Response to Reviewer #1

We thank the reviewer for the constructive comments and feedback. Below, we provide point-by-point responses to each comment.

Original review comments are in **black,** and our responses are in **blue.** *Changes to be introduced in the revised version of the manuscript are shown in italics.*

General comment:

This is the first time that such detailed, micro-scale flood damage data has been made publicly available. The dataset is accessible through the provided links and is clearly described and well-organized in the attached Excel sheets, with separate documentation for both commercial and residential buildings. In its current form, the data can be used in a variety of contexts, including cross-validation of flood damage models, both residential and commercial, for other regions and countries, improving existing models, and identifying overlooked damage mechanisms.

It is noteworthy that several Italian universities have collaborated to develop a common survey methodology and a standardized dataset for post-event flood damage data collection. I hope that the aim is to maintain this effort over time, thereby creating a longitudinal dataset that supports the continuous improvement and adaptation of damage models to reflect the evolving physical and economic vulnerabilities of exposed assets, as well as to enable their validation.

Specific comments:

- **Section 2.2: Data collection**

Data collection is reported to have started immediately after the event. In contrast, previous studies that relied on post-event data collection typically began 6–8 months later, or even beyond that. This delay was intentional, allowing people time to reconstruct their buildings so that when surveys were conducted, most or all buildings would have been fully reconstructed, enabling a more accurate assessment of the original damage.

This difference in timing should be considered a limitation of the current dataset. Since data were collected immediately, the reported damage may not capture the full cost of damages that become apparent only during or after reconstruction efforts.

> We thank the reviewer for raising this point about the timing of the data collection campaign. We agree that collecting data immediately after the event limited the possibility to capture certain types of damages, especially indirect ones, and led to an underrepresentation of damages to inaccessible and closed buildings.
>
> In Section 5.2 (Data limitations) of the manuscript, the partial coverage of all damaged assets is already acknowledged; however, we will further expand this section with the

consideration on the potential underestimation of the indirect damages (e.g., damage caused by the humidity) for the surveyed assets, due to the timing of the survey.

Nevertheless, it is worth mentioning that the immediate post-event survey was essential to record accurate information that would have been difficult to obtain later. For instance, the visible flood marks at the time of the survey enabled precise water depth measurements. Moreover, collecting testimonies shortly after the event ensured that direct damage reports were not affected by memory omissions, since damages were clearly observable in the field at that time.

The following edits and the new sentence will be incorporated in the revised version of the manuscript:

> Line 287: *Despite the high quality of the datasets,* ***three*** *main limitations should be highlighted.*
>
> Line 289: *Second, the timing of the survey may have resulted in a potential underestimation of indirect damages, such as those caused to furniture, coating and plasters by humidity.*

- **Section 2.4: Technical validation**

It is very good that the collected data is also reviewed by an external team, in addition to the original data collection team. However, it is still unclear whether this review is conducted only on the paper forms before they are entered into KoboToolbox, or if it also includes forms that have already been digitized in the platform. If the review is limited to the paper forms, it is important to also double-check the digital entries. In previous studies, reviewing the digitized data has proven useful in identifying additional typing errors.

The validation process was conducted in three steps. The first step involved only the data collected using paper forms, as during this phase the teams checked the data before entering it into KoboToolbox. The second step was carried out by a single reviewer, who performed a quality check directly in KoboToolbox after all data had been digitized on the platform. Thus, this second step was conducted on the entire dataset. The third step took place after the second review, during which all teams were asked to correct and complete missing information in KoboToolbox. Thus, only the first step was limited to the data originally collected on paper forms.

In the revised manuscript, we will clarify that the complete datasets were involved in the second validation step, as follows:

> Lines 140-141: *Second, an independent review was conducted, i.e. by a surveyor external to the field team, focusing on data coherence after all the paper forms were digitalized in KoboToolbox (Phase 3).*

Minor comments:

1. Introduction and case study

- Figure 1. Avoid using a yellow dot to represent the municipality of Catarino, as it is very similar to the one representing the economic activities. Please change for another colour.
- Figure 1. We are missing the representation of the three surveyed municipalities within the Misa River basin. Would it be possible to include them in the figure showing the basin? This would help illustrate which part of the basin was surveyed.
- Figure 1. In the legend, specify as in the caption that the economic and residential buildings in the municipalities are the surveyed ones.

