

Reviewer #1

R1.C1: *This paper provides a detailed description of the development of an expert based flood content damage model called INSYDE. It seems to be a follow up paper on the structure damage version of INSYDE, a model that seems to have found quite widespread use in the literature. The paper is well written and describes the development process well. The methods are solid but not very innovative and have been around in the grey literature for a long time (e.g. US Army Corps of Engineers). This paper goes in quite some detail describing the methods and adds much needed validation and is therefore definitely a useful addition to the scientific literature. That being said I have concerns about the validation results and more importantly the analysis of the results.*

Reply: We thank the Reviewer for the positive assessment of the manuscript. While we acknowledge the existence of various kind of damage models in the literature, we want to underline that INSYDE-content explicitly applies the synthetic approach to the context of content damage in residential buildings, a domain in which models are scarce. The proposed model was indeed designed following a component-wise, probabilistic and “what-if” logic, which builds on and extends the original INSYDE framework (Dottori et al., 2016). Its transparent and modular structure allows the users to clearly understand all assumptions and input variables, facilitating local adaptation based on available knowledge and data. In this respect, INSYDE-content overcomes limitations of both empirical models (typically fitted to specific case studies) and other synthetic models, which often function as black boxes without allowing modifications or explicit access to internal assumptions. While this aspect was briefly introduced in the original manuscript (L80-83), we have clarified and strengthened this point in the revised version of the introduction and conclusions.

R1.C2: *Figure 4 shows that for detached and semi-detached houses the variation in observed damages is much larger than the variation in predicted damages. My first impression is that the model always roughly predicts the same damage regardless of the circumstances (the blue dots are a nearly horizontal line). I think it may not be so bad because the log-log scale masks some of the variation. However, more information is required so readers can actually tell the model performance. For example, I currently cannot see if the variation in observed values is just based on some large outliers or whether there is some more fundamental problem whereby the observed losses have much more variation than the modelled losses. Also is there even any correlation between modelled and observed losses?*

Reply: We thank the Reviewer for the valuable comment. As also discussed in Molinari et al. (2020), claim data at the building scale are often affected by significant uncertainty and potential bias. For this reason, traditional flood damage models, as well as INSYDE-content, are typically more reliable when applied at aggregated spatial scales, rather than at the level of individual buildings.

This intrinsic variability in the observed data makes the interpretation of building-scale validation results particularly challenging. In our original manuscript (L368-374), we emphasized the importance of providing uncertainty ranges in the predicted damage values as a way to enhance the informative content of the model compared to purely deterministic approaches.

Nonetheless, we acknowledge the Reviewer’s request for a clearer assessment of model performance. In the revised version of the manuscript, we have included quantitative error metrics (mean absolute error and mean bias, calculated based on the median of the predicted values per building, compared to the observed losses) to complement the visual inspection of the plots. However, we have also added a caveat to caution against over-interpreting these metrics, since observed values should not be considered absolute ground truth in damage modeling, due to the inherent limitations in claim data quality.

As also noted by the Reviewer, the apparent flatness of the predicted damage values for detached and semi-detached houses is partially explained by the use of log-log axes, which compress the visual

perception of variability. Furthermore, from a theoretical perspective, a limited dispersion in predicted losses is to be expected for these building types, given their relatively homogeneous characteristics and the shallow inundation depths recorded during the two flood events. In contrast, greater variability is observed for multi-family residential buildings, which generally exhibit broader heterogeneity in both exposure and vulnerability, leading to capturing a damage prediction variability of the same order of that for observed losses. In the revised version of the manuscript, we have therefore included comments on these aspects regarding the interpretation of the results shown in Figure 4 (Figure 5 in the revised version).

R1.C3: *I understand that there is unexplained uncertainty in the model predictions as indicated by the uncertainty ranges in figure 4. However, if the model typically predicts more or less the same mean how do I know such a complicated model adds any value compared to a simple mean value as prediction?*

Reply: As also noted in our response to R1.C2, the apparent lack of variation in the predicted values is partly an effect of the log-log scale used in the plots, which visually compresses differences in magnitude. Moreover, for the two case studies presented, the variability in observed inundation depth was relatively limited, leading to a correspondingly limited spread in the predicted central estimates. However, we believe that the added value of our approach lies in several aspects beyond the mean prediction. First, INSYDE-content is a probabilistic, component-based model that explicitly propagates uncertainty from input variables to output damage estimates. The resulting prediction intervals offer critical information to decision-makers, allowing them to understand the sensitivity of outcomes to input assumptions and to assess risk under uncertainty, an aspect that cannot be captured by a simple mean-based model.

