse representation across different company sizes and sectors. This would further strengthen the generalizability of the findings. Moreover, comparative analyses across multiple extreme flood events in different geographical regions and socio-economic contexts, for instance, in Belgium and the Netherlands in the case of the 2021-event would allow for broader generalization of findings. Finally, future research could explore the interrelations between different types of damages, for example by applying multi-level models, to better understand how direct, and indirect damages interact."*

We hope that the reviewer is satisfied with the changes proposed. Again, we thank for the valuable comments that helped us to improve the manuscript.

References

Askr, H., Abdel-Salam, M., and Hassanien, A. E.: Copula entropy-based golden jackal optimization algorithm for high-dimensional feature selection problems, Expert Systems with Applications, 238, 121582, https://doi.org/10.1016/j.eswa.2023.121582, 2024.

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.

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.

Sucar, L. E.: Probabilistic Graphical Models: Principles and Applications, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-61943-5, 2021.

Thieken, A. H., Bessel, T., Kienzler, S., Kreibich, H., Müller, M., Pisi, S., and Schröter, K.: The flood of June 2013 in Germany: how much do we know about its impacts?, Nat. Hazards Earth Syst. Sci., 16, 1519–1540, https://doi.org/10.5194/nhess-16-1519-2016, 2016.

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.

Reviewer #2:

This manuscript quantifies drivers of damages to companies by rare flood events via 3 data-driven techniques, which ultimately lead to a Bayesian Network. This study could have potential, but its possible novelty is currently hidden behind a rather complicated and untransparent chain of calculations. In particular, the justification of using the 3 data-driven models is unclear. Why not less? Why not more? Why these? This could easily be arbitrary. And what does the Bayesian Network add to the variable importance analysis via those 3 models? I raise more questions below. I believe these should be addressed before the paper can be reconsidered for publication.

The authors would like to thank the reviewer for acknowledging the significance of our study and for providing valuable and constructive feedback. The comments were extremely helpful in improving the quality of the manuscript and will be acknowledged. We provide detailed responses to each comment below, with reviewer comments shown in blue and our responses in black. All references cited in our responses are listed at the end of this letter.

In the revised manuscript, we will introduce the variable selection section with a clear rationale for choosing the three machine learning techniques, as follows (P6/L156-165):

"2.2 Variable Selection

Flood damage processes vary by region, flood type, and asset type (Mohor et al., 2020; Sairam et al., 2019; Wagenaar et al., 2018). Since our analysis focuses on flash floods and covers direct and indirect damages, we use a data-driven approach to identify which variables strongly influence these diverse outcomes. We adopt three feature selection approaches that are robust to multicollinearity and capable of capturing nonlinear relationships and interactions. To this end, we employ three complementary machine learning techniques: Elastic Net (EN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). EN efficiently handles multicollinearity and performs variable selection through regularization; RF captures nonlinear relationships and complex interactions via ensemble decision trees; and XGBoost, a gradient boosting algorithm, provides high predictive accuracy and models intricate dependencies. By combining the strengths of these methods, we assume to ensure a comprehensive assessment of variable importance. To mitigate potential biases from relying on a single model, we aggregate the variable importance scores across all three methods to derive a final ranking.''

In this study, BNs are employed to complement the machine learning models by providing a probabilistic framework for analyzing multivariate dependencies and scenario-based inference. Whereas EN, RF, and XGBoost primarily emphasize predictive accuracy and variable ranking, BNs explicitly capture conditional dependencies among variables. We will revise the section to better explain this motivation as follows (P9/L226-233):

"Bayesian networks (BNs) are probabilistic graphical models that represent dependencies among multiple variables and enable multivariate predictive density estimation (Sucar, 2021). In this study, BNs are employed to complement the machine learning models by providing a probabilistic framework for analyzing multivariate dependencies and scenario-based inference. Whereas EN, RF, and XGBoost primarily emphasize predictive accuracy and variable ranking, BNs explicitly capture conditional dependencies among the variables. This is particularly valuable in flood damage analysis, where damage outcomes result from complex interactions between hazard intensity, company characteristics, and preparedness measures. Moreover, BNs are able to estimate posterior probabilities of damages given partial evidence (e.g., observed water depth or company preparedness), thereby offering a transparent and interpretable tool for risk assessment under uncertainty".