We thank the Reviewer for the suggestions; we will updated Figure 1 accordingly:

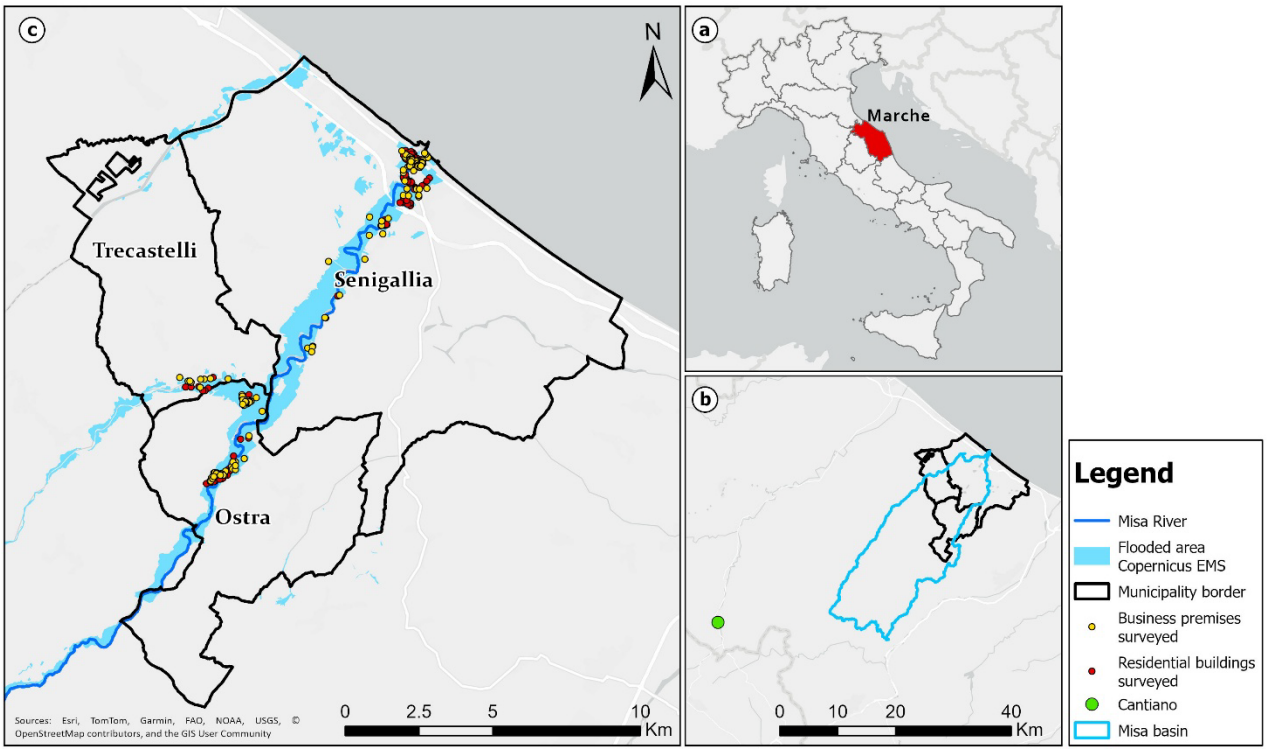


Figure 1. (a) Marche region shown in red; (b) location of the three municipalities of Ostra, Senigallia, and Trecastelli surveyed within the Misa Basin, and the municipality of Cantiano, indicated with a green dot; (c) points representing the buildings surveyed within the three municipalities, and the flooded area of the 2022 event, as provided by the Copernicus Emergency Management Service (EMS).

2. Methods

In the excel data dictionary:

Economic activities

Form A:

- It is not very clear to me the differences among the 4 variables representing building elevation (ΔQ, hg, h1 and h2).

  In the PDF of FORM A, there is a picture illustrating what these variables refer to. Specifically, the figure shows a reference survey point, indicated by a black horizontal dashed line, from which all measurements are taken. This is also the point where the external water depth is measured. To clarify the four measurements and support other researchers and practitioners in collecting these values in the field, we will provide an additional measurement guide uploaded with the datasets in Zenodo. This guide is attached to this response (see page 6).