Second, beyond estimating damage, the model provides a transparent and flexible structure to represent damage mechanisms explicitly. This feature allows, for instance, for scenario testing and the evaluation of building-scale mitigation strategies, which would alter specific input parameters and therefore result in different damage outcomes. This functionality makes INSYDE-content particularly suitable not only for risk estimation but also for supporting risk reduction planning.

In the revised version of the manuscript, we have better highlighted these aspects in the discussion and conclusions, including a clearer explanation of how the probabilistic design and modular structure of the model enhance its applicability and usefulness beyond average damage prediction.

R1.C4: *Also very common error metrics are missing such as Mean Absolute Error, correlation coefficient or R2, so it's nearly impossible to assess how the model is doing from the information presented in the paper. Not all these metrics are needed but at least more information. Table 4 only gives an aggregated comparison, so basically gives a bias value. In one region there seems to be some bias but the authors do not really explain where this bias might be coming from. Lastly, I would expect an in depth analysis and discussion of the model performance in the paper based on the validation. That analysis is missing, making the validation not very useful in its current form.*

Reply: As discussed in our response to R1.C2, our original decision not to include commonly used error metrics was driven by the recognition that claim data at the building scale are themselves often affected by considerable uncertainty and potential bias. Consequently, they cannot be considered a definitive benchmark for validation purposes in the traditional sense. For flood damage models, quantitative validation against building-level claims must be indeed interpreted with caution, as it often reflects discrepancies not only in model performance but also in data quality and reporting practices (Molinari et al., 2020; Di Bacco et al., 2024). That said, we understand the Reviewer's request for a more detailed assessment. In the revised version of the manuscript, we have therefore included a set of standard error metrics, computed using the median predicted value per building, to provide a clearer quantitative basis for comparison with observed claims, and we have added a broader reflection in the discussion section on the limitations and interpretability of validation results in the context of flood damage modeling.

R1.C5: *Some of the input variables for the model validation seem sampled whereas others seem observed and the current text is very unclear about what is sampled and what is observed. This makes it even more difficult to interpreted the validation results.*

Reply: We agree that a clearer distinction between observed and sampled input variables can be useful for better understanding the setup of the validation exercises and interpreting the results. To address this, in the revised version of the manuscript we have included a summary table (new Table 4) listing the model input variables used in the two validation case studies and making the corresponding text more explicit.

R1.C6: *The word “to” in the title doesn’t read well, maybe you can replace it with “for”? Or another solution.*

Reply: In the revised version of the manuscript, we have updated the title to improve readability. The new title now reads as: “INSYDE-content: a synthetic, multi-variable flood damage model for household contents”.

Reviewer #2

R2.C1: *The paper presents a new model for estimating flood damage to household contents. In contrast to damage to residential buildings, damaged household items are neglected in many (scientific) flood damage models or are estimated using simple approaches such as a lumped share of the estimated building damage. In practical loss estimation applications such as cost-benefit-analysis, where a loss estimation for all sectors and damage types is needed, further approaches exist, e.g. specific stage-damage functions for contents. These approaches are often not well documented or published. With INSYDE-contents the authors propose a detailed flood damage assessment of household contents based on 11 typical household items, their mean replacement values and the estimated number of damaged items, which depends on characteristics of the flood event and the affected buildings. The paper also adds insights on the model's performance and validation using two real world data sets. So, I think the paper and the model presented provide a valuable contribution to the scientific literature on flood loss modelling. Still, the paper could be further improved with regard to the following aspects.*

Reply: We thank the Reviewer for the positive evaluation of our work and for recognizing the contribution of the INSYDE-content model to the advancement of flood loss modelling. We appreciate the constructive comments and suggestions, which we address point by point in the following responses and that we have incorporated in the revised version of the manuscript.

