ENVIRONMENTAL RESEARCH LETTERS

PURPOSE-LED PUBLISHING™

Part of the ENVIRONMENTAL RESEARCH SERIES

LETTER • OPEN ACCESS

Varying flood exposure due to uncertain data of flood hazard and population distribution

To cite this article: Wendi Shao *et al* 2025 *Environ. Res. Lett.* **20** 114029

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
11 July 2025

REVISED
16 September 2025

ACCEPTED FOR PUBLICATION
6 October 2025

PUBLISHED
14 October 2025

LETTER

Varying flood exposure due to uncertain data of flood hazard and population distribution

Wendi Shao[1,2,3,6], Jiaqi Dong[1,2,6], Jingwei Li[1,2], Mengmeng Li[1,2], Ju Shen[1,4], Yijing Wu[1,5] and Shiqiang Du[1,2,*]

[1] School of Environmental and Geographical Sciences, Shanghai Normal University, Shanghai 200234, People's Republic of China
[2] Yangtze River Delta Urban Wetland Ecosystem National Field Scientific Observation & Research Station, Shanghai 200234, People's Republic of China
[3] Xiangshan NO.2 High School, Ningbo 315731, People's Republic of China
[4] Shanghai Emergency Management and Chemicals Registration Center, Shanghai 200020, People's Republic of China
[5] Shanghai Survey and Mapping Institute, Shanghai 200063, People's Republic of China
[6] Co-first authors.
[*] Author to whom any correspondence should be addressed.

E-mail: shiqiangdu@shnu.edu.cn

Keywords: flood exposure, uncertainty, spatial patterns, data selection

Supplementary material for this article is available online

Abstract

Gridded population and flood hazard data are crucial for flood exposure assessments. However, current assessments incorporate uncertainties related to data selection, yet the mechanisms through which subjective data selection propagate uncertainties in exposure models remain poorly understood. To address this gap, this study conducted a comparative assessment of flood exposure in China using five population datasets and five flood hazard datasets. Furthermore, it explored the absolute and relative impacts of data uncertainties on 100 year return period flood exposure and discussed the underlying causes. Results exhibit substantial variations in flood exposure when different data combinations are employed. Specifically, there is a significant difference of 333 million individuals within the exposure range, with the highest estimate being 2.82 times the lowest one. Overall, the exposure variation was primarily from differences in flood hazards rather than population patterns, but their relative importance differed spatially depending on factors of slope, altitude, and artificial surface coverage. Despite the differences, all 25 data combinations revealed a disproportional larger share of population in floodplains, which was 2.28–3.49 times the share of floodplains. These findings are significant for understanding the uncertainties of flood exposure and can shed lights on informed policies for risk management.

1. Introduction

Flooding is among the most devastating natural hazards (UNDRR 2020). In the context of climate change and socioeconomic development, the frequency of flood events and associated losses have shown significant increases (Willner *et al* 2018, Jongman 2021). A proximate explanation for the changing flood risk arises from the accumulation of assets to meet socioeconomic demands within floodplains (Du *et al* 2018, Rogers *et al* 2025). Flood exposure dynamics reflect the interactions between resource development in floodplains and flood risk mitigation, which draws widespread attention as a critical concern for global sustainability (Shi *et al* 2020, Ward *et al* 2020).

Data-driven flood exposure assessments are largely dependent on two critical types of gridded data: flood hazard and population (Du *et al* 2018, Rogers *et al* 2025). With advancements in flood modeling algorithms and computational capabilities, simulated flood hazard data have become increasingly accessible (Aerts *et al* 2020). Meanwhile, emerging big Earth data and dasymetric mapping techniques have revolutionized population distribution modeling, significantly enhancing both spatial precision and temporal resolution of exposure datasets (Rudari *et al* 2015, Nardi *et al* 2019, Karra *et al* 2021, Tellman *et al* 2021). This methodological paradigm shift has facilitated the operationalization of multi-scale flood exposures. However, the predominant reliance on singular flood hazard dataset and/or population dataset often obscures the compounding uncertainties inherent to data selection (Lindersson *et al* 2020). Notably, flood hazard data derived from flood models introduce endogenous uncertainties (MacManus *et al* 2021), while spatial allocation algorithms for demographical statistics introduce additional uncertainties (Smith *et al* 2019, Láng-Ritter *et al* 2025). These uncertainties could accumulate and propagate through exposure analyses, thereby amplifying the overall uncertainty in flood exposure assessment, which could further lead to misinformed policies of risk management (Quante *et al* 2024).

A growing body of literature has demonstrated the uncertainties in flood exposure assessment. For instance, Huang and Wang (2020) conducted a comparative study of U.S. flood exposure using multiple floodplain datasets (i.e. RFCON, GAR, JRC, and FEMA), highlighting the significant impacts of flood data selection on exposure. Bernhofen *et al* (2021) used a model-independent geomorphological river flood susceptibility map to estimate exposed population and found the river threshold plays a key role in determining the exposure. Bernhofen *et al* (2022) evaluated the applicability of global datasets for flood exposure assessment in countries like Colombia, England, Ethiopia, India, and Malaysia and uncovered substantial divergence under different datasets. Moghim *et al* (2023) compared the simulation results of multiple models (HEC-RAS and LISFLOOD-FP) at watershed scales and found that both model selection and model parameter settings significantly affect flood hazard outcomes. Lindersson *et al* (2021) compared a hydrogeomorphic floodplain map (GFPLAIN) with two flood hazard maps (JRC and GAR) and found the data consistency was influenced by climatic humidity, river volume, topography, and coastal proximity.