- It would be valuable for future work to include an additional sediment variable representing large objects (e.g., tanks, cars, rubble from other buildings), as these objects could cause additional damage to building structures upon impact.

  We thank the Reviewer you for the suggestion. This aspect was included in the survey forms under the field *"other"*, where surveyors could report the presence of large objects.

Residential buildings

- In the dictionary of the database, it is not very clear the distinction of B, C and D forms, specify there too that B is for the housing unit, C is for the common areas and D for attached buildings.

  We will add a sentence to each worksheet of the data dictionary to clarify the distinction between the forms.

Form B, C, D:

- This form in the floor section includes a variable 'damage due to high velocity' how is this collected? Based on people perception? How do you double check this assumption?

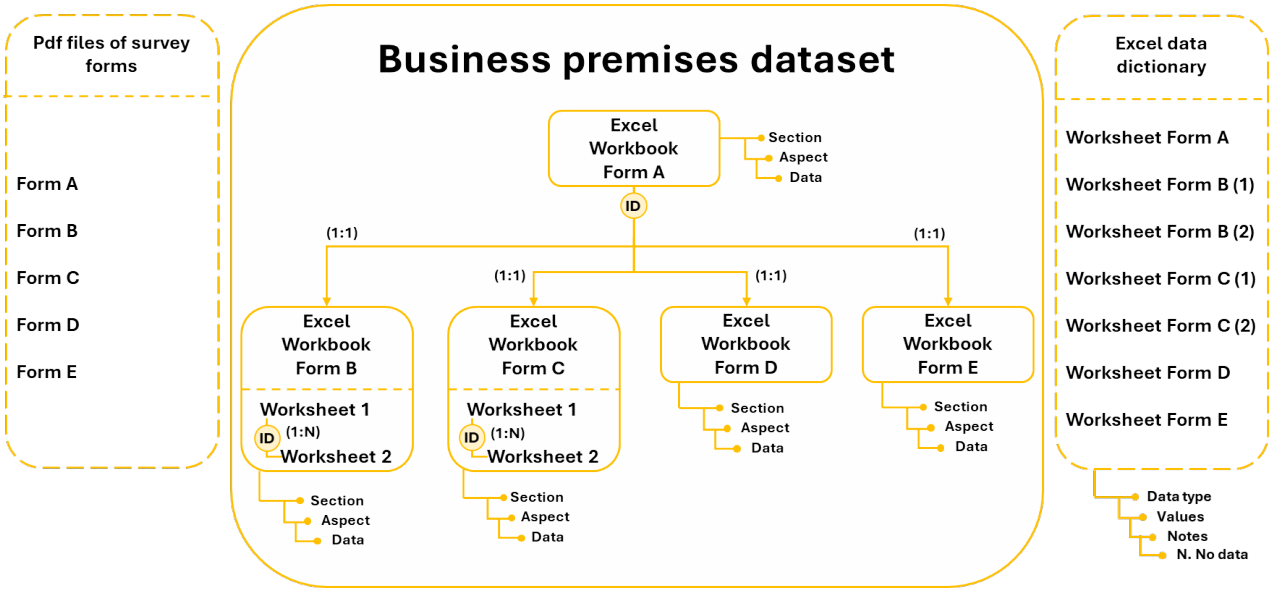
  Floodwater velocity was assessed based on the interpretation of people's descriptions of how the flood water propagated during the event. While we did not have independent measurements, these narrative-based data offer valuable localized, though qualitative, information. Such information is often the only possible source of insight into local flow conditions in the absence of instrumental data and can serve to validate hydrodynamic models. Nevertheless, consistency of the reported water velocity was assessed by comparing information collected from nearby buildings.

  No changes will be made in the revised version of the manuscript.

3. Data records

- Figure 3: Since all the forms are connected to Form A in a 1:1 relationship, please indicate the 1:1 connection for Forms C and D in the sketch as well.

We have corrected Figure 3 according to the Reviewer suggestion.