Title: I suggest a different word than "deciphering" because that's not what is being done in this study.

We thank the reviewer for raising this point. Our analysis is a process of uncovering the drivers of direct and indirect damages to companies, which involve complex interrelationships, and extracting insights that are not immediately apparent from the raw data. This is directly reflected in our methodology and findings:

1. Multivariate modeling (Figure 5): Variable importance scores capture the combined effects of hazard, preparedness, and company characteristics on multiple damage types, explicitly accounting for interdependencies rather than isolated relationships.
2. Bayesian network analysis (Figures 7): The BN models conditional probabilities, showing how hazard intensity, company size, and precautionary measures collectively shape damage outcomes.
3. Multivariate probabilistic estimation (Figure 8): Probabilistic outcomes highlight how small and micro-companies are disproportionately affected under extreme conditions, reflecting complex interactions between hazard severity and company characteristics.

We feel that, in this context, "deciphering" is an appropriate term, as it conveys the analytical effort required to reveal and quantify underlying mechanisms that are not directly observable. We believe it best captures both the analytical depth and the objectives of our study. Therefore, we would like to keep the title as is.

L44-52: The message needs to be streamlined here with regard to rare/high-impact events.

We will revise the explanation as follows (P2/L45-53):

"*The severity of indirect damages can be equally significant and, in the case of rare and high-impact flood events, may even exceed direct damages (Koks et al., 2015; Pfurtscheller and Vetter, 2015; Sieg et al., 2019). For instance, Pfurtscheller and Vetter (2015) reported that indirect damages are often underestimated by companies, despite sometimes exceeding direct damages during rare flood events. Using an Input-Output (IO) model, Li et al. (2018) showed that business interruptions and operational restrictions in Shanghai's manufacturing firms can propagate along interlinked value chains, with indirect damages under extreme storm flood scenarios reaching up to $60 billion. Similarly, Sieg et al. (2019) employed a supply-side IO model and identified the manufacturing, and financial sectors vulnerable to indirect damages. It should be noted, however, that not all studies classify business interruptions or operational restrictions as indirect damages. The definition of indirect damage varies across the literature. In this study, we specifically focus on business interruptions and restrictions as a key component of indirect flood damages. Altogether, these studies underscore that indirect damages, especially during low-probability, high-impact flood events, can be substantial and warrant systematic investigation to better understand the processes*".

L107, 109, 119 and elsewhere: Consider something like "rare" in place of "unprecedented", because there now is a precedent.

We thank the reviewer for this helpful comment. In lines L106 and L108, we will replace the word "unprecedented" with "rare" as suggested. In L118 we will revise the text as follows:

"The July 2021 flood in Germany has been widely described as extraordinary in terms of its hydrological magnitude, spatial extent, exceeding the scale and severity of previously recorded floods in the affected regions (Mohr et al., 2023; Thieken et al., 2023; Zander et al., 2023) *and it caused an estimated €33.1 billion in direct damages and €7.1 billion in indirect damages (Trenczek et al., 2022)".*

L141, L214: The analyses for each damage type could have been combined, as they are also internally related, via a multivariate regression. Why employ this more elegant solution making optimal use of all information (by not considering the responses as independent)?

We thank the reviewer for this question. Although there is to some extent interdependency across the addressed damage types, we analyzed them individually for two main reasons. First, it allowed us to capture asset-specific processes and identify distinct drivers for each category (e.g., buildings, equipment, goods & stock, business interruption), which can behave differently in a rare flood event and can also vary across company sizes and sectors. Second, the dataset had varying levels of completeness across damage types. By analyzing them separately, we were able to make use of larger subsamples, rather than restricting the analysis to the smaller set of companies with complete data across all damage types.