Regarding impacts of the selection of population data on flood exposure assessments, Hinkel *et al* (2021) found that different population data between LandScan and GRUMP could cause variation in exposed population by a factor of 1.7. Smith *et al* (2019) evaluated the impact of different population datasets on flood exposure in countries like Ghana, Haiti, Mexico, and Sri Lanka, and found variations in flood exposure between the population datasets of HRSL, LandScan, and WorldPop. Building on the identified uncertainty cascades in exposure analysis, Smith *et al* (2019) further advocated the implementation of uncertainty-aware metrics in flood management policy instruments. Similarly, Mohanty and Simonovic (2021) used population datasets of WorldPop, GPW, LandScan, and GHS to compare flood exposure assessments in Canada, revealing significant discrepancies in exposure estimates. These studies suggested that the selection of flood hazard and population data could bring about significant uncertainties for flood exposure assessments. However, they have not examined the effects of different dataset choices on flood exposure in China, nor have they analyzed the underlying mechanisms driving these impacts.

China ranks among the most flood-prone countries worldwide, characterized by its unique monsoon-driven extreme precipitation regimes, complex riverine systems, and populated floodplains—all contributing to its status as a hotspot for flood risk assessments (UNDRR 2020). In recent years, flood hazards have shown an increasing trend in China (IPCC 2022). Meanwhile, China has been undergoing a rapid rural-urban transition, which is likely contributing to increased flood exposure (Du *et al* 2018, Rentschler *et al* 2023, Bai and Shi 2025). Nevertheless, current research exhibits persistent epistemic shortcomings in understanding the uncertainties of flood exposure therein. Notably, while Aerts *et al* (2020) represent an exception through the comparative analysis of flood dataset interoperability in flood exposure assessment, their study remains constrained by a critical issue: uncertainties from intertwined effect of flood hazards and population patterns are not systematically assessed.

Therefore, the interplay between flood hazard and population data requires rigorous uncertainty quantification to inform evidence-based flood risk governance in China. To do so, this study analyzes uncertainties in flood exposure under the joint influence of different population and flood datasets. It compares the relative importance of population and flood hazard dimensions in influencing flood exposure and investigates the factors potentially contributing to these uncertainties. Such an analysis could provide a scientific basis for understanding flood risk and achieving sustainable development goals.

2. Data

Five representative datasets of flood hazards were collected to compare differences in flood exposure (table S1). The flood hazard datasets include GAR (Rudari *et al* 2015), GLOFRIS (Ward *et al* 2013), CAMA-UT (Yamazaki *et al* 2011), JRC (Dottori *et al* 2016), and ECMWF (Balsamo *et al* 2015), which have been widely applied in flood exposure assessment and flood risk management (Leyk *et al* 2019, Li *et al* 2025). These datasets were representative outcomes derived from different flood simulation methods, climate data, and modeling scales. For example, the

GAR was based on hydrological observation data for flood simulation, while the other datasets mainly used climate reanalysis data (Bernhofen *et al* 2018). The modeling scales were also different across these datasets (supplementary data: table S1), which potentially impacted the flood inundation results (Bernhofen *et al* 2021).

Similarly, five population datasets were collected for the same year (2010) to ensure data comparability (table S2), including the WorldPop (Tatem 2017), LandScan (Lebakula *et al* 2025), GPW (Gridded Population of the World v4.11) (Doxsey-Whitfield *et al* 2015), China Grid Population Distribution Dataset (hereafter referred to as CnPop) (Wang and Wang 2022), and the Point of Interest-based population data (hereafter referred to as PoiPop) (Ye *et al* 2019). They were representative outcomes of different demographic and auxiliary data, as well as different methods of population dasymetric mapping (Leyk *et al* 2019). For instance, the GPW4 spatialized population uses an area-weighted method, the LandScan employs a dasymetric model based on machine learning to disaggregate population data, and the CnPop allocates population numbers using a multi-factor weighting approach. The WorldPop and the PoiPop use random forests for population disaggregation, with the distinction that the WorldPop leveraged various auxiliary variables (e.g. slope, impervious surface, nighttime light), while the PoiPop utilizes POI data and multi-source remote sensing data to disaggregate county-level census data to geographic grids.

3. Methods

3.1. Assessing model agreement index (MAI) among flood datasets

To further investigate the discrepancies among flood hazard datasets, *MAI* is employed following Trigg *et al* (2016). The *MAI* ranges from 0 (completely inconsistent) to 1 (completely consistent). It is calculated as follows:

$$\mathrm{MAI} = \frac{\sum_{i=2}^{n}\left(\frac{1}{n}\right)a_i}{A} \tag{1}$$

where n is the number of flood hazard datasets, which is 5 in this study; A is the number of flooded pixels by either of the flood hazard datasets; a_i represents the number of pixels identified as flooded by a number of flood hazard datasets i.

3.2. Assessing flood exposure

Flood exposure is defined as the population presence in a 100 year return period floodplain, which has gained widespread adoption in contemporary risk assessments (Du *et al* 2018, Abedin *et al* 2024). The flood-exposed population is calculated as follows:

$$\mathrm{population_{exposure}} = \mathrm{population_{total}} \cap \mathrm{hazard_{flood}} \tag{2}$$

where $Population_{exposure}$ represents the flood-exposed population, $Population_{total}$ is the total population within a spatial unit, and $Hazard_{flood}$ refers to the flood inundation extent, which is defined as areas where the flood depth exceeds 0 cm (Du *et al* 2018).

3.3. Quantifying differences in flood exposure

Three indicators are employed to quantify the differences in flood exposure across various data combinations: absolute exposure difference, relative exposure difference, and coefficient of variation. Absolute exposure difference refers to the disparity between the highest and lowest flood-exposed populations. Similarly, the relative exposure difference is quantified as the range between the highest and lowest flood-exposed population ratios (Hierink *et al* 2022).

The coefficient of variation captures the dispersion of flood-exposed populations across different data combinations, eliminating the influence of different measurement scales and data units (Brown 1998). It is calculated as follows:

$$c_{\mathrm{v}} = \frac{\sqrt{\sum_{i=1}^{n}\left(\mathrm{ep}_i - \mathrm{mean_{ep}}\right)^2}}{\sqrt{n} \times \mathrm{mean_{ep}}} \tag{3}$$

where c_{v} is the coefficient of variation of flood-exposed populations across different data combinations; ep_i is the flood-exposed population under data combination i; $mean_{ep}$ is the mean of flood-exposed populations across all data combinations; and n is the number of data combinations.

