Doctoral Dissertation

EFFECT OF COMPOUNDING BIAS FROM GLOBAL DATASETS ON LOCAL TSUNAMI EXPOSURE ASSESSMENT:
A CASE STUDY IN TSUNAMI-PRONE CITY OF BANDA ACEH, INDONESIA

Rajuli Amra

June 2025

Graduate School of Engineering,
The University of Osaka

ABSTRACT

Information on assets and populations potentially exposed to tsunamis is essential for risk mitigation prior to tsunami occurrence. This information is obtained by intersecting the extent of tsunami inundation with exposure datasets. The accuracy of tsunami inundation is strongly influenced by elevation data. Together, elevation data and exposure datasets are referred to as essential variables. However, these essential variables are often not available in data-scarce regions, leading to utilization of 'best available' global datasets. This research analyses the applicability of global datasets for tsunami exposure assessment at a local scale. The study evaluates 11 digital elevation models, 8 impervious surface layers, and 4 gridded population datasets. The evaluation focuses on how these datasets can improve tsunami inundation model performance and how their compounding biases affect the exposure estimates in the tsunami-prone city of Banda Aceh, Indonesia. This city is known for being heavily affected by the 2004 Indian Ocean Tsunami.

To minimize variability in global elevation model assessment, this research introduces a sequential validation approach that assesses how well the models reproduce a historical event while accounting for their inherent biases. In parallel, the global exposure datasets are compared against high-resolution local data to evaluate intrinsic and analytical biases. The least biased dataset variants from both assessments are selected and further used to analyse how biases compound when elevation models and exposure datasets are integrated. Using the least biased dataset variants, the study projects tsunami exposure estimates for future tsunamigenic events in the area of interest using cross-uncertainty assessment, which integrates probabilistic inundation with confidence levels of the global datasets used. Based on this cross-uncertainty assessment, the research evaluates how existing mitigation strategies can mitigate future tsunamigenic events. This research is presented through six chapters.

Chapter 1 explains research objectives and research locations.

Chapter 2 examines the biases of global impervious surface and gridded population datasets.

Chapter 3 analyses the sensitivity of global elevation models in modelling historical tsunamis.

Chapter 4 calculates the compounding biases from global datasets and performs the cross-uncertainty tsunami assessment.

Chapter 5 provides an evaluation of existing mitigation systems against future tsunamigenic scenarios.

Chapter 6 presents conclusions and recommendations based on the results obtained.

The results show global exposure datasets exhibit bias by overestimating tsunami-exposed built-up areas while underestimating tsunami-exposed populations. Additionally, elevation data contributes more to underestimation than population datasets. When used together, these biases compound, nearly doubling the underestimation of tsunami-exposed populations compared to when each dataset is used separately.

ACKNOWLEDGMENT

First of all, I would like to offer my sincere thanks to Professor Susumu Araki for his role in supervising my PhD research. I really appreciate his patience with me even though I took longer than required to finish this research. This research would have been impossible to complete without the help and knowledge sharing from Professor Christian Geib from The German Aerospace Center (DLR) and Dr. Gareth Davies from Geoscience Australia. I would also like to thank Professor Syamsidik from the Tsunami and Disaster Mitigation Research Center (TDMRC) for sharing data relevant to this research location. I also thank The Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan for sponsoring this PhD program. Last but not least, I would like to offer my sincere gratitude to Professor Aoki Shin-Ichi for recommending me for the MEXT scholarship and for checking on my condition during the most challenging moments of my PhD journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

1.1 Recent Advances in Tsunami Science

Several mega-tsunamis have occurred over the last two decades (Fig. 1.1). These tsunamis claimed hundreds to thousands of lives. For instance, the 2004 Indian Ocean Tsunami (IOT) claimed over 230,000 lives across 14 countries, the 2011 Tohoku Tsunami caused approximately 20,000 deaths, and the 2018 Palu Tsunami led to over 4,300 casualties (UNDRR, 2019). In addition, these tsunamis have led to significant economic losses, with estimates ranging from $10 billion to $300 billion USD (Rafliana et al., 2022). This highlights the profound economic impact, along with the tragic loss of life. Despite their significant drawbacks, tsunamis have served as catalysts for rapid advancement of tsunami science.

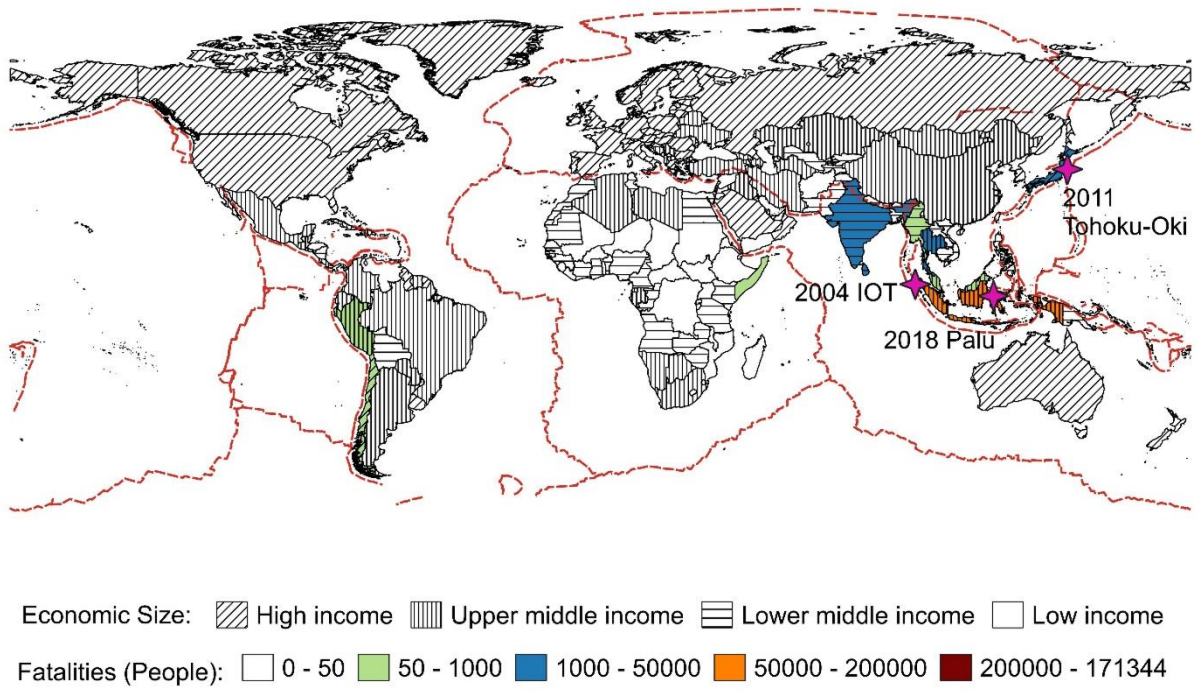


Fig. 1.1 Spatial distribution of tsunami fatalities between 2000 and 2024. These data were curated by Shaw et al. (2025) and UNDRR (2016). The hatched lines represent the classification of economic power (World Bank, 2024). The magenta star markers show the epicenters of the three most catastrophic tsunamis on the timeline.

In the aftermath of the 2004 Indian Ocean Tsunami (IOT), significant advancements were made in the field of tsunami physical studies, particularly in the areas of modelling and reconstructing tsunami source models (Harig et al., 2022; Fujii et al., 2021; Satake, 2014; Tsushima et al., 2014). This event also catalysed research in disaster risk

reduction (DRR), focusing on hazard exposure mapping, evacuation strategies, tsunami-resilient land use management, and post-disaster reconstruction efforts (Suppasri et al., 2015). Furthermore, the 2004 IOT enhanced international collaboration in near-real-time tsunami monitoring by establishing the Indian Ocean Tsunami Warning and Mitigation System (IOTMWS), which is managed by the Intergovernmental Oceanographic Commission (IOC) (Hettiarachchi, 2018). The approach to tsunami monitoring has evolved from relying solely on near-distance sources and local observations to incorporating far-field tsunami sources and global observations (Imamura et al., 2019; Bernard and Titov, 2015).

Building upon these advancements, the 2011 Tohoku Tsunami catalysed further methodological innovations. Notably, probabilistic tsunami hazard assessment (PTHA) has been extensively developed to address uncertainties in hazard assessment that cannot be quantified through conventional scenario-based tsunami hazard assessment or worst-case scenarios (Davies et al., 2018; Grezio et al., 2017; De Risi and Goda, 2017). PTHA systematically considers all potential tsunami generation mechanisms and associated uncertainties in tsunami sources and parameters, thereby providing comprehensive risk information for long-term planning and coastal management in regions susceptible to potential tsunamis (Behrens et al., 2021).

While the 2004 Indian Ocean Tsunami (IOT) and the 2011 Tohoku Tsunami have significantly advanced the understanding of earthquake-generated tsunamis, the 2018 Palu tsunami introduced new complexities to tsunami science. Unlike its predecessors, the 2018 Palu tsunami was triggered by a more intricate combination of source mechanisms involving both an earthquake and a submarine landslide (Behrens et al., 2021). This combination produced catastrophic results, even though the fault mechanism initially suggested that the tsunami would not be highly destructive (Lahcene et al., 2021). The extensive damage was subsequently confirmed by comparing the Normalized Difference Vegetation Index (NDVI) values before and after the tsunami event (Fig. 1.2). NDVI values range from 0 to 1, where 0 indicates completely non-vegetated areas and values between 0.5 and 1 represent areas with healthy, dense vegetation cover.

Prior to the tsunami occurrence, NDVI values within the coastal areas of Palu City were predominantly in the range of 0.5 to 1, indicating that these areas had partial vegetation cover (Fig. 1.2b). Following the tsunami event, NDVI values decreased dramatically to predominantly 0-0.5, demonstrating that significant land cover changes had occurred as vegetation was destroyed and areas were transformed into water bodies (Fig. 1.2c). Additionally, land cover change detection analysis revealed that the vegetated sandpit located at the river mouth was completely removed, which contributed to the collapse of the bridge (Fig. 1.2d and 1.2e).

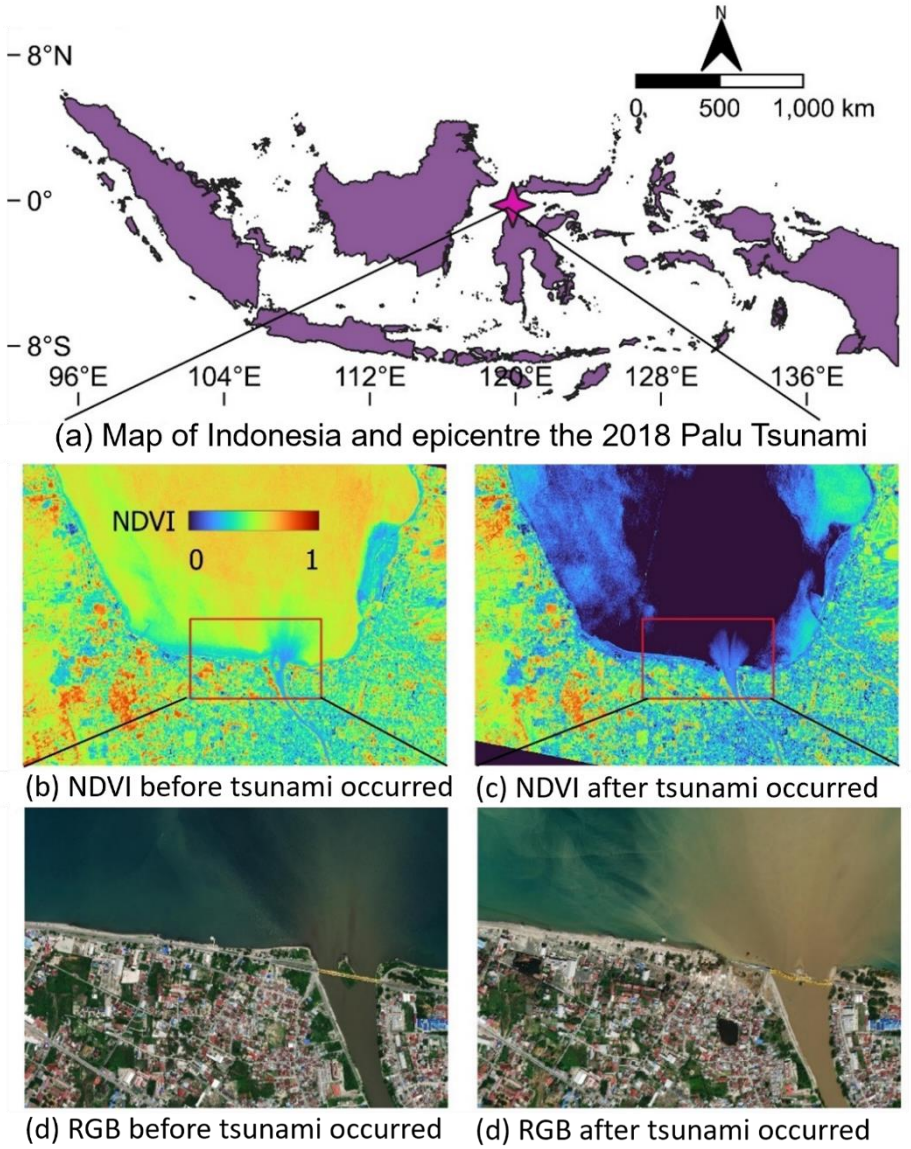


Fig. 1.2. Palu City before and after the tsunami on September 28, 2018. The normalized difference vegetation index (NDVI) shows land cover changes processed from Planet Scope (Planet, 2025). The RGB images display physical infrastructure damage, retrieved from the Maxar Open Data Program (Maxar, 2025).

Following the Palu tsunami, research on landslides as tsunamigenic sources has gained wide attention (Cecioni et al., 2023; Heidarzadeh and Mulia, 2023; Somphong et al., 2022; Nakata et al., 2020). The 2018 Palu tsunami also improved tsunami damage assessment studies by expanding the fragility function database (mathematical relationships that estimate damage probability based on hazard intensity) for compound tsunamigenic sources–cases where multiple mechanisms such as earthquakes and landslides contribute to tsunami generation (Lahcene et al., 2021; Mas et al., 2020; Imamura et al., 2019).

Despite advancements in scientific understanding, the 2018 Palu tsunami underscored persistent challenges in disaster risk reduction (DRR) within this vulnerable region, which experienced significant tsunami events in 1927, 1968, and 1996 (Ho et al., 2021). These challenges pertain to the accurate prediction of tsunami intensity, including its extent and height, as well as the projection of expected casualties (Iuchi et al., 2023; Rafliana et al., 2022). Previous research has emphasized that the precise modelling of tsunami hazards and the assessment of exposure are critically dependent on two fundamental types of input data: elevation models and exposure datasets (Raduszynski and Numada, 2023; Tonini et al., 2021; León et al., 2019; Ehrlich et al., 2018). However, as Behren et al. (2021) noted, these essential data are often unavailable, outdated, or inaccurate, particularly in low- and middle-income countries, where the majority of tsunami events occur (see Fig. 1).

1.2 Challenges in Tsunami Exposure Assessment

Given the challenges introduced by these essential input data, the following subsections systematically examine these challenges and present the research gaps that must be addressed by future studies.

1.2.1 Limitations in Elevation Models

In addition to tsunami source scenarios, numerical models, and computational capabilities, the precision of tsunami inundation estimates is significantly affected by the quality of topographic and bathymetric data (Gibbons et al., 2022; Sugawara, 2021; Griffin et al., 2015). Ideally, high-resolution topographic data obtained through airborne light detection and ranging (LiDAR) can adequately represent terrain features, thereby facilitating the creation of detailed inundation maps (McClean et al., 2020).

However, the availability of LiDAR topographic data is limited, particularly in middle- and lower-income countries (Pronk et al., 2024). Simultaneously, two of the three most catastrophic tsunamis over the past two decades have occurred in middle-income countries (see Fig. 1.1). As a result, most inundation hazard assessments depend on the "best-available" elevation datasets, including global digital elevation models (DEMs) (Hawker et al., 2018).

Unfortunately, global DEMs frequently exhibit inaccuracies, which can lead to incorrect flood predictions. For example, the widely used Shuttle Radar Topography Mission (SRTM) (Farr et al., 2007) often underestimates flood coverage, leading to an underestimation of the population at risk of coastal flooding (Hinkel et al., 2021). This discrepancy may be attributed to errors associated with SRTM, such as biases related to elevation and vegetation (Yamazaki et al., 2017).

Recent research has addressed inherent biases in SRTM. These investigations have led to the development of improved variants of SRTM, such as NASADEM (NASA JPL, 2020), CoastalDEM (Kulp and Strauss, 2018), and multi-error-removed improved terrain DEM (MERIT) (Yamazaki et al., 2017). The latter was created by integrating data from SRTM and ALOS World 3D – 30 m (AW3D30) (Tadono et al., 2016). In addition to SRTM, efforts to mitigate errors in the relatively new Copernicus DEM – 30 m (COP30) (Fahrland et al., 2022), have resulted in two other enhanced variants: the forest and building removed DEM (FABDEM) (Hawker et al., 2022), and DiluviumDEM (Dusseau et al., 2023). These improved elevation datasets are collectively referred to as error-reduced DEMs.