R2.C2: *Introduction (line 44 - 59): While it is acknowledged that the authors present the relevant literature, not many insights about the existing approaches are provided. Please be clearer about the weaknesses and strengths of the models mentioned. And discuss later in the paper, what your model contributes in comparison to the existing approaches.*

Reply: In the original manuscript, we intended to provide a concise but critical overview of existing flood damage models for building contents. In particular, we noted the limited transferability of empirical models (the ones mentioned in L44-59) due to the strong regional dependence of content types and distributions (L74-76). We also discussed the only existing Italian empirical model (Carisi et al., 2018), noting that it estimates content damage indirectly via a regression on building damage. As stated in the manuscript, this approach overlooks key factors such as the spatial distribution of contents within buildings, thereby limiting its ability to capture content-specific vulnerability. Similarly, we mentioned the synthetic model by Nofal et al. (2020), emphasizing how its simplified assumptions (such as contents placement within a hypothetical single-family wooden building) hinder real-world applicability due to the absence of content and building variability representation.

Moreover, we devoted a dedicated subsection to comparing the outputs of INSYDE-content with those of Carisi et al. (2018) to empirically illustrate the differences in model behavior.

We hope that the clarifications provided above demonstrate that the original manuscript already addressed these aspects in both the Introduction and the comparative analysis. Therefore, we respectfully consider the present discussion to be adequate in its current form.

R2.C3: *Methodology: Since the model development is an important part of the whole paper, I think it should be presented in more detail. It doesn't become clear in my view, why these 11 items (lines 100-102) were selected.*

Reply: As mentioned in the original manuscript, the selection of the 11 household content items was based on a survey of real cases, conducted through the analysis of real estate listings with photographic evidence. These items represent the essential and most frequently occurring contents typically found in residential dwellings, and were selected to ensure relevance, generalizability and applicability of the model to a wide range of household types. We have clarified this point in the revised manuscript.

R2.C4: *Later, the sampling procedure that led to 60 buildings and the sample itself could be better described.*

Reply: The sample of 60 buildings was derived by applying strict selection criteria to a broader dataset of approximately 500 real estate listings examined. The selection was based on the completeness and consistency of the available information, including key geometrical attributes, architectural layouts, interior characteristics and sufficient photographic documentation to enable full model parameterization. Only listings that met all these requirements were retained, in order to ensure the reliability of the input data used in the model setup. As this point is already covered in L111-118 of the original manuscript, we consider the existing description of the sample selection to be sufficiently comprehensive and therefore believe that no further details are available to provide.

R2.C5: *Furthermore, Table 3 and the analyses behind it, should be better explained.*

Reply: Table 3 summarizes, in a synthetic way, the relationships between the variables considered in the model, which are described in detail in the Supplementary Material 1. However, to improve clarity and aid readers' understanding, in the revised manuscript we have included a more detailed description of this table, by leveraging an example for one content item (see also reply to comment R2.C6).

R2.C6: *A lot of material is presented in the Supplement, but I would prefer to see at least one example how the damage function was derived for one item in the main text.*

Reply: We appreciate the Reviewer's suggestion and, accordingly, we have included in the main text of the revised manuscript an example how the damage function for beds was derived, including its fragility curve and the dependence on relevant variables. This inclusion also supports a better interpretation of Table 3 (see reply to comment R2.C5).

R2.C7: *Along the same lines, the methods in section 2.2 could benefit from some more details on the data and the methods used.*

Reply: The methodology and data used in Section 2.2 follow the approach adopted in the INSYDE 2.0 model for buildings (Di Bacco et al., 2024). To improve clarity and address the Reviewer's comment, we have revised the manuscript by including some concise details on the development of the synthetic datasets used in this study (see revised version of Section 2.2.1).

R2.C8: *Altogether, I think the paper could benefit from a flow chart or another image showing the different stages of the model development and evaluation as well as the data sets involved.*

Reply: In the revised paper, we have included a flow chart in the Introduction to provide a clear visual summary of the main stages of model development and evaluation and updated the corresponding text accordingly.