We will add the following lines in the revised manuscript (P6/L141-146):

"We analyzed them individually for two main reasons. First, this approach allowed us to capture loss-specific processes and identify distinct drivers for each category (e.g., buildings, equipment, goods & stock, business interruption), which can behave very differently during a rare flood event and can also vary across company sizes and sectors. Second, the dataset had varying levels of completeness across damage types: some companies reported only building damages, while others provided data on equipment or business interruption. By analyzing them separately, we were able to make use of larger and more reliable subsamples, rather than restricting the analysis to the smaller set of companies with complete data across all damage types".

L143: Across what scale where the missing data imputed, i.e. how far were they apart on average.

We will clarify this in the revised manuscript. We used the Gower distance to calculate similarity between observations, which is ideal for a dataset with different types of variables (continuous, nominal, and ordinal) (Kowarik and Templ, 2016). For rows with missing data, the average distance to their 5 nearest neighbors was approximately 0.09, indicating that imputation was performed among relatively similar observations.

The following lines will be added in the revised manuscript (P6/L149-153):

"We used the Gower distance to calculate similarity between observations, which is ideal for a dataset with different types of variables, i.e. continuous, ordinal and nominal (Kowarik and Templ, 2016). We calculated the average Gower distance between each row with missing data and its 5 nearest neighbors. The mean of these distances across all rows with missing values was approximately 0.09, indicating that imputation was performed among observations that were relatively similar in terms of their characteristics."

L155: J(beta) is not in the equation.

P7/L174: $J(\beta)$ will be replaced with $Obj(\beta)$ to match the notation used in Equation 1.

L157: What does use of the MAE as objective function imply about the nature of the residuals given a response which is between 0 and 1 or counts between 0 and 540?

We thank the reviewer for this insightful question. The use of MAE as the objective function implies that residuals are treated symmetrically, without giving extra weight to large deviations. This is particularly suitable for our flood damage data, where responses range from 0 to 1 or 0 to 540 days, ensuring that both small and large errors are proportionally considered.

Our choice of MAE was based on two main considerations:

- MAE is robust to outliers, which is particularly important as flood damage data often contain extreme values. Unlike Mean Squared Error (MSE), which disproportionately penalizes large errors, MAE treats all deviations proportionally, providing a more stable and representative measure of overall model performance.
- Second, for the Permutation Variable Importance (PVI) analysis, MAE provides a direct and interpretable measure of error. The performance loss is in the same units as our response variables—relative loss (0 to 1) and duration (0 to 540 days). This allows for a clear, tangible assessment of how each variable's permutation affects predictive accuracy. While other metrics are appropriate for specific distributional assumptions, MAE is model-agnostic and ideal for generating comparable PVI scores across our different modelling approaches.

We will include the following text in the revised manuscript (P7/L176-179):

"*The use of MAE as the objective function treats residuals symmetrically, ensuring that both small and large errors are proportionally considered. This metric is robust to outliers and provides an interpretable measure of error in the same units as the response variables i.e., relative loss (0–1) and duration (0–540 days)".*

L159: It's not entirely true that the model cannot handle nonlinearities – it can do so via transformations or in Generalised Linear Model form.

We agree that Elastic Net can, in principle, handle nonlinear relationships if appropriate transformations of the predictors are included. However, in our study, we applied Elastic Net in its standard linear form without additional transformations, and therefore it primarily captures linear associations between predictors and the response. Nonlinear effects were instead captured by the Random Forest and XGBoost models, which can model complex nonlinear relationships directly.

We will rephrase the sentence in the revised manuscript as follows (P7/L180-182):

"*Elastic Net is a powerful linear model that is effective in handling multicollinearity. However, in its standard application without explicit transformations as used in this study, it primarily captures linear associations and cannot model complex nonlinear relationships directly*".

L201: What are the implications of combining the variable importance across the 3 models?

By combining the variable importance scores across Elastic Net, Random Forest, and XGBoost, we are combining the complementary strengths of each model. This approach mitigates potential biases that could arise from relying on a single model and provides a more robust and comprehensive assessment of the variable importance. The aggregated ranking reflects variables that consistently influence predictions across multiple modelling frameworks, offering greater confidence in identifying key drivers of flood damages.

