We thank the reviewer for this thorough and constructive overall assessment, which has substantially improved the manuscript. We address each point in turn.

On the value of the study and its relationship to USGS ShakeMap/PAGER (see also Comments 1–2, 19): We agree that USGS ShakeMap and PAGER already incorporate observed ground motions when local station data are available (Worden et al., 2010); we cite this work directly to make the point explicit. This study does not use, benchmark against, or make any claim about the USGS ShakeMap or PAGER products, and we have removed the term "ShakeMap" throughout the manuscript, including in the title, to eliminate this source of confusion. The two-stage scenario-based/station-updated comparison presented here is instead implemented independently within REDAS (developed under the REDACt project) for the 30 October 2020 Samos earthquake. The contribution of the study is not the general fact that observational updating reduces bias, which is indeed well established, but the specific, quantified demonstration that this updating step moves the {Moderate+Extensive+Collapse} exceedance probability across an operationally meaningful threshold (from the 0–10% range to the 15–30% range) in a real deep-basin urban setting, with direct implications for response-tier escalation during the Golden Hours. We have revised the Introduction and Conclusions to state this contribution explicitly rather than leaving it implicit.

On the single-event scope and title (see also Comment 40): We agree this is a single-event case study, and the manuscript has been revised throughout to use singular, case-specific language (e.g., "this deep-basin urban environment" rather than "deep-basin urban environments") wherever the claim is scoped to the Samos/Bayraklı case. We also propose removing 'ShakeMaps' from the title, replacing it with 'ground-motion parameter maps,' to avoid any implication of a comparison against the published USGS ShakeMap product, which this study does not undertake. The proposed revised title is detailed in our response to Comment 1, and we kindly submit it for the handling editor's consideration.

On the description of the calibration methodology (see also Comments 8, 10, 11): We agree this would benefit from a more explicit description. We have added an explicit description of the bias-correction procedure to Section 4.1: a single multiplicative bias factor is computed for each ground-motion parameter and spectral period, chosen to minimize the absolute error between observed and predicted values at the four strong-motion stations, following the same logic as the phantom-station approach implemented in the ELER rapid loss estimation platform (Demircioglu et al., 2010; Hancilar et al., 2010), and is then applied uniformly to rescale the scenario-based ground-motion field over the entire grid. We have also added logarithmic-ratio results alongside the existing absolute-difference metrics (Comment 10), and a full description of the finite-fault source characterization and the specific input parameters required by each of the four GMPEs (Comments 6, 8), so that the calibration workflow is now fully traceable from source characterization through to the final calibrated field.

.

On overgeneralization beyond the single case study (see also Conclusions): We agree that conclusions must be scoped to what a single-event case study can support. The Conclusions section already states explicitly that "the analysis is based on a single earthquake case study within a deep-basin setting; therefore, extrapolation to other tectonic and urban contexts should be undertaken cautiously," and any broader statements (e.g., that similar mechanisms "may occur" in other deep-basin environments) are phrased as hypotheses for future work rather than established findings. We have reviewed the manuscript throughout to ensure that claims outside the immediate Samos/Bayraklı case are consistently hedged in this way.

On the general lack of references (see also Comments 14, 21, 28, 30, 34, 45, 47, 54, 55, 61, and others): We thank the reviewer for this observation and have conducted a systematic pass through the manuscript to identify and fill citation gaps. New references have been added covering: the USGS ShakeMap/PAGER observational-updating mechanism (Worden et al., 2010; Wald et al., 2010), the REDAS/REDACt platform (REDACt-Project, 2023), the ELER calibration methodology (Demircioglu et al., 2010), basin-depth modelling uncertainty (Day et al., 2008), intermediate-period amplification evidence (Cetin et al., 2022; Demirel et al., 2022), the international applicability of HAZUS fragility functions (Farahani et al., 2020; Choi et al., 2019; Nguyen et al., 2025; Boutaraa and Chaibedra, 2025; Delves et al., 2026), and the "Golden Hours" / multi-class damage classification literature (Li and Xu, 2025; Li et al., 2023), among others detailed in our responses to the individual comments below.

On AI-assisted language editing and residual "AI slop" and repetition (see also Comments 24, 26, 29, 31): We thank the reviewer for this comment regarding language quality. As stated in the Acknowledgements section, AI-based tools were used solely for language editing and clarity improvement; the scientific content, methodology, analyses, and conclusions are entirely the work of the author. In response to this comment, the manuscript has been carefully and thoroughly re-read in full, and all passages identified as repetitive or imprecise have been revised to improve conciseness, precision, and natural academic flow. For example, vague expressions such as "optimistic bias" and "persistent methodological gap" have been replaced with precise descriptions, and redundant or circular sentences have been removed. The revised manuscript presents the scientific content in a clearer and more carefully edited form.

1. ShakeMap refers to a specific product created and published by USGS. It appears that the author has used ShakeMap to refer to a map of ground shaking from an earthquake event. Please only use ShakeMap when referring to the product, and use another terminology, e.g., shaking maps, when referring to a generic map of ground shaking.

We thank the reviewer for this important clarification. We confirm that this study does not use, reference, or benchmark against the official USGS ShakeMap product. All ground-motion representations discussed in this manuscript are generated within the REDAS rapid damage estimation platform, using GMPE-based estimates (Stage 1) and their station-based bias-corrected counterparts (Stage 2).
Given that the term "ShakeMap" is not needed to describe the USGS product anywhere in this study, we have removed it entirely and replaced every generic use, including in the title, abstract, methodology, results, and figure captions, with "ground-motion parameter maps," qualified as appropriate by "scenario-based," "station-calibrated," or "spatially distributed." This terminology is already used consistently elsewhere in the manuscript (e.g., Section 3, Section 4.3) and its extension throughout removes any risk of confusion with the USGS product.

As a direct consequence of this terminology revision, we propose the following change to the manuscript title:

Original title: "Operational damage misclassification in scenario-based ShakeMaps: evidence from station-updated ground-motion fields in a deep-basin urban environment"
Proposed revised title: "Operational damage misclassification in scenario-based ground-motion parameter maps: evidence from station-updated strong-motion observations in a deep-basin urban environment"

We kindly ask the handling editor to consider this title revision during the revision process, and we will implement it in the revised manuscript upon approval.

2. Introduction through line 84: The author has not introduced their study event in the first part of the Introduction, instead indicating that this issue occurs broadly. However, both ShakeMap and PAGER estimates produced by USGS immediately after an earthquake do in fact account for observed ground motions, and are only constrained by the availability of local station data (see e.g., Wald et al. Challenges in Rapid Ground Motion Estimation for the Prompt Assessment of Global Urban Earthquakes; https://doi.org/10.1785/0120100101). As a result, please clarify whether the author is referencing only the Samos earthquake as missing observational data in the produced ShakeMaps, and evidence that observational data were not included.

We thank the reviewer for this comment, and for pointing us to this reference. We note that the DOI provided corresponds to Worden et al. (2010), "A Revised Ground-Motion and Intensity Interpolation Scheme for ShakeMap" (Bulletin of the Seismological Society of America, 100(6), 3083–3096), which describes precisely how the USGS ShakeMap system combines observed ground-motion data with GMPE-based estimates through a weighted-average interpolation scheme. We agree entirely with the reviewer that USGS ShakeMap and PAGER already incorporate observational updating whenever local station data are available, and we have cited Worden et al. (2010) directly in support of this point.

To clarify, this study does not claim that USGS ShakeMap or PAGER fail to incorporate observational data, nor does it use or benchmark against these products. The two-stage comparison presented here, scenario-based (Stage 1) versus station-updated (Stage 2), is a framework implemented independently within REDAS (developed under the REDACt project), applied specifically to the 30 October 2020 Samos earthquake. Stage 1 is constructed by design to represent the ground-motion estimate obtainable from earthquake source parameters and GMPEs alone, deliberately excluding the AFAD strong-motion recordings; Stage 2 incorporates those recordings through bias correction. This is a controlled analytical comparison defined for this case study, not a claim about what data were available to any operational system during the actual 2020 event. The contribution of this study is not that observational updating reduces bias — which is well established — but the quantified demonstration that, in this deep-basin setting, this updating moves the cumulative {Moderate+Extensive+Collapse} exceedance probability across an operationally meaningful threshold (from the 0–10% range to the 15–30% range), with direct implications for emergency response prioritization during the critical early hours.

We have revised the Introduction accordingly, adding a passage that (i) introduces REDAS and its origin under the REDACt project; (ii) acknowledges that operational systems such as ShakeMap and PAGER already perform this type of observational updating, consistent with Worden et al. (2010); and (iii) clarifies that this study aims to demonstrate the contribution of station-based updating to disaster management decision-making through an independent, two-stage implementation within REDAS for the Samos earthquake.