R2.C9: *Table 1: The equation for SA is given as $SA = F(SA; NF)$. Do you mean FA as independent variable here?*

Reply: We thank the Reviewer for spotting this typo. The correct formulation is indeed $SA = f(FA; NF)$, and we have corrected this in the revised manuscript.

R2.C10: *Table 1: Why do you use FA (instead of SA) in the equation for HU?*

Reply: Because FA is the independent variable, while SA is derived from FA through an empirical functional relationship (i.e., HU depends on SA implicitly).

R2.C11: *line 176/177: rephrase ("by the same authors" is a bit confusing here)*

Reply: This has been fixed in the revised manuscript.

R2.C12: *Table 3: As already mentioned above, the rationales and analyses behind the equations in Table 3 need more explanation.*

Reply: See reply to comments R2.C5 and C6.

R2.C13: *lines 217: "total actualized losses" - please check term*

Reply: In the revised manuscript we have replaced "total actualized losses" with "total losses (adjusted to year 2023 values)".

R2.C14: *Figure 2: This figure and the methods behind it need more explanation in my view. Also, the two data sets should be better described in the paper (briefly, but still in more detail than is currently the case).*

Reply: As also noted in response to comment R2.C7, the data and methods underlying the results shown in Figure 2 follows the approach developed for the INSYDE 2.0 model for buildings (Di Bacco et al., 2024). In the revised manuscript, we have improved the description in Section 2.2.1 to provide additional details regarding the data used in the analysis.

R2.C15: *Figure 4: These results should be analyzed and discussed in much more detail. How come that the estimates for (semi-)detached house do not show much variability (in contrast to the estimates for apartments)? The authors should present more in depth analysis of these results, including common metrics for errors or model performance (RMSE, MAE etc.), and they should discuss potential weaknesses of their model. How could the model be further improved to better capture the variability of the observed damage/claims?*

Reply: This concern was also raised by Reviewer 1 (see comment R1.C2). We report here the same response for completeness.

As also discussed in Molinari et al. (2020), claim data at the building scale are often affected by significant uncertainty and potential bias. For this reason, traditional flood damage models, as well as INSYDE-content, are typically more reliable when applied at aggregated spatial scales, rather than at the level of individual buildings.

This intrinsic variability in the observed data makes the interpretation of building-scale validation results particularly challenging. In our original manuscript (L368-374), we emphasized the importance of providing uncertainty ranges in the predicted damage values as a way to enhance the informative content of the model compared to purely deterministic approaches.

Nonetheless, we acknowledge the Reviewer's request for a clearer assessment of model performance. In the revised version of the manuscript, we have included quantitative error metrics (mean absolute error and mean bias, calculated based on the median of the predicted values per building, compared to the observed losses) to complement the visual inspection of the plots. However, we have also added a caveat to caution against over-interpreting these metrics, since observed values should not be considered absolute ground truth in damage modeling, due to the inherent limitations in claim data quality.

As also noted by the Reviewer, the apparent flatness of the predicted damage values for detached and semi-detached houses is partially explained by the use of log-log axes, which compress the visual perception of variability. Furthermore, from a theoretical perspective, a limited dispersion in predicted losses is to be expected for these building types, given their relatively homogeneous characteristics and the shallow inundation depths recorded during the two flood events. In contrast, greater variability is observed for multi-family residential buildings, which generally exhibit broader heterogeneity in both exposure and vulnerability, leading to capturing a damage prediction variability of the same order of that for observed losses. In the revised version of the manuscript, we have included comments on these aspects regarding the interpretation of the results in Figure 4.

R2.C16: *In general, I think that the results could be better interpreted and discussed. In a merged section "results and discussion" there's always the risk that the discussion is too short. The authors should expand theirs.*

Reply: The decision to merge the results and discussion was made to ensure a more concise and coherent presentation, avoiding unnecessary repetition across sections. We believe that the key findings and their implications have already been discussed in sufficient detail throughout the combined "Results and discussion" section, with no additional elements identified that would meaningfully enrich the discussion and justify a stand-alone section.

R2.C17: *Supplement 1 is very helpful and detailed. It will enable others to apply the model, too, which is much appreciated.*

Reply: Thank you.