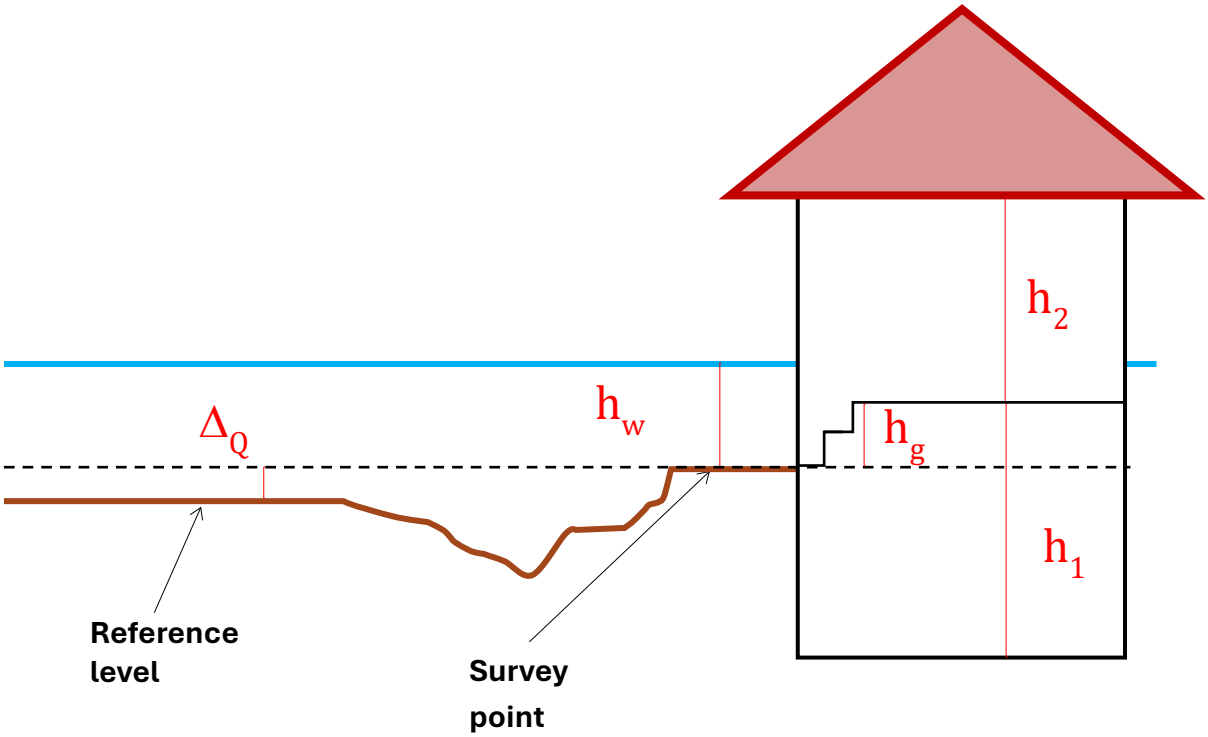


Figure 1. Sketch of the cross section of a building depicting the measurements taken in the field (ΔQ, h_w, h_g, h_1, h_2), reference level and survey point. Case with ΔQ and h_g positive, and h_1 negative.