Original manuscript (Lines 54–61):
"Accordingly, near-real-time damage estimation systems that become operational immediately after an event have emerged as a cornerstone of contemporary disaster management. Such platforms aim to provide decision-makers with early, spatially explicit estimates of earthquake-induced ground shaking and potential impacts, thereby supporting the prioritization of response actions during the initial phase. Representative implementations include HAZUS (FEMA, 2025), ELER (Hancilar et al., 2010), PAGER (Wald et al., 2010), SELENA (NORSAR and ICG 2026), OpenQuake (Silva et al., 2013), AFAD-RED (AFAD, 2013), and REDAS (REDAS, 2020).
Despite their operational value, the reliability of rapid post-event damage products is fundamentally constrained by the accuracy of the underlying ground-shaking representation and the resulting damage estimation outputs."

Revised manuscript:
The passage has been retained with the following changes:

(i) "contemporary disaster management" replaced with "early-phase earthquake emergency response and disaster risk management" (see response to Comment 16);
(ii) the undefined citation "(REDAS, 2020)" replaced with "(REDACt-Project, 2023)"; and
(iii) the following passage inserted immediately after this sentence:

REDAS is developed under the REDACt project (Rapid Earthquake Damage Assessment ConsorTium), funded by the European Union through the Black Sea Cross-Border Cooperation Programme (REDACt-Project, 2023). We note that operational systems such as ShakeMap and PAGER already incorporate observed ground-motion data whenever local station recordings are available, combining them with GMPE-based estimates through weighted-average interpolation (Worden et al., 2010). Building on this conceptually similar logic, this study aims to demonstrate the contribution of station-based updating to disaster management decision-making through an independent, two-stage implementation within REDAS for the 30 October 2020 Samos earthquake.

3. Line 110: Suggest adding map of earthquake epicenter, faulting regime, and damaged areas to help orient readers.

We thank the reviewer for this helpful suggestion. Figure 3 has been updated to include the epicenter and fault trace of the 30 October 2020 Samos earthquake together with its focal mechanism (strike, dip, rake, and Mw, sourced from the AFAD Focal Mechanism Solution), alongside the existing spatial distribution of observed building damage in the Mansuroğlu Neighborhood. This updated figure now orients the reader within the regional seismotectonic context while directly linking it to the observed damage pattern discussed throughout the manuscript.

Original manuscript (Figure 3):
Showed only the location of the Samos earthquake epicenter (as a red star, without fault or focal mechanism information) alongside the spatial distribution of observed building damage in the Mansuroğlu Neighborhood.

Revised manuscript:
Figure 3 has been updated to additionally show the fault trace and the AFAD focal mechanism (Mw 6.6; Plane 1: Strike 95°/Dip 43°/Rake −87°; Plane 2: Strike 270°/Dip 46°/Rake −91°; source: AFAD Focal Mechanism Solution, https://deprem.afad.gov.tr/event-detail/39422), together with the location of Mansuroğlu Town relative to İzmir and the epicenter. The building-damage panel is retained and its readability improved
(see also our response to Comment 52).

4. Line 113: Suggest adding more information about "structural deficiencies", including the type of structures commonly in the area, type of structures impacted and the most, common deficiencies that resulted in damage.

We thank the reviewer for this suggestion. The predominant structural typology in the study area consists of reinforced-concrete moment-resisting frame buildings of 7–9 stories that commonly lack shear walls; these characteristics, together with the specific structural deficiencies documented after the 30 October 2020 Samos earthquake, are described in full detail in Section 3.3. We have revised the passage at Line 113 to briefly identify these structural characteristics at first mention and to direct readers explicitly to Section 3.3 for the complete discussion, rather than leaving "structural deficiencies" undefined at this point in the text.

Original manuscript (Lines 113–115):
Although structural deficiencies related to design, detailing, and construction practices contributed to the observed damage (METU EERC, 2020; DAUM, 2020) one of the primary factors governing the the spatial concentration of damage in Bayraklı has been widely attributed to the combined influence of local site response effects and structural vulnerability.

Revised manuscript:
Although structural deficiencies related to design, detailing, and construction practices contributed to the observed damage (METU EERC, 2020; DAUM, 2020), these deficiencies primarily involved reinforced-concrete moment-resisting frame buildings of 7–9 stories that commonly lack shear walls. The predominant building typologies and associated structural characteristics of the study area are described in detail in Section 3.3. The spatial concentration of damage in Bayraklı has been widely attributed to local site response effects, in particular the resonance between basin-amplified ground motions and the natural periods of the dominant mid-rise building stock.
(Note: this revision also resolves the circularity issue raised separately in Comment 35, since the two contributing factors, structural deficiencies and site response, are now presented sequentially rather than conflated in a single clause.)

5. Line 115: Please add reference for how the effect of "structural deficiencies" was dissociated from the "influence of local site response effects".

We thank the reviewer for this comment. As described in our response to Comment 4, the passage at this location has been restructured so that structural deficiencies and local site response effects are now presented as two distinct, sequential statements rather than a single combined clause. To support the specific claim regarding local site response, we have added a direct citation to Çetin et al. (2022) and Demirel et al. (2022), both of which document the resonance between basin-amplified ground motions and the natural periods of the mid- to high-rise reinforced concrete building stock in Bayraklı during the 30 October 2020 Samos earthquake.

Original manuscript (Lines 113–115):
Although structural deficiencies related to design, detailing, and construction practices contributed to the observed damage (METU EERC, 2020; DAUM, 2020) one of the primary factors governing the the spatial concentration of damage in Bayraklı has been widely attributed to the combined influence of local site response effects and structural vulnerability.

Revised manuscript (building on the Comment 4 revision):
...The spatial concentration of damage in Bayraklı has been widely attributed to local site response effects, in particular the resonance between basin-amplified ground motions and the natural periods of the dominant mid-rise building stock **(Cetin et al., 2022; Demirel et al., 2022)**.
(Both references are already present in the bibliography; no new entries required; only the in-text citation is added at this specific claim.)

6. Section 3: Please describe how the earthquake source was characterized for generating the shaking maps: e.g., point source, planar source, 3D geometry, etc. Additionally, please provide details about the values of source parameters used in each of the 4 GMPEs.

We thank the reviewer for this comment. We have added a description of the earthquake source characterization to Section 3, and a table detailing the specific source-parameter requirements of each of the four GMPEs to Section 3.1, as detailed below.

Revised manuscript
Section 3 (Methodology, main body) — new paragraph added:

The 30 October 2020 Samos earthquake occurred offshore of the northern coast of Samos Island at 11:51 UTC. In this study, the earthquake source was characterized using a finite-fault model, with source parameters derived from the AFAD earthquake catalogue (AFAD | deprem.gov.tr, 2026). The rupture plane was defined by a strike of N270°E and a dip of 46°, consistent with the east–west trending normal-faulting mechanism reported for this event, with a seismogenic width of 15 km. The epicenter was located at 37.902°N, 26.794°E, with a hypocentral depth of 14.9 km and a moment magnitude of Mw 6.6. The location and focal mechanism of the earthquake are shown in **Error! Reference source not found.**.

Section 3.1 ("Selection of Ground Motion Prediction Equations") — this paragraph and Table did not exist in the original submission and have been newly added in their entirety

All four GMPEs require moment magnitude (M_w) and V_{s30} as common input parameters, but differ in how they encode style-of-faulting, distance dependence, and basin-depth effects, as summarized in Table 1. ASK14, CB14, and CY14 encode the normal-faulting mechanism through binary reverse- and normal-faulting indicators (Frv, Fnm), whereas BSSA14 uses an alternative unspecified/reverse/normal-slip classification (U, RS, NS). Similarly, distance dependence is captured differently across models: BSSA14 relies solely on the Joyner–Boore distance (Rjb), while ASK14, CB14, and CY14 additionally require the rupture distance (Rrup) and the strike-normal distance (Rx) to capture hanging-wall and directivity effects. Site response is represented through V_{s30} in all models, combined with basin-depth proxies $Z_{1.0}$ (ASK14, BSSA14, CY14) or $Z_{2.5}$ (CB14).

Table 1. Input parameters required by each of the four adopted NGA-West2 GMPEs.

