Journal: NHESS
Title: **FLEMOflash - Flood Loss Estimation MOdels for companies and households affected by flash floods**
Author(s): Singh et al.
MS No.: egusphere2025-1512
MS Type: Research Article
Iteration: First review

The paper introduces **FLEMOflash**, a novel multivariate probabilistic Flood Loss Estimation Model tailored for flash floods. The model builds on survey data collected after flash flood events in 2002, 2016, and 2021 in Germany, encompassing both affected companies and households. FLEMOflash employs a data-driven feature selection approach alongside Bayesian networks to derive probabilistic loss estimates.
The topic clearly falls within the scope of the journal, and the manuscript is generally well written and well organised. However, I have concerns regarding some of the underlying assumptions of the model, which, in turn, raise doubts about its validity for reliably estimating flash flood damage. I believe the authors should provide a more robust justification for their hypotheses to strengthen the credibility and robustness of their results.
Below, I first present general concerns, followed by more specific comments.

General concerns

1) My first concern relates to the criteria used for identifying flash flood events (and then data to be implemented to derive the model). Specifically, I find the use of average slope as a proxy problematic. While slope may influence local flow velocity, it does not adequately capture the main defining characteristic of flash floods — their rapid onset and short lead times. This concern is further supported by the reported warning lead times in Tables 1 and 2, which range from 0 to 240 hours and 0 to 168 hours, respectively. These values appear inconsistent with typical flash flood dynamics, where lead times are often just a few hours.
Additionally, the use of a low-resolution DEM may not provide the accuracy needed to derive reliable slope estimates at the point observation scale.
Have the authors considered using the *concentration time* of the river basin where the observations are located as a more physically meaningful proxy for flash flood potential? This could provide a better indication of response time and be more consistent with established hydrological understanding of flash flood processes.

2) My second concern relates to the set of explanatory variables used in the model. One of the primary damage mechanisms in flash flood events is structural damage, which is strongly influenced by the physical vulnerability of affected buildings. However, the model does not appear to include variables that capture this aspect, such as construction material, number of floors, or level of maintenance — all of which significantly affect a building's susceptibility to structural damage.
While I understand that the set of variables was likely constrained by the information collected through the survey, I would like to know whether the authors considered integrating ancillary data to address these critical gaps. For example, building-level data from national censuses or geoportals could provide valuable proxies for physical vulnerability.
Inclusion of such information could improve the explanatory power and practical relevance of the model, particularly in contexts where decisions rely on nuanced understanding of asset-specific vulnerabilities.

3) A third concern regards obtained results, especially in terms of damage mechanisms. I would have expected to observe a significant influence of flow velocity or, at least, of the hydrodynamic force associated with the flow but this is not the case.

4) All the concerns mentioned above converge in the results obtained, particularly in the relative loss estimates provided by the model. These estimates range between 0.2 and 0.5, even for high water depths (around or above 2 meters). Such values are comparable to those typically produced by models for riverine

floods (see, e.g., FLEMO ps), which raises doubts about the model's ability to capture the distinctively more destructive nature of flash floods.

Specific comments

Table 1 and Table 2 → the meaning of some variables is not clear. For instance, does emergency plan refer to the existence of a municipal emergency plan or a company emergency plan? Which is the meaning of the precaution indicator? Which are the emergency measures considered? I suggest including an explanatory table in the supplementary material

Line 102-104
"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" → while this could be a good option for spatially correlated variables such as velocity and warning lead time (after verifying that the distance between points is limited), it may lead to misleading assumptions for other missing variables. For example, variables such as *in*, *sp*, and *sec* (for companies) or *ke*, *fa*, and *b* (for buildings) are not necessarily spatially correlated. It would be helpful if the authors could provide a more thorough discussion on this point, particularly addressing the potential limitations and implications of their imputation strategy for these types of variables.

Section 3.1
The meaning of two CPTs in the table (d-con, wd-hs) should also be discussed. Moreover, I think this section should be expanded discussing results for all damage components (i.e. companies BUI, EQU, GNS and household CON), even without reporting all the CPTs.

Figure 5 → I think that results explanation will be supported if each CPT is identified with a letter

Line 289- 296
"The integration of knowledge about emergency action into the FLEMOflash model alongside water depth and contamination provides a comprehensive understanding of how preparedness can mitigate loss during flash floods. Knowledge about 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" → Knowing what to do does not necessarily imply that individuals will take action. Do the authors have any insight into why this variable appears to be significant in the model, potentially even more so than the actual implementation of protective measures (me, mu)?

Minor comments

Line 58-59
The conventional multivariate flood loss estimation models often employ decision tree-based approaches to assess the role of different variables in influencing flood loss → Multivariate synthetic models also exist

Line 72
The objective of this study is to build a novel Flood Loss Estimation MOdel affected by flash floods (FLEMO*flash*) → check grammar

Line 254
The FLEMO*flash* model with the best performance, identified in Fig 3 → Which one is it? i.e., To which combinations of predictors, bins and neighbours correspond?

Line 256
C-GUI → Do authors mean C-BUI?

Line 256-257
"For households (P:BUI and P:CON), the losses are significantly underestimated by the SDF-P" → I cannot appreciate that

Line 276 -278
"The CPT suggests that low water depths 275 (< 0.28 m) are most likely associated with low loss (< 0.05), while high water depths (> 0.15m) with high loss (> 0.24)" → I would replace 0.05 with 0.17 and 0.15 with 1.5

Line 284-288
"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)" → it seems numbers are incorrect, please check or explain better

Line 292 → which is Figure 5e? see comment above