3.4. Analyzing the relative importance of flood hazard and population uncertainty

To compare the impacts of flood dimension and population dimension on flood exposure variability, flood exposure difference is quantified under each dimension (Bernhofen *et al* 2022). First, we anchor the analysis to a specific flood hazard dataset (denoted as i) and systematically evaluate population exposure variability by integrating all available population datasets. This iterative calculation yields the exposure range metric df_1, representing the uncertainty spectrum associated with population datasets. Second, this workflow is replicated using each of the alternative flood datasets, conducting parallel computations to derive the corresponding exposure metrics $[df_2, df_3, \ldots, df_n]$. Similarly, the calculation process can be used to determine the flood exposure variability under the population dimension. The calculation can be conducted as follows:

$$df_{\dim} = \frac{\sum_{i=1}^{n} df_i}{n} \tag{4}$$

where $df_{\dim}$ represents the flood exposure variability under the hazard dimension (or population dimension); df_i is the flood exposure variability based on dataset i; and n is the number of datasets under the specific dimension.

To assess the relative importance of uncertainties in flood hazard and population dimensions to the exposed population assessment, the relative importance index (RII) is employed following Ran *et al* (2022). This index compares the coefficient of variation in the flood hazard dimension with that of the population dimension. It can be calculated as follows:

$$RII = \frac{c_{v_{flood}}}{c_{v_{population}}} \tag{5}$$

where c_{vflood} and $c_{\mathrm{vpopulation}}$ are the coefficients of variation of exposed populations in the hazard and population dimensions, respectively. A larger RII indicates that uncertainties in the flood hazard dimension have a greater influence on flood exposure variability, and vice versa.

4. Results and discussion

4.1. Uncertainties in flood exposure in China

At the national scale, data selection exerts profound impacts on the outcomes of flood exposure assessments. Utilizing 25 distinct data combinations, the population exposed to a 100 year flood area ranges from 183 million to 516 million individuals, a disparity corresponding to exposure ratios varying between 13.61% and 38.27% of the total population. This results in a variation of up to ∼333 million people, or a 2.82-fold difference (figure 1). These findings coincide with Aerts *et al* (2020) and underscore the critical limitation of flood exposure assessments that rely solely on single dataset, which is inherently subject to substantial uncertainties and may lead to either overestimation or underestimation of actual flood exposure.

Despite the uncertainties associating with exposed population assessment, our results consistently exhibit a phenomenon that exposed population ratio is much higher than the floodplain's ratio to total terrestrial area, as the former remains 2.28–3.49 folds of the latter across all the combinations of flood hazard and population data (figure 1). This indicates a disproportionate distribution of population in floodplains, a phenomenon also found in other studies (Du *et al* 2018, Devitt *et al* 2023). This study proves that such a phenomenon is robust to uncertainties associating with data selection in China, which poses a critical challenge for China to achieve flood risk management in a changing climate (Du *et al* 2019, Ding *et al* 2022).

At the basin scale, the average of relative exposure differences across all dimensions and data combinations is 33.9% in eastern coastal regions, which is higher than that in central basins (23.1%) and western basins (20.7%). At the sub-basin scale, higher values of relative exposure difference are predominantly concentrated in the eastern coastal regions and central regions (figure 2(a)). A number of 238 sub-basins exhibit an average relative difference exceeding 30.0% (seen as obvious difference), accounting for 45.0% of all sub-basins; particularly in eastern region, 88 sub-basins show obvious differences, with a much higher ratio (67.7%) than the national average (45.0%). These results suggest a concentration of relative exposure differences in eastern China.

4.2. Variations in the impacts of flood hazard and population dimensions on flood exposure in China

At the national scale, the selection of flood hazard data has a more pronounced influence on both absolute and relative differences in flood exposure, compared to the choice of population data. With a fixed population dataset and varying flood hazard data, absolute flood exposure differences range from 239 million to 314 million people, corresponding to relative differences of 17.71% to 23.34%. In contrast, when scanning different population data with a fixed flood hazard dataset, the absolute differences of flood exposure range from 24.61 million to 93.69 million people, with relative exposure differences between 1.83% and 6.95%, which are notably smaller than those caused by variations in flood hazard data. At the basin and sub-basin scales, the flood hazard factors also show an overall higher impact than that of population data on flood exposure differences (figure 2(b)). The results prove the rationality of previous studies that intuitively only consider the impacts of variations in flood hazard data on exposed population to some extent (Aerts *et al* 2020, Huang and Wang 2020).

However, the impacts of flood hazard and population dimensions on flood exposure in China exhibit significant spatial variation (figures 2(c) and (d)). In terms of flood hazard dimension, the spatial pattern of flood exposure uncertainty is consistent with the situation when both dimensions are considered together (figure 2(c)). At the basin scale, the average relative difference in the eastern coastal basins is 28.5%, compared to 18.6% in the central basins and 13.7% in the western basins; at the sub-basin scale, among the 107 sub-basins with relative differences exceeding 30.0% under hazard factors, 56.1% are located in the eastern coastal regions.