We will include the following text in the revised manuscript (P6/L163-165):

"*By combining the strengths of these methods, we ensure a comprehensive assessment of variable importance. To mitigate potential biases from relying on a single model, we aggregate the variable importance scores across all three methods to derive a final ranking*."

Eq9, L219, Appendix: It's conditional probabilities, not fractions in Bayes Rule! I.e. X_i|E and E|X_i.

The notation will be updated throughout the manuscript and in the Supplementary information.

L222: Why not leave it discrete rather than introducing another layer of assumptions?

Thank you for this constructive feedback. We will revise the manuscript to clarify the rationale for using a continuous distribution as follows (P10/L250-254)):

"The posterior probability of flood damage given the observed evidence E is discrete in nature. However, this discrete representation is limited by the binning of the data and does not allow precise estimates or a meaningful characterization of predictive uncertainty. To address this, we derived a continuous distribution of direct and indirect damages by fitting a probability distribution based on weighted sampling of the empirical damage data, following the approach of Schoppa et al., (2020). This allows for a more precise representation of uncertainty and predictions at finer scales beyond the original bins"

The following figure was developed to illustrate this process and is provided for reference, but will not be included in the manuscript:

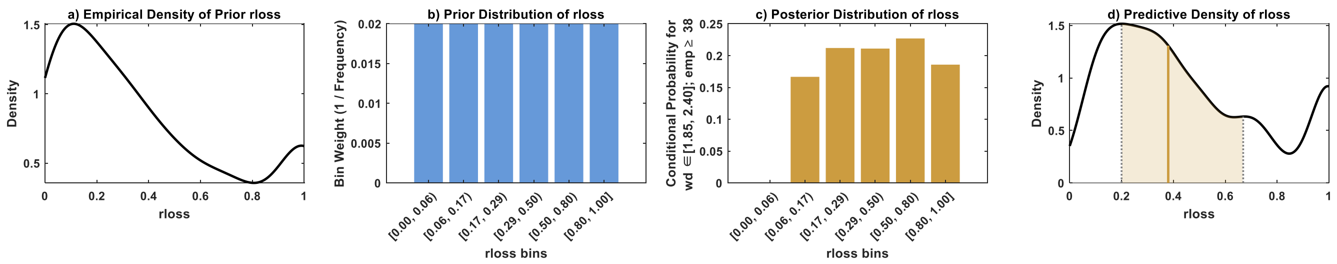

Figure: Visualizations of the prior, posterior, and predictive distributions of rloss (a) Empirical kernel density estimate of the prior rloss based on collected data (b) Prior distribution of rloss represented as bin weights (inverse frequency) across discretized intervals (c) Posterior distribution of rloss conditioned on wd ∈ [1.85, 2.40) and emp ≥ 38 (d) Predictive distribution of rloss generated by resampling 1000 values using the prior bin weights and the posterior probabilities. The solid vertical line indicates the median (50th percentile), while the dotted vertical lines represent the 25th and 75th percentiles, representing the predictive uncertainty. The shaded area highlights the interquartile range.

L228: How do the five models relate to the Bayesian Network?

Based on the Line indicate (L228), we interpret that the reviewer meant 3 models (EN, RF and XGB) and not 5.

In this study, BNs are employed to complement the machine learning models by providing a probabilistic framework for analyzing multivariate dependencies and scenario-based inference. Whereas EN, RF, and XGBoost primarily emphasize predictive accuracy and variable ranking, BNs explicitly capture conditional dependencies among variables. We will revise the section to better explain this motivation as follows (P9/L226-233):

"Bayesian networks (BNs) are probabilistic graphical models that represent dependencies among multiple variables and enable multivariate predictive density estimation (Sucar, 2021). In this study, BNs are employed to complement the machine learning models by providing a probabilistic framework for analyzing multivariate dependencies and scenario-based inference. Whereas EN, RF, and XGBoost primarily emphasize predictive accuracy and variable ranking, BNs explicitly capture conditional dependencies among the variables. This is particularly valuable in flood damage analysis, where damage outcomes result from complex interactions between hazard intensity, company characteristics, and preparedness measures. Moreover, BNs are able to estimate posterior probabilities of damages given partial evidence (e.g., observed water depth or company

preparedness), thereby offering a transparent and interpretable tool for risk assessment under uncertainty".

