➢ Referee #1

Overall comments

The authors present a study titled "Integrating flood-induced population movements into future flood damage estimates in Japan." The manuscript introduces a novel framework for quantifying and predicting population relocation in response to flood events, incorporating climate change effects, to evaluate future population distribution and flood risk in Japan. Overall, I find the study to be well-constructed and the manuscript clearly written. The integration of statistical causal inference with long-term land-use and population projections offers a valuable contribution to socio-hydrological modeling. However, I have identified several critical areas regarding spatial resolution, flood typology, and the modeling of relocation "pull" factors that warrant further attention. I am recommending major revisions to address these methodological limitations and scope concerns. Pending these revisions, I believe the manuscript would be a strong candidate for publication in Natural Hazards and Earth System Sciences (NHESS).

Response:

We sincerely thank the reviewer for their careful reading of our manuscript and constructive evaluation of our study. We are grateful for the positive comments on the novelty of the framework and its contribution to socio-hydrological modeling. In response to the reviewer's comments, we revised the manuscript to clarify the scope of the flood types considered, added a discussion of spatial-resolution limitations, expanded the limitations regarding relocation destinations and pull factors, and clarified the treatment of flood-induced population movements in the proposed framework. We also revised the terminology throughout the manuscript to specify more clearly that the future damage assessment focuses on fluvial flood damage. Detailed responses to each specific comment are provided below.

Comment 1

Methodological Detail (Line 295): The authors cite Yanagihara et al. (2024) for the fluvial flood damage model. However, the current manuscript lacks sufficient detail regarding this model's internal mechanics. Please include a concise summary of the model's core equations or assumptions to ensure the reader can follow the assessment without referring exclusively to prior work.

Response:

We thank the reviewer for this helpful comment. We agree that the original manuscript did not provide sufficient detail on the fluvial flood damage model used to estimate future damage costs. We added a concise summary of the model structure and key assumptions so that readers can understand the assessment without relying exclusively on Yanagihara et al. (2024). The added text describes the inundation model, the condition under which inundation occurs, how rainfall–runoff processes are represented, the limitation that pluvial flooding processes and inland water drainage facilities are not explicitly simulated, and the procedure used to estimate damage costs.

Original text	Revised text
Future damage costs were estimated using an existing fluvial flood damage model (Yanagihara et al., 2024) that considered rainfall changes associated with climate change and land use changes associated with population changes across Japan for the baseline period, the near future, and the end of the 21st century. The costs were evaluated as expected annual values derived from damage calculations for return periods of 30, 50, 100, and 200 years. Maximum inundation depths were calculated using a fluvial flood inundation analysis on a 250 m grid, and corresponding damage costs were estimated based on these depths using depth-specific damage rates. The effects of rainfall changes were incorporated by adjusting the input rainfall used in the fluvial flood inundation analysis. (p. 12, lines 296–302 in the original manuscript)	Future fluvial flood damage costs were estimated using an existing fluvial flood damage model (Yanagihara et al., 2024) that considered rainfall changes associated with climate change and land use changes associated with population changes across Japan for the baseline period, the near future, and the end of the 21st century. In this model, flood inundation is simulated using a two-dimensional unsteady flow model uniformly applied to rivers and floodplains across Japan. Inundation occurs when simulated floods exceed the flood control safety level assigned to each river section. Although rainfall is applied over each river basin and converted into runoff in the model, the model does not explicitly simulate pluvial flooding processes or inland water drainage facilities. The costs were evaluated as expected annual values derived from damage calculations for return periods of 30, 50, 100, and 200 years. Maximum inundation depths were calculated using a fluvial flood inundation analysis on a 250 m grid. Damage costs were

	then estimated by overlaying inundation depths with land-use and asset-value data and applying depth-specific damage rates for each asset category. The effects of rainfall changes were incorporated by adjusting the input rainfall used in the fluvial flood inundation analysis.

Comment 2

Flood Typology and Title: The study focuses exclusively on fluvial flooding. Given the significance of pluvial and coastal flooding in Japan, the title should be more specific. In the title, I suggest changing "future flood damage" to "future fluvial flood damage" with corresponding adjustments in the Abstract and throughout the text to avoid over-generalization.