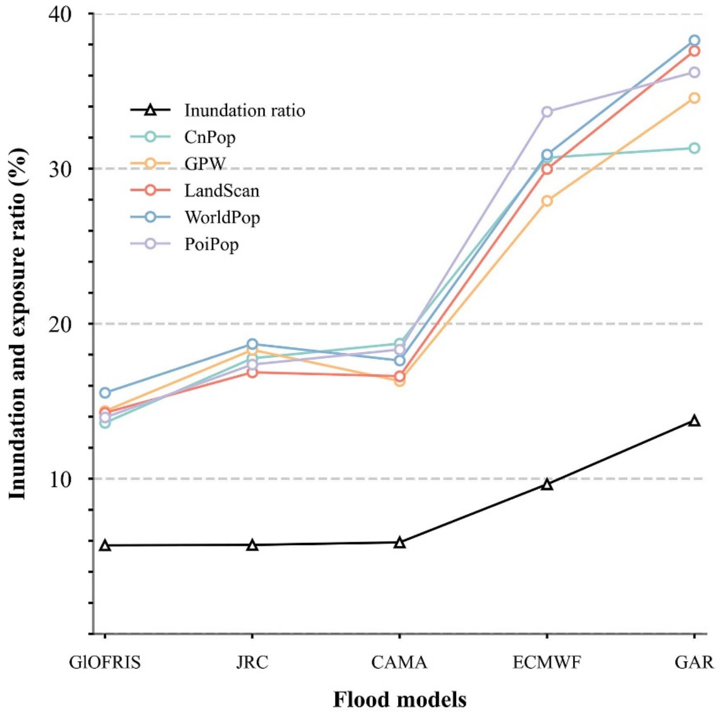

Figure 1. The ratios of flood inundation areas to the total area and of the exposed population to the total population.

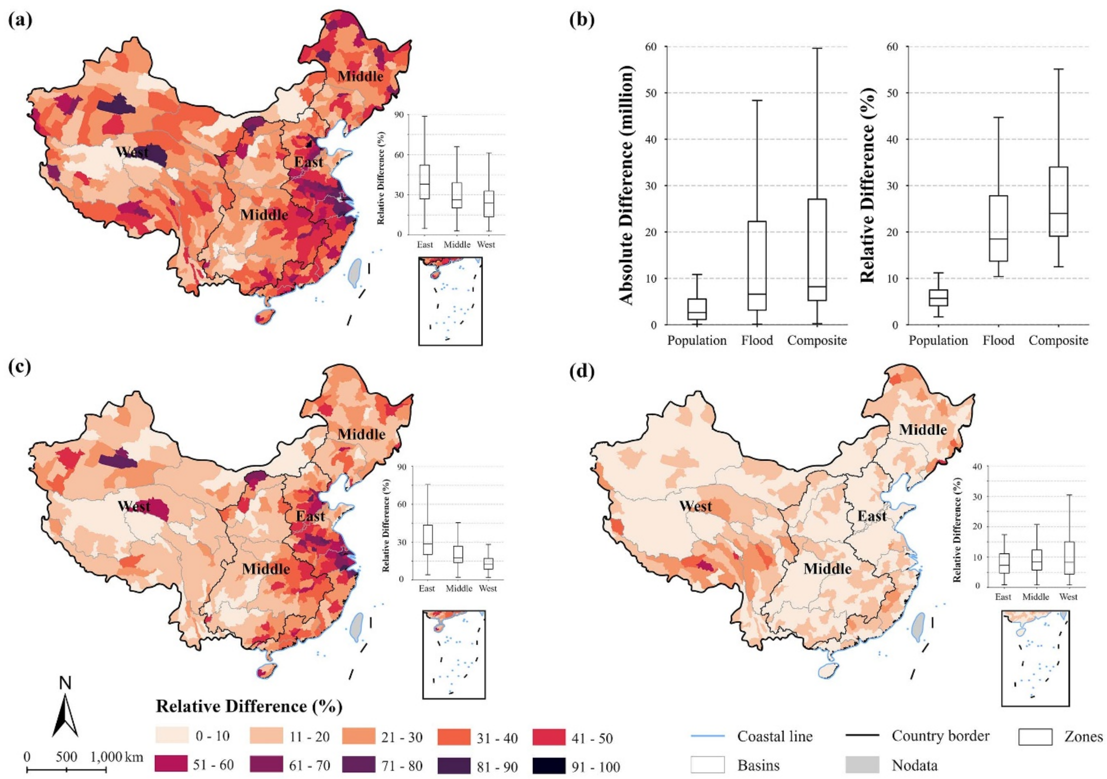

Figure 2. Relative exposure differences under influences of flood hazard and population data. The composite effects of flood hazard and population data (a), a statistic of the effects (b), and the individual effect of flood hazard (c) and population data (d).

However, regarding population dimensions, higher relative exposure differences are concentrated in the western regions. The average relative exposure difference under population factors in the western basins is 7.0%, while it is 6.0% and 5.1% in the eastern coastal and central regions, respectively. At the sub-basin scale (figure 2(d)), of the 45 sub-basins with relative exposure differences exceeding 30.0% under population factors, 70.0% are located in the western regions. The results imply that the uncertainties of population pattern should not be neglected when assessing flood exposure (Smith *et al* 2019, Mohanty and Simonovic 2021, Bernhofen *et al* 2022, Láng-Ritter *et al* 2025).

According to the RII results, defined as the ratio of the coefficient of variation for the flood hazard dimension to that of the population dimension, the flood hazard dimension has a more significant impact on flood exposure variability overall, which again assures the findings from the comparison between relative exposure differences under the influence of flood hazard and population data. At both the national and basin scales, the RII values are greater than 1, indicating that the selection of flood datasets has a larger effect on flood exposure outcomes. At the sub-basin scale, 82.04% of the sub-basins (434 in total) had RII values greater than 1, suggesting that the variability in the flood hazard dimension outweighs that in the population dimension (figure 3). However, 95 sub-basins have RII values less than 1, mainly concentrated in the western regions, highlighting regional differences in RII values and implying that in regions like western China the uncertainties of population pattern play a key role in shaping flood exposure.

4.3. Underlying mechanisms by which data selection influences flood exposure

The influence of population datasets on exposure should be a result of different patterns, as the total population has been adjusted to the same value for all datasets. In other words, the main source of uncertainty in the population dimension comes from the different methods of dasymetric mapping in the population datasets. Flood exposure is concentrated in fewer flood pixels when using WorldPop, PoiPop, and LandScan as these datasets result in higher exposure density. These datasets use settlement areas as important auxiliary variables during the dasymetric mapping of population (Swanwick *et al* 2022), leading to more concentrated population distributions compared to other population datasets. In contrast, GPW and CnPop, which employ simpler population disaggregation methods (Doxsey-Whitfield *et al* 2015, Wang and Wang 2022), results in more evenly distributed population patterns, as indicated by flatter cumulative curves of exposed population (figure S1).