L230-242: This part is redundant – see above. The function of the 3 models, despite factor selection is unclear. And why 3 models and not more or less?

In the revised manuscript, we will introduce the variable selection section with a clear rationale for choosing the three machine learning techniques, as follows (P6/L156-165):

"2.2 Variable Selection

Flood damage processes vary by region, flood type, and asset type (Mohor et al., 2020; Sairam et al., 2019; Wagenaar et al., 2018). Since our analysis focuses on flash floods and covers direct and indirect damages, we use a data-driven approach to identify which variables strongly influence these diverse outcomes. We adopt three feature selection approaches that are robust to multicollinearity and capable of capturing nonlinear relationships and interactions. To this end, we employ three complementary machine learning techniques: Elastic Net (EN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). EN efficiently handles multicollinearity and performs variable selection through regularization; RF captures nonlinear relationships and complex interactions via ensemble decision trees; and XGBoost, a gradient boosting algorithm, provides high predictive accuracy and models intricate dependencies. By combining the strengths of these methods, we assume to ensure a comprehensive assessment of variable importance. To mitigate potential biases from relying on a single model, we aggregate the variable importance scores across all three methods to derive a final ranking.''

Results & discussion: Too much time is spent describing univariate results. And the bivariate correlations defeat the purpose of multivariate analysis.

We thank the reviewer for this comment. While some univariate correlations are described in the manuscript, these serve primarily to provide context on the dataset and highlight potential relationships between influencing factors and damage types. The main focus of our study is on multivariate relationships, which are addressed extensively through:

1) Multivariate model variable importance: Figure 5 and the corresponding discussion present variable importance scores derived from multivariate models, capturing the combined effects of hazard, preparedness, and company characteristics on all five types of flood damage. This approach identifies key drivers while accounting for interdependencies among variables.
2) Bayesian network analysis: Figures 7 and 8, along with the associated discussion, illustrate the probabilistic dependencies among the top influencing factors and damage types. The BN explicitly models conditional relationships in a multivariate framework, highlighting interactions between hazard variables (e.g., water depth, flow velocity), company-specific factors (e.g., size, employees), and preparedness measures (e.g., precaution, emergency measures).
3) Multivariate probabilistic damage estimation: Figure 8 shows the distributions of damage outcomes under different hazard, exposure, and preparedness scenarios, estimated using the Markov blankets of the Bayesian networks. These results reflect the combined influence of multiple variables and quantify uncertainty, demonstrating the practical implications of multivariate analysis.

The discussion explicitly addresses how hazard intensity, company characteristics, and precautionary measures jointly influence direct and indirect damages (e.g., business interruption and restriction durations). For instance, small and micro-companies are disproportionately affected under extreme flood conditions, highlighting the interplay of hazard severity and company size. Precautionary measures substantially mitigate business restriction durations, particularly under higher water depths, illustrating the multivariate impact of preparedness and hazard exposure.

In the revised manuscript, we will rewrite the results in a more tangible way to emphasize these multivariate findings as follows (P23/L503-518):

"The duration of business interruption varies with velocity and company size. Micro-companies (1–9 employees) show a consistent pattern under low and moderate flow conditions, with median interruption durations of around 22 days. Under torrential flow conditions, the interruption duration rises sharply to nearly 60 days. Small companies (10–49 employees) exhibit a similar trend, though their interruption duration under torrential flow is slightly lower. Medium and large companies (>49 employees) demonstrate greater resilience, with median interruption durations ranging from about 11 to 33 days across all flow conditions. The results indicate that small companies, especially micro-companies, are disproportionately affected by the 2021 unprecedented flood event.