GMPE	Mw	Vs30	Style-of-faulting encoding	Distance dependence	Basin-depth proxy
ASK14 (Abrahamson et al., 2014)	✓	✓	Frv, Fnm	Rrup, Rjb, Rx	$Z_{1.0}$
BSSA14 (Boore et al., 2014)	✓	✓	U, RS, NS	Rjb	$Z_{1.0}$
CB14 (Campbell and Bozorgnia, 2014)	✓	✓	Frv, Fnm	Rrup, Rjb, Rx	$Z_{2.5}$
CY14 (Chiou and Youngs, 2014)	✓	✓	Frv, Fnm	Rrup, Rjb, Rx	$Z_{1.0}$

7. Line 269: How was the dataset of 962 buildings collected: were they randomly sampled within a selected region, or were they all the damaged buildings within a region, or were they a subsample of the damaged buildings within a selected region? How was the region selected?

We thank the reviewer for this comment. The building inventory for the Mansuroğlu Neighborhood (Bayraklı District, İzmir) was obtained from the Republic of Türkiye Ministry of Environment, Urbanization and Climate Change, General Directorate of Construction Affairs. The dataset was not randomly sampled; it comprises all buildings for which detailed structural and damage information was available in this official inventory for the study area. The analysis region was defined by the administrative boundaries of the Mansuroğlu Neighborhood and the spatial extent of this ministry-provided database. The Bayraklı district was selected because it experienced disproportionately severe structural damage despite being located approximately 70 km from the epicenter — a circumstance that makes it particularly well-suited for examining the role of ground-motion representation in deep-basin environments.

Original manuscript (Section 3.3, near Line 269):
The building inventory for the Mansuroğlu Neighborhood (Bayraklı District, Izmir) was obtained from the Republic of Türkiye Ministry of Environment, Urbanization and Climate Change, General Directorate of Construction Affairs. The dataset was subsequently harmonized with the HAZUS building classification scheme by considering key attributes such as structural system type, number of stories, and level of engineering design.

Revised manuscript
The dataset was not randomly sampled; it comprises all buildings for which detailed structural and damage information was available in the official ministry inventory for the Mansuroğlu Neighborhood. The analysis region was defined by the administrative boundaries of the neighborhood and by the spatial extent of this ministry-provided database

8. Line 316: How were the grid estimates associated with station locations? Additionally, since the scenario based levels are obtained using GMPEs, why were grid based values used instead of calculating them directly at the station locations?

We thank the reviewer for this comment. In the REDAS framework, ground-motion parameter maps are generated on a regular spatial grid, where GMPE-based estimates are computed at each grid node using the corresponding source-to-site distance and site characterization parameters. The scenario-based estimates reported in Table 2 for each station location correspond to the values taken from the nearest grid node.

Regarding the second part of the comment, a grid was used instead of computing GMPE estimates directly at the station locations, because the four AFAD stations are unevenly spaced and too few to represent how ground motion varies across the study area. The grid provides the continuous spatial coverage needed for the exceedance-based damage assessment in Section 4.3. Station recordings are then used to derive period-dependent bias correction factors, which are subsequently applied to rescale the GMPE-based grid field across the entire study area, as described in Section 4.1.

Original manuscript (Section 4.1, existing passage):
"...scenario-based (Stage 1) estimates extracted from the nearest grid centroids, and (iii) station-updated (Stage 2) estimates after bias correction..."

Revised manuscript
New sentence added:" A grid was used instead of computing GMPE estimates directly at the station locations, because the four AFAD stations are unevenly spaced and too few to represent how ground motion varies across the study area. The grid provides the continuous spatial coverage needed for the exceedance-based damage assessment in Section 4.3. "

9. Figure 5: How were H1 and H2 components defined at the stations? Did they align with geographic N and E, or with respect to the direction of earthquake rupture?

We thank the reviewer for this comment. No rotation of the horizontal components was applied in this study. H1 and H2 correspond directly to the two recorded horizontal components as archived by AFAD, aligned with the geographic N–S and E–W directions, respectively, rather than being rotated with respect to the fault rupture direction. We have clarified this in the text accompanying Figure 6.

Revised manuscript:
The following sentence has now been appended to that paragraph to address the present comment: "H1 and H2 denote the two recorded horizontal components, aligned with the geographic N–S and E–W directions, respectively; no rotation with respect to the fault rupture direction was applied.

10. Eq 3, 4: What was the reason for choosing to use absolute difference instead of the ratio? Typically it's the log of spectral accelerations that are generated by GMPEs, and the standard deviation is also given for log(Sa). similarly the fragility functions are generally defined on the log scale (using lognormal distribution) on Sa. As a result, ratios of spectral acceleration are more representative of the differences that absolute values.

In the present study, absolute differences in spectral acceleration are adopted rather than log-ratios, with the aim of expressing the bias as a direct shift in spectral demand at specific periods, which is more intuitive and easier to interpret for practitioners who work with operational intensity thresholds and damage states. At the same time, both absolute differences and log-ratios are widely used in earthquake engineering, and a log-scale formulation is fully consistent with the statistical treatment of GMPE residuals and lognormal fragility functions. Following the reviewer's suggestion, logarithmic ratios between observed and modeled spectral accelerations, defined as "ln" format, have now been computed and reported in Table 2, so that the bias can be examined consistently in both absolute and logarithmic terms.

Revised manuscript:
Table 2 — new row added for each station:
Improvement – Log-ratio (%)" reported alongside the existing "Improvement – Absolute (%)" row, for all four stations and all spectral periods (PGA, S0.3, S0.6, S1.0, S3.0).

New paragraph added, immediately following the existing absolute-difference explanation:
For logarithmic measures, the pre-calibration logarithmic error (LE_{est}) was defined as the natural logarithm of the ratio between observed and scenario-based ground-motion intensities, $LE_{est} = \ln(I_{obs}/I_{est})$, whereas the post-calibration logarithmic error (LEcal) was computed analogously for the calibrated ground-motion parameter maps, $L_{Ecal} = \ln(Iobs/I_{cal})$. The relative reduction in logarithmic error due to calibration was expressed in an analogous manner to Eq. (5) by replacing the absolute errors with the absolute values of LE_{est} and LE_{cal}.

11. **Line 341**: What is the process of calibration? Paragraph 29: Please include references for the statements in the paragraph.

The calibration procedure has been described in Section 4.1 as a bias-correction step following the phantom-station approach of the ELER platform (Demircioglu et al., 2010; Hancilar et al., 2010); the full description is given in the Revised manuscript below.

Revised manuscript:
Not present in the original submission, the following paragraph is newly added in its entirety to Section 4.1: Calibration is implemented here as a bias-correction step, following the phantom-station approach of the ELER rapid loss estimation platform (Demircioglu et al., 2010; Hancilar et al., 2010). A single multiplicative bias factor is computed for each ground-motion parameter and spectral period by minimizing the absolute error between the observed and scenario-based (Stage 1) values at the four strong-motion stations. This factor is then applied uniformly to rescale the scenario-based grid, producing the station-calibrated ground-motion field (Stage 2) used for the hazard analysis in Section 4.2. For the risk and damage analysis in Section 4.3, the ground-motion parameter value at each building's own coordinates was obtained by interpolating from the surrounding grid nodes.

12. **Line 46**: What is "multi-level building damage classification framework"?

We thank the reviewer for this comment. The term "multi-level building damage classification framework" has been clarified in the revised manuscript by explicitly referencing the HAZUS damage states — Slight, Moderate, Extensive, and Collapse — which define the levels of structural performance used in this study.

Original manuscript (Lines 45–49):
Multi-level building damage classification frameworks enable response strategies to be formulated on a rational basis by providing search-and-rescue teams with an initial, model-based representation of damage severity across the affected area, while also facilitating the deployment of personnel and equipment with appropriate capacity and characteristics to the affected areas.

Revised manuscript:
Multi-level building damage classification frameworks, such as the Slight, Moderate, Extensive, and Collapse damage states defined in the HAZUS methodology (FEMA, 2025), enable response strategies to be formulated by providing search-and-rescue teams with an initial, model-based representation of damage severity across the affected area. They also facilitate the deployment of personnel and equipment with appropriate capacity and characteristics to locations where higher damage levels are expected.

13. **Line 47**: Sentence intent is unclear: … enable response strategies to be formulated on a rational basis …

We thank the reviewer for this comment. The phrase "on a rational basis" has been removed, as its intent was unclear. The sentence has been restructured to directly state that these frameworks enable response strategies to be formulated by providing search-and-rescue teams with an initial, model-based representation of damage severity across the affected area.

Original manuscript
Multi-level building damage classification frameworks enable response strategies to be formulated **on a rational basis** by providing search-and-rescue teams with an initial, model-based representation of damage severity across the affected area, while also facilitating the deployment of personnel and equipment with appropriate capacity and characteristics to the affected areas.