Response:

We thank the reviewer for this important suggestion. We agree that the original title and terminology could have implied a broader scope than intended. We therefore revised the title by changing "future flood damage" to "future fluvial flood damage." We also revised the terminology throughout the Abstract and main text to specify "fluvial flood damage costs" where the future damage assessment is discussed. At the same time, we did not replace all instances of "flood" with "fluvial flood," because the observed flood-related data used to estimate FIPMs do not fully separate fluvial and pluvial components. We therefore used "fluvial" when referring to the future damage assessment, while retaining "flood" or "flooding" when referring to observed flood events, flood exposure, and flood-related damage records.

Original title	Revised title
Integrating flood-induced population movements into future flood damage estimates in Japan	Integrating flood-induced population movements into future fluvial flood damage estimates in Japan

Comment 3

Spatial Resolution: The use of a 250 m flood grid is relatively coarse for a national-scale study seeking municipal-level precision. This resolution likely misses localized flood risks and micro-topographical features that influence individual relocation decisions. This should be addressed as a limitation in Section 6.3.

Response:

We thank the reviewer for this helpful comment. We agree that the 250 m spatial resolution of the fluvial flood inundation analysis may not fully capture localized flood hazards and micro-topographic features that can influence FIPMs. We have therefore added this point to the limitations.

Added text
Second, the 250 m spatial resolution of the fluvial flood inundation analysis may limit the representation of localized flood hazards and micro-topographic features. Although this resolution is appropriate for maintaining consistency in a nationwide assessment, it may miss small-scale variations in inundation depth and flow velocity that can influence FIPMs. Therefore, this uncertainty should be considered when interpreting fine-scale spatial variations in FIPMs. Future studies should use higher-resolution inundation data, where available, to better assess fine-scale relationships between flood exposure and population movements.

Comment 4

Multimodal Flooding Impacts: The observational dataset used to quantify FIPMs includes residents affected by various water-related disasters (pluvial, coastal, etc.), yet the predictive model only simulates fluvial events. The authors should discuss the net effect of this discrepancy near Line 375 and in the limitations. Specifically, how does the exclusion of pluvial and coastal flood projections affect the estimated impact of FIPMs the results of the study?

Response:

We thank the reviewer for raising this important point. We agree that the relationship between the observational flood-related data used to quantify FIPMs and the predictive fluvial flood damage model required clearer explanation. We revised the Methods to clarify that the observed flood data used in the DiD analyses do not fully separate fluvial and pluvial components. For the grid-cell-level analysis, the flood inundation maps provide estimates of inundated areas and inundation depths during the target heavy rainfall events but do not distinguish between fluvial and pluvial flooding. For the municipality-level analysis, the household damage indicators were calculated to represent household damage associated with both fluvial and pluvial flooding. We also expanded the limitations to discuss how this treatment of flood types may affect the results. The future damage assessment focuses primarily on fluvial flood damage, and coastal flooding is outside the scope of the predictive damage assessment. As described by Yanagihara et al. (2024), the fluvial flood inundation model used for future projections partly includes the effect of pluvial flooding through basin-scale rainfall–runoff processes, but it does not explicitly account for inland water drainage facilities. Therefore, the estimated impacts of FIPMs may be uncertain in areas where drainage-related pluvial flooding or coastal flooding strongly influences population movements. We clarified that the results should be interpreted as an assessment of the impact of FIPMs on future fluvial flood damage, rather than as a comprehensive assessment of all flood types. We note that the text around Line 375 in the original manuscript referred to the IDMC dataset, which was used only for validation and not for the DiD analyses. Therefore, here we focused on clarifying the relationship between the flood-related data used in the DiD analyses and the predictive fluvial flood damage model; the interpretation of the IDMC dataset is addressed separately in our response to Comment 6.

Original text	Revised text
For grid cells in the treatment group, the maximum inundation depth was calculated based on values provided in the flood inundation maps. Details regarding the total population data at the 500 m grid-cell level, data on areas affected by water-related disasters, and flood inundation maps are provided in Sect. S2.	For grid cells in the treatment group, the maximum inundation depth was calculated based on values provided in the flood inundation maps. The flood inundation maps provide estimates of inundated areas and inundation depths during the target heavy rainfall events and do not distinguish between fluvial and pluvial flooding.