The relative impact of population data is associated with basin area, average slope, and average elevation as these factors are significantly negatively correlated with the RII ($p < 0.001$; table S3). This suggests that in sub-basins with larger areas, steeper slopes, and higher elevations, population data variability plays a greater role in shaping flood exposure, which are the typical characteristics of many sub-basins in mountainous western China. The findings indicate higher demands on dasymetric mapping of population to conduct a reliable flood exposure analysis in regions with larger areas, steeper slopes, and higher elevations (Baynes *et al* 2022, Lu and Weng 2024), particularly in the rural as indicated by Láng-Ritter *et al* (2025).

Both the total flood area and the spatial pattern play a role in shaping the exposure differences caused by flood hazard data. First, the differences in total flood area play a key role in the variability of flood exposure ratios. For example, the flood exposure ratios associating with GAR and ECMWF remain 1.72–2.48 times those of GLOFRIS, JRC, and CAMA (figure 2(a)); consistently, the flood extents of GAR and ECMWF are exactly 1.63–2.41 times of the extents of GLOFRIS, JRC, and CAMA. On the other hand, the spatial pattern of flood extents is another reason causing flood exposure variability. The JRC, CAMA, and GLOFRIS exhibit similar total flood areas, at 5.71%–5.90% of the total terrestrial area. However, the differences in flood exposure associating with the three flood hazard datasets could be as large as 1.63 times as much given a fixed population dataset. The notable differences of flood exposure are likely a result of the variations in flood patterns, as the MAI between these flood datasets is generally less than 0.4, and even below 0.1 in some sub-basins (figure 4).

The relative impact of flood hazard data is associated with indicators reflecting urbanization, according to the correlations between indicators and the RII ($p < 0.001$; table S3). The RII is significantly positively correlated with the artificial surface coverage, the average nighttime light index, nighttime light coverage ratio, average GDP, and the proportion of flat areas within the sub-basin, suggesting that an increase in these indices could amplify the impacts of flood hazard. These indices show high values in the eastern coastal regions (Rentschler *et al* 2023), which is consistent with the higher relative importance of the hazard data in these areas and agrees with the findings of Aerts *et al* (2020). Even slight differences in flood inundation extent in these regions can lead to significant differences in flood exposure, making a careful selection of flood hazard maps crucial for assessing flood exposure in these areas. Therefore, there is an urgent need for localized modeling to accurately reflect flood inundation extents and reduce uncertainties (Yamazaki *et al* 2011, Ward *et al* 2013, Balsamo *et al* 2015, Rudari *et al* 2015, Dottori *et al* 2016, Khoshkonesh *et al* 2024).

4.4. Limitations and future prospects

Like all data-driven analyses, this study is subject to inherent limitations that necessitate cautious interpretation of its findings. First, it only uses the maximum flood extents as a proxy for floodplains for calculating population exposure but neglects the impacts of different flood depths. Although the definition of floodplain has been widely used in flood

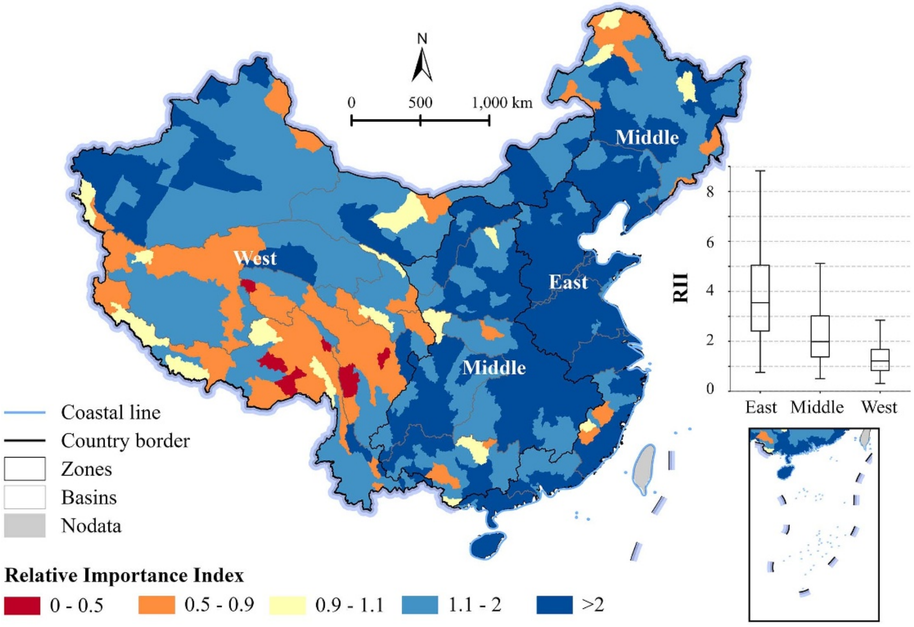

Figure 3. The Relative Importance Index (RII) characterizing the contributions of flood hazard dimension and the population dimension to exposed population assessment.