Revised manuscript:
Multi-level building damage classification frameworks, such as the Slight, Moderate, Extensive, and Collapse damage states defined in the HAZUS methodology (FEMA, 2025), enable response strategies to be formulated by providing search-and-rescue teams with an initial, model-based representation of damage severity across the affected area. They also facilitate the deployment of personnel and equipment with appropriate capacity and characteristics to locations where higher damage levels are expected.

14. Line 49: Please add a reference. Perhaps it has the same reference as Li and Xu, 2025; Li et al., 2023, in which case, suggest adding it again to clarify.

We thank the reviewer for this comment, and confirm the reviewer's suggestion is correct: the same reference used earlier in this paragraph for the "Golden Hours" concept, Li and Xu (2025), applies here as well. We have added this reference at the end of the sentence to clarify that the deployment of personnel and equipment according to damage severity is directly supported by this study, which addresses how multi-class damage classification frameworks enable differentiated response strategies — from prioritizing rescue operations in severely damaged areas to efficiently allocating resources in less affected zones.

Original manuscript
...while also facilitating the deployment of personnel and equipment with appropriate capacity and characteristics to the affected areas.

Revised manuscript:
They also facilitate the deployment of personnel and equipment with appropriate capacity and characteristics to locations where higher damage levels are expected **(Li and Xu, 2025)**.

15. Line 50: What is meant by time gain of 6-12 hours?

We thank the reviewer for this comment. The phrase "time gain of 6–12 hours" was ambiguous, as it did not specify the point of comparison. The sentence has been revised to clarify that this refers to the advantage of completing damage assessment within the first 6–12 hours following an earthquake, compared to conventional field-based approaches, which typically take longer to deliver comparable information.

Original manuscript
Previous studies have demonstrated that a time gain of approximately 6–12 hours in post-earthquake damage assessment can substantially improve rescue outcomes, highlighting the critical importance of rapid and automated damage classification approaches (Erdik et al. 2011;Li and Xu 2025)

Revised manuscript:
Previous studies have shown that early damage assessment, completed within the first 6–12 hours following an earthquake, can substantially improve search-and-rescue outcomes compared to conventional field-based approaches (Erdik et al. 2011; Li and Xu 2025).

16. Line 54: What is “contemporary disaster management”?

We thank the reviewer for this comment. The term "contemporary disaster management" was intended to refer broadly to current earthquake emergency response practices, but its scope was ambiguous. It has been replaced with "early-phase earthquake emergency response and disaster risk management", which more precisely reflects the operational context addressed in this study.

Original manuscript
Accordingly, near-real-time damage estimation systems that become operational immediately after an event have emerged as a cornerstone of contemporary disaster management.

Revised manuscript:
Accordingly, near-real-time damage estimation systems that become operational immediately after an event have emerged as a cornerstone of early-phase earthquake emergency response and disaster risk management.

17. Line 54: What are “Such platforms”?

We thank the reviewer for this comment. The phrase "Such platforms" referred to the near-real-time damage estimation systems mentioned in the preceding sentence. To avoid ambiguity, this has been replaced with "These rapid damage estimation platforms", making the reference explicit.

Original manuscript
...a cornerstone of contemporary disaster management. Such platforms aim to provide decision-makers with early, spatially explicit estimates of earthquake-induced ground shaking and potential impacts, thereby supporting the prioritization of response actions during the initial phase.

Revised manuscript:
…a cornerstone of early-phase earthquake emergency response and disaster risk management. These rapid damage estimation platforms aim to provide decision-makers with early, spatially explicit estimates of earthquake-induced ground shaking and potential impacts, thereby supporting the prioritization of response actions during the initial phase.

18. Line 55: What are “spatially explicit estimates”?

The phrase 'spatially explicit estimates' has been replaced with wording that specifies these are spatially distributed estimates of ground-shaking intensity and structural impacts across the affected area.

Original manuscript
Such platforms aim to provide decision-makers with early, spatially explicit estimates of earthquake-induced ground shaking and potential impacts, thereby supporting the prioritization of response actions during the initial phase

Revised manuscript:
These rapid damage estimation platforms aim to provide decision-makers with early, spatially distributed estimates of earthquake-induced ground shaking intensity and potential structural impacts, thereby supporting the prioritization of response actions during the initial phase.

19. Line 56: I do not believe that HAZUS has capability for near-real-time damage estimation system. Please verify for all listed systems.

We thank the reviewer for this important observation, and confirm it is correct. The original wording could be misread as implying that all listed systems, including HAZUS and OpenQuake, operate as near-real-time platforms. We have revised the manuscript to describe all listed tools generally as "earthquake damage estimation platforms" and to explicitly clarify that only some of them (PAGER, ELER, AFAD-RED, and REDAS) are used in operational rapid impact estimation workflows, whereas HAZUS, SELENA, and OpenQuake are primarily applied in scenario-based studies and do not operate as fully automated near-real-time systems.

Original manuscript
Representative implementations include HAZUS (FEMA, 2025), ELER (Hancilar et al., 2010), PAGER (Wald et al., 2010), SELENA (NORSAR and ICG 2026), OpenQuake (Silva et al., 2013), AFAD-RED (AFAD, 2013), and REDAS (REDAS, 2020

Revised manuscript:
Representative implementations of earthquake damage estimation platforms include HAZUS (FEMA, 2025), ELER (Hancilar et al., 2010), PAGER (Wald et al., 2010), SELENA (NORSAR and ICG 2026), OpenQuake (Silva et al., 2013), AFAD-RED (AFAD, 2013), and REDAS (REDACt-Project, 2023). Among these, systems such as PAGER, ELER, AFAD-RED, and REDAS are used in operational rapid impact estimation workflows, whereas HAZUS, SELENA, and OpenQuake are primarily applied in scenario-based studies and do not operate as fully automated near-real-time systems.

20. Line 59: What is meant by “reliability”? Does it refer to accuracy, or the real-time availability and uptime of the products?

We thank the reviewer for this comment. "Reliability" in this context refers to the accuracy of the damage estimates and how well the underlying ground-shaking representation reflects actual site conditions, not to system uptime or availability. To remove this ambiguity, the sentence has been revised to explicitly pair "accuracy and reliability" and to specify what constrains them.

Original manuscript
Despite their operational value, the reliability of rapid post-event damage products is fundamentally constrained by the accuracy of the underlying ground-shaking representation and the resulting damage estimation outputs.

Revised manuscript
Despite their operational value, the accuracy and reliability of rapid post-event damage products are fundamentally constrained by how well the underlying ground-shaking fields represent actual site conditions and are translated into damage-estimation outputs.

21. Line 61: Please provide a reference.

We thank the reviewer for this comment. A reference has been added. The description of scenario-based operational frameworks that combine GMPE-based ground-motion fields with building fragility functions is consistent with the methodology described in Erdik et al. (2011), which has been cited accordingly.

Original manuscript
...most operational systems rely on scenario-based, event-specific ShakeMaps derived from regional site parameters (e.g., V_{s30}, $Z_{1.0}$, $Z_{2.5}$) and Ground Motion Prediction Equations (GMPEs), combined with fragility functions representing the building stock.

Revised manuscript
..most operational systems rely on scenario-based, event-specific ground-motion parameter maps derived from regional site parameters (e.g., V_{s30}, $Z_{1.0}$, $Z_{2.5}$; for definitions see Section 3.2) and Ground Motion Prediction Equations (GMPEs), combined with fragility functions representing the building stock (Erdik et al. 2011).

22. Line 62: Please define the terms: V_{s30}, $Z_{1.0}$, and $Z_{2.5}$.

We thank the reviewer for this comment. The terms V_{s30}, $Z_{1.0}$, and $Z_{2.5}$ are defined in detail in Section 3.2 of the manuscript. To improve readability at first mention, a brief parenthetical note has been added directing readers to Section 3.2 for the full definitions.

Original manuscript
….regional site parameters (e.g., V_{s30}, $Z_{1.0}$, $Z_{2.5}$) and Ground Motion Prediction Equations (GMPEs)..

Revised manuscript
...regional site parameters (e.g., V_{s30}, $Z_{1.0}$, $Z_{2.5}$; for definitions see Section 3.2) and Ground Motion Prediction Equations (GMPEs)..

23. Line 64: What is meant by "inevitably introduces both epistemic and aleatory uncertainties". Please clarify and add reference.

We thank the reviewer for this comment. The sentence has been revised to explicitly define both terms: epistemic uncertainty arises from incomplete knowledge of source parameters, site conditions, and model limitations, whereas aleatory uncertainty reflects the inherent randomness in ground-motion variability.