To the best of our knowledge, the application of error-reduced DEMs has been predominantly limited to sea-level rise projections and glacier assessments (Seeger et al., 2023; Chen et al., 2022). This limitation underscores the necessity of evaluating the performance of these new DEMs in broader coastal flooding contexts, particularly concerning tsunami inundation. Moreover, when uncertainties in DEM accuracy are compounded with uncertainties in tsunami source data, the resultant effect may substantially affect the reliability of the predictions. This critical intersection of uncertainties and their impact on exposure estimates remains largely unexplored in the current literature, highlighting the urgent need for targeted research in this area.

1.2.2 Limitations in Land Cover Roughness Models

In the absence of high-resolution topographic data, the effects of terrain attributes can be represented by integrating "bare-earth" topography, which excludes buildings and vegetation elevations, with friction models (Gibbons et al., 2022; Sadashiva et al., 2022). To investigate detailed tsunami behaviours, prior research has developed several friction models, including the structure resolve model (SRM), drag force model (DFM), and individual drag force model (iDFM) (Fukui et al., 2022). However, these models require relatively high-resolution elevation datasets ranging from 0.5 m to 5 m. Consequently, the application of a simple bottom friction model remains advantageous owing to its compatibility with moderate-resolution data (10 – 100 m).

The implementation of the bottom friction model in inundation modelling involves selecting Manning's roughness coefficients (n-values) to represent the effects of terrain features such as buildings and vegetation (Bricker et al., 2015). These coefficients may be spatially uniform or vary based on land cover maps, also referred to as land cover roughness (LCR) models. Previous studies have indicated that high-resolution LCR models can produce more accurate inundation results (Kaiser et al., 2011; Gayer et al., 2010).

However, the extent to which the spatial resolution of LCR models influences inundation estimates has not been thoroughly examined. Although Laso Bayas et al. (2015) explored the effect of various LCR models derived from multiresolution land cover maps on tsunami simulations, the low accuracy of the utilized land cover maps ultimately impeded a definitive conclusion. This knowledge gap necessitates further investigation using accurate land cover maps to establish reliable relationships between the LCR model resolution and inundation prediction accuracy.

1.2.3 Limitations in Tsunami Hazard Assessment

In addition to the challenges associated with accurately modelling the extent of inundation, the characteristics of exposure data—comprising input data, methodology, and spatial and temporal resolution—also significantly impact the assessment of inundation exposure (Behrens et al., 2021). Oktari et al. (2025) emphasized that information regarding tsunami exposure plays a crucial role in enhancing the preparedness of coastal communities.

The assessment of tsunami exposure entails the integration of tsunami intensity measures, such as flow depth and inundation extent, with critical societal exposure variables, including built-up areas, buildings, and population density (Ehrlich et al., 2018). Consequently, the precision of these societal variables ultimately determines the accuracy of exposure information. However, Behrens et al. (2021) observed that in numerous developing countries, exposure data are often unavailable or outdated, necessitating the use of the "best available" global open data for exposure assessments.

With the advancement of Earth observation systems, numerous global-scale exposure datasets, such as land cover maps, built-up layers, and gridded population datasets, have become publicly available (Table 1.1). These datasets were developed using diverse input data, methodologies, and spatiotemporal resolutions. When datasets with low temporal resolutions are integrated with others for long-term disaster assessment, they may introduce bias owing to geolocation misalignment across datasets and methodological inconsistencies (Sleeter et al., 2017). Bonatz et al. (2024) specifically emphasized that methodological inconsistencies, such as the definition of built-up areas, resulted in substantial differences in the exposed population, with variations of up to 65%.

Variations in the distribution of populations across built-up areas or buildings result in numerous discrepancies in the exposed population (Kuffer et al., 2022). Smith et al. (2019) highlighted those existing open datasets face challenges in accurately representing concentrations of exposure, as the total exposed population is dispersed over extensive areas. Although several studies have been conducted to assess bias in global exposure datasets, most have focused on the country level. In contrast, Bernhofen et al. (2022) emphasized that national-level bias may not accurately reflect behaviour at the local scale.

Given the identified limitations, it is essential to evaluate global exposure products and their inherent biases before utilizing them for local risk assessment, particularly in regions with limited data. Furthermore, because elevation models and exposure datasets each possess distinct biases, it is critical to comprehend how these biases may compound when used in conjunction. However, to our knowledge, research

specifically addressing the combined effects of these biases on tsunami exposure assessments remains limited.

Table 1.1. Global exposure datasets

Dataset	Resolution	Epoch	Input datasets	Reference
Built-up and impervious surfaces				
GHS–BUILT	100 m	1975 - 2030	Landsat; Sentinel-2	Pesaresi et al, 2024
GUF	12 m	2012	TanDEM; TerraSAR-X	Esch et al., 2017
GAIA	30 m	1985 - 2018	Landsat; Sentinel-1	Gong et al., 2020
GAUD	30 m	1992 - 2020	Landsat; NTL	Zhao et al., 2022
GISA	30 m	1972 - 2019	Landsat 5, 7, 8	Huang et al., 2021
WSF-2015	10 m	2015	Landsat-8; Sentinel-1	Marconcini et al., 2020
WSF-2019	10 m	2019	Sentinel-2; Sentinel-1	Marconcini et al., 2021
WSF-Evolution	30 m	1985 - 2015	Landsat-5; Landsat-7	Marconcini et al., 2021
GISD30	30 m	1985 - 2020	Landsat 4, 5, 7, 8	Zhang et al., 2022
Population distribution				
GHS–POP	100 m	1975 - 2030	Census	Pesaresi et al, 2024
HRSL	30 m	2016	Census	Tiecke et al., 2017
WorldPop	100 m	2000 - 2020	GPWv4; Census	Lloyd et al., 2019
LandScan Global	1 km	2000 - 2022	Census	Lebakula et al., 2025
GlobPop	1 km	1990 - 2020	Products assimilation	Liu et al., 2024
Land cover map				
ESRI Annual	10 m	2017 - 2022	Sentinel-2	Karra et al., 2021
ESA World cover	10 m	2020 - 2021	Sentinel-2; Sentinel-1	Zanaga et al., 2021
GLC	10 m	2015 - 2019	Proba-V	Buchhorn et al., 2020
Dynamic World	10 m	2016 - 2024	Sentinel-2	Brown et al., 2022
GLC-2015	30 m	2014	Multi Global LC fusion	Li et al., 2023
GLC_FCS30D	30 m	1985 - 2022	Landsat 5, 7, 8, 9	Zhang et al., 2024
Glance	30 m	2001 – 2020	Landsat 5, 7, 8, 9	Friedl et al., 2022

1.3 Area of interest (AOI)

The Area of Interest (AOI) of this study is situated in Banda Aceh (BNA), a prominent coastal city located at the northern extremity of Sumatra, Indonesia (Fig. 1.3). The BNA comprises nine subdistricts (Fig. 1.3b) and is characterized by a low-lying coastal plain with an average elevation of less than 10 m above sea level (Meilianda et al., 2019). The city is positioned along the tectonically active Sunda-Andaman subduction zone and was significantly impacted by the 2004 Indian Ocean Tsunami (IOT) (Fig. 1.3c). The 2004 IOT was induced by a catastrophic earthquake with a moment magnitude (Mw) of 9.2 (Yanagisawa et al., 2010; Koshimura et al., 2009). The consequent tsunami resulted in approximately 73,000 fatalities within the BNA, representing 30% of its total population in 2004 (Syamsidik et al., 2023).

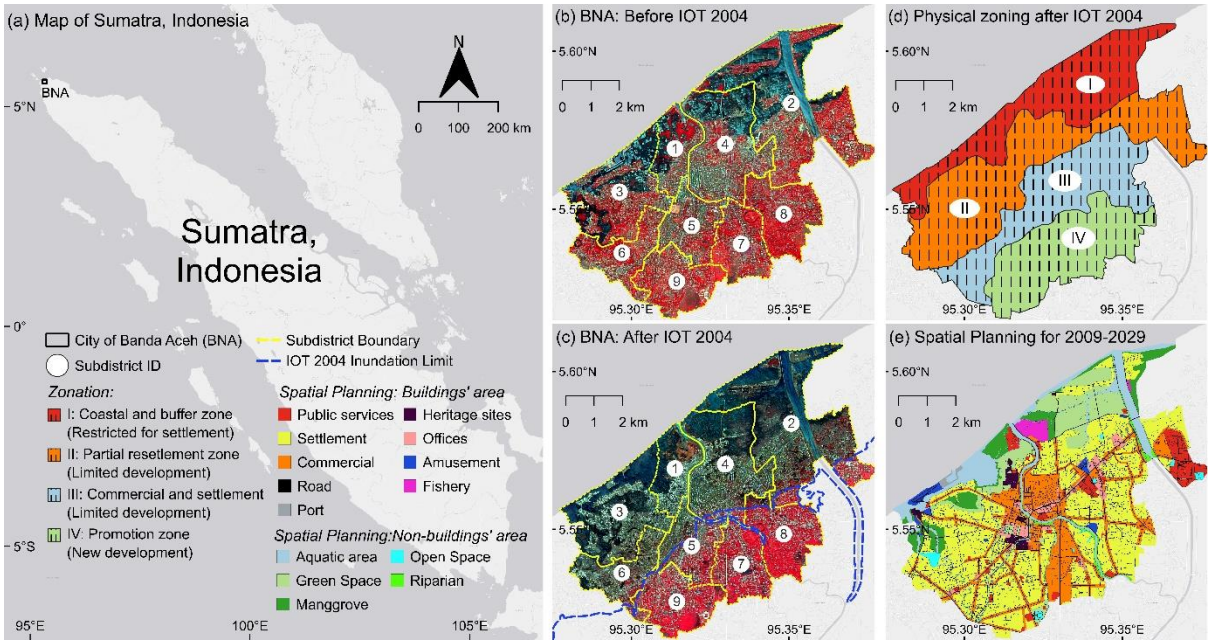


Fig. 1.3. Area of interest (AOI). (a) Map of Sumatera (Basemap: ESRI Light Gray). (b) BNA before the 2004 IOT, yellow lines and circled numbers indicate the subdistricts boundary and ID, respectively (Basemap: SPOT 5 05-08-2004). (c) BNA after the 2004 IOT, blue line indicates the inundation limit, manually digitized from SPOT-5 image (Basemap: SPOT 5 29-01-2005). (d) Physical zoning, and (e) Spatial land use planning map of 2009-2029 (Banda Aceh Municipality, 2018).

In response to the 2004 IOT disaster, BNA City was reconstructed according to a master plan that incorporated historical tsunami data through four designated zoning schemes (Fig. 1.3d). Zone I was designated as a restricted development area, Zones II and III were identified as limited development zones, and zone IV was earmarked as a promoted development zone. However, local residents expressed opposition to this zoning system because of its restrictions on their living areas (Takabatake, 2022). As a result, the local municipality revised the initial zoning plan by transforming Zones I, II, and III into ecozones and traditional city center areas, while retaining the concept of the newly promoted development zone (Zone-IV) (Banda Aceh Municipality, 2018).This revision was translated into a spatial land use planning map, as shown in Fig. 1.3e.

Despite these planning efforts, previous research documented substantial urban expansion within the coastal zone after the 2004 IOT event. Meilianda et al. (2019) found a 30.7% increase in built-up areas by 2017 relative to pre-tsunami baselines, while Syamsidik et al. (2023) reported an additional 31% growth in building stock between 2017 and 2021. This rapid coastal development raises critical questions

about adherence to the prescribed land use planning framework and potential exposure to future tsunami hazards.

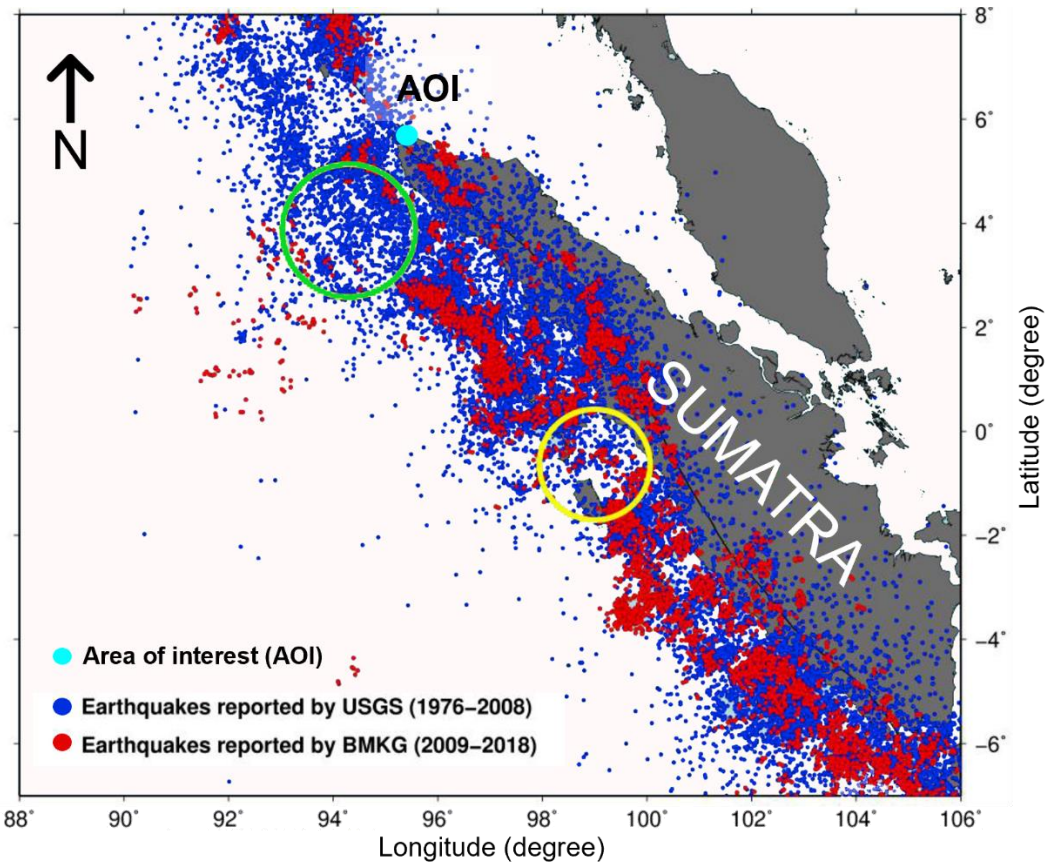


Fig. 1.4. Spatial distribution of seismic events from 1976 to 2018 revealing two distinct seismic gap zones along the Sumatran coast denoted by coloured circles. The green circle shows the closest seismic gap zone within the area of interest (AOI). Modified from Jihad et al. (2020)

The 2004 IOT also triggered ongoing post-seismic processes, including long-term slow-slip events (SSEs) that generated microseismicity and tectonic tremors (Sarkawi et al., 2024). The US Geological Survey documented over 6,000 earthquakes with magnitudes Mw ≥ 4.5 along the Sunda megathrust during the first decade following 2004 (USGS, 2017). Although these earthquakes occurred randomly along the Sunda subduction zone, Jihad et al. (2020) identified seismic gap zones with reduced activity along the Sumatran coast (Fig. 1.4). The decrease in seismic activity reflects an ongoing stress accumulation mechanism, implying that the gap zone retains the potential to generate large earthquakes. Through the analysis of seismic data (Mw ≥ 2.0) from 1976 to 2018, which were compiled from both the USGS and the Indonesian Meteorology**,** Climatology, and Geophysics Agency (BMKG), Jihad et al. (2020) determined that these gap zones could potentially generate tsunamigenic events up

to a magnitude of Mw 8.7. Critically, the seismic gap zone closest to the BNA (indicated by the green circle in Fig. 1.4) poses a direct threat to our AOI.

1.4 Research Formulation

Considering the substantial urban transformation and persistent seismological hazards, it remains essential to evaluate the long-term exposure changes, the alignment of urban development with land use planning, and the efficacy of the current mitigation system in addressing potential tsunamigenic events within the AOI. However, given the existing challenges in exposure assessment owing to the availability of essential datasets, several fundamental research questions emerge:

1. Two decades after the 2004 IOT event, what are the current tsunami exposure estimates within the AOI?
2. In the absence of local exposure datasets, how accurate are exposure estimates using global datasets? How do the biases inherent in global datasets affect the magnitude and variation of tsunami exposure estimates?
3. To what extent has the two decades of post-disaster urban development adhered to prescribed spatial land use planning?
4. How adequate are existing mitigation measures against the identified Mw 8.7 tsunamigenic threat?

Addressing these questions is essential for developing evidence-based disaster risk reduction strategies and informing future urban planning decisions in tsunami-prone coastal environments with limited data availability. Therefore, to address these research questions and bridge the identified knowledge gaps, this study aims to:

1. Assess the sensitivity of global error-reduced Digital Elevation Models (DEMs) in modelling tsunami inundation due to the absence of accurate elevation information. In addition, this study examines the performance of DEMs when integrated with multiresolution land cover roughness (LCR) models to accurately simulate tsunami inundation.
2. Quantify the biases inherent in global built-up and gridded population datasets when projecting assets and populations vulnerable to tsunami hazards and evaluate their accuracy in comparison to our high-resolution exposure models.