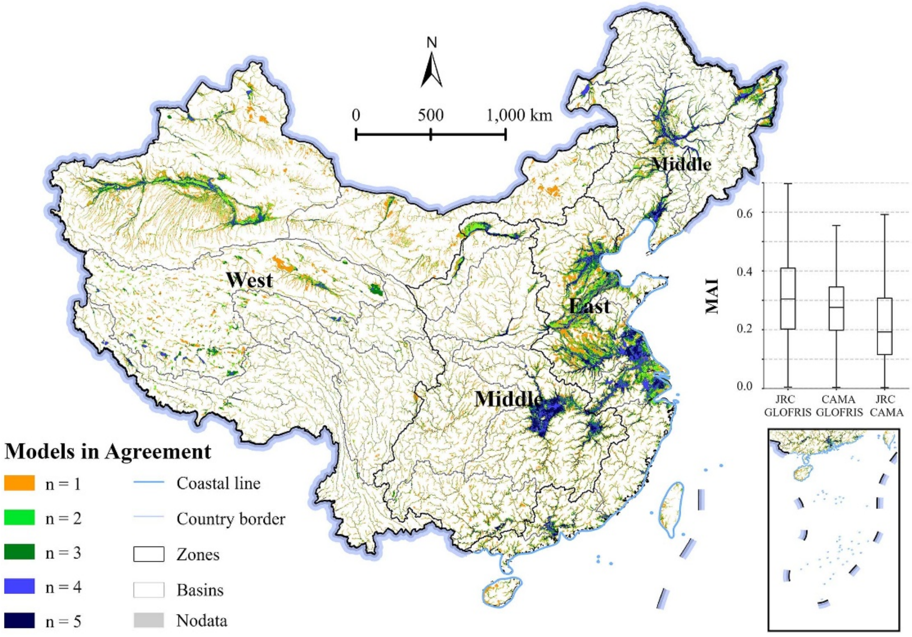

Figure 4. Consistency of five flood hazard datasets (n indicates the number of datasets in which the pixel is flooded).

exposure assessment (Samoray *et al* 2024), the varying risk associated with different flood depths is also important (Mohanty and Simonovic 2021). Future research could help to clarify the uncertainties of population exposure across various flood depths when using different flood datasets. Second, this study focused on the flood exposure for a single year but neglected uncertainties regarding temporal changes in exposed population over a period, which is also critical for guiding flood management policies (Mazumder *et al* 2022, Xu and Qiang 2024). Future studies could explore the dynamics in flood-exposed populations under different combinations of data products. Third, although representative data were used for the comparative study, they were not validated against real values. Regarding population data, future research could utilize finer-scale population data to develop more realistic exposure assessment models (Ceola *et al* 2014, Tellman *et al* 2021), and even to reflect seasonal dynamics of the population and exposure by leveraging mobile phone signaling and other big data (Luan *et al* 2024). Regarding flood hazard data, future studies could employ localized hydrodynamic modelling (Schubert *et al* 2024, Haces-Garcia *et al* 2025) to simulate historical events (Tellman *et al* 2021) and possible flood processes under future climate parameters (Rogers *et al* 2025) to produce reliable flood hazard data in assessing population exposure, particularly for areas with a relatively high importance of flood hazard data. Fourth, this study did not account for flood protection levels, which may lead to an overestimation of populations exposed to flood maps that do not incorporate flood protection levels (Wu *et al* 2024). With urbanization and economic development there is an increasing focus on flood protection improvement. Incorporating flood protection data into flood simulations could yield more accurate exposure estimates (Wang *et al* 2021).

5. Conclusion

To understand the potential impacts of data uncertainty on flood exposure assessment, this study systematically investigated China's flood-exposed population across five sets of flood hazard data and five sets of population data. The results showed significant differences in flood exposure estimates across different data combinations and yielded a 100 year flood population ranging from 183 to 516 million, with a disparity of up to 333 million or 2.82 folds. Overall, the uncertainties of flood exposure were primarily from the impact of flood inundation extents rather than the impacts of population data discrepancies. Spatially, the influence of flood inundation differences was more pronounced in the east, while population factors had a relatively greater impact on exposure in western sub-basins. The relative importance of flood hazard and population data to flood exposure was significantly associated with factors of slope, elevation, and artificial surface coverage. Despite the differences, all data combinations revealed a disproportional distribution of population in floodplains as the exposed population ratio remained 2.28–3.49 folds of the floodplains' share to the total lands. These findings highlight the importance of incorporating data uncertainties into flood exposure assessment. A robust understanding of flood exposure can guide policymakers towards sound decisions on flood risk management and promote the achievement of sustainable development goals.

Data Availability Statement

All datasets used in this study are publicly available online. For the flood hazard maps, the CAMA-UT dataset is available from http://hydro.iis.u-tokyo.ac.jp/~yamadai/CaMa-Flood_v3.6/, the GAR dataset is available from the Global Risk Data Platform www.undrr.org/publication/global-assessment-report-disaster-risk-reduction-2015, the GLOFRIS is available from www.wri.org/data/aqueduct-floods, the JRC flood maps is available from https://data.jrc.ec.europa.eu/dataset/jrc-floods-floodmapgl_rp50y-tif, and the ECMWF flood maps is available from https://apps.ecmwf.int/datasets/. For the population datasets, the WorldPop dataset is available from https://hub.worldpop.org/geodata/listing?id=69, the LandScan is available from https://landscan.ornl.gov/, the GPW is available from www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-gpwv4-popcount-r11-4.11, the CnPop is available from www.resdc.cn/DOI/DOI.aspx?DOIid=32, and the PoiPop could be obtained from https://person.zju.edu.cn/yangxuchao#710315. Other data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant Nos. 42101251 and 41871200).

Author contributions

Wendi Shao
Data curation (equal), Formal analysis (equal), Investigation (equal), Methodology (equal), Writing – original draft (equal)

Jiaqi Dong 0009-0000-1159-0139
Methodology (equal), Software (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Jingwei Li
Funding acquisition (supporting), Supervision (supporting), Writing – review & editing (supporting)

Mengmeng Li 0000-0001-5093-386X
Validation (equal), Writing – review & editing (equal)

Ju Shen
Methodology (equal), Software (supporting), Visualization (supporting), Writing – review & editing (supporting)

Yijing Wu
Formal analysis (equal), Methodology (equal), Software (equal), Visualization (equal)

Shiqiang Du 0000-0002-9787-186X
Conceptualization (lead), Project administration (lead), Resources (lead), Supervision (lead), Writing – original draft (equal), Writing – review & editing (lead)