Original manuscript
...it inevitably introduces both epistemic and aleatory uncertainties

Revised manuscript
it inevitably introduces uncertainty of two kinds: epistemic uncertainty, arising from incomplete knowledge of source parameters, site conditions, and model limitations; and aleatory uncertainty, which reflects the inherent randomness in ground-motion variability.

24. Line 64: What is meant by "persistent methodological gap"?

We thank the reviewer for this comment. The phrase "persistent methodological gap" referred to the systematic difference between scenario-based ground-motion projections and actual post-event observations. To avoid ambiguity, this phrase has been replaced with "systematic discrepancy," and the surrounding text restructured to improve clarity.

Original manuscript
Consequently, a persistent methodological gap may emerge between pre-event scenario-based projections and post-event damage patterns governed by the actual ground-motion field and observed impacts.

Revised manuscript
Consequently, a systematic discrepancy may emerge between pre-event scenario-based projections and post-event damage patterns governed by the actual ground-motion field and observed impacts.

25. Line 67: What is meant by “distort intermediate-period demand”?

We thank the reviewer for this comment. The term "distort" was used to describe the systematic deviation of intermediate-period spectral demand from GMPE-based predictions due to basin-induced amplification effects. To avoid ambiguity, this has been replaced with "substantially alter... relative to GMPE-based predictions," which more precisely describes the direction and reference point of this deviation. As part of restructuring the Introduction to reduce redundancy (see also our response to Comment 26), this clarified sentence now appears later in the Introduction, where the same concept is developed in full.

Original manuscript
In complex deep-basin urban environments, where amplification effects may distort intermediate-period demand, this gap can become particularly pronounced, complicating the balance between timeliness and representational accuracy in early-phase decision support.

Revised manuscript
In deep-basin urban environments, where basin-induced amplification may substantially alter intermediate-period spectral demand relative to GMPE-based predictions, scenario-based ground-shaking representations tend to underestimate spectral demand relative to recorded observations.

26. Line 67: What are the “complications” to the balance?

We thank the reviewer for this comment. The phrase "complicating the balance between timeliness and representational accuracy" has been removed, as its meaning was not sufficiently precise. The surrounding text has been restructured to directly describe the systematic discrepancy between scenario-based projections and post-event observations in deep-basin environments, conveying the intended meaning without the ambiguous phrasing.

Original manuscript
...this gap can become particularly pronounced, complicating the balance between timeliness and representational accuracy in early-phase decision support.

Revised manuscript
...scenario-based ground-shaking representations tend to underestimate spectral demand relative to recorded observations, which may in turn lead to underestimation of exceedance probabilities associated with severe damage states.

27. Line 69: Please further describe “system performance in terms of statistical prediction accuracy and bias reduction”

We thank the reviewer for this comment. The phrase has been clarified to explicitly state that performance has been evaluated by quantifying residuals between predicted and observed ground-motion intensities and by assessing reductions in systematic bias in shaking and loss estimates.

Original manuscript
Beyond these methodological constraints, existing research has predominantly evaluated system performance in terms of statistical prediction accuracy and bias reduction in ground-shaking and loss estimation (e.g., Silva et al. 2025).

Revised manuscript
Beyond these methodological constraints, existing research has predominantly evaluated the performance of rapid impact estimation frameworks by quantifying residuals between predicted and observed ground-motion intensities and by assessing reductions in systematic bias in shaking and loss estimates.

28. Line 71: Please add references for "previous studies".

We thank the reviewer for this comment. References have been added to support this statement. Silva et al. (2025) is cited for the demonstrated benefits of seismic monitoring in improving ground-shaking estimation and loss assessment, and Gülerce et al. (2022) is cited for the evaluation of GMPE performance against recorded strong motions from this same earthquake.

Original manuscript
While previous studies have demonstrated that observational calibration improves ground-motion and damage model reliability, limited attention has been paid to….

Revised manuscript
While previous studies have demonstrated that observational calibration improves the accuracy of ground-motion and damage estimates (Silva et al. 2025; Gülerce et al., 2022), limited attention has been paid to...

29. Line 74: What is meant by "optimistic bias"?

We thank the reviewer for this comment. The term "optimistic bias" has been replaced with a more explicit description: scenario-based ground-shaking representations tend to underestimate spectral demand relative to recorded observations, which may in turn lead to underestimation of exceedance probabilities associated with severe damage states.

Original manuscript
In deep-basin urban environments, where basin-induced amplification may alter intermediate-period spectral demand, optimistic bias in scenario-based ground-shaking representations may systematically underestimate exceedance probabilities associated with severe damage states.

Revised manuscript
In deep-basin urban environments, where basin-induced amplification may substantially alter intermediate-period spectral demand relative to GMPE-based predictions, scenario-based ground-shaking representations tend to underestimate spectral demand relative to recorded observations, which may in turn lead to underestimation of exceedance probabilities associated with severe damage states.

30. Line 75: Please provide reference.

We thank the reviewer for this comment. A reference to Silva et al. (2025) has been added. While Silva et al. (2025) primarily address the benefits of seismic monitoring for reducing bias and uncertainty in ground-shaking and loss estimation, their findings provide direct support for the premise that scenario-based representations lacking observational updating may lead to systematic underestimation of impact metrics — a bias that, in an operational context, may translate into governance-level misclassification during the early response phase.

Original manuscript
Such underestimation does not merely constitute a modelling discrepancy; rather, it may assign affected areas to a lower operational response regime than warranted, thereby inducing governance-level misclassification during the early phase of emergency management.

Revised manuscript
Such underestimation is not merely a modelling discrepancy; rather, it may assign affected areas to a lower operational response regime than warranted, a condition referred to in this study as governance-level misclassification during the early phase of emergency management (Silva et al. 2025).

31. Line 76: It is unclear what the sentence is intended to say. It looks repetitive.

We thank the reviewer for this observation, which we agree is well taken. The sentence in question repeated the definition of governance-level misclassification already conveyed in the preceding sentence. It has been removed, and the definition has been integrated directly into the first sentence to improve clarity and eliminate redundancy.

Original manuscript
….thereby inducing governance-level misclassification during the early phase of emergency management. Operational misclassification refers to the assignment of affected areas to a lower response category due to underestimation of cumulative damage exceedance probabilities.

Revised manuscript
..a condition referred to in this study as governance-level misclassification during the early phase of emergency management (Silva et al. 2025).

32. Line 79: what is "response category"?

We thank the reviewer for this comment. This sentence has been removed as part of the restructuring addressed in our response to Comment 31. The concept of response categorization is now conveyed through the explicit description of the transition between monitoring-oriented and mobilization-oriented response regimes, which provides sufficient operational context without requiring a separate, undefined term

Original manuscript
Operational misclassification refers to the assignment of affected areas to a lower response category due to underestimation of cumulative damage exceedance probabilities.

Revised manuscript
This sentence has been removed.

33. Line 80: What are "operational response thresholds", and is the author referring specifically to Turkey or globally?

We thank the reviewer for this comment. The term "operational response thresholds" is now defined explicitly in Section 3 (Methodology) as the intensity or damage-probability levels at which emergency response escalates from monitoring to active search-and-rescue and resource mobilization. We clarify that this notion is used in a generic sense and is not specific to Türkiye: threshold-based escalation of emergency actions is a common feature of earthquake emergency response frameworks worldwide, although the specific numerical thresholds and categories vary by country and institution.

Original manuscript
The explicit linkage between ShakeMap-driven damage exceedance probabilities and operational response thresholds has received comparatively limited attention in the disaster risk science literature.

Revised manuscript
….exceedance probabilities therefore provide a more robust basis for informing threshold-based emergency response decisions, here referred to as operational response thresholds (Erdik et al., 2011). In this context, 'operational response thresholds' refer to the intensity or damage-probability levels at which emergency response is escalated from monitoring to active search-and-rescue and resource mobilization. This notion is not specific to Türkiye; threshold-based escalation of emergency actions is a common feature of earthquake emergency response frameworks worldwide, although the specific numerical thresholds and categories vary by country and institution.

34. Line 101: Are these Mw the ones published immediately after the event, or the Mw as of 2026 by the different agencies? If the latter, why would there be epistemic uncertainty of 0.4 points 5 years after the event?

We thank the reviewer for this important clarification request. The Mw values cited are the initial magnitudes reported by the respective agencies immediately following the 30 October 2020 Samos earthquake, as consistently referenced in the literature (e.g., Demirel et al., 2022), not updated or reconciled values. The manuscript has been revised to state this explicitly. The variability among these initial agency reports reflects the epistemic uncertainty inherent in rapid, near-real-time source characterization — precisely the context in which these values are discussed in this study.