3. Analyse compounding bias effects on tsunami exposure estimates to understand cumulative uncertainties in assessment methodologies.
4. Evaluate exposure progression by analysing the evolution of exposure over the past two decades and project anticipated exposure from prospective tsunamigenic scenarios.
5. Assess the current mitigation adequacy by incorporating compounding bias into a probabilistic tsunami hazard assessment to critically evaluate existing mitigation systems within the AOI.

1.5 Thesis Structures

To address the aforementioned research objectives systematically, this study is divided into the following chapters.

- *Chapter I: Introduction*. This section describes the research background and objectives.
- *Chapter II: Bias in Global Exposure Datasets*. This chapter examines global exposure data and their biases in assessing the changes in exposure within our AOI two decades after the 2004 IOT. In this section, high-resolution datasets for land cover, built-up areas, and gridded populations were created and then compared with global datasets.
- *Chapter III: Sensitivity of Digital Elevation Models and Land Cover Roughness.* This section offers an in-depth evaluation of global elevation models and their effectiveness in reconstructing the 2004 IOT. Additionally, this chapter provided an evaluation how the spatial resolution of LCR models impacts their performance when integrated with these elevation models
- *Chapter IV: Compounding Bias and Cross-Uncertainty Exposure Assessment.* This chapter examines how the compounding bias from elevation models and global exposure datasets affects tsunami exposure estimates within our area of interest. In this section, a tsunami hazard assessment is conducted to anticipate future tsunamigenic events by employing hypothetical tsunami scenarios. The results from the probabilistic simulations were combined with bias assessments from the DEMs and LCR models to create a cross-uncertainty tsunami hazard map. Utilizing this hazard map along with the evaluated exposure datasets, the population exposure estimates in BNA were analysed.

— *Chapter V: Exposure and Mitigation System.* This chapter evaluates the effectiveness of tsunami shelters against the number of populations exposed to future tsunamigenic events. This chapter also evaluates the effectiveness of spatial planning against land cover evolution.
— *Chapter VI: Discussion.* This chapter outlines the conclusions and limitations of this study. Additionally, based on these research findings, some suggestions to improve tsunami disaster preparedness and risk reduction were offered.

CHAPTER II: BIAS IN GLOBAL EXPOSURE DATASETS

2.1 Evaluated Global Datasets

To conduct long-term exposure evolution detection after the 2004 Indian Ocean Tsunami, only global datasets with temporal coverage from 2004 to 2014 were selected. As a result, out of 21 global datasets listed in Table 1.1 (see chapter 1), this chapter reviewed 12 global exposure datasets, including eight built-up areas or impervious surface layers and four gridded population datasets covering the Banda Aceh (BNA) region. A visual comparison of each dataset is displayed in Fig. 2.1. Sections 2.2 and 2.3 provided detailed explanations of each dataset.

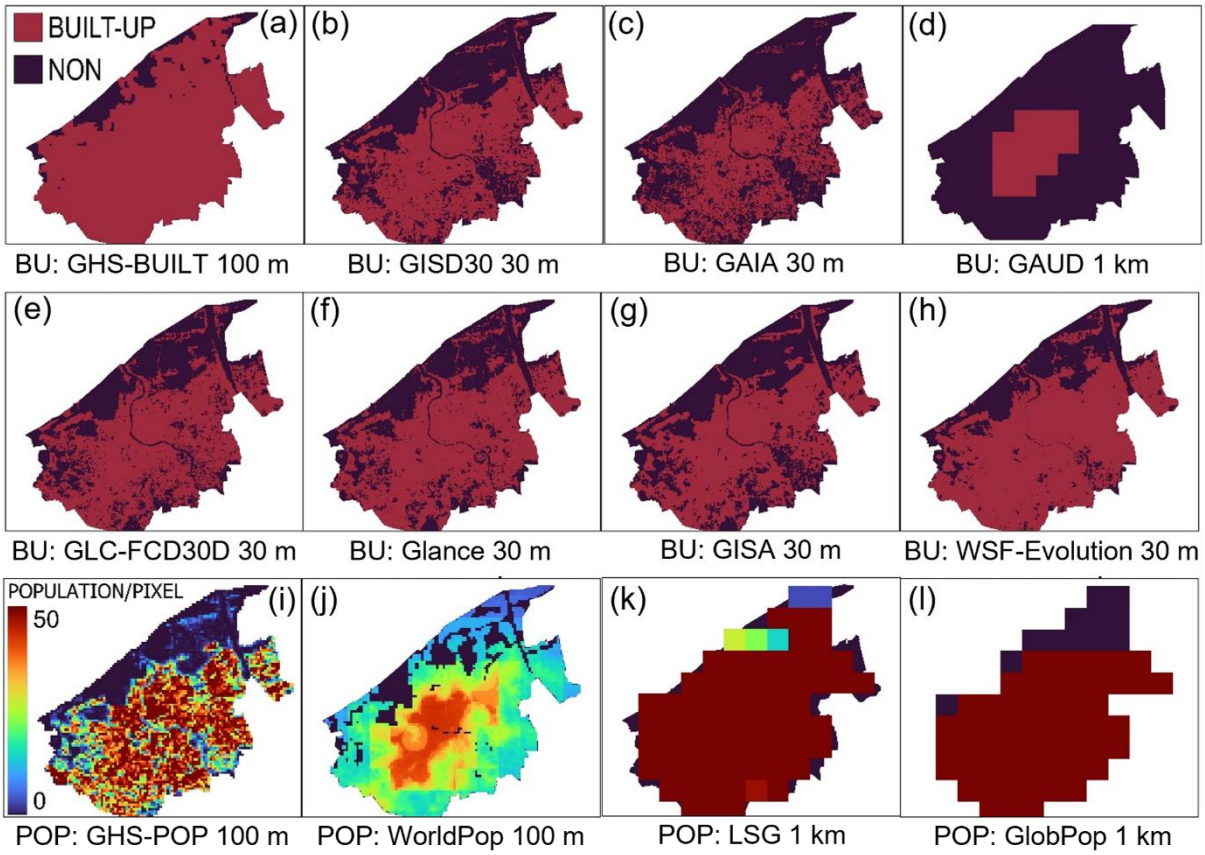


Fig. 2.1. Visual comparison of subset of global datasets covering BNA region for epoch 2004, with BU indicating built-up and impervious layers, while POP indicates gridded population datasets.

2.2 Built-up and Impervious Dataset

The resolution of the evaluated built-up and impervious surface layers ranges from 30 m to 1 km. Additionally, two land cover maps were included by exclusively analysing their built-up classes. A brief description of the datasets used is as follows:

a. *GHS-BUILT: Global Human Settlement Layer – Built Surface (Fig. 2.1a).* The GHS-BUILT layer was developed utilizing Symbolic Machine Learning (SML) based on a linear regression algorithm (Pesaresi et al., 2024). The input data integrated spatial-temporal interpolation of five-year observed collections of multisensor, multi-platform satellite images. Specifically, Landsat (MSS, TM, ETM sensor) data support the 1975, 1990, 2000, and 2014 epochs, while a Sentinel-2 (S2) image composite supports the 2018 epoch. Additionally, data from Facebook settlement delineation, Microsoft, and Open Street Map (OSM) building delineation were incorporated. The GHS-BUILT raster layer provides a built-up fraction using the total area count in square meters per each 100x100 grid size. This fraction is categorized into three groups: residential built-up surfaces, non-residential built-up surfaces, and a combination of both. This study utilized the combination of categories, acknowledging that human activities are not solely concentrated in residential buildings. The accuracy assessment indicated that the F1 score of GHS-BUILT reached 0.82 for total validation points.

b. *GISD30: Global 30 m Impervious-Surface Dynamic (Fig. 2.1b).* GISD30 was developed utilizing Harmonized Landsat imagery spanning from 1984 to 2020, specifically from Landsat 4, 5, 7, and 8, and was classified using the random forest algorithm (Zhang et al., 2022). The training and validation datasets were sourced from other global impervious surface land cover map products. The GISD30 raster layer is provided at a temporal resolution from 1985 to 2020, with five-year intervals, and is categorized into two classes: built-up and non-built-up. The accuracy assessment indicated that the overall accuracy achieved a score of 90.1%.

c. *GAIA: Global Artificial Impervious Area (Fig. 2.1c)*. Similar to the GISD30, the GAIA raster layer was provided at five-year intervals from 1985 to 2018 (Gong et al., 2020). In order to detect impervious surfaces, GAIA excluded all potential areas of vegetation, water, and bare land, subsequently generating temporally consistent artificial impervious maps through the analysis of Landsat time series data. To enhance accuracy, two ancillary datasets—nighttime light (NTL) and Sentinel-1 Synthetic Aperture Radar—were incorporated as input data. The accuracy assessment indicated that the overall accuracy achieved a score of 90%.

d. *GAUD: Global Annual Urban Extents Dataset (Fig. 2.1d)*. GAUD was developed utilizing harmonized nighttime light (NTL) time-series composites by integrating multi-source NTL observations, thereby providing a comprehensive and consistent record of the nightscape (Zhao et al., 2022). Areas emitting light were subsequently identified as built-up regions. The raster layer is available at a 1 km resolution, encompassing the period from 1992 to 2020.

e. *GLC FCS30D: Global 30 m Land-Cover Dynamics with Fine Classification System (Fig. 2.1e).* The GLC FCS30D land cover map encompasses the period from 1985 to 2022 (Zhang et al., 2024). Utilizing Landsat imagery as the input dataset and employing local adaptive classification models, it achieved a mean accuracy of 81.91%. This land cover map comprises ten land classes, of which only the impervious surface class was selected for further analysis.

f. *Glance: Global Land Cover Estimation (Fig. 2.1f).* The Glance dataset provides a global annual land cover map for the period from 2001 to 2020, with a spatial resolution of 30 m (Friedl et al., 2022). The land cover classification was conducted using a random forest classifier in conjunction with Continuous Change Detection and Classification algorithms. The input data comprised Landsat imagery, which was categorized into seven land classes, with the "developed/built-up" class being utilized for further analysis (Table 2.1). This land cover product achieved an overall accuracy of 79%.

g. *GISA: Global Impervious Surface Area version 2.0 (Fig. 2.1g).* GISA 2.0 was developed through the integration of various built-up area layers, including GISA 1.0, GAIA, GAUD, and GHS-BUILT (Huang et al., 2022). By employing a random forest classifier to assess the concordance of each input dataset, GISA 2.0 attained an F1-score of 0.93.

h. *WSF: World Settlement Footprints Evolution (Fig. 2.1h).* WSF provides information on human settlement from 1984 until 2015 at 30 m resolution. Leveraging Landsat imagery as the input dataset, WSF employed adaptive thresholding method to extract temporal mean values from several spectral indices, including Normalized Different Vegetation Index (NDVI), Normalized Different Building Index (NDBI), and Modified Normalized Different Water Index (MNDWI). These spectral indices were then used as input for image classification using random forest algorithm.

Table 2.1 Glance classification scheme

Land cover class	Descriptions
Water	Perennial water bodies encompass rivers, channels, ponds, storage basins, and seas.
Developed (Built-up areas)	Areas with high levels of use; land occupied by structures, including any land that is functionally linked to developed or constructed activities
Barren	Landscapes featuring natural formations of soil, sand, or rocks, where vegetation covers less than 10% of the surface
Tree Cover	Regions where the tree canopy covers more than 30% of the area. It's crucial to recognize that locations where trees have been cleared, like clear-cuts, are categorized according to their present condition, such as barren, sparsely vegetated, or covered with shrubs or grasses.
Herbaceous	The region is mainly dominated by herbaceous plants, with the total vegetation surpassing 10%. Trees cover less than 30% of the area, while shrubs make up under 10% of the land.
Shrublands	Regions where tree coverage is below 30% of the area, yet the total vegetation exceeds 10%, with shrubs making up more than 10% of the vegetation.
Ice/Snow	Areas where snow and ice consistently blanket more than half of the terrain throughout the year

2.3 Gridded Population Dataset

The resolution of the gridded population datasets ranges from 100 m to 1 km. A brief description of the datasets used is as follows:

a. *GHS-POP: Global Human Settlement Layer – Population (Fig. 2.1i)*. The GHS-POP layer was developed by disaggregating the population count into each built-up pixel cell of the GHS-BUILT layer at a 100 m spatial resolution (Pesaresi et al., 2024). The population data spanning from 1975 to 2030 were sourced from CIESIN for the Gridded Population of the World, version 4.11 (GPWv4.11), the UN World Population Prospects 2022, and the UN World Urbanization Prospects 2018. These population estimates were subsequently disaggregated within the residential built-up surfaces category at the district-level administrative boundaries (see Section 2.2). The administrative boundary data were derived from the Database of Global Administrative Areas (GADM) level 2 (https://gadm.org/). Láng-Ritter et al. (2025) identified that GHS-POP underestimated the total population in rural areas by 84% in 2000. However, Leyk et al. (2019) suggested that this gridded dataset might be suitable for use in urban areas due to its detailed urban settlement data.

b. *WorldPop: World Population Dataset (Fig. 2.1j).* WorldPop population grid datasets are generated using a Random Forest–based approach, distributing population estimates from GPWv4 into built-up layers at a 100 m spatial resolution, covering the period from 2000 to 2020 (Lloyd et al., 2019). The built-up layers are compiled and harmonized from multiple sources, including GSH-BUILT, OSM, and Global Urban Footprint (Esch et al., 2017). Opdyke and Fatima (2024) asserted that this dataset overestimates urban exposure estimates in Australia by 21%.

c. *LSG: LandScan Global (Fig. 2.1k)*. LSG disaggregated the census data into 1 km grid cells for the period from 2000 to 2022 by utilizing a likelihood coefficient between the auxiliary data and the population counts (Lebakula et al., 2025). Compared with other datasets, LSG is considered to represent the nighttime population rather than the residential population. Additionally, this dataset demonstrates the distribution of the working and traveling population, particularly in urban areas where there is a concern of population overestimation (Liu et al., 2024). In rural areas, LSG underestimates the population by 68% (Láng-Ritter et al., 2025).

d. *GlobPop*: *Global Gridded Population Dataset (Fig. 2.1l)*. GlobPop is a relatively new gridded population dataset with a 1 km spatial resolution, developed by harmonizing other global datasets, including LSG, GHS-POP, and WorldPop (Liu et al., 2024). In comparison to actual population dynamics, GlobPop demonstrates superior temporal accuracy relative to the other three datasets.

2.4 Bias Assessment Method

Two levels of bias assessment were conducted: intrinsic bias and analytical bias. Intrinsic bias was considered as the bias originating from the data source, which may arise due to limitations in input data, methodology, or validation data. Although the intrinsic bias assessment for each global built-up and gridded population layer was analyzed prior to public release, most were validated against global-scale data. Consequently, re-evaluation of biases in both dataset types against actual local data is crucial.

To evaluate the intrinsic bias, total built-up areas and population estimates from global datasets were compared against local datasets. To illustrate the processes involved in the intrinsic bias assessment, a flowchart is presented in Fig. 2.2. All the processes

were performed using GIS applications. Initially, all global exposure datasets were clipped within the area of interest (AOI). Then, total built-up areas and population within the AOI extent were derived and compared with the local datasets. The calculated biases were expressed in equations 1 and 2, with positive values indicating overestimation and negative values representing underestimation:

$$Relative\ bias\ in\ built-up\ areas = \frac{BU_{Global,T} - BU_{Local,T}}{BU_{Local,T}}\ x\ 100\% \quad (1)$$

$$Relative\ bias\ in\ population = \frac{POP_{Global,T} - POP_{Local,T}}{POP_{Local,T}}\ x\ 100\% \quad (2)$$

where $BU_{Global,T}$ and $BU_{Local,T}$ represent the total built-up areas within the AOI extent derived from global and local built-up datasets, respectively, for the year *T* (where T = 2004 or 2014). Similarly, $POP_{Global,T}$ and $POP_{Local,T}$ denote the total population within the same extent derived from global and local gridded population datasets, respectively, for the year *T*.