References

Abedin J, Zou L, Yang M, Rohli R, Mandal D, Qiang Y, Akter H, Zhou B, Lin B and Cai H 2024 Deciphering spatial-temporal dynamics of flood exposure in the United States *Sustain. Cities Soc.* **108** 105444

Aerts J P M, Uhlemann-Elmer S, Eilander D and Ward P J 2020 Comparison of estimates of global flood models for flood hazard and exposed gross domestic product: a China case study *Nat. Hazards Earth Syst. Sci.* **20** 3245–60

Bai X and Shi P 2025 China's urbanization at a turning point—challenges and opportunities *Science* **388** eadw3443

Balsamo G, Albergel C, Beljaars A, Boussetta S, Brun E, Cloke H, Dee D, Dutra E, Muñoz-Sabater J and Pappenberger F 2015 ERA-interim/land: a global land surface reanalysis data set *Hydrol. Earth Syst. Sci.* **19** 389–407

Baynes J, Neale A and Hultgren T 2022 Improving intelligent dasymetric mapping population density estimates at 30 m resolution for the conterminous United States by excluding uninhabited areas *Earth Syst. Sci. Data* **14** 2833–49

Bernhofen M V, Cooper S, Trigg M, Mdee A, Carr A, Bhave A, Solano-Correa Y T, Pencue-Fierro E L, Teferi E and Haile A T 2022 The role of global data sets for riverine flood risk management at national scales *Water Resour. Res.* **58** e2021WR031555

Bernhofen M V, Trigg M A, Sleigh P A, Sampson C C and Smith A M 2021 Global flood exposure from different sized rivers *Nat. Hazard Earth Syst.* **21** 2829–47

Bernhofen M V, Whyman C, Trigg M A, Sleigh P A, Smith A M, Sampson C C, Yamazaki D, Ward P J, Rudari R and Pappenberger F 2018 A first collective validation of global fluvial flood models for major floods in Nigeria and Mozambique *Environ. Res. Lett.* **13** 104007

Brown C E 1998 *Applied Multivariate Statistics in Geohydrology and Related Sciences* (Springer) pp 155–7

Ceola S, Laio F and Montanari A 2014 Satellite nighttime lights reveal increasing human exposure to floods worldwide *Geophys. Res. Lett.* **41** 7184–90

Devitt L, Neal J, Coxon G, Savage J and Wagener T 2023 Flood hazard potential reveals global floodplain settlement patterns *Nat. Commun.* **14** 2801

Ding W, Wu J, Tang R, Chen X and Xu Y 2022 A REVIEW OF FLOOD RISK in China during 1950–2019: urbanization, socioeconomic impact trends and flood risk management *Water* **14** 3246

Dottori F, Salamon P, Bianchi A, Alfieri L, Hirpa F A and Feyen L 2016 Development and evaluation of a framework for global flood hazard mapping *Adv. Water Resour.* **94** 87–102

Doxsey-Whitfield E, MacManus K, Adamo S B, Pistolesi L, Squires J, Borkovska O and Baptista S R 2015 Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4 *Pap. Appl. Geogr.* **1** 226–34

Du S, Cheng X, Huang Q, Chen R, Ward P J and Aerts J C J H 2019 Brief communication: rethinking the 1998 China floods to prepare for a nonstationary future *Nat. Hazards Earth Syst. Sci.* **19** 715–9

Du S, He C, Huang Q and Shi P 2018 How did the urban land in floodplains distribute and expand in China from 1992–2015? *Environ. Res. Lett.* **13** 034018

Haces-Garcia F, Ross N, Glennie C L, Rifai H S, Hoskere V and Ekhtari N 2025 Rapid 2D hydrodynamic flood modeling using deep learning surrogates *J. Hydrol.* **651** 132561

Hierink F *et al* 2022 Differences between gridded population data impact measures of geographic access to healthcare in sub-Saharan Africa *Commun. Med.* **2** 117

Hinkel J *et al* 2021 Uncertainty and bias in global to regional scale assessments of current and future coastal flood risk *Earth's Future* **9** e2020EF001882

Huang X and Wang C 2020 Estimates of exposure to the 100-year floods in the conterminous United States using national building footprints *Int. J. Disaster Risk Reduct.* **50** 101731

IPCC 2022 Climate change 2022: impacts, adaptation and vulnerability *Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed H-O Pörtner *et al* p 3056

Jongman B 2021 The fraction of the global population at risk of floods is growing *Nature* **596** 37–38

Karra K, Kontgis C, Statman-Weil Z, Mazzariello J C, Mathis M and Brumby S P 2021 Global land use/land cover with Sentinel 2 and deep learning *2021 IEEE Int. Geoscience and Remote Sensing Symp. IGARSS* (IEEE) pp 4704–7

Khoshkonesh A, Nazari R, Nikoo M R and Karimi M 2024 Enhancing flood risk assessment in urban areas by integrating hydrodynamic models and machine learning techniques *Sci. Total Environ.* **952** 175859

Láng-Ritter J, Keskinen M and Tenkanen H 2025 Global gridded population datasets systematically underrepresent rural population *Nat. Commun.* **16** 2170

Lebakula V *et al* 2025 Landscan global 30 arcsecond annual global gridded population datasets from 2000 to 2022 *Sci. Data* **12** 495

Leyk S, Gaughan A E, Adamo S B, de Sherbinin A, Balk D, Freire S, Rose A, Stevens F R, Blankespoor B and Frye C 2019 The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use *Earth Syst. Sci. Data* **11** 1385–409

Li L, Zhou C, Li J, Li M, Wu Y, Gong Z and Du S 2025 Increasing population exposed to both seismicity and flooding in low-income countries *Geomatics Nat. Hazards Risk* **16** 2493217

Lindersson S, Brandimarte L, Mård J and Di Baldassarre G 2020 A review of freely accessible global datasets for the study of floods, droughts and their interactions with human societies *WIREs Water* **7** e1424