(p. 4, lines 92–95 in the original manuscript)	Therefore, the grid-cell-level analysis uses the estimated inundation without separating fluvial and pluvial components. Details regarding the total population data at the 500 m grid-cell level, data on areas affected by water-related disasters, and flood inundation maps are provided in Sect. S2.
To represent flood magnitude, we used the following municipality-level indicators: the proportion of households affected below floor level by flooding, the proportion of households affected above floor level by flooding, and the proportion of households that were completely destroyed (including washed away) by flooding. These indicators were expressed as percentages of the total households, following the method of Okamoto et al. (2023). (pp. 5–6, lines 134–138 in the original manuscript)	To represent flood magnitude, we used the following municipality-level indicators: the proportion of households affected below floor level by flooding, the proportion of households affected above floor level by flooding, and the proportion of households that were completely destroyed (including washed away) by flooding. These indicators were calculated to represent household damage associated with both fluvial and pluvial flooding. They were expressed as percentages of the total households, following the method of Okamoto et al. (2023).

Added text
Third, the treatment of flood types requires careful interpretation. This study focuses primarily on future fluvial flood damage. As described by Yanagihara et al. (2024), the fluvial flood inundation model used for future projections partly includes the effect of pluvial flooding through basin-scale rainfall–runoff processes, but it does not explicitly account for inland water drainage facilities. To maintain consistency with this modeling framework, the flood data used in the DiD analyses were treated without separating fluvial and pluvial components. Thus, neither the observational data nor the predictive model fully separates fluvial and pluvial components, although their treatment of flood types is not identical. Coastal flooding is outside the scope of the predictive damage assessment. This limitation may introduce uncertainty in areas where drainage-related pluvial flooding or coastal flooding strongly influences population movements. Therefore, the results should be interpreted as an assessment of the impact of FIPMs on future fluvial flood damage, rather than as a comprehensive assessment of all flood types.

Comment 5

"Pull" Factors and Relocation Dynamics: The current framework assumes that populations relocate proportionally to existing density. In reality, migration is driven by specific "pull" factors such as employment hubs and family networks—as seen in historical precedents like Hurricane Katrina in the USA, where displacement from New Orleans was highly concentrated in specific regions (e.g., Houston) rather than dispersed uniformly. This simplification should be explicitly noted as a limitation in Section 6.3.

Response:

We thank the reviewer for this important comment and for providing a concrete example. The original manuscript noted that the framework does not explicitly consider the specific destinations of population movements and instead uniformly scales population distributions. However, we agree that the implications of this simplification and the role of destination-specific pull factors were not sufficiently explicit. We therefore revised the limitation to provide examples of pull factors and clarify that the current framework cannot represent cases in which post-flood population movements are concentrated in specific destination areas.

Original text	Revised text
Third, while integrating FIPMs into future population projections, the current framework does not explicitly consider the specific destinations of population movements. Instead, population distributions were uniformly scaled by applying adjustment factors across areas to ensure consistency with overall national or municipal totals. Recent studies have developed population migration models capable of explicitly incorporating pull factors influencing migration (Niu, 2022). Future research should utilize these models to explicitly consider the destinations of population movements, thereby enhancing the methodological framework. (p. 18, lines 436–441 in the original manuscript)	Fifth, while integrating FIPMs into future population projections, the current framework does not explicitly consider the specific destinations of population movements. Instead, population distributions were uniformly scaled by applying adjustment factors across areas to ensure consistency with overall national or municipal totals. This approach does not represent destination-specific pull factors, such as employment opportunities and family or social networks. Therefore, the framework cannot capture cases in which post-flood population movements are concentrated in specific destination areas. Recent studies have developed population migration models capable of explicitly incorporating pull factors influencing migration (Niu, 2022). Future research should utilize these models to explicitly consider the destinations of population movements, thereby enhancing the

	methodological framework.

Comment 6

Temporal Resolution and Validation Discrepancy: Since the model relies on 5-year census data, it inherently captures long-term permanent displacement while overlooking short-term movements. This temporal mismatch may explain the significant discrepancy between the model's 2020 estimate (~16,000 movements) and the IDMC data (~212,000 movements). This distinction between "forced displacement" and "permanent migration" needs a more thorough treatment in the Discussion.