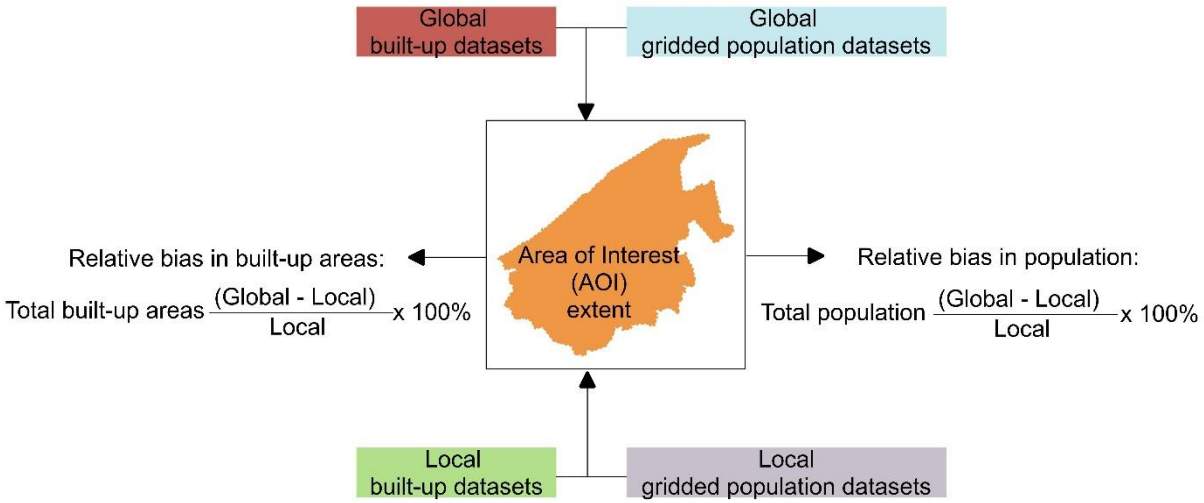


Fig. 2.2. Flowchart to perform the intrinsic bias assessment

To assess built-up areas and populations exposed to tsunami inundation, the global exposure datasets were overlaid with tsunami inundation extents. Total built-up areas and populations within the inundation extent were then referred to as exposure estimates. Biases introduced by this exposure estimates were referred to as analytical

bias. To identify the analytical bias, exposure estimates derived from global exposure datasets were compared against the local datasets, where the 2004 IOT inundation limit served as the inundation extent reference.

The analytical bias assessment consisted of two analyses. The first analysis focused on biases in exposure estimates during the 2004 IOT event. This analysis exclusively compared the 2004 temporal exposure datasets between global and local datasets (equations 3 and 4). The second analysis focused on biases in exposure estimate evolution over the 2004-2014 period. In this analysis, additional exposure estimates for the 2014 temporal period were evaluated and compared with the 2004 temporal estimates to assess how biases changed during the 10-year interval (equations 5 and 6). Subsequently, the rates of exposure evolution from global datasets were compared with those derived from local datasets. To summarize all processes involved in the analytical bias assessment, a flowchart illustration is presented in Fig. 2.3. All processes can be expressed mathematically using the following equations:

$$Relative\ bias\ in\ exposed\ built-up\ areas = \frac{EBU_{Global,2004} - EBU_{Local,2004}}{EBU_{Local,2004}}\ x\ 100\% \quad (3)$$

$$Relative\ bias\ in\ exposed\ population = \frac{EPOP_{Global,2004} - EPOP_{Local,2004}}{EPOP_{Local,2004}}\ x\ 100\% \quad (4)$$

$$Rate\ of\ exposure\ evolution\ \ for\ built-up\ areas = \frac{EBU_{(BU\ datasets,2014-2004)}}{EBU_{(BU\ datasets,2004)}}\ x\ 100\% \quad (5)$$

$$Rate\ of\ exposure\ evolution\ \ for\ populations = \frac{EPOP_{(POP\ datasets,2014-2004)}}{EPOP_{(POP\ datasets,2004)}}\ x\ 100\% \quad (6)$$

where $EBU_{Global,2004}$ and $EBU_{Local,2004}$ represent the total built-up areas within the 2004 IOT inundation extent derived from the global and local built-up datasets, respectively, for the year 2004. Similarly, $EPOP_{Global,2004}$ and $EPOP_{Local,2004}$ denote the total population within the same extent derived from the global and local gridded population datasets, respectively, for the year 2004. $EBU_{(BU\ datasets,2014-2004)}$ and $EPOP_{(POP\ datasets,2014-2004)}$ represent the evolution of built-up areas and population within the 2004 IOT inundation extent over the 2004-2014 period, where *BU datasets* and *POP datasets* can be either global or local sources. Meanwhile, *BU datasets, 2004* and

POP datasets, 2004 are built-up and population datasets derived from either global or local sources, but exclusively for the year 2004.

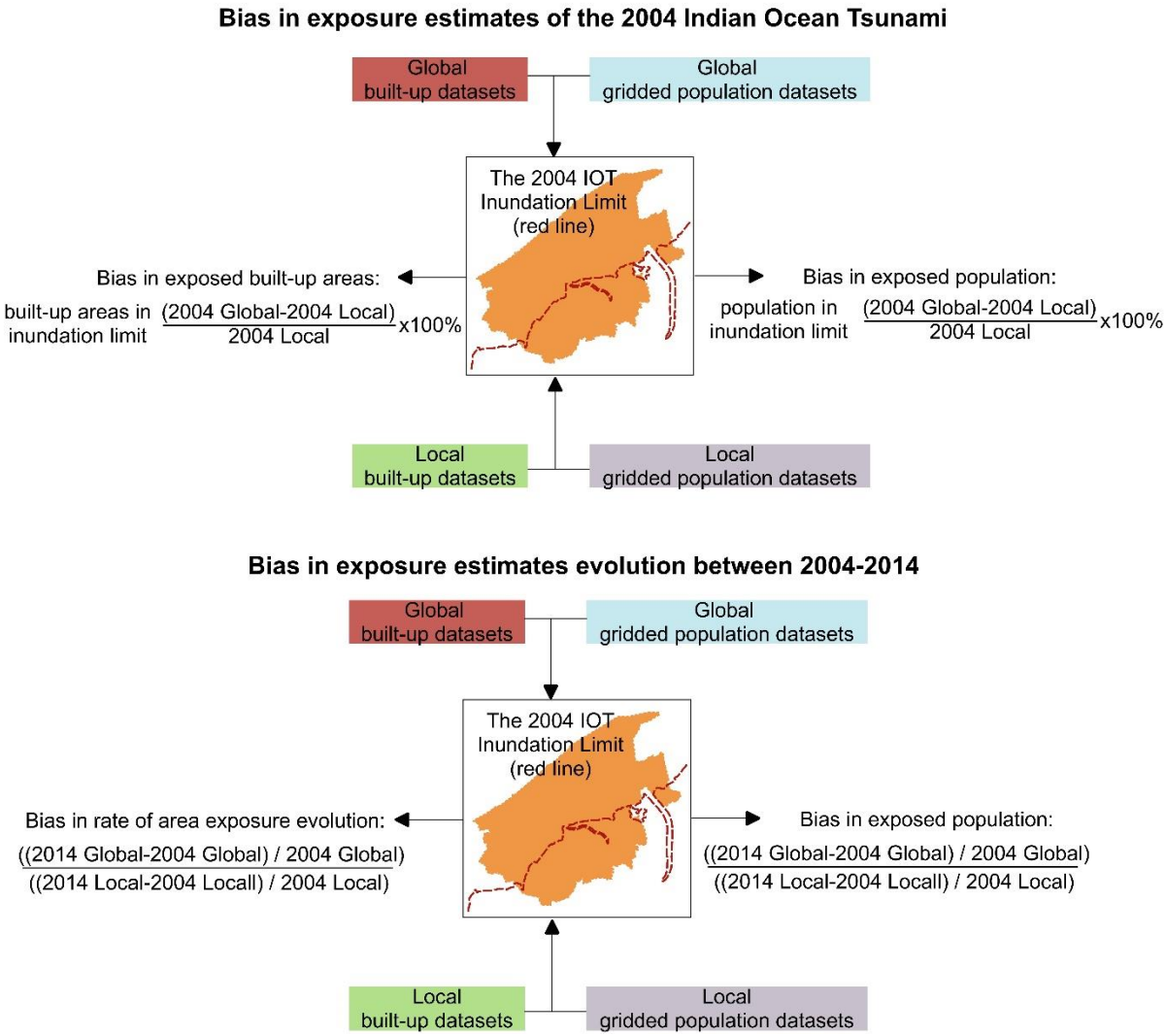


Fig. 2.3. Flowchart to perform the analytical bias assessment

2.5 Generating Local Exposure Datasets

Similar to the global exposure datasets, the local exposure datasets consisted of built-up area and gridded population layers. This study developed the exposure layers at 5-meter resolution, covering three temporal periods: 2004, 2014, and 2024. The 2004 and 2014 datasets were used to evaluate the global exposure datasets and assess exposure evolution following the 2004 Indian Ocean Tsunami (IOT). Meanwhile, the 2024 datasets were subsequently used to assess exposure estimates for future hypothetical tsunamigenic events.

2.5.1 Local Built-Up Area Datasets

To generate the local built-up area datasets, the land cover classification was performed using satellite imagery as input through Object-Based Image Analysis with a supervised Random Forest algorithm (OBIA-RF) (Kotaridis & Lazaridou, 2020; Phiri et al., 2018). Satellite images are optical images which consist of multiple spectral bands (e.g., visible bands: red, green, blue; near-infrared; shortwave infrared) that capture electromagnetic radiation across different wavelengths, providing both radiometric information (reflectance values across these various spectral bands) and textural information (spatial patterns and relationships between neighbouring pixels) (Shim, 2014).

OBIA-RF utilizes both radiometric and textural information extracted from the satellite images, where OBIA first segments the image into homogeneous objects based on pixel similarity and then RF classifies these objects into different land cover categories using the combined spectral-textural features (Hermosilla et al., 2022; Sideris et al., 2020). With adequate prior knowledge, such as labelled ground truth points, OBIA-RF is regarded as a reliable method for land cover mapping (Geiβ et al., 2017). It is important to note that a satellite image is released publicly at several different radiometric levels (Young et al., 2017):

- Digital Number (DN). This level represents the raw, uncalibrated electromagnetic signal detected by the satellite sensor. As illustrated in Fig. 2.4, solar radiation from the sun travels through space, reflects off the Earth's surface, and reaches the satellite sensor. The sensor records these incoming photons as discrete integer values for each pixel area. These integer values are also referred to as the raw "counts" or DN values. DN values are unitless and sensor-specific, representing the fundamental radiometric data before any calibration or atmospheric processing is applied.
- Top-of-Atmosphere (ToA) radiance. At this level, solar radiation has travelled from the sun to the Earth's surface and back to the satellite, but the values are now calibrated to represent physically meaningful radiance units ($W \cdot m^{-2} \cdot sr^{-1} \cdot nm^{-1}$). As illustrated in the diagram (Fig. 2.4), the measurement conceptually occurs at the top of the atmospheric boundary (blue hatched). The "At-sensor Radiance" represents the total electromagnetic energy detected by

the satellite sensor, which includes both the target radiance reflected from the surface (solid arrows from the surface object) and path radiance contributed by atmospheric scattering events (dotted arrows from the red scattering point). In simpler terms, radiance (L_λ) measures the amount of electromagnetic energy flowing through a unit area (pixel size, m²), into a unit solid angle of the sensor's view (sr), per unit wavelength (nm), and can be expressed as:

$$L_\lambda = \frac{DN}{Gain} + Bias \quad (7)$$

where *Gain* and *Bias* are calibrating coefficients and are provided in the ancillary XML metadata files that accompany satellite image downloads from operational data sources.

— Top-of-Atmosphere (ToA) reflectance. This level converts the ToA radiance into reflectance values by normalizing for solar illumination conditions. As shown in the diagram (Fig. 2.4), this accounts for the incoming solar radiation (yellow sun) and geometric factors such as solar zenith angle. ToA reflectance removes the influence of varying solar irradiance and illumination geometry, making it possible to compare measurements across different times and locations. ToA reflectance (ρ) is dimensionless (typically expressed as a fraction between 0 and 1 or as a percentage) and represents the proportion of incident solar radiation that is reflected back to the sensor, expressed as:

$$\rho_\lambda^{ToA} = \frac{\pi L_\lambda}{Cos(\theta_s) Es_\lambda} \quad (8)$$

where π equals to 3.14, $cos(\theta_s)$ is the cosine of the solar zenith angle, and Es_λ is solar irradiance. Both solar zenith angle and irradiance values are provided in the ancillary XML metadata files that accompany satellite image downloads from operational data sources. Although satellite images at TOA level still contain atmospheric effects, they can be used in remote sensing for environmental monitoring using spectral indices such as Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), etc (Roy et al., 2016).

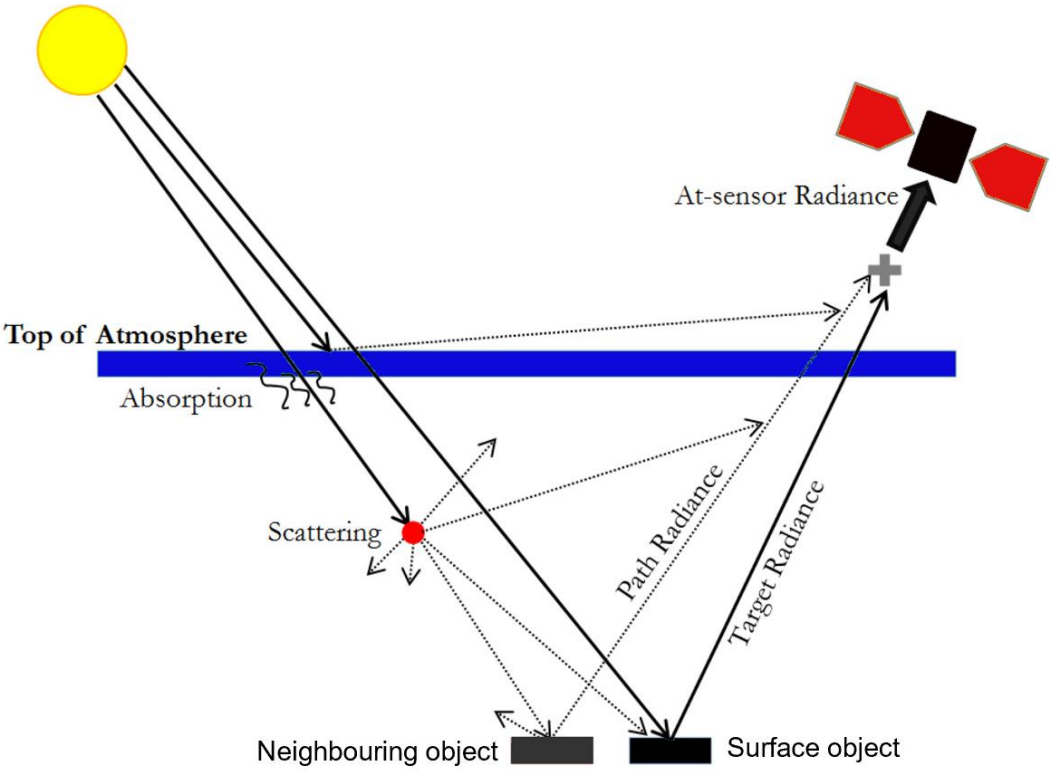


Fig. 2.4. Diagram illustrates the radiative transfer processes and radiometric measurement levels in satellite remote sensing. Modified from Fernández-Pacheco et al. (2018)

– Bottom-of-Atmosphere (BoA) reflectance. This represents the final, atmospherically corrected surface reflectance values. As illustrated in the diagram (Fig. 2.4), the atmospheric correction process removes the effects of atmospheric absorption (where solar radiation is absorbed within the atmosphere) and scattering (represented by the red scattering point and associated dotted path radiance arrows) that occur as solar radiation travels through the atmosphere. The BoA reflectance can be conceptually expressed as:

$$\rho_{\lambda}^{BoA} = \frac{\rho_{\lambda}^{ToA} - \rho_{\lambda}^{atm}}{T_{\lambda}^{down} + T_{\lambda}^{up}} \tag{9}$$

where ρ_{λ}^{atm} is atmospheric path reflectance (scattering by air molecules and aerosols), while T_{λ}^{down} and T_{λ}^{up} are the total fraction of sunlight reaches the surface after passing through the atmosphere and then successfully travels back up to the satellite, respectively. These calibration coefficients are provided in the ancillary XML metadata files that accompany satellite image downloads from

operational data sources. It should be noted that the conversion from ToA to BoA reflectance is complex and typically requires sophisticated atmospheric correction algorithms such as the 6S algorithm, which account for various atmospheric parameters including aerosol optical depth, water vapor content, and atmospheric gases (Yang et al., 2021). BoA reflectance is essential for accurate land cover classification and mapping applications, as it represents the true spectral characteristics of Earth's surface objects without the confounding effects of atmospheric variability (Tu et al., 2022).

2.5.1.a Input Satellite Images

The inputs for the land cover classification were multiresolution satellite images covering the period of 2004-2024. All images were cloud-free and derived from two sensors: SPOT 5 for the 2004-2014 period and PlanetScope (PS) for the 2024 period. The SPOT 5 images consist of four spectral bands at 10 m resolution and one panchromatic band at 5 m resolution (Table 2.2). These images were obtained from the SPOT World Heritage (SWH) program and were provided in Level 0 image file format, which are radiometrically corrected to top-of-atmosphere (ToA) radiance level but geometrically uncorrected in planar earth coordinates (Nosavan et al., 2020; Frazier & Hemingway, 2021). Therefore, additional pre-processing steps were required to prepare these images for analysis, including geometric correction, and radiometric calibration to obtain surface reflectance values suitable for land cover classification.

To correct the geometric distortions in the SPOT 5 images, orthorectification was performed using the SPOT SWH Carto-services (https://swh-2a-carto.fr/), provided by the French national space agency or *Centre National D'études Spatiales* (CNES). The orthorectification process corrects geometric distortions in satellite imagery by removing perspective effects and transforming the image to an accurate overhead view with proper map coordinates (Shean et al., 2016). It should be noted that the orthorectification was performed to both multispectral bands and panchromatic band (PAN) (see Table 2.2). At this stage, the SPOT 5 images has transformed into level 1─geometrically orthorectified at ToA radiance radiometric level (Crawford et al., 2023).