Lindersson S, Brandimarte L, Mård J and Di Baldassarre G 2021 Global riverine flood risk—how do hydrogeomorphic floodplain maps compare to flood hazard maps? *Nat. Hazards Earth Syst. Sci.* **21** 2921–48

Lu W and Weng Q 2024 An ANN-based method for population dasymetric mapping to avoid the scale heterogeneity: a case study in Hong Kong, 2016–2021 *Comput. Environ. Urban Syst.* **108** 102072

Luan X, Paryzat H, Chu J, Shu X, Gu H, Tong D and Li B 2024 Different roads take me home: the nonlinear relationship between distance and flows during China's spring festival *Humanit. Soc. Sci. Commun.* **11** 1356

MacManus K, Balk D, Engin H, McGranahan G and Inman R 2021 Estimating population and urban areas at risk of coastal hazards, 1990–2015: how data choices matter *Earth Syst. Sci. Data* **13** 5747–801

Mazumder L T, Landry S, Alsharif K and Eduful M K 2022 Changing demographics and the environmental equity of coastal floodplain in Tampa, Florida *Int. J. Disaster Risk Reduct.* **79** 103186

Moghim S, Gharehtoragh M A and Safaie A 2023 Performance of the flood models in different topographies *J. Hydrol.* **620** 129446

Mohanty M P and Simonovic S P 2021 Understanding dynamics of population flood exposure in Canada with multiple high-resolution population datasets *Sci. Total Environ.* **759** 143559

Nardi F, Annis A, Di Baldassarre G, Vivoni E R and Grimaldi S 2019 GFPLAIN250m, a global high-resolution dataset of Earth's floodplains *Sci. Data* **6** 180309

Quante L, Willner S N, Otto C and Levermann A 2024 Global economic impact of weather variability on the rich and the poor *Nat. Sustain.* **7** 1419–28

Ran Q, Wang J, Chen X, Liu L, Li J and Ye S 2022 The relative importance of antecedent soil moisture and precipitation in flood generation in the middle and lower Yangtze River basin *Hydrol. Earth Syst. Sci.* **26** 4919–31

Rentschler J, Avner P, Marconcini M, Su R, Strano E, Vousdoukas M and Hallegatte S 2023 Global evidence of rapid urban growth in flood zones since 1985 *Nature* **622** 87–92

Rogers J S, Maneta M M, Sain S R, Madaus L E and Hacker J P 2025 The role of climate and population change in global flood exposure and vulnerability *Nat. Commun.* **16** 1287

Rudari R, Silvestro F, Campo L, Rebora N, Boni G and Herold C 2015 *Improvement of the global flood model for the GAR 2015* (United Nations Office for Disaster Risk Reduction (UNISDR)) p 69 (available at: www.undrr.org/publication/improvement-global-food-model-gar-2015)

Samoray C, Hino M, Siders A R, Agopian A and Mach K J 2024 Housing amenity and affordability shape floodplain development *Land Use Policy* **144** 107216

Schubert J E, Mach K J and Sanders B F 2024 National-scale flood hazard data unfit for urban risk management *Earth's Future* **12** e2024EF004549

Shi P *et al* 2020 Disaster risk science: a geographical perspective and a research framework *Int. J. Disaster Risk Sci.* **11** 426–40

Smith A, Bates P D, Wing O, Sampson C, Quinn N and Neal J 2019 New estimates of flood exposure in developing countries using high-resolution population data *Nat. Commun.* **10** 1814

Swanwick R H, Read Q D, Guinn S M, Williamson M A, Hondula K L and Elmore A J 2022 Dasymetric population mapping based on US census data and 30 m gridded estimates of impervious surface *Sci. Data* **9** 523

Tatem A J 2017 WorldPop, open data for spatial demography *Sci. Data* **4** 170004

Tellman B, Sullivan J A, Kuhn C, Kettner A J, Doyle C S, Brakenridge G R, Erickson T A and Slayback D A 2021 Satellite imaging reveals increased proportion of population exposed to floods *Nature* **596** 80–86

Trigg M A *et al* 2016 The credibility challenge for global fluvial flood risk analysis *Environ. Res. Lett.* **11** 094014

UNDRR 2020 *Human Cost of Disasters: An Overview of the Last 20 Years (2000-2019)* (Centre for Research on the Epidemiology of Disasters (CRED), United Nations Office for Disaster Risk Reduction (UNDRR)) p 17 (available at: www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019)

Wang C and Wang J 2022 Kilometer grid dataset of China's historical population spatial distribution (1990–2015) *National Tibetan Plateau / Third Pole Environment Data Center.* (https://doi.org/10.12078/2017121101)

Wang D, Scussolini P and Du S 2021 Assessing Chinese flood protection and its social divergence *Nat. Hazards Earth Syst. Sci.* **21** 743–55

Ward P J *et al* 2020 Review article: natural hazard risk assessments at the global scale *Nat. Hazard Earth Syst.* **20** 1069–96

Ward P J, Jongman B, Weiland F S, Bouwman A, van Beek R, Bierkens M F, Ligtvoet W and Winsemius H C 2013 Assessing flood risk at the global scale: model setup, results, and sensitivity *Environ. Res. Lett.* **8** 044019

Willner S N, Otto C and Levermann A 2018 Global economic response to river floods *Nat. Clim. Change* **8** 594–8

Wu Y, Li J, Wu H, Duan Y, Shen H and Du S 2024 Sustainable urban planning to control flood exposure in the coastal zones of China *Landscape Ecol.* **39** 141

Xu J and Qiang Y 2024 Environmental justice implications of flood risk in the contiguous United States—a spatiotemporal assessment of flood exposure change from 2001 to 2019 *Cartogr. Geogr. Inf. Sci.* **51** 492–512

Yamazaki D, Kanae S, Kim H and Oki T 2011 A physically based description of floodplain inundation dynamics in a global river routing model *Water Resour. Res.* **47** W04501

Ye T *et al* 2019 Improved population mapping for China using remotely sensed and points-of-interest data within a random forests model *Sci. Total Environ.* **658** 936–46