Original manuscript
The moment magnitude of the earthquake was reported as Mw = 6.6 by the Disaster and Emergency Management Authority of Türkiye (AFAD, 2020), Mw = 6.9 by the Kandilli Observatory and Earthquake Research Institute (KOERI, 2020) and Mw = 7.0 by the United States Geological Survey (USGS, 2020)

Revised manuscript
The moment magnitude of the earthquake was initially reported as Mw = 6.6 by the Disaster and Emergency Management Authority of Türkiye (AFAD, 2020), Mw = 6.9 by the Kandilli Observatory and Earthquake Research Institute (KOERI, 2020), and Mw = 7.0 by the United States Geological Survey (USGS, 2020).

35. Line 115: The sentence structure is circular as the author mentions: although structural deficiency contributed, one primary factor was local site effects + "structural vulnerability".

We thank the reviewer for this observation, which we agree is correct. This circularity has been addressed as part of the same revision described in our responses to Comments 4 and 5. The paragraph now presents structural deficiencies and local site response as two distinct, sequential factors rather than conflating them in a single circular clause: the discussion of structural deficiencies is followed by an explicit cross-reference to Section 3.3 for detailed building typology information, while local site response — specifically the resonance between basin-amplified ground motions and the natural periods of the dominant mid-rise building stock — is presented as the primary factor governing the spatial concentration of damage, with supporting references (Cetin et al., 2022; Demirel et al., 2022).

36. Line 124: What is meant by "seismic hazard assessment based solely on standardized code-based approaches", especially in the context of author's focus on ShakeMaps?

We thank the reviewer for this comment. The phrase has been revised to specify exactly what is meant: generic site amplification factors prescribed in standardized seismic design codes, as opposed to site-specific characterization that accounts for basin geometry and local soil conditions. This distinction is now made explicit and connected directly to the scenario-based ground-motion representations discussed throughout the manuscript.

Original manuscript
Consequently, seismic hazard assessment based solely on standardized code-based approaches is insufficient; instead, detailed microzonation studies are imperative….

Revised manuscript
Consequently, ground-motion estimation based solely on standardized code-based site amplification factors is insufficient in such environments. Site-specific characterization and detailed seismic microzonation studies, which systematically characterize spatial variations in soil conditions, site amplification, and ground-motion response at the local scale, are necessary to capture local site effects and basin-induced amplification accurately in scenario-based ground-motion representations (Demirel et al., 2022)

37. Line 125: Please explain "detailed microzonation studies".

We thank the reviewer for this comment. The term has been clarified in the revised manuscript: seismic microzonation refers to the systematic characterization of spatial variations in soil conditions, site amplification, and ground-motion response at the local scale, providing the site-specific data necessary to improve ground-motion representations in complex urban environments such as the Izmir Bay basin. A reference to Demirel et al. (2022) has been retained to support this statement.

Original manuscript
..instead, detailed microzonation studies are imperative to adequately capture local site effects and basin-induced ground-motion amplification (Demirel et al., 2022).

Revised manuscript
Site-specific characterization and detailed seismic microzonation studies, which systematically characterize spatial variations in soil conditions, site amplification, and ground-motion response at the local scale, are

necessary to capture local site effects and basin-induced amplification accurately in scenario-based ground-motion representations (Demirel et al., 2022).

38. Line 139: The previous statement does not support the inclusion of "in combination with local site amplification effects".

We thank the reviewer for correctly identifying this inconsistency. The preceding statement notes that recorded ground motions remained below code-prescribed design levels, which primarily supports the role of structural deficiencies rather than site amplification. The phrase "in combination with local site amplification effects" has therefore been removed to avoid this logical contradiction.

Original manuscript
This observation suggests that structural vulnerabilities and variations in code implementation and construction quality contributed significantly to the observed damage, in combination with local site amplification effects.

Revised manuscript
This observation suggests that structural vulnerabilities and variations in code implementation and construction quality contributed significantly to the observed damage patterns in the Bayraklı district

39. Line 142: The author has described the structural deficiencies, but then point to the relevance of "basin-related amplification characteristics". It is unclear how the structural deficiencies reason given by the author instead points to site amplification as the reason for excessive damage.

We thank the reviewer for correctly identifying this missing logical connection. A transitional sentence has been added to clarify that while basin-induced amplification did not independently exceed code-based design levels, it amplified intermediate-period spectral demand sufficiently to push structurally deficient mid-rise frames towards severe damage and collapse. This makes the interaction between the two contributing factors explicit and provides a clearer motivation for the exceedance-based assessment framework adopted in this study.

Original manuscript
.underscores the substantial role of structural and construction-related deficiencies (Askan et al., 2021). These basin-related amplification characteristics are particularly relevant for exceedance-based damage assessment frameworks.

Revised manuscript
..underscores the substantial role of structural and construction-related deficiencies (Askan et al., 2021). While these basin-related amplification characteristics did not, by themselves, exceed code-based design levels, they amplified intermediate-period spectral demand sufficiently to push structurally deficient mid-rise frames towards severe damage and collapse.

40. Line 146: Since the author has only analyzed one event, suggest changing the statement to "… this deep-basin urban environment…"

We thank the reviewer for this suggestion, which we have adopted. The reference to the study context has been expressed in the singular form — "this deep-basin urban setting" — which accurately reflects that the analysis is based on a single case study.

Original manuscript

In such settings, scenario-based ShakeMaps that do not fully capture basin-induced effects may underestimate exceedance probabilities associated with operationally critical damage thresholds. Consequently, deep-basin urban environments provide a suitable context for examining how representational bias in ground-shaking fields may translate into governance-level misclassification during the early response phase.

Revised manuscript
In such a deep-basin urban setting, scenario-based ground-motion parameter maps that do not fully capture basin-induced effects may underestimate exceedance probabilities associated with operationally critical damage thresholds, and thus provide a suitable context for examining how representational bias in ground-shaking fields may translate into governance-level misclassification during the early response phase.

41. Line 164: Please define M, E, C.

We thank the reviewer for this comment. The abbreviations M, E, and C are now defined at their first occurrence, denoting Moderate, Extensive, and Collapse damage states, respectively.

Original manuscript
the primary decision-relevant metric is defined as the exceedance probability of cumulative damage states (e.g., {M+E+C}), rather than absolute damage counts..

Revised manuscript
.the primary decision-relevant metric is defined as the exceedance probability of cumulative damage states (e.g., {M+E+C}, where M = Moderate, E = Extensive, and C = Collapse), rather than absolute damage counts..

42. Line 165: Please provide supporting reference. This seems counter-intuitive as emergency responders would rather prefer to have absolute damage counts, but it is in fact the alleatory uncertainty from modeling limitations that restricts providing counts, and instead probabilities are provided.

We thank the reviewer for this comment. We acknowledge that, intuitively, emergency responders might prefer absolute damage counts. REDAS can perform damage and loss estimation at the city, district, and town levels, which makes it possible to identify the areas where damage is likely to be concentrated and thereby guide disaster management authorities in prioritising their response. Within this framework, however, rapid post-event loss estimation involves substantial uncertainty, both in the underlying ground-motion representation and in the damage model itself, consistent with the general uncertainty challenges discussed in Erdik et al. (2011). This makes exact building-level counts unreliable within the short operational timeframe available in the first hours after an earthquake. Exceedance probabilities of aggregated damage states are less sensitive to this uncertainty and therefore provide a more stable basis for defining response thresholds. The manuscript has been revised accordingly.

Original manuscript
…the primary decision-relevant metric is defined as the exceedance probability of cumulative damage states (e.g.,{M+E+C}), rather than absolute damage counts, as these probabilities directly inform operational response thresholds during the early phase of emergency management.

Revised manuscript
…the primary decision-relevant metric is defined as the exceedance probability of cumulative damage states (e.g., {M+E+C}, where M = Moderate, E = Extensive, and C = Collapse), rather than absolute damage counts. This choice reflects the fact that modelling uncertainties in rapid post-event assessment make exact building-level counts unreliable; exceedance probabilities therefore provide a more robust basis for informing threshold-based emergency response decisions, here referred to as operational response thresholds (Erdik et al., 2011).

43. Line 166: What is meant by “theoretical damage distributions”?

We thank the reviewer for this comment. The term "theoretical damage distributions" has been replaced with "model-based damage distributions," which more accurately reflects that these are computational estimates derived by combining ground-motion parameter maps with fragility curves, rather than purely theoretical constructs.

Original manuscript
Based on the resulting ground-motion representation, theoretical damage distributions were derived by combining the ShakeMaps with fragility curves representing different building classes within the study area.