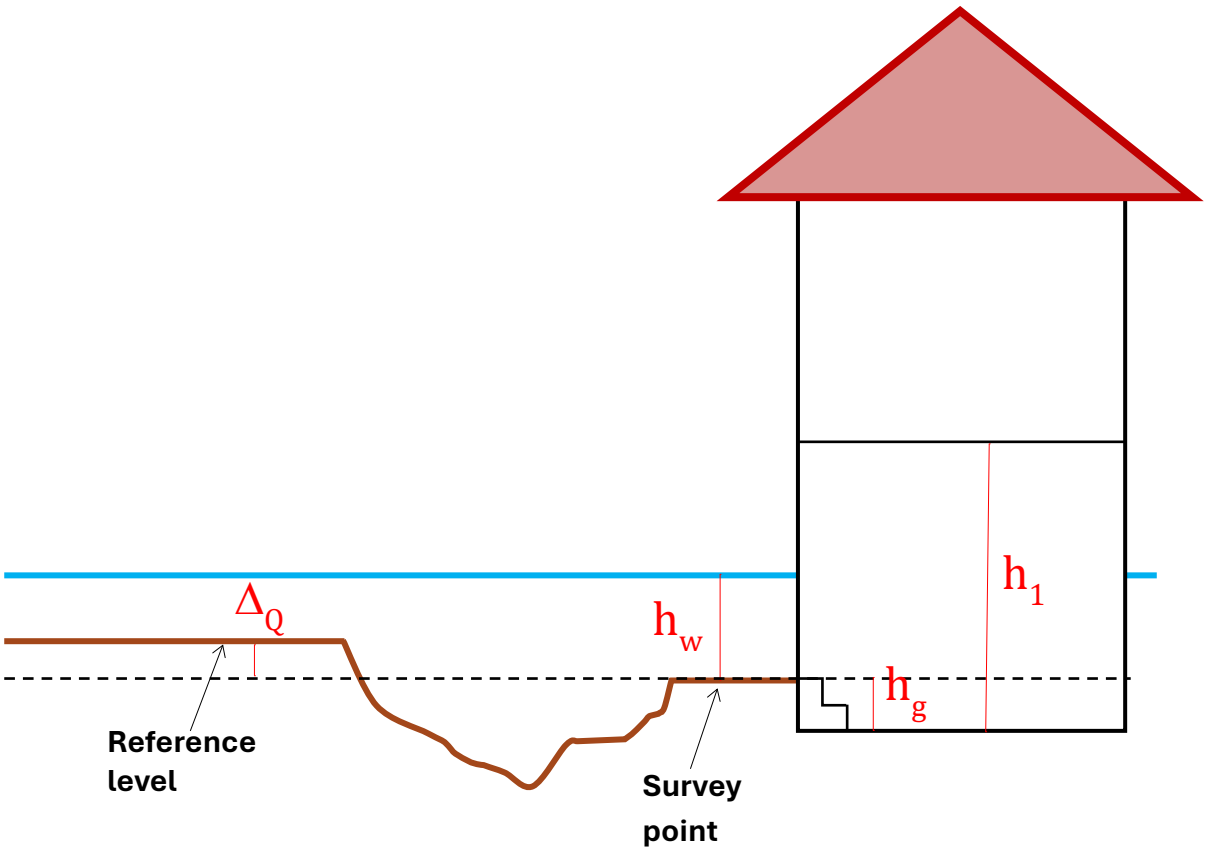


Figure 2. Sketch of the cross section of a building depicting the measurements taken in the field (ΔQ, h_w, h_g, h_1), reference level and survey point. Case with ΔQ and h_g negative, and h_1 positive.

Variables

- Δ_Q indicates the height difference between the elevations of the survey point and the reference level. The value is positive when the survey point is located at a higher elevation relative to the reference level point, and negative when it is below it. Figure 1 depicts the case in which this measure is positive; Figure 2 depicts the case in which this measure is negative. The measurement is taken in the field by two surveyors.
- h_w indicates the external water depth outside the building, measured at the survey point.
- h_g indicates the height of the first floor with respect to the survey point. This measurement is taken, for example, when the first floor is accessed via stairs. It is positive when the first floor is higher than the survey point, and negative when it is lower. Figure 1 depicts the case in which this measure is positive; Figure 2 depicts the case in which this measure is negative.
- h_1 represents the total height of the first floor measured relative to the h_g level. It is positive when the first floor is above the h_g level, and negative when it is below. Negative values clearly indicate that the first floor is a basement or semi-basement level. Figure

1 depicts the case in which h_1 is negative; Figure 2 depicts the case in which h_1 is positive.

- $\boldsymbol{h_2}$ represents the total height of the second floor measured relative to the h_g level if h1 is negative, and relative to h_1 if h_1 is positive. h_2 is always positive, as it refers to floors above ground level.

Survey Point and Reference Level

- The **survey point** serves as the primary spatial reference from which vertical distances to the building floors and reference level are recorded relative to this point. Specifically, it is the exact location on the ground where the external water depth h_w is measured.
- The **reference level** is a fixed elevation benchmark defined locally for each building, typically corresponding to a flat area adjacent to that building. This allows, by using a Digital Terrain Model (DTM), for all measured heights to be accurately converted into absolute elevations (e.g., the water surface elevation at the building location, provided in *FORM_A.xls*, was determined by summing Δ_Q, h_g, and the ground elevation of the reference level).