Response to Reviewer 4

We thank Reviewer for the thorough and candid assessment. The comments are substantive and have helped us identify concrete areas for improvement. We respond to the general and specific comments in turn.

General comments: The manuscript analyses emotional responses on social media following the 2025 Myanmar earthquake, framed as a transboundary disaster in this work. The authors combine large language model for emotion classification and named entity recognition to then construct an emotion–entity coupled network, aiming to model the flow and transformation of emotions beyond typical aggregate sentiment analysis. The topic is interesting and relevant to the NHESS readership, particularly in the context of understanding societal perception of disasters based on social media data and NLP based methods. Such analyses can potentially provide valuable insights for disaster communication, management, and risk awareness. The dataset is large, and the analytical workflow involving LLMs and network modelling demonstrates substantial technical effort. However, the application and interpretation of these methods remain largely qualitative. Much of the analysis relies on visualisations and descriptive network representations, with no quantitative validation of the inferred emotional dynamics. As a result, the link to natural hazards science and disaster risk reduction remains mostly conceptual and is not sufficiently articulated to be actionable or informative for hazard practice.

General Comments Response: The reviewer's overarching concern is that the analytical interpretation remains largely qualitative and visual, with insufficient quantitative validation, and that the connection to actionable natural hazards practice is not adequately articulated. We accept these as legitimate criticisms and address them through the specific responses below. We would note, however, that the emotion–entity coupling framework we propose is methodological in nature, and its primary contribution lies in demonstrating a new analytical pathway for structurally situated emotion analysis rather than in producing operationally prescriptive outputs in this single application. That said, we agree that the manuscript as submitted did not do enough to translate the analytical findings into concrete disaster risk reduction implications, and we will address this directly in the revision.

Specific comments:

1. Quantitative and statistical validation: The manuscript presents hypotheses and interpretations that are primarily supported through visual inspection of networks and figures. The authors should provide some statistical or empirical basis to evaluate these hypotheses, some quantitative metrics will demonstrate value beyond descriptive visualisations.

   Response1: Accepted. As also noted by Reviewer 2, the absence of quantitative validation metrics is a genuine methodological gap. In the revision, we will report precision, recall, and F1 scores for both the emotion classifier and the NER model on held-out test data, as well as inter-annotator agreement (Cohen's κ) for the hybrid annotation procedure. For the network-level findings, we will supplement the visual representations with quantitative descriptive statistics—including degree distribution summaries, emotion proportion by entity category, and dominant emotion frequency counts—to provide an empirical basis for the interpretive claims.

2. Transboundary perspective and novelty: Given that the majority of posts originate from regions not directly affected by the earthquake, the authors should clarify what additional insights are revealed compared to existing studies of domestic or locally affected disasters. It would be useful to explicitly contrast the findings with patterns reported in past work.

Response2: This is a pertinent question that the manuscript did not address with sufficient directness. The transboundary framing is not merely a contextual backdrop but is analytically consequential in several respects: the geographic decoupling between the affected population and the reacting public means that emotional engagement is mediated entirely through discourse rather than personal experience or proximity-based risk, which produces a distinctive emotional topology. Specifically, we observe that empathic and solidarity-oriented emotions cluster around institutional and historical referents (Chinese rescue teams, Wenchuan earthquake memory) rather than around direct victim identification—a pattern less prominent in domestic disaster studies where personal risk salience and local community networks shape the affective landscape differently. In the revision, we will add an explicit comparative paragraph in the discussion that contrasts our findings with patterns reported in prior studies of domestic Chinese disaster discourse, to make the transboundary contribution legible.

3. Introduction structure and event description: The Introduction should be rewritten to be more concise and focused. A clearer description of the earthquake event itself (damage, impacts, timeline, response actions) is needed, particularly where these aspects are later referenced in the analysis. Fig 1 has some details but that is not sufficient.

   Response3: We agree on both counts. The Introduction will be substantially condensed in the revision, eliminating the redundant passages already identified by other reviewers. Regarding the earthquake event description: we will add a dedicated paragraph providing a factual account of the disaster's key parameters—magnitude, epicenter location, casualty figures, timeline of major response actions, and cross-border impacts—supported by appropriate references. This contextual grounding is necessary for readers to interpret the temporal patterns observed in the data and is currently underprovided.

4. Redundancy between Introduction and Literature Review: The literature review is thorough and well explained. However, there is considerable redundancy between the Introduction and Literature Review. I suggest combining or restructuring these sections, as several arguments introduced early on are difficult to follow without the context provided later in the literature review.

   Response4: We will restructure these two sections in the revision. The Introduction will be tightened to focus on the research problem and gap, while the Literature Review will be reorganized around three clearly bounded thematic threads. Overlapping argumentative content—particularly the repeated critique of aggregate sentiment approaches—will be consolidated into a single location. We will consider whether a partial merger of the two sections, following the convention of some NHESS submissions, better serves the paper's logical progression.

5. Discussion and conclusion sections: Similar redundancy appears between the Discussion and Conclusion. These sections could be substantially shortened and made clearer. In particular, the authors should provide more concrete examples of practical insights gained from the analysis, beyond methodological improvements. That is how the public sentiment will be utilized by govt, aid agencies ? Especially in such a transboundary event?

   Response5: The redundancy between these two sections will be addressed by restructuring the Discussion to focus on theoretical and methodological implications and the Conclusion to synthesize broader significance rather than recapitulate findings. More importantly, we will add a dedicated paragraph in the Discussion articulating concrete operational implications: for

governmental agencies, emotion-entity mapping can identify which institutional actors are serving as trust anchors or anxiety amplifiers in real time, informing targeted communication responses; for international aid organizations operating in transboundary contexts, tracking which geographical and organizational entities are associated with positive versus negative emotional clusters can guide strategic messaging to sustain public engagement and counter misinformation. These implications will be framed specifically in relation to the transboundary scenario, where official communication channels face the additional challenge of cross-jurisdictional coordination.