Revised manuscript
Based on the resulting ground-motion representation, model-based damage distributions were derived by combining the spatially distributed ground-motion parameter maps with fragility curves representing different building classes within the study area.

44. Line 211-216: The author has listed three sources of V_{s30}. It is unclear which source was ultimately used, and if multiple sources were used, how they were combined.

We thank the reviewer for correctly identifying that the data integration procedure was not sufficiently described. The Vs30 distribution was derived in two steps: the USGS topographic slope-based dataset was used as the regional baseline, and values at locations where locally derived estimates were available from geotechnical borehole logs were subsequently used to refine it. The Izmir Metropolitan Municipality shear-wave velocity contour maps were used solely for validation purposes, to confirm the spatial consistency of the adopted model, and were not blended into the final Vs30 field. This procedure has been clarified in the revised manuscript.

Original manuscript
The spatial distribution of Vs30 was derived through multi-source data integration. As a baseline, the 'Topographic Slope as a Proxy for Seismic Site Conditions' dataset provided by the USGS (USGS, 2023) was adopted. This regional model was subsequently refined using locally derived Vs30 estimates obtained from available geotechnical borehole logs within the study area. The resulting Vs30 distribution was further evaluated through comparison with equivalent shear-wave velocity contour maps produced for the Izmir Metropolitan Municipality…

Revised manuscript
The spatial distribution of Vs30 was derived through a two-step process. First, the USGS topographic slope-based Vs30 dataset (USGS, 2023) was adopted as the regional baseline. This baseline was then refined by replacing values at locations where locally derived Vs30 estimates were available from geotechnical borehole logs within the study area. The resulting Vs30 distribution was subsequently validated through comparison with the shear-wave velocity contour maps produced for the Izmir Metropolitan Municipality as part of the Izmir Earthquake Scenario and Master Plan (Izmir Metropolitan Municipality and Boğaziçi University 1999), which confirmed the spatial consistency of the adopted site condition mode.

45. Line 230: Please provide a reference.

We thank the reviewer for this comment. References have been added to support this statement. Day et al. (2008) directly addresses basin depth effects on long-period ground-motion modelling, and Askan et al. (2021) provides evidence of such deviations in the context of the 2020 Samos earthquake and the İzmir Bay basin.

Original manuscript
It is well recognized that, particularly in regions characterized by deep and complex basin geometries, uncertainties associated with the empirical estimation of Z1.0 and Z2.5 may lead to systematic deviations in the modelling of long-period ground-motion demands.

Revised manuscript
It is well recognized that, particularly in regions characterized by deep and complex basin geometries, uncertainties associated with the empirical estimation of $Z_{1.0}$ and $Z_{2.5}$ may lead to systematic deviations in the modelling of long-period ground-motion demands (Askan et al., 2021; Day et al., 2008).

46. Line 233: How is "intermediate-period" defined?

We thank the reviewer for this comment. The term "intermediate-period" is now explicitly defined as T = 0.6–1.0 s, corresponding to the fundamental vibration periods of the dominant mid-rise reinforced concrete building stock in the study area. This definition is consistent with the period range identified as critical throughout the manuscript.

Original manuscript
In deep-basin environments, uncertainties in sediment depth proxies may disproportionately influence intermediate-period spectral demand, thereby altering cumulative damage exceedance probabilities even in the absence of extreme peak ground acceleration values.

Revised manuscript
In deep-basin environments, uncertainties in sediment depth proxies may disproportionately influence intermediate-period spectral demand, here defined as T = 0.6–1.0 s, thereby altering cumulative damage exceedance probabilities even in the absence of extreme peak ground acceleration values.

47. Line 234: Please provide reference.

We thank the reviewer for this comment. A reference to Cetin et al. (2022) has been added. This study directly demonstrates that intermediate-period spectral amplification in the İzmir Bay basin substantially increased seismic demand on mid-rise buildings despite relatively modest PGA values, which supports the statement made in the manuscript.

Original manuscript
In deep-basin environments, uncertainties in sediment depth proxies may disproportionately influence intermediate-period spectral demand, thereby altering cumulative damage exceedance probabilities even in the absence of extreme peak ground acceleration values.

Revised manuscript
In deep-basin environments, uncertainties in sediment depth proxies may disproportionately influence intermediate-period spectral demand, here defined as T = 0.6–1.0 s, corresponding to the fundamental vibration periods of the dominant mid-rise building stock in the study area, thereby altering cumulative damage exceedance probabilities even in the absence of extreme peak ground acceleration values (Cetin et al., 2022).

48. Eq 1: Why is velocity represented by m/sn instead of m/s? What does "n" refer to?

We thank the reviewer for pointing out this unit inconsistency. The unit "m/sn" where "sn" is a Turkish abbreviation for seconds, was inadvertently used in Equation (1) and in Figure 2, and has now been corrected to "m/s" throughout the manuscript, including within the figure itself.

49. Line 241: What does it mean by "characterized by predominant site periods in the range of 1.0–3.0 s"?

We thank the reviewer for this comment. The phrase has been clarified to explicitly state that this refers to the fundamental natural periods of the soil deposits at a given location, which control the frequency content of site amplification.

Original manuscript
Investigations conducted for the 2020 Samos earthquake indicate that districts such as Bayraklı and Karşıyaka, characterized by predominant site periods in the range of 1.0–3.0 s, exhibited pronounced amplification of spectral accelerations, particularly within the 0.5–1.5 s period range (Cetin et al., 2022).

Revised manuscript
Investigations conducted for the 2020 Samos earthquake indicate that districts such as Bayraklı and Karşıyaka, where the fundamental natural periods of the soil deposits range from 1.0 to 3.0 s, exhibited pronounced amplification of spectral accelerations, particularly within the 0.5–1.5 s period range (Cetin et al., 2022).

50. Line 245: Which "observations" is the author referring to?

We thank the reviewer for this comment. The phrase "these observations" has been made explicit: it refers to the pronounced amplification of spectral accelerations in the 0.5–1.5 s period range observed in Bayraklı and Karşıyaka during the 2020 Samos earthquake. An additional reference to Askan et al. (2021) has also been included to further support the statement regarding GMPE limitations in deep sedimentary environments.

Original manuscript
These observations reported in the literature suggest that standard Ground Motion Prediction Equations (GMPEs) may not adequately represent basin-induced amplification effects and long-period ground-motion characteristics, especially under deep sedimentary basin conditions.

Revised manuscript
These findings, specifically the pronounced amplification of spectral accelerations in the 0.5–1.5 s period range observed in Bayraklı and Karşıyaka during the 2020 Samos earthquake (Cetin et al., 2022), suggest that standard GMPEs may not adequately represent basin-induced amplification effects and long-period ground-motion characteristics in deep sedimentary environments (Askan et al., 2021).

51. Line 248: What does it mean by “was adopted through bias correction, not as an optional enhancement, but as a methodological requirement”

We thank the reviewer for this comment. The sentence now makes this reasoning explicit: GMPE-based predictions alone were shown to substantially underestimate intermediate-period spectral demand in this deep-basin setting, which is why observational updating using observed strong-motion station data was treated as a methodological requirement rather than an optional enhancement.

Original manuscript
Consequently, in this study, the refinement of ground-motion estimates initially derived from theoretically defined site and basin parameters using observed strong-motion station data was adopted through bias correction, not as an optional enhancement, but as a methodological requirement to improve hazard and risk representation.

Revised manuscript
Consequently, in this study, bias correction using observed strong-motion station data is essential rather than optional, given that GMPE-based predictions alone were shown to substantially underestimate intermediate-period spectral demand in this deep-basin setting.

52. Figure 3: The panel on the left is difficult to read. Usage of green color both for the background and the damage state makes it difficult to distinguish them. The resolution appears low to differentiate among buildings.

We thank the reviewer for this observation, which we agree is correct. The figure has been fully regenerated: the background is now a high-resolution satellite basemap (replacing the flat green polygon fill), and the building-damage points retain their original color legend, which is now clearly distinguishable against the satellite imagery. The neighborhood boundary and building density are now legible at building-level scale.

Original manuscript (Figure 3):
Left panel used a flat green polygon fill for the Bayraklı/İzmir background, with damage-state points overlaid in a similarly green-toned palette, making the neighborhood boundary and individual buildings difficult to distinguish at the printed resolution.
Revised manuscript:
Figure 3 has been regenerated with a high-resolution satellite basemap (© Google Maps 2015, retrieved via QGIS QuickMapServices) showing individual buildings and streets, with the Mansuroğlu Neighborhood boundary and building-level damage points overlaid using the original, clearly distinguishable color legend.