Table 2.2. Input datasets for image analysis

Period	Scene ID	Acquisition	Sensor	Resolution
2004	004-S5-255-339 005-S5-255-339	05/08/2004	SPOT 5 Spectral radiometric wavelength (in µm): Near infrared (NIR): 0.78-0.89 Red: 0.63-0.69 Green: 0.52-0.60 Short wave infrared (SWIR): 1.55-1.75 Panchromatic (PAN): 0.48-0.71	Spatial: NIR: 10 m Red: 10 m Green: 10 m SWIR: 10 m PAN: 10 m
2005	004-S5-255-339 006-S5-255-339	29/01/2005		
2014	004-S5-255-339 005-S5-255-339	18/05/2014		
2024	20240629_032604_67_24af_3B	29/06/2024	Planet Super Dove Spectral radiometric wavelength (in µm): NIR: 0.84 – 0.88 Red: 0.65 – 0.680 Green: 0.55 – 0.58 Blue: 0.46-0.52	Spatial: NIR: 3 m Red: 3 m Green: 3 m Blue: 3 m

To obtain images at surface reflectance level, the atmospheric correction was performed using the 6s algorithm. This process was executed using the Python extension of Orfeo Toolbox (pyOTB) (Grizonnet et al., 2017). PyOTB sequentially converted the radiometric values from ToA radiance to ToA reflectance, then to BoA (surface) reflectance. Fig. 2.5 presents the transformation of SPOT 5 images from sensor measurements (ToA radiance) to physically meaningful surface reflectance values. The analysis was performed using a random sample of approximately 1,000 pixels per spectral band, with boxplots illustrating the distribution of digital values across the green (0.50-0.59 µm), red (0.61-0.68 µm), NIR (0.78-0.89 µm), and SWIR (1.58-1.75 µm) spectral ranges. Instead of the 0-1 range, the surface reflectance values were scaled by 1,000 to reduce quantization error. At this stage, the SPOT 5 images has transformed into level 2.

Fig. 2.6 summarizes the product from each pre-processing steps for the SPOT 5 images. As shown in the illustration, there is a progressive brightening from ToA radiance to ToA reflectance and then to surface reflectance. The radiance image appears darkest as it represents the raw digital numbers converted to physical units of energy measured at the sensor (Fig. 2.6a). The TOA reflectance image shows increased brightness after normalization for solar illumination conditions and Earth-Sun distance, eliminating variations due to seasonal and daily solar geometry (Fig. 2.6c).

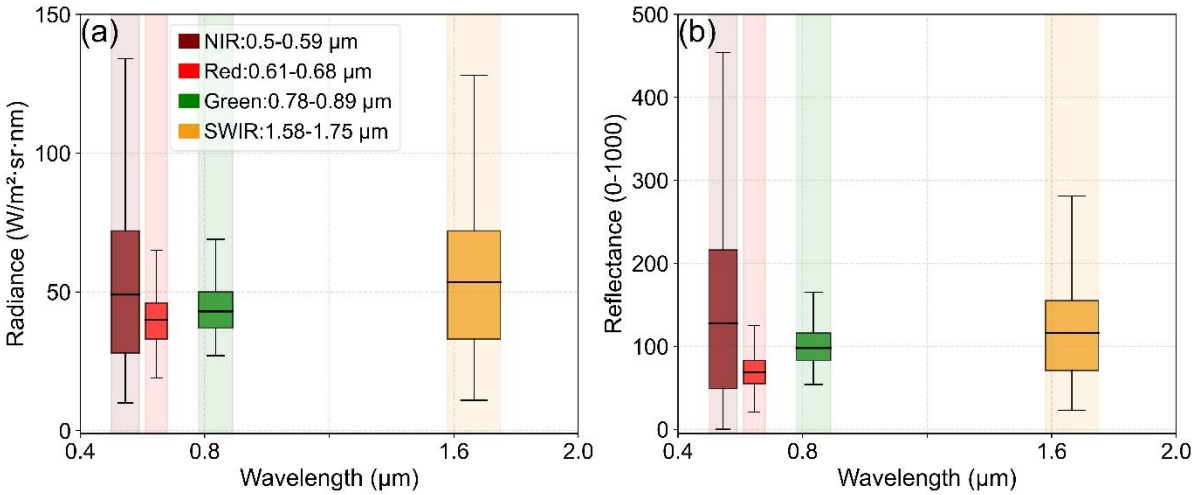


Fig. 2.5. Comparison band-to-band radiometric properties of SPOT 5 image at: (a) ToA radiance, and (b) BoA (surface) reflectance.

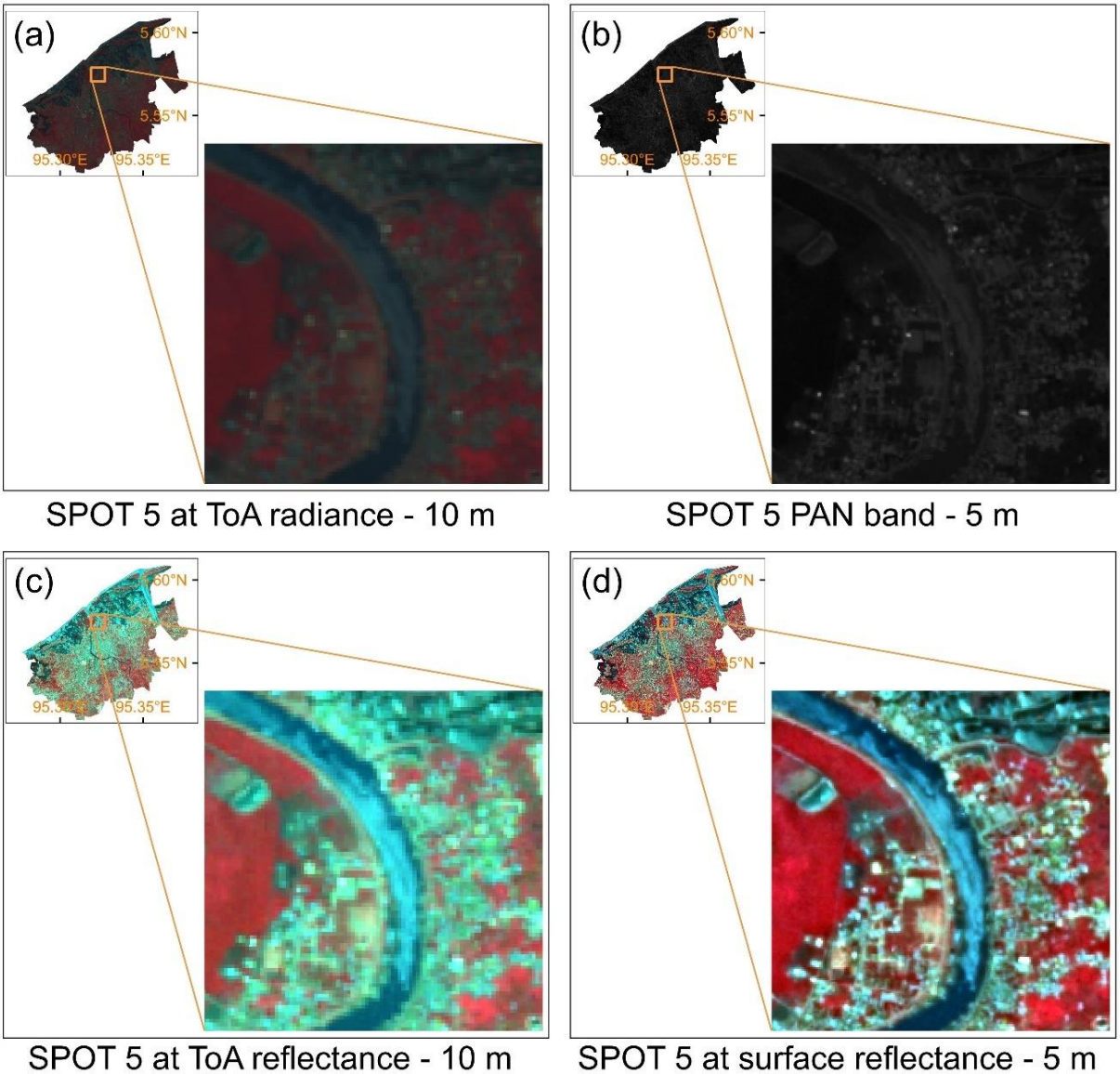


Fig. 2.6. A visual comparison of the 2004 SPOT 5 image at different radiometric levels.

The surface reflectance image appears brightest following atmospheric correction, which removes the effects of atmospheric scattering and absorption that attenuate the signal during its passage through the atmosphere (Fig. 2.6d). The pansharpening process resulted the surface reflectance to be more sharped. In addition, the pansharpening process enhanced the spatial detail of the surface reflectance image by incorporating the higher resolution information from a panchromatic band (Fig. 2.6b), resulting in improved edge definition and feature clarity while preserving the spectral characteristics of the multispectral data.

Meanwhile, PS images consist of visible bands (i.e., red, green, and blue) and NIR band (Planet, 2023). Each band has spatial resolution at 3 m and was obtained from Planet Explorer (www.planet.com/explorer). The PS data products are delivered as orthorectified surface reflectance (level 2), having undergone both geometric correction (orthorectification) and atmospheric correction to remove scattering and absorption effects, with reflectance values are scaled by 10,000. To match the spatial resolution of the SPOT 5 images, the PS datasets were resampled at 5 m resolution.

2.5.1.b Harmonization of Multisensor Input Images

Fig. 2.7 presents the false colour band composites of the analysis-ready input images, which were used for land cover classification (Fig. 2.7a, 2.7c, and 2.7d) and identification of the 2004 IOT inundation limit (Fig. 2.7a and 2.7b). The false colour composites use NIR, red, and green bands instead of the standard red-green-blue combination (natural colour). This could enhance the contrast between vegetation, water, and built-up areas, facilitating a better visual comparison between pre- and post-tsunami conditions within the AOI.

However, as described in Table 2.2, the spectral wavelengths between the SPOT-5 and PS datasets differ significantly. Moreover, while both sensors provide 4-band multispectral imagery, the SPOT-5 sensor does not include a blue spectral band. Conversely, PS does not provide an SWIR spectral band. This disparity can affect downstream products such as land cover maps, leading to inconsistency and introducing bias in land cover change analysis (Leach et al., 2019). Besides the radiometric differences, satellite images derived from different sensors potentially have slight spatial misalignment at the pixel level. Indeed, spatial misalignment can also occur within the same sensor, especially for images acquired at different periods

(Behling et al., 2014). It should be noted that spatial misalignment even at the sub-pixel level can affect the reliability of change detection maps (Sundaresan et al., 2007).

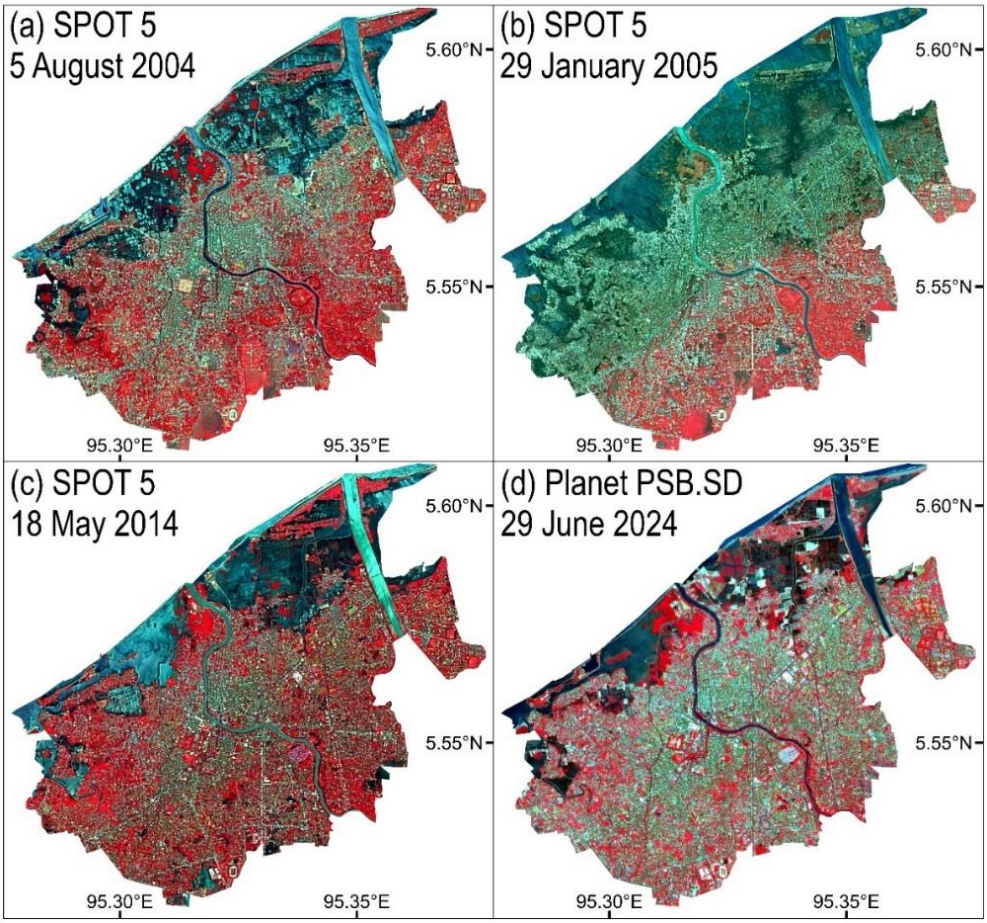


Fig. 2.7. Input images at 5 m resolution for the land cover classification and the 2004 IOT inundation limit detection.

Therefore, spatial and spectral harmonization of the input images prior to land cover classification is essential. The harmonization was performed in two sequential steps. First, spatial misalignment between the multisensor images was detected and minimized through image-to-image co-registration. Second, radiometric inconsistencies in the spatially aligned images were assessed and corrected through radiometric normalization.

The image-to-image co-registration was performed using AROSICS, a Python-based, open-source tool designed for the automatic detection and correction of spatial misalignments between cross-sensor images at sub-pixel level (Scheffler et al., 2017).

It could work effectively regardless of the images' resolution, spectral band characteristics, cloud cover coverage, and temporal land cover dynamics.

AROSICS estimates misalignment by generating a grid of tie points (TPs) across the image overlap, spaced at a user-defined interval (e.g., 100 pixels). For each tie point (TP), X/Y offsets at a given geographical position (also known as X/Y displacement) between target and reference images are calculated within a local window of 256×256. The method uses phase correlation in the frequency domain to first determine integer pixel shifts. These integer shifts are then refined to achieve sub-pixel accuracy. Following displacement calculation, each shift must pass through sequential validation tests:

1. Reliability Check. This ensures the detected shift is statistically significant and not caused by noise or ambiguous patterns (Rogass et al., 2013). AROSICS measures the sharpness of the peak in the cross-power spectrum (Fourier domain) and calculates a reliability percentage. Tie points (TPs) with reliability values below 30% are rejected, as they likely represent false matches (e.g., in cloudy or homogeneous areas).
2. Structural Similarity Index (SSIM) Test. This verifies whether the estimated shift actually improves local image alignment. AROSICS compares the structural similarity between the reference and shifted target patch using the Mean SSIM (MSSIM) metric (Zhou Wang et al., 2004). TPs are discarded if the MSSIM decreases after applying the shift, indicating poor correction.
3. RANdom SAmple Consensus (RANSAC) detection (Fischler & Bolles, 1981). This removes globally inconsistent TPs that deviate from the expected geometric transformation. AROSICS fits an affine transformation model to all candidate shifts and iteratively flags outliers (e.g., shifts caused by clouds or terrain distortions). The algorithm automatically adjusts the outlier threshold to exclude approximately 10% of TPs, ensuring robust results.

Validated TPs and their final X/Y displacement vectors are then used to warp the target image via affine transformation. Approximately 1,000 TPs were utilized to calculate and correct the spatial misalignment between PS (reference) and SPOT 5 scenes (target images). Fig. 2.8 shows TPs distribution and their X/Y calculated shift displacement for the 2004 SPOT 5 image. Before co-registration, AROSICS detected

the maximum geolocation offset of 55 m relative to the PS image. The spatial misalignment was reduced to a maximum of 5 m after the co-registration was applied. Overall, the geolocation offsets in the multitemporal all input images ranged from a minimum of 10 m (observed in SPOT 5 2014) to a maximum of 55 m (observed in SPOT 5 2004) (Table 2.3). The co-registration process effectively reduced the geolocation misalignment by 85.6%–92.2%.