The analysis of business restriction duration, based on the Markov blanket, further shows that companies without precautionary measures experience the longest restrictions. For instance, the median restriction duration for companies without precautions increases from roughly 102 days when water depth < 1 meter to about 210 days for water depth > 2 meters. Implementing medium precautionary measures (see variable precaution in Table A1) results in a modest reduction in restriction duration, particularly for deeper water, where the 75th percentile decreases from 368 days to 330 days. A more substantial reduction is observed in companies with strong precautionary measures, where the median restriction durations remain below 150 days. For shallow water depth (< 1 meter), effective precautionary measures reduce the 75th percentile to 178 days, compared to 238 days in companies without precautions. These results highlight the effectiveness of precautionary measures in reducing business restriction durations".

L394: Purpose of sentence unclear.

The sentence was intended to justify the use of a multivariate approach. We will revise the sentence as follows (P16/L396-397):

"*Furthermore, significant correlations exist between several influencing factors, underscoring the importance of a multivariate modelling approach*".

L373: Who's expert knowledge?

We thank the reviewer for the comment. The original mention of "expert knowledge" referred to the combined input of the study authors, who have domain expertise in flood risk assessment and damage modeling. To clarify and simplify, we will revise the sentence as follows (P16/383-384):

"*Based on data availability, 19 potentially relevant influencing factors were selected, covering hazard characteristics, emergency measures, precautionary actions, and company characteristics (Table 1)."*

Fig6: What is it's function for the manuscript?

As described in the manuscript (P19/L426–430), Figure 6 presents the kernel density estimations of the top four influencing factors considered for multivariate probabilistic damage modelling across five different types of damage. The violin plots illustrate the probability density of scaled variables (ranging from 0 to 1), with quartile lines indicating central tendencies and variability. The presence of skewed distributions and multimodal characteristics highlights the complexity of flood damage relationships across different damage types.

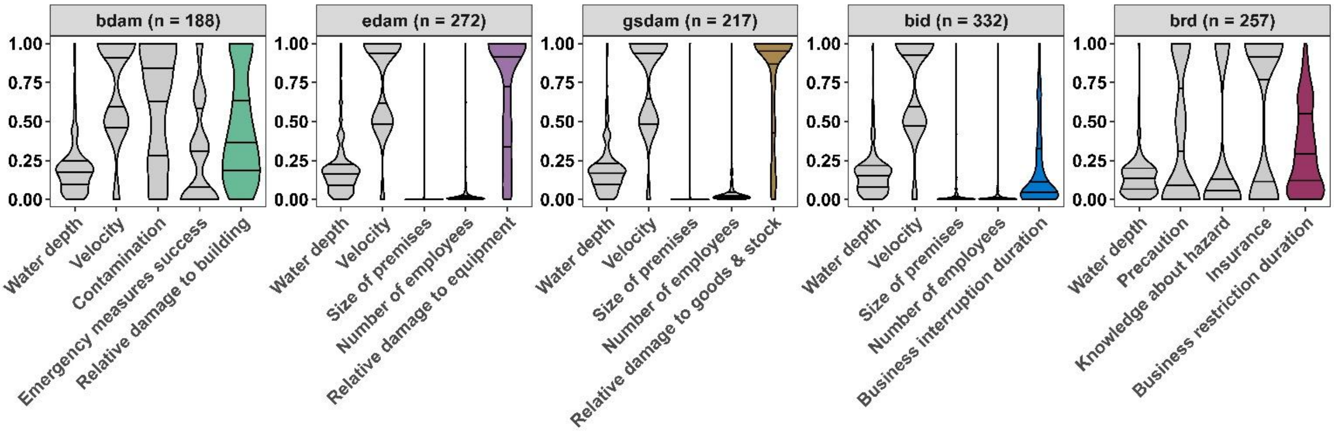

Figure 6: Kernel density estimations of influencing factors and damage types, with all variables scaled between 0 and 1. The lines in the violin plots indicate the quartiles.

Fig7: The directions matter here, no? And some of them are not intuitive!

We thank the reviewer for this question. In the Bayesian network, arrow directions indicate conditional dependencies between variables but do not imply causality. Some directions may seem unintuitive because the structure is derived from a score-based learning algorithm that optimizes overall network fit to the data, rather than reflecting causality.