53. Line 261: What does “harmonization” refer to?

We thank the reviewer for this comment. The term "harmonized" referred to the process of reclassifying each building in the inventory according to HAZUS structural typology categories, based on available attributes such as structural system type, number of stories, and assumed seismic design level. To avoid ambiguity, the term has been replaced with an explicit description of this reclassification procedure.

Original manuscript
The dataset was subsequently harmonized with the HAZUS building classification scheme by considering key attributes such as structural system type, number of stories, and level of engineering design.

Revised manuscript
Each building was subsequently reclassified according to the HAZUS building classification scheme based on its structural system type, number of stories, and assumed level of engineering design.

54. Line 264: Please provide reference.

We thank the reviewer for this comment. HAZUS fragility functions were selected because they offer a comprehensive, standardized, and internationally validated set of structural fragility models covering a wide range of building typologies. HAZUS models have been extensively tested and applied in diverse seismic risk contexts worldwide, including Iran (Farahani et al., 2020), South Korea (Choi et al., 2019), Vietnam (Nguyen et al., 2025), Algeria (Boutaraa and Chaibedra, 2025), and data-scarce regions such as Burundi (Delves et al., 2026), serving as a methodological foundation for fragility curve development in international applications (FEMA, 2025).

Original manuscript
Although originally developed for the United States building stock, HAZUS fragility functions were adopted in this study due to their broad representation of structural typologies and their widespread use in international seismic risk assessments.

Revised manuscript
Although originally developed for the United States building stock, HAZUS fragility functions were adopted in this study because they offer a comprehensive, standardized, and internationally validated set of structural fragility models covering a wide range of building typologies. HAZUS models have been extensively tested and applied in diverse seismic risk contexts worldwide, including Iran (Farahani et al., 2020), South Korea (Choi et al., 2019), Vietnam (Nguyen et al., 2025), Algeria (Boutaraa and Chaibedra, 2025), and data-scarce regions such as Burundi (Delves et al., 2026), serving as a methodological foundation for fragility curve development in international applications (FEMA, 2025).

55. Line 278: Please provide a reference.

We thank the reviewer for this comment. The height distribution statistics were derived directly from the building inventory obtained from the Geographic Data Sales Portal of the Republic of Türkiye Ministry of Environment, Urbanization and Climate Change. A reference to this data source has been added.

Original manuscript
In terms of height distribution, the inventory is dominated by mid-rise buildings (4–7 stories, 48.3%) and high-rise buildings (≥8 stories, 36.7%), whereas low-rise buildings (1–3 stories) constitute 13.5% of the stock.

Revised manuscript
In terms of height distribution, the inventory is dominated by mid-rise buildings (4–7 stories, 48.3%) and high-rise buildings (≥8 stories, 36.7%), whereas low-rise buildings (1–3 stories) constitute 13.5% of the stock (SGB, 2026).

56. Line 291: Suggest including the fragility curves for all selected building types to help orient readers about their relative performance.

We thank the reviewer for this suggestion. A figure (Figure 5) with two panels has been added to Section 3.3. The left panel shows the HAZUS spectral-displacement-based fragility curves for the Complete damage state at the Moderate-Code design level for the three spectral-displacement-based building types adopted in this study (C1L, C1M, C1H), allowing direct comparison of their relative structural performance. The right panel shows the intensity-based fragility curves (Slight, Moderate, Extensive, and Complete damage states) for the BI000 classification, used for buildings whose structural system could not be reliably characterized. Together, these panels illustrate the relative vulnerability of all building classification categories used in this study.

Original manuscript:
No fragility curve figure was included in Section 3.3.

Revised manuscript:
A new two-panel figure has been added to Section 3.3:

- **(a)** Probability of exceeding the Complete damage state as a function of spectral displacement (Sd, cm) for C1L, C1M, and C1H building types at the Moderate-Code design level, based on the HAZUS fragility parameters implemented in REDAS.
- **(b)** Probability of exceeding each of the four damage states (Slight, Moderate, Extensive, Complete) as a function of macroseismic intensity for the BI000 classification.

57. Line 301: Please describe “intensity-based fragility functions”.

We thank the reviewer for this comment. The term has been clarified: unlike spectral-demand-based fragility functions, intensity-based functions relate the probability of exceeding a damage state directly to macroseismic intensity rather than spectral demand parameters, and are therefore more suitable for buildings whose structural behavior cannot be reliably characterized through spectral-displacement-based methods.

Original manuscript

These structures were therefore evaluated using intensity-based fragility functions under the BI000 classification.

Revised manuscript

These structures were therefore evaluated using intensity-based fragility functions under the BI000 classification, which relate the probability of exceeding a damage state directly to macroseismic intensity rather than spectral demand parameters, and are therefore more suitable for buildings whose structural behavior cannot be reliably characterized through spectral-displacement-based methods.

58. Line 319: What is meant by “local representational differences”?

We thank the reviewer for this comment. The phrase has been made explicit: it refers to the station-level discrepancies between GMPE-based predictions and observed ground-motion parameters.

Original manuscript

...establishes the basis for interpreting how local representational differences may influence spatial ShakeMaps and exceedance-based damage estimates in subsequent sections.

Revised manuscript

….shows how the differences between GMPE-based predictions and the observed values affect the ground-motion parameter maps and the damage estimates discussed in the following sections.

59. Line 320: How was the radius established: was it based on the outer extents of the neighborhood, or based on a weighted centroid?

We thank the reviewer for this comment. The 3.0 km radius was defined with respect to the approximate geometric centroid of the Mansuroğlu Neighborhood. This threshold was selected to prioritize strong-motion stations most likely to capture the local ground-motion characteristics of the study area and to maximize the influence of spatially proximate observations on the bias correction procedure. A clarification has been added to the manuscript.

Original manuscript

All AFAD strong-motion stations located within a 3.0 km radius of the Mansuroğlu Neighborhood (n = 4) were included in the comparison, ensuring full utilization of the available observational data in the immediate vicinity of the study area.

Revised manuscript

All AFAD strong-motion stations within a 3.0 km radius of the Mansuroğlu Neighborhood (n = 4) were included in the comparison. This radius was chosen because stations within this distance are close enough to represent the local ground-motion characteristics of the study area, which gives them the strongest influence on the bias correction.

60. Figure 5: Can the spectral acceleration plots be enlarged to improve clarity?

We thank the reviewer for this suggestion. The four spectral acceleration plots (stations 3513, 3514, 3520, 3522) have been regenerated. Axis labels and legends have also been enlarged for improved readability.

Original manuscript (Figure 5):

The four spectral acceleration plots were arranged in a single narrow column alongside the station-location map, resulting in small, difficult-to-read plots.

Revised manuscript:
In the revised manuscript (Figure 6), the figure has been updated with the four spectral acceleration plots (3513, 3514, 3520, 3522) rearranged below the station-location map, each significantly enlarged, with larger axis labels and clearly positioned legends.

61. Figure 5: What is "Intensity" in the table?

We thank the reviewer for this comment. The "Intensity" values listed in the table of Figure 5 refer to macroseismic intensity estimates derived from recorded peak ground-motion parameters (PGA and PGV) using the empirical relationships of Wald et al. (1999). This has been clarified in the text accompanying the figure.

Revised manuscript
In the revised manuscript (Figure 6), Intensity values represent macroseismic intensity estimates derived from PGA and PGV using the empirical relationships of Wald et al. (1999).

62. Table 1: Please define the terminology used in the table in the table caption, e.g., S_T, SHA, etc.

We thank the reviewer for this comment. An explanatory note has been added directly below the table: SHA refers to Seismic Hazard Analysis; S0.3, S0.6, S1.0, and S3.0 denote spectral acceleration values at periods of 0.3, 0.6, 1.0, and 3.0 s, respectively; and Improvement (%) refers to the relative reduction in absolute error following station-based calibration.

Revised manuscript
In the revised manuscript (Table 2), SHA: Seismic Hazard Analysis; S0.3, S0.6, S1.0, and S3.0: spectral acceleration at T = 0.3, 0.6, 1.0, and 3.0 s, respectively; Improvement (%): relative reduction in absolute error following station-based calibration.

63. Table1 : Which directional component at each station location is listed in the table?

We thank the reviewer for this comment. The directional component has been clarified: the values listed under "Observed Data" represent the geometric mean of the two horizontal components (E–W and N–S) recorded at each AFAD strong-motion station. This is consistent with the horizontal component representation adopted in the NGA-West2 GMPE framework used in Stage 1 of the analysis.

Revised manuscript
..Improvement (%): relative reduction in absolute error following station-based calibration. Observed ground-motion parameters represent the geometric mean of the two horizontal components (E-W and N-S) recorded at each station.