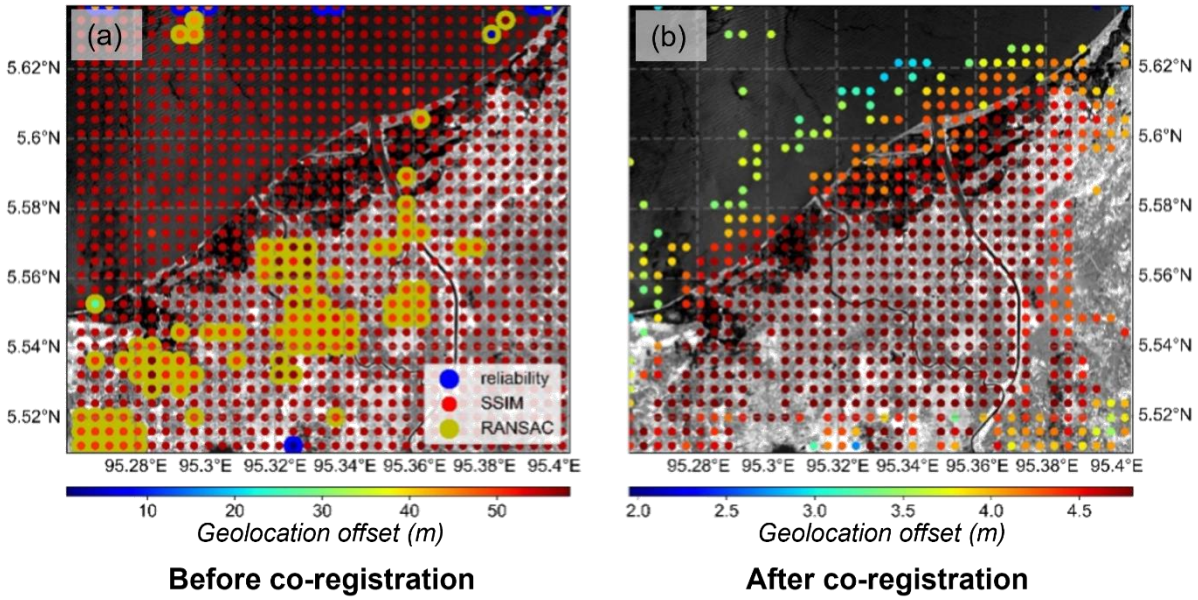


Fig. 2.8. TP distribution used to calculate spatial offset between the 2004 SPOT 5 and 2024 PS images. The colour bars indicate X/Y offsets in meters for each TP: (a) before and (b) after co-registration.

Table 2.3. Spatial misalignment of input images, relative to the 2024 PS image

Target image	Spatial geolocation offset (detected maximum X/Y displacement)	
	Before co-registration	After co-registration
SPOT 5 2004	55 m	5.0 m
SPOT 5 2005	25 m	2.7 m
SPOT 5 2014	10 m	2.5 m

Next, to address the radiometric inconsistency across input co-registered images, the radiometric normalization was applied using the Iterative Reweighted Multivariate Alteration Detection transformation algorithm (IR-MAD) (Leach et al., 2019). IR-MAD yields optimal results when the acquisition gap between the reference and target images is minimal, such as a few months. However, given that the SPOT 5 sensor ceased operations in 2015 while the PS constellation began producing imagery in 2016, it would be difficult to use either sensor as the radiometric reference for the other.

Therefore, as an alternative, the Landsat products were selected as the radiometric reference (Crawford et al., 2023). Landsat produces a long record and consistent time-series dataset which have been widely used in many studies for normalizing radiometric differences in multitemporal and inter-sensor datasets (Ju et al., 2025; Tu et al., 2022; Houborg & McCabe, 2018). The Landsat satellite has been operating since 1972 (Landsat 1) and most recently launched Landsat 9 in 2021. The U.S. Geological Survey (USGS) initiated collection-based processing of the entire image archive, which was processed as collection 1 in 2016. By 2020, the USGS reprocessed images in collection 1 archive and included additional new images to create the collection 2.

For this study, the Landsat Collection 2 Level 2 was employed as the radiometric reference, with selected scenes having narrow temporal gaps with the input datasets (SPOT 5 and PS images), as detailed in Table 2.4. It should be noted that neither SPOT 5 nor PS images were spatially co-registered to Landsat products. This is because the SPOT 5 images have already been co-registered to the PS images, while the spatial misalignment between PS and Landsat products was approximately 6 m, which was considered sufficient for change detection analysis (Wegmueller et al., 2021; Leach et al., 2019).

Table 2.4. Landsat scenes used as references for radiometric normalization

Target images		Reference images	
Sensor	Acquisition	Sensor	Acquisition
SPOT 5	5 August 2004	Landsat 5 TM Scene ID: LT51310562004356BKT00	21 December 2004
SPOT 5	29 January 2005	Landsat 5 TM Scene ID: LT51310562004356BKT00	21 December 2004
SPOT 5	18 May 2014	Landsat 8 OLI Scene ID: LC81310562014047LGN01	16 February 2014
PS	29 June 2024	Landsat 9 OLI-2 Scene ID: LC09_L2SP_131056_20240611_02	27 June 2024

The normalized different vegetation index (NDVI) values before and after radiometric normalization were used as indicator for identifying improvements in radiometric consistency. The NDVI is a widely used vegetation index that measures the density

and health of vegetation by calculating the ratio between spectral bands of NIR (ρ_{NIR}) and red (ρ_{RED}) (equation 10) (Houborg & McCabe, 2016). Its values range from -1 to +1, where values closer to +1 indicate dense, healthy vegetation, values near 0 represent bare soil or non-vegetated surfaces, and negative values typically indicate water bodies or clouds. NDVI is particularly useful for radiometric normalization assessment because vegetation areas should maintain consistent index values across different sensors and acquisition dates when radiometric differences are properly corrected (Leach et al., 2019).

The NDVI values were exclusively extracted at invariant pixels between the reference and target images. IR-MAD transformation identifies invariant pixels, applying a no-change probability threshold of 95% for invariant pixel selection. These invariant pixels are subsequently utilized in a regression model to adjust the radiometry of the target image to align with the reference image. As an example, Fig. 2.9 shows the detected invariant pixels between the SPOT 5 images before and after the 2004 IOT event and the 2004 Landsat 5 TM at 30 m resolution (radiometric reference). Number of invariant (unchanged) pixels detected (N) was 213 for the 2004 SPOT 5 (Fig. 2.9a). Due to land cover change caused by the tsunami occurrence, this number reduced to 133 for the 2005 SPOT 5 (Fig. 2.9b). This indicates that invariant pixel detection was precise since the majority of unchanged pixels (i.e., yellow coloured markers) were located in areas not affected by the tsunami event.

The inconsistency in NDVI values between the reference and target images was measured using root mean squared error (RMSE) (equation 11). Fig. 2.10 shows the NDVI value distribution before and after radiometric normalization was applied (derived from invariant pixels in Fig. 2.9). Before normalization, differences in NDVI values were observed with an RMSE of 0.26 for the 2004 SPOT 5 (Fig. 2.10a) and 0.33 for the 2005 SPOT 5 (Fig. 2.10c). After normalization, the discrepancy in NDVI values reduced, with an RMSE of 0.02 for both datasets (Fig. 2.10b and 2.10d). Overall, for all input datasets, the radiometric normalization improved radiometric consistency by 87.5%–92.3%, as detailed in Table 2.5.

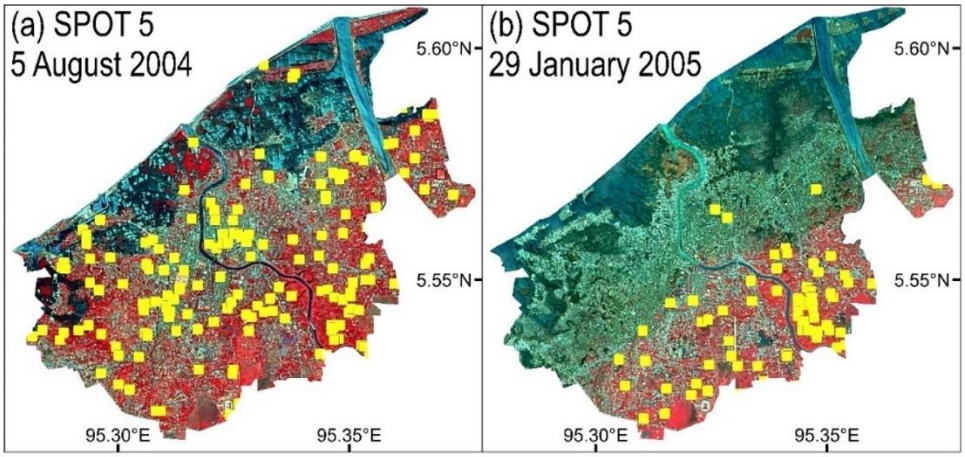


Fig. 2.9. Distribution of detected invariant pixels (yellow markers) between the reference image (2004 Landsat 5 TM) and SPOT 5 images at: (a) four months before and (b) one month after the 2004 IOT.

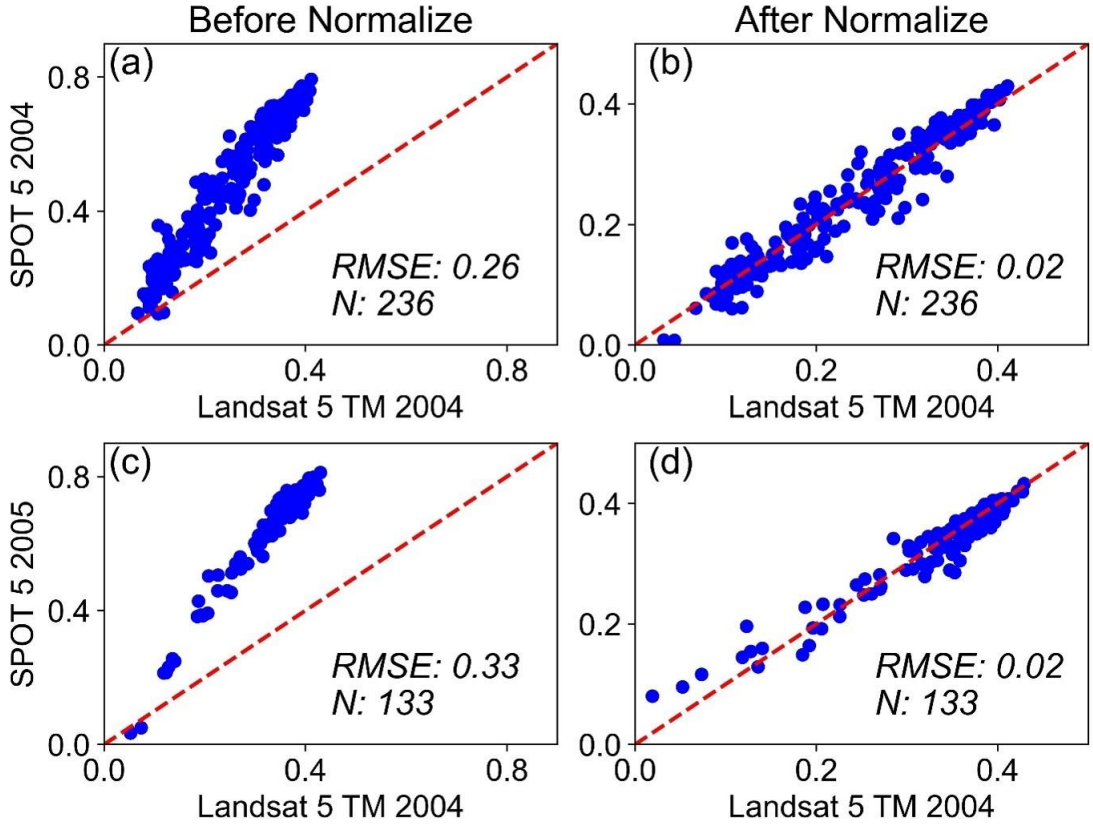


Fig. 2.10. Comparison of NDVI values before and after radiometric normalization, derived from invariant pixels in Fig. 2.9

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (10)$$

$$RMSE_{NDVI} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Reference\ image\ NDVI_i - Target\ image\ NDVI_i)^2} \quad (11)$$

Table 2.5. Comparison of NDVI values between input and reference images (in RMSE)

Input datasets	Total invariant pixels (N)	Before normalization	After normalization
SPOT 5 2004	236	0.26	0.02
SPOT 5 2005	133	0.33	0.02
SPOT 5 2014	200	0.32	0.04
PS 2024	220	0.28	0.01

2.5.1.c The 2004 IOT Inundation Limit

A combination of manual digitization and semi-automated technique were performed to identify the inundation extent of the 2004 IOT event. The semi-automated approach involved comparing the normalized difference water index (NDWI) layers from images taken before (the 2004 SPOT 4; Fig. 2.11a) and after the tsunami (the 2005 SPOT 4; see Fig. 2.11b). The NDWI is a water detection index that measures the presence and extent of surface water by calculating the ratio between bands of green (ρ_{GREEN}) and NIR (ρ_{NIR}) (equation 12) and is widely used to define the flood extent (Hese & Heyer, 2016).

$$NDWI = \frac{\rho_{GREEN} - \rho_{NIR}}{\rho_{GREEN} + \rho_{NIR}} \quad (12)$$

Theoretically, NDWI values ranges from −1 to +1 based on the spectral characteristics of land surface features. Since water has strong absorption properties in the NIR spectrum, it produces high NDWI values (Koshimura et al., 2020). The tsunami inundation extraction analysis is based on the principle that NDWI values will be higher in flooded areas due to water presence and will show a sharp decline in dry areas at the inundation boundary. This fundamental concept forms the basis for establishing NDWI threshold values to delineate tsunami-affected zone (Samela et al., 2022).

The trial-and-error thresholding method was applied and an NDWI threshold of +0.15 was obtained to distinguish wet areas. Subsequently, a post-classification step was

performed on the post-tsunami NDWI layer (Fig. 2.11b) through manual digitization of regions that were either missed or incorrectly classified by the semi-automated method. This integrated approach helps reduce bias in determining the tsunami extent, with the obtained boundary for the 2004 IOT event shown in Fig. 2.11c.

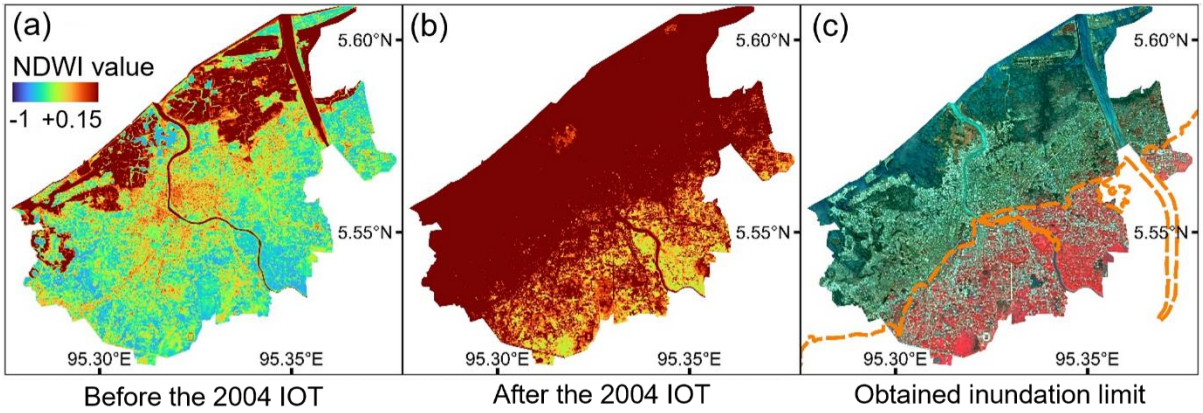


Fig. 2.11. NDWI layers before and after 2004 IOT event, with the red colour indicating water/wet areas. The orange line shows inundation limit, derived from manually digitization of post-tsunami NDWI layer.

2.5.1.d Land cover classification

The OBIA-RF for land cover classification process was executed using pyOTB, which involved image segmentation, zonal statistics, model training, and image classification. Considering variation in four-bands composition between SPOT 5 and PS sensors (see Table 2.2), only three of similar spectral bands were used (i.e., NIR, red, and green). The accuracy of land cover classification is determined by number of bands from input images (Zeferino et al., 2020; Gašparović et al., 2018). Given this, four additional bands from spectral indices were included as environmental features.

These spectral indices included vegetation (NDVI) and water (NDWI) indices which has been explained in sections 2.5.1.b and 2.5.1.c. The other two spectral indices were the second modified soil adjusted vegetation index (MSAVI2) and the visible-red near-infrared built-up index (VrNIRBI). The MSAVI2 was included to better distinguish vegetation from soil backgrounds, particularly in areas with sparse vegetation cover, as it reduces soil brightness influences that can affect NDVI measurements (equation 13) (Gholizadeh et al., 2018). Meanwhile, the VrNIRBI was incorporated to enhance the identification of built-up areas and impervious surfaces (equation 14) (Montero et al., 2023).

$$MSAVI2 = \frac{2 * \rho_{NIR} + 1 - \sqrt{(2 * \rho_{NIR} + 1)^2 - 8(\rho_{NIR} - \rho_{RED})}}{2} \quad (13)$$

$$VrNIRBI = \frac{\rho_{RED} - \rho_{NIR}}{\rho_{RED} + \rho_{NIR}} \quad (14)$$

Fig. 2.12 presents the pipeline process for the land cover classification. First, the three spectral and four environmental indices bands were stacked to form a seven-band input image. It should be noted that the scale of radiometric value between spectral and environmental indices bands were different. Given this, all radiometric values were rescaled to 0 and 1. Next, the image segmentation was performed, resulting in polygon features with unique identifiers. Each polygon was characterized by its radiometric properties through zonal statistics, which calculated the mean and standard deviation values for all seven bands within each segment.