Response:

We thank the reviewer for this important comment. We agree that the distinction between event-based forced displacement and longer-term population or migration changes should be explained more explicitly, because this temporal difference partly explains the discrepancy between our estimate and the internal displacement number from the IDMC dataset. We also note that this study uses two types of demographic data: five-year census data for the grid-cell-level analysis and annual resident registration data for the municipality-level analysis. We have clarified this point in the methodological framework section. We have revised the Discussion to clarify that the IDMC dataset captures event-based forced displacement, including both short-term and longer-term movements, whereas our estimates mainly capture longer-term changes in population distribution and registered migration that remain visible in demographic datasets. We also clarified that our estimates should not be interpreted as the total number of people displaced by flood events, nor as strictly permanent migration.

Original text	Revised text
FIPMs were quantified at both the grid-cell and municipality levels utilizing available demographic data to capture population movements occurring within and between municipalities in response to varying flood magnitudes. Detailed methodologies for quantifying FIPMs are presented in Sect. 3, while methodologies for the subsequent processes of the framework are described in Sect. 4. (p. 4, lines 77–80 in the original manuscript)	FIPMs were quantified at both the grid-cell and municipality levels using available demographic data to capture population movements occurring within and between municipalities in response to varying flood magnitudes. Specifically, FIPMs were quantified using two types of observed datasets that combine flood-related and demographic information. Within-municipality population changes were estimated using flood inundation maps for areas affected by Typhoon Hagibis (2019) and the heavy rainfall event of July 2020, together with 500 m grid-cell population data. Inter-municipality population movements were estimated using municipality-level flood damage records and annual migration data. These data were used to estimate empirical relationships between flood magnitude and

	FIPMs, rather than to directly extrapolate the observed events themselves. The areas covered by the flood inundation maps used in the 500 m grid-cell-level analysis and the treated municipalities used in the municipality-level DiD analysis are shown in Figs. S1 and S2, respectively. Detailed methodologies for quantifying FIPMs are presented in Sects. 2.2 and 2.3, while methodologies for the subsequent processes of the framework are described in Sects. 2.4–2.6.
This discrepancy is largely attributable to the fact that the IDMC dataset includes displacements caused by various types of flooding, capturing all forced movements regardless of duration. In contrast, our estimates are based on total population data at five-year intervals, excluding individuals who returned after displacement events within these intervals. Similarly, our municipality-level estimates rely on annual net migration rates derived from resident registration data, thus excluding short-term evacuees who returned without registering a new address. (pp. 16–17, lines 385–389 in the original manuscript)	This discrepancy is largely attributable to differences in what our estimate and the IDMC internal displacement number represent. First, in terms of flood types, the IDMC dataset may include displacement caused by various flood-related hazards, including coastal flooding. However, displacement corresponding specifically to the flood types represented in our estimates could not be extracted from the IDMC dataset, as noted above. Second, the IDMC dataset captures event-based forced displacement, including both short-term and longer-term movements, irrespective of whether displaced people later returned. In contrast, our estimates are based on total population data at five-year intervals and annual resident registration data, and therefore primarily capture longer-term changes in population distribution and registered migration. Individuals who were temporarily displaced but returned within the census interval, or who did not register a new address, are not captured in our estimates. Thus, our estimates should not be interpreted as the total number of people displaced by flood events, nor as strictly permanent migration, because the available data do not directly identify whether movements are permanent. Rather, they should

	be interpreted as longer-term population changes that remain visible in demographic datasets and can affect future population distribution, land use, and fluvial flood damage estimates.

Comment 7

Code and Data Availability: The current statements stating "available upon request" do not appear to meet the journal's submission guidelines (https://www.natural-hazards-and-earth-system-sciences.net/submission.html). The authors must deposit the underlying research data and code in a public repository or provide a formal explanation as to why it cannot be made publicly accessible.

Response:

We thank the reviewer for this important comment. We agree that the original "available upon request" statement was insufficient under the journal's policy on code and data availability. We are currently preparing a public repository for the code and data while checking which materials can be redistributed. In the revised manuscript, we will provide repository information for materials that can be made public and update the "Code availability" and "Data availability" sections accordingly. For materials that cannot be redistributed because of licensing or other restrictions, we will provide a formal explanation.