6. Use of highly abstract or philosophical language: The manuscript frequently employs theoretical or philosophical terminology that may be difficult to interpret, particularly for readers with an engineering or applied hazards background like me. The authors should aim to use clearer, more direct language and more explicitly relate their analysis to post disaster activities.

   Response6: We acknowledge that the manuscript's theoretical register is pitched toward a social science readership and will require recalibration for the NHESS audience. In the revision, we will systematically review and simplify passages employing terms such as "latent cognitive infrastructures," "semantic-affective topology," and "affective publics," replacing or supplementing them with more direct formulations that foreground the analytical and applied dimensions. This does not require abandoning the theoretical framing, but it does require making the bridge between theoretical constructs and empirical observations more explicit and accessible.

7. Overall length and clarity: The manuscript would benefit significantly from tightening. Reducing repetition and focusing on key contributions would improve readability and impact.

   Response7: Accepted. The revision will involve a systematic pass to eliminate redundant argumentation across all sections, with particular attention to the Introduction, Discussion, and Conclusion. Our target is a meaningfully shorter manuscript that concentrates analytical energy on the core methodological contribution and empirical findings.

Technical corrections and minor comments

1. Title clarity: The title is difficult to interpret without reading the abstract and introduction. A clearer, more descriptive title would improve accessibility.

   T1–Response: We will revise the title to be more directly descriptive of the study's analytical content and empirical scope, reducing reliance on theoretical shorthand that requires prior familiarity to decode.

2. Terminology consistency: Please use consistent terminology throughout (e.g. “named entity” vs “key entity”).

   T2–Response: We will conduct a systematic review of terminology throughout the manuscript to ensure consistency in all key technical and conceptual terms. This includes, but is not limited to, standardizing "named entity" versus "key entity." We will identify all instances of terminological inconsistency—including variant phrasings for constructs such as the coupled network, emotion classification categories, and entity subcategories—and enforce uniform usage throughout.

3. Simplify complex phrasings such as “latent cognitive infrastructures of digital disaster” should be rewritten in simpler English.

T3–Response: We will conduct a thorough revision of the manuscript's language to replace or clarify overly abstract and philosophical formulations throughout the text. This includes, but is not limited to, expressions such as "latent cognitive infrastructures of digital disaster." All similarly opaque constructions will be rewritten in clearer, more direct English that remains accessible to readers with an applied natural hazards background, while preserving the necessary conceptual precision.

4. "Fine-grained" emotion annotation: But the analysis appears to rely on single-label emotion classification. The authors should clarify why this is considered fine-grained.
   T4–Response: The reviewer correctly notes that single-label classification sits in tension with the "fine-grained" descriptor. We will clarify in the revision that "fine-grained" refers to the use of seven discrete emotion categories (as opposed to the binary or three-class valence schemes prevalent in prior disaster sentiment studies) rather than to multi-label or continuous annotation. This distinction will be made explicit at the point of first use.
5. Annotation examples: Providing concrete examples of emotion and entity annotations (e.g. a small table) would greatly improve transparency.
   T5–Response: We will add a small illustrative table presenting concrete examples of posts alongside their assigned emotion labels and extracted named entities, to improve transparency regarding the annotation output.
6. Figures and presentation:
   Figure 5 should be split into labelled subpanels, with clearer annotations and a more detailed caption.
   Figure 6 needs axes, a labelled y-axis and clear indication of units or normalisation.
   Figures 8–10 are very difficult to interpret in static form. While they may be effective in an interactive dashboard, the manuscript should extract and present more interpretable summaries or quantitative insights from these networks.
   T6–Response: Regarding Figure 5: we will split it into clearly labeled subpanels (a–d for continuous variable distributions; e–f for categorical distributions; g for the correlation heatmap) with a more detailed caption explaining what each panel shows and what the key takeaway is.
   Regarding Figure 6: we will add a labeled y-axis with explicit indication of the unit (post count or normalized proportion) and ensure all axes carry appropriate labels and scale markers.
   Regarding Figures 8–10: we accept that dense static network visualizations impose substantial interpretive burden. In the revision, we will supplement these figures with summary tables reporting key quantitative network properties (top nodes by degree centrality, dominant emotion per entity category, most frequent entity co-occurrence pairs), so that the analytical insights can be read independently of the visual rendering.
7. Keyword filtering: The keywords used for data collection can be more explicitly documented.
   T7–Response: We will add a supplementary table listing all keywords used for data collection, including both Chinese-language terms and their English translations, with a brief note on the selection rationale.
8. Spatial context: Maps showing the geographical distribution of users would help contextualise the spatial patterns discussed in the text.

T8–Response: We appreciate this suggestion. However, the spatial distribution of users is not a primary analytical dimension of the current study, and the province-level distribution has already been presented in Figure 5. Adding a dedicated map visualization would extend the paper's scope beyond its core focus on emotion-entity coupling dynamics. We will nonetheless enhance the caption and descriptive text accompanying Figure 5 to more explicitly contextualize the spatial participation patterns—particularly the notable contribution from Yunnan Province given its geographical proximity to Myanmar—and acknowledge geospatial analysis as a productive direction for future research.

We hope these responses demonstrate that the concerns raised by Reviewer are taken seriously and are addressable within a substantive revision. We remain committed to producing a manuscript that meets the methodological and presentational standards appropriate for the NHESS readership.