The ground truth labels were then spatially joined with the segmented polygons based on their location, creating labelled polygon sample datasets with radiometric properties for each land cover class. To address the uneven sample distribution across land cover classes (Ma et al., 2017), a proportional class-weighted sampling approach was implemented by selecting 70% of labelled polygons as training datasets. The remaining 30% served as validation sets to assess classification accuracy.

Following Teluguntla et al. (2018), the ground truth labels were obtained by manually annotating very high-resolution historical basemap imagery from Google Earth Pro using photointerpretation techniques. The basemap imagery was selected to closely match the acquisition dates of the input satellite images. The ground truth labels consisted of approximately 3,000 randomly distributed points for each temporal period, divided into five land cover classes: water, bare land, low-vegetation, high-vegetation, and built-up areas (Fig. 2.13). To ensure accurate label assignment, the annotated points were validated using the classification scheme outlined in Table 2.6.

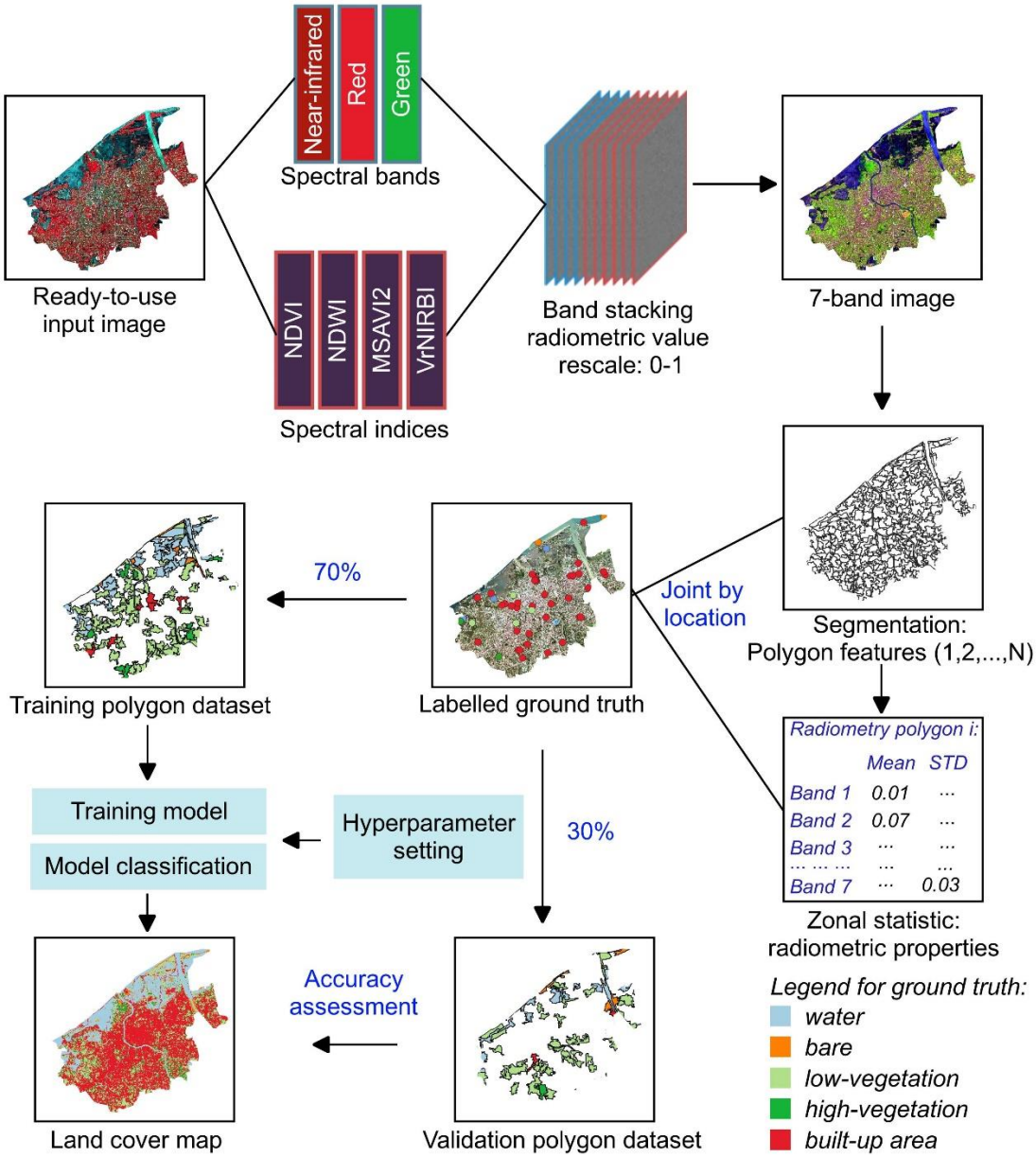


Fig. 2.12. The pipeline process for land cover classification using OBIA-RF method.

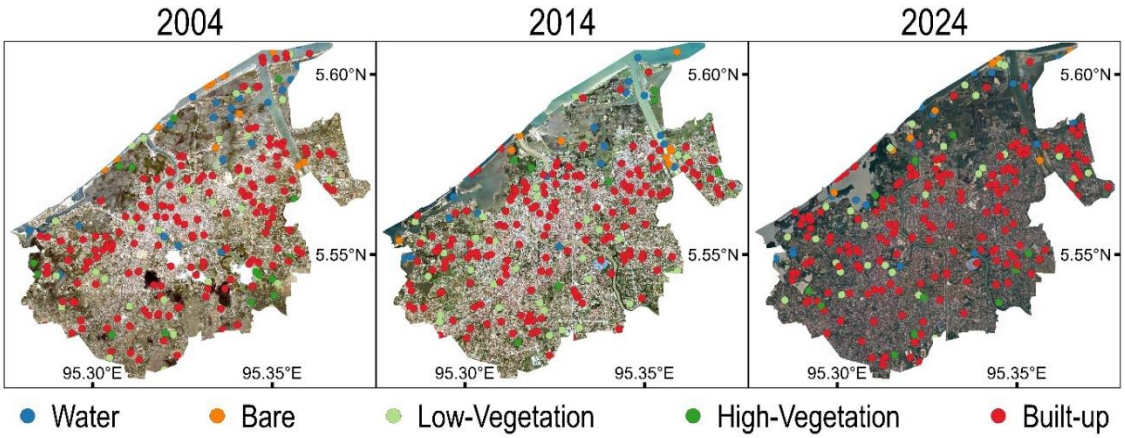


Fig. 2.13. Distribution of ground truth labelled points (10% subset) for three temporal periods. Background: Google Earth Pro very high resolution imagery.

Table 2.6. Classification scheme

Class	Description (Montero et al., 2023)	Examples
Water	• Water is present in image • Include intertidal zone • Contain no vegetation • NDWI value: more than 0	Shallow coastal areas, river, ponds
Bare	• Contain no vegetation and water • Does not include intertidal zone • Does not include unpaved roads • MSAVI2 value: -1 to 0.2	Beach, soil, sand, and reclaimed land
Low-vegetation	• Homogeneous green covered area • Human planted crops • Less dark than surrounding vegetation • less dense canopy vegetation • Small trees in a city park, pedestrian or highway median. • NDVI value: 0.2 to 0.7	Grass, paddy field, corn, and shrub
High-vegetation	• Dense canopy vegetation • Darker than surrounding vegetation • High trees in a city park • Included flooded vegetation • NDVI value: more than 0.7	Mangrove and forest
Built-Up (BU)	• All man-made structures • VrNIRBI value: more than 0	Road, building, impervious surface

The land cover model was then trained using the prepared training dataset based on RF classifier, where each polygon's radiometric features (mean and standard deviation values from seven bands) served as input variables and the corresponding land cover labels as target variables. The performance of land cover classification using RF classifier is highly dependent on its hyperparameters. It should be noted that RF is a nonparametric ensemble learning algorithm (Breiman, 2001).

The RF model's architecture is depicted in Fig. 2.14. This ensemble method employs bagging to produce numerous base predictors, with the ultimate prediction determined by majority voting across all individual models. According to bagging theory, model accuracy improves with an increasing number of predictors, though this relationship holds only up to a certain point (Chowdhury, 2024). In RF, the bagging process generates many decision trees (NT), each trained on a bootstrapped subset of the original training data (sampled with replacement). The trees are developed to their full depth (DT)—the maximum possible path length from the root to any leaf node—enabling them to model intricate data relationships. However, excessively deep trees may overfit the training data, necessitating careful optimization of the maximum depth to achieve a trade-off between model complexity and generalization.

During tree construction, each node split is determined by evaluating a randomly chosen subset of features, using metrics such as Gini impurity or entropy to identify the optimal partition. To enhance robustness and prevent overfitting, the minimum samples required to split a node is imposed (VPS). This parameter defines the smallest number of data points that a node must contain to be eligible for further splitting. Lower values permit more granular splits but increase susceptibility to noise, whereas higher values encourage broader, more generalizable decision boundaries. The final classification is derived by aggregating the predictions of all trees through majority voting. A grid search method from Python's Scikit-learn (Version 1.5) was employed to determine the optimal RF parameters (Liao et al., 2024). Table 2.7 lists the optimal parameters with respect to the training datasets.

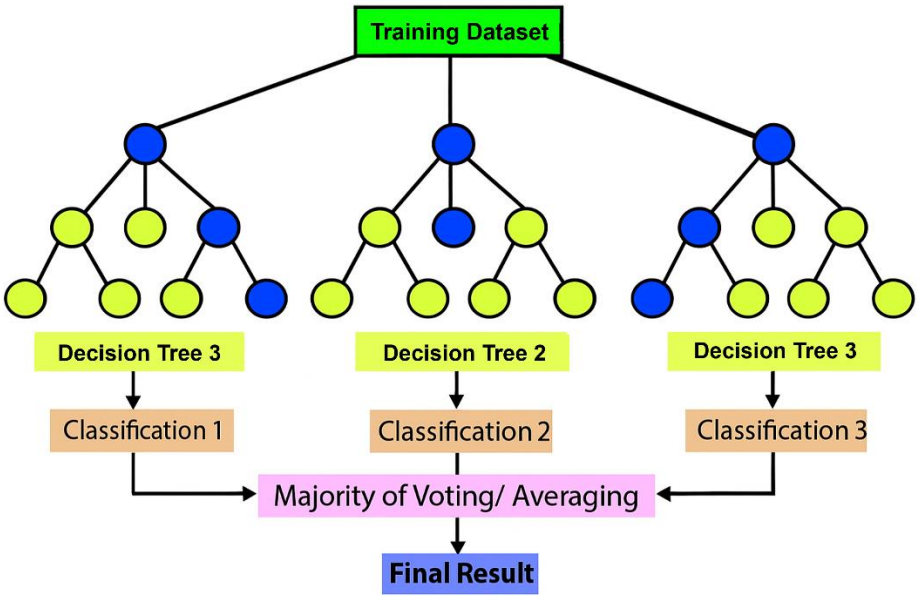


Fig. 2.14. The model's architecture of random forest classifier.

Table 2.7. Optimal parameter for random forest classifier

Land cover model	Sample sizes	**DT**	**VPS**	**NT**
2004	2219	16	5	200
2014	2229	16	2	300
2024	2240	16	2	100
Parameters	Random variable			
Maximum depth (**DT**)	2, 4, 8 ,16, 32			
Minimum sample to split (**VPS**)	2, 5, 10, 25, 50			
Number of Trees (**NT**)	1, 10, 50, 100, 200, 300, 500, 1000			

To minimize potential misclassification between built-up and bare classes owing to similarities in the spectral characteristics (Geiβ et al., 2017), a manual post-classification correction was performed. Additionally, the generated land cover maps were reviewed by two local experts to ensure their accuracy.

Fig. 2.15 presents the results of land cover classification. The accuracy of land cover maps was evaluated using a confusion matrix against validation points, represented by three accuracy indicators: F1 score, overall accuracy, and kappa index (Table 2.8). The accuracy assessment showed that the generated land cover maps were relatively accurate, with overall accuracies of 95.2% for the 2004 land cover (LC-04; Fig. 2.15a), 92.7% for the 2014 land cover (LC-14; Fig. 2.15b), and 94.5% for the 2024 land cover (LC-24; Fig. 2.15c).

Table 2.8. The confusion matrix and land cover classification accuracy

LC-04 (Land cover in 2004, before the 2004 IOT event)

		Truth Label					Accuracy Metrics (*P: Precision; R: Recall)*				
*Land cover**		*I*	*II*	*III*	*IV*	*V*	P	R	F1	Overall Accuracy	Kappa Index
Predicted Label	*I*	146	0	0	0	1	0.99	0.99	0.99	95.2%.	0.92
	II	2	38	0	0	16	0.88	0.68	0.77		
	III	0	0	142	2	8	0.92	0.93	0.93		
	IV	0	0	5	62	0	0.97	0.93	0.95		
	V	0	5	7	0	523	0.95	0.98	0.97		

LC-14 (Land cover in 2014, 10 years after the 2004 IOT event)

		Truth Label					Accuracy Metrics				
Land cover		*I*	*II*	*III*	*IV*	*V*	P	R	F1	Overall Accuracy	Kappa Index
Predicted Label	*I*	127	0	2	0	5	0.92	0.95	0.93	92.7%	0.88
	II	1	26	2	0	15	0.87	0.59	0.7		
	III	7	2	119	7	10	0.88	0.82	0.85		
	IV	0	0	8	70	1	0.91	0.89	0.9		
	V	3	2	4	0	537	0.95	0.98	0.96		

LC-24 (Land cover in 2024, 20 years after the 2004 IOT event)

		Truth Label					Accuracy Metrics				
Land cover		*I*	*II*	*III*	*IV*	*V*	P	R	F1	Overall Accuracy	Kappa Index
Predicted Label	*I*	*130*	*0*	*0*	*0*	*0*	0.98	1	0.99	94.5%	0.91
	II	*2*	*24*	*1*	*0*	*12*	0.75	0.62	0.68		
	III	*0*	*3*	*129*	*7*	*6*	0.9	0.89	0.89		
	IV	*0*	*0*	*3*	*69*	*0*	0.9	0.96	0.93		
	V	*0*	*5*	*11*	*1*	*524*	0.97	0.97	0.97		

** I: water; II: bare; III: low-vegetation; IV: high-vegetation, and V: built-up areas*

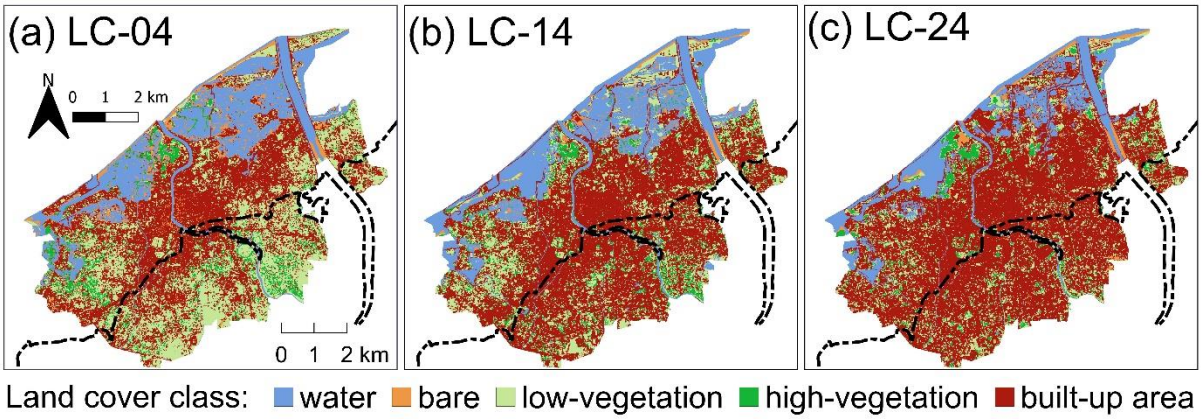


Fig. 2.15. Local land cover map for temporal period of: (a). 2004, (b) 2014, and (c) 2024. The black dashed line represents the 2004 IOT inundation limit.

Table 2.9 provides the percentage of each land cover type within the total AOI between 2004-2024. The land cover change analysis revealed that the built-up areas have expanded from 38.73% of the total BNA region in 2004 to 50.9% in 2014 and further rose to 62.81% in 2024—a total expansion of 1,442.35 ha. This growth corresponded with decreased coverage of other land cover classes, particularly low-vegetation regions which declined by nearly 50% from approximately 1,750 ha in 2004 to 910.74 ha by 2024. The high-vegetation and bare areas also showed decreasing trends, though less dramatically. Notably, bare areas slightly increased by 10 ha compared to 2014, likely due to ongoing reclamation processes, further evidenced by the concurrent reduction in inland water coverage.