We will clarify this in the revised manuscript as follows (P19/L437-440):

"The direction of the arrows represents conditional dependencies between variables but does not imply causality (Schröter et al., 2014). Some directions may appear unintuitive because the structure is derived from a score-based learning algorithm that optimizes the overall network fit to the data, not necessarily causality."

Conclusion: Too short and doesn't add sufficient novelty.

In the revised manuscript, we will emphasize the novelty, limitations and future scope in the conclusions section as follows (P23-24/L520-539):

"The July 2021 flood in Germany highlighted the significant vulnerability of companies to extreme floods, with both direct and indirect damages resulting in substantial financial costs. A central question of this study was whether the influencing factors behind flood damage during the extreme July 2021 event differ from those in earlier floods from 2002 to 2016. Our findings indicate that core hazard related variables, including water depth, flow velocity, and contamination, remain consistent predictors of damage across different events. Similarly, company characteristics such as size of the premises and number of employees continue to play an important role. What sets the 2021 flood apart is the elevated importance of emergency preparedness and behavioral responses, particularly in shaping indirect damages such as business restriction duration, while the sector was not that important. A novel insight from this study is the demonstrated link between knowledge about flood hazard and amount of precaution taken, highlighting its relevance in reducing business restriction duration. Small and micro-companies that implemented very good precaution measures experienced notably shorter restriction durations.

While the study has deciphered the drivers of company damages during the 2021 flood event, it does have some limitations. First, the sample size for some company categories, particularly large companies, was small, which limits the generalization of findings. Second, survey participation was voluntary, which may have introduced self-selection bias. Although 431 responses are a notable sample size given the challenges of post-disaster data collection, future studies should aim for more diverse representation across different company sizes and sectors. This would further strengthen the *generalizability of the findings. Moreover, comparative analyses across multiple extreme flood events in different geographical regions and socio-economic contexts, for instance, in Belgium and the Netherlands in the case of the 2021-event would allow for broader generalization of findings. Finally, future research could explore the interrelations between different types of damages, for example by applying multi-level models, to better understand how direct, and indirect damages interact. Overall, the results underscore the critical role of emergency preparedness and risk communication during extreme events, serving as essential complements to structural protection measures that may be less effective under extraordinary conditions."*

We hope that the reviewer is satisfied with the changes proposed. Again, we thank for the valuable comments that helped us to improve the manuscript.

References:

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

Kowarik, A. and Templ, M.: Imputation with the R Package VIM, J. Stat. Soft., 74, https://doi.org/10.18637/jss.v074.i07, 2016.

Li, W., Wen, J., Xu, B., Li, X., and Du, S.: Integrated Assessment of Economic Losses in Manufacturing Industry in Shanghai Metropolitan Area Under an Extreme Storm Flood Scenario, Sustainability, 11, 126, https://doi.org/10.3390/su11010126, 2018.

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

Pfurtscheller, C. and Vetter, M.: Assessing entrepreneurial and regional-economic flood impacts on a globalized production facility, J Flood Risk Management, 8, 329–342, https://doi.org/10.1111/jfr3.12102, 2015.

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.

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.

Sieg, T., Schinko, T., Vogel, K., Mechler, R., Merz, B., and Kreibich, H.: Integrated assessment of short-term direct and indirect economic flood impacts including uncertainty quantification, PLoS ONE, 14, e0212932, https://doi.org/10.1371/journal.pone.0212932, 2019.

Sucar, L. E.: Probabilistic Graphical Models: Principles and Applications, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-61943-5, 2021.

Thieken, A. H., Bubeck, P., Heidenreich, A., Von Keyserlingk, J., Dillenardt, L., and Otto, A.: Performance of the flood warning system in Germany in July 2021 – insights from affected residents, Nat. Hazards Earth Syst. Sci., 23, 973–990, https://doi.org/10.5194/nhess-23-973-2023, 2023.

Trenczek, J., Lühr, O., Lukas, E., and Viktoria, L.: Damage from flash floods and floods in July 2021 in Germany. An ex-post analysis. Project report, Prognos AG, Düsseldorf, 2022.

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.