Table 2.9. Land cover class area estimates

Year	Land cover	Land-class fraction (areas and percentage)				
		Water	Bare	Low-vegetation	High-vegetation	Built-up areas
2004	LC-04	1174.83 ha	320.54 ha	1750.04 ha	374.61 ha	2288.61 ha
		19.88%	5.42%	29.62%	6.34%	38.73%
2014	LC-14	1224.48 ha	137.38 ha	1168.85 ha	370.16 ha	3007.76 ha
		20.72%	2.32%	19.78%	6.26%	50.9%
2024	LC-24	866.67 ha	147.52 ha	910.74 ha	272.6 ha	3710.96 ha
		14.67%	2.5%	15.41%	4.61%	62.81%

2.5.2 Local Gridded Population Datasets

To construct the gridded population layers, the binary dasymetric mapping was adopted (Swanwick et al., 2022). First, the built-up class was extracted from each land cover map, resulting in built-up layers as shown in Fig. 2.16a-2.16c. These built-up layers were then intersected with subdistrict polygon boundaries to calculate the total built-up pixels within each polygon. Next, the population-to-built-up pixel ratio was calculated for each subdistrict. Finally, these ratio values were assigned to each built-up pixel, while non-built-up pixels were assigned a value of zero. Table 2.10 presents the population data for the AOI from 2004-2024, and the generated gridded population layers are shown in Fig. 2.16d-2.16f.

While some studies narrowly define built-up areas to include only buildings for more granular exposure assessment (Bonatz et al., 2024; Tiecke et al., 2017), this approach may oversimplify population distribution during tsunami events, as people are not always inside buildings. Therefore, this study employed a broader definition encompassing all man-made surfaces, including roads, buildings, and impervious areas, to better reflect the distribution of human activity.

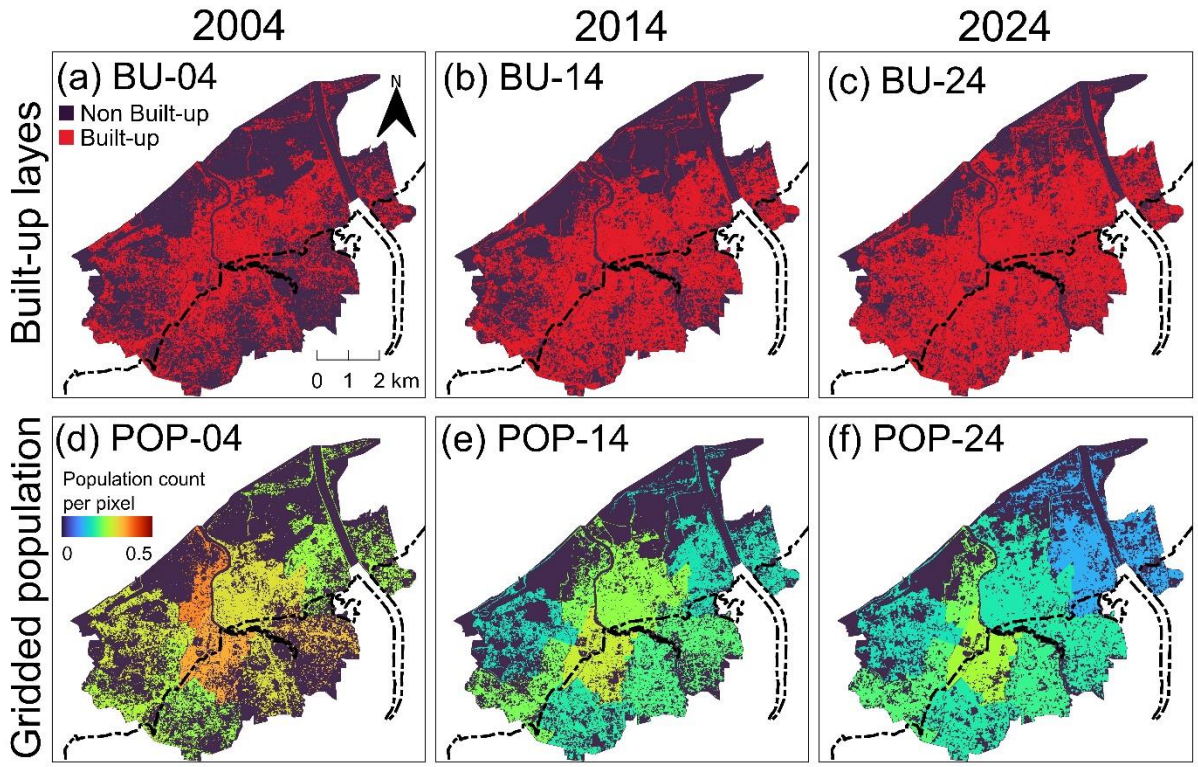


Fig. 2.16. Local exposure datasets for temporal period 2004-2024: (a-c) built-up area and (d-e) gridded population layers. The black dashed line represents the 2004 IOT inundation limit.

Table 2.10. Population census data by subdistrict within the AOI (2004-2024).

ID	Subdistrict	2004	2005	2014	2018	2024
1	Kuta Raja	20,217	5,122	12,831	13,632	14,943
2	Syiah Kuala	42,776	35,514	35,702	37,938	34,545
3	Meuraxa	31,218	5,657	18,979	20,561	25,916
4	Kuta Alam	55,062	43,113	49,545	52,645	44,836
5	Baiturrahman	37,449	36,783	35,249	37,445	34,111
6	Jaya Baru	22,005	11,348	24,481	26,525	27,239
7	Lueng Bata	18,360	18,254	24,581	26,119	25,802
8	Ulee Kareng	17,510	17,388	25,170	26,745	27,926
9	Banda Raya	19,071	19,015	22,961	24,878	26,651
	Total	263,668	192,194	249,499	265,111	261,969
	Source	JICA (2005)		BPS (2024)		

See Fig. 1.3 to refer to the subdistrict ID distribution

2.6 Bias Assessment of Global Built-up Area Datasets

Table 2.11 presents the estimation of built-up areas within the AOI and the 2004 IOT inundation limit. By leveraging the local built-up layers as a reference, the analysis revealed that global built-up area datasets exhibited both overestimation and underestimation of total built-up areas in 2004 and 2014 (Fig. 2.17a). For the 2004 temporal period, GHS-BUILT demonstrated the most severe overestimation at 2,620.09 ha, exceeding the actual built-up area estimates by twofold. While this overestimation persisted in 2014, GHS-BUILT's overestimation decreased to 69% (2,077.14 ha).

Conversely, GLD FCD30D and GAUD underestimated the estimation of built-up areas in 2004 by 55% and 29%, respectively. By 2014, both datasets exhibited overestimation patterns. This aligns with the behaviour of other datasets, which showed increases in total built-up areas between 2004-2014. However, GAIA was an exception, providing the same estimation of built-up areas in both 2004 and 2014. This resulted in GAIA underestimating built-up areas in 2014.

The observed biases in the built-up area estimations subsequently influenced the total built-up area exposed to the tsunami (Fig. 2.17b). The local dataset estimated that approximately 1,535.71 ha of built-up areas were affected by the 2004 IOT. Meanwhile, the global datasets overestimated the inundated areas by 6% to 98%. GHS-BUILT demonstrated the greatest overestimation by projecting about 3,045.3 ha of tsunami-exposed areas, nearly twice amount projected by the local dataset. Similar to the bias pattern in estimating total built-up areas, GLD FCD30D and GAUD underestimated

tsunami-exposed areas. The tsunami-affected areas were estimated at 976.1 ha by GLD FCD30D and 850.9 ha by GAUD, which are 37% and 45% lower, respectively, than the local dataset.

Table 2.11. Estimation of built-up areas derived from global built-up datasets

Dataset	Resolution (m)	Built-up areas within the AOI (ha)		Built-up areas within the 2004 IOT inundation limit (ha)	
		2004	2014	2004	2014
Local (reference)	5	2288.61	3007.76	752.9	1,260.8
GAIA	30	2,729.3	2,729.3	1,101.4	1,101.4
GISA	30	3,440.1	3,960.3	1,308.5	1,564.2
GISD30	30	3,381.5	3,931.3	1,370.2	1,647.2
Glance	30	3,768.8	3,920.6	1,498.7	1,550.8
GLC FCD30D	30	1,627.7	4,088.2	660.5	1,670.2
WSF	30	4,302.5	4,554.2	1,734.9	1,794.3
GHS-BUILT	100	4,908.7	5,084.9	1,863.4	1,869.4
GAUD	1,000	1,021.0	3,403.5	170.2	1,616.7

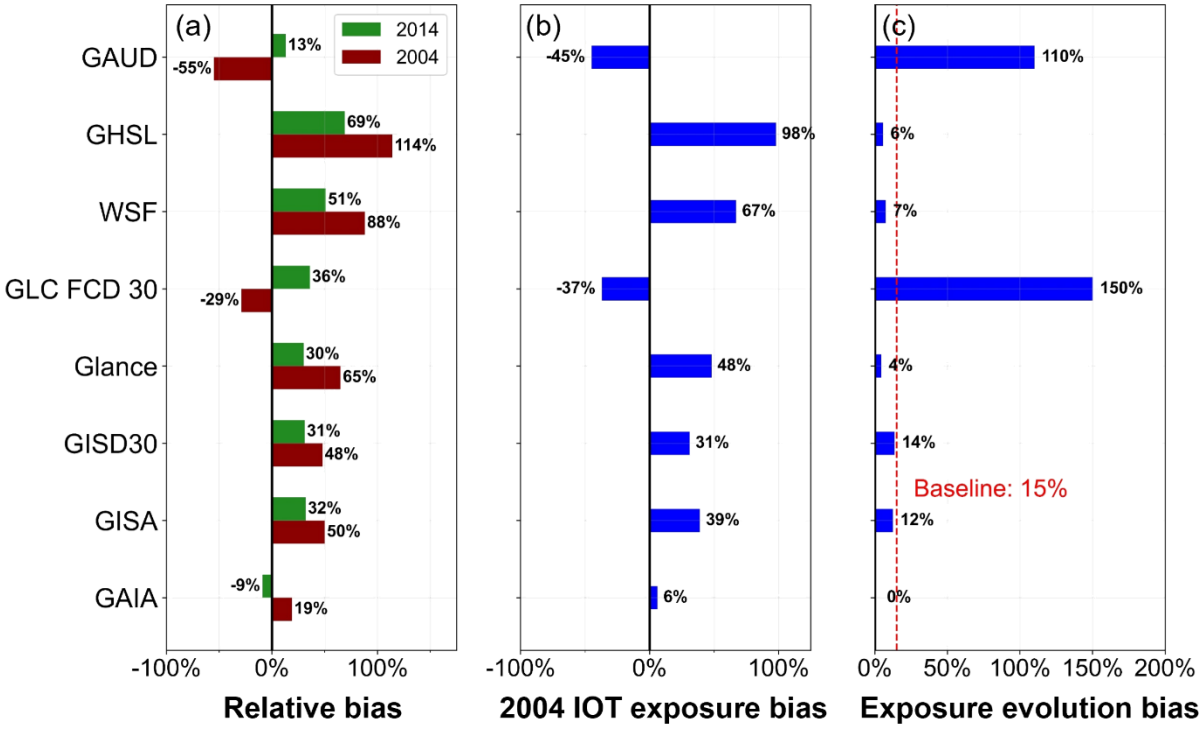


Fig. 2.17. Bias assessment for global built-up area datasets: (a) relative bias, (b) bias in the 2004 IOT exposure, and (c) bias in the exposure evolution between 2004-2014. The red line represents the reference dataset for built-up area expansion within the 2004 IOT limit.

These biases in built-up area estimation also influenced projections of built-up area evolution between 2004-2012 within the 2004 IOT inundated areas (Fig. 2.17c). The local dataset revealed that built-up areas had expanded within the inundated areas by

719.15 ha, representing approximately a 15% increase compared to before the 2004 IOT. Although the global built-up datasets captured this post-tsunami expansion trend (except for GAIA), they largely underestimated the actual expansion rate, particularly Glance with showing only a 4% expansion rate. In contrast, both GLC FCD30 and GAUD overestimated the post-tsunami expansion trend by 150% and 110%, respectively. This is likely related to their behaviour of underestimating total built-up areas in 2004, which ultimately led to a substantial increase by 2014.

2.7 Bias Assessment of Global Gridded Population Datasets

Table 2.12 presents the estimation of population within the AOI and the 2004 IOT inundation limit. The local datasets which were developed based on actual census data showed that the population in 2014 was 5% lower than in 2004. This indicates that the population did not fully recover to the pre-tsunami level even a decade after the 2004 IOT event. This decreasing pattern was also captured by WorldPop. In contrast, both GlobPop and LSG demonstrated an increase in population by 10% and 6%, respectively. However, GHS-POP was an exception, as it provided the same population estimates for both 2004 and 2014.

By using the local datasets as a reference, the analysis revealed that that global datasets underestimated the actual population, except for LSG (Fig. 2.18a). In 2004, the underestimation ranged between 29% and 47%, with WordPop exhibiting the greatest bias, underestimating 123,813 people. Although WorldPop showed a similar underestimation rate in 2014, its absolute value declined to 117,917. This is related to the fact that the total population projected by WorldPop also decreased in 2014. In contrast, LSG demonstrated an overestimation pattern by projecting 16,531 additional people in 2004, compared to the reference dataset. By 2014, the overestimation substantially increased to 48,650 people. This is because LSG showed an increase in the total population during the 2004-2014 period while the reference dataset projected the opposite.

Ultimately, biases in total population estimates influenced the projections of tsunami-exposed populations. Consistent behaviours were observed, with LSG exhibiting an overestimation of 2% and other datasets showing an underestimation ranging from 35% to 51% (Fig. 2.18b). Furthermore, WorldPop exhibited the greatest

underestimation by capturing only approximately 86,239 people exposed to the 2004 IOT. This estimation was nearly twice lower than the reference dataset.

Table 2.12. Estimation of population derived from global gridded population datasets

Dataset	Resolution (m)	Total population within the AOI		Total population within the 2004 IOT inundation limit	
		2004	2014	2004	2014
Local (reference)	5	263,668	249,499	176,003	144,525
WorldPop	100	139,855	131,582	86,239	75,308
GHS-POP	100	175,242	175,242	98,973	98,973
GlobPop	1,000	187,339	205,787	114,882	112,401
LSG	1,000	280,199	298,149	179,329	166,485

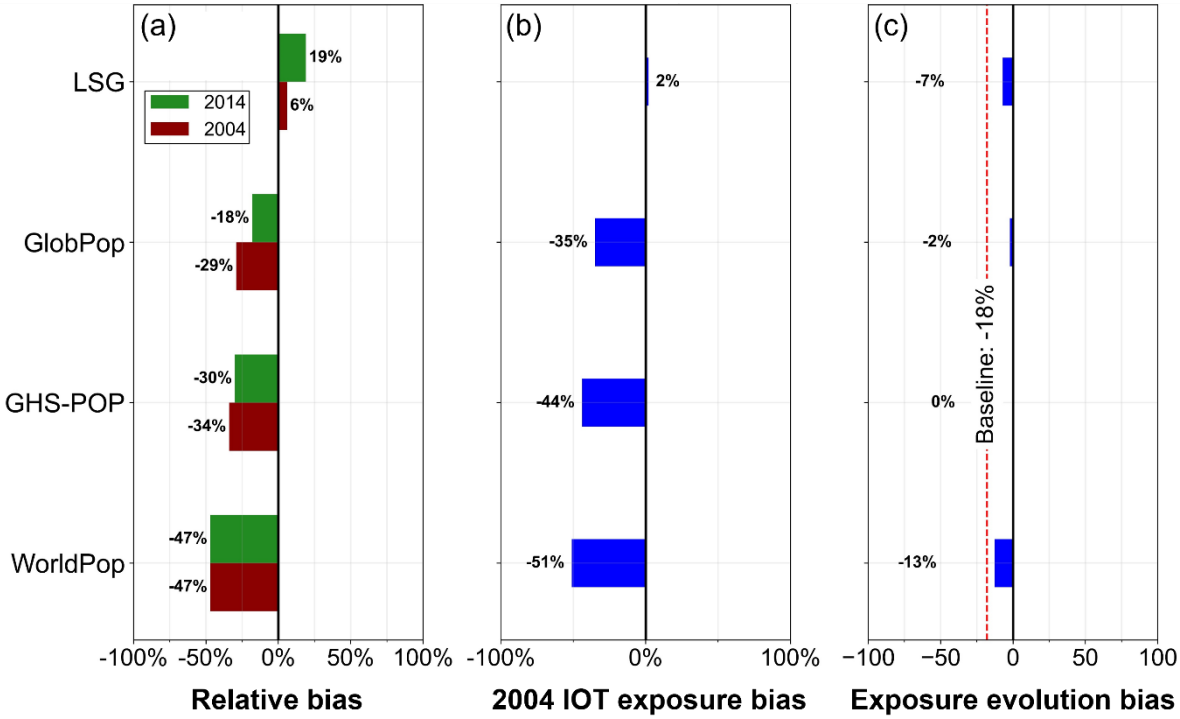


Fig. 2.18. Bias assessment for global gridded population datasets: (a) relative bias, (b) bias in the 2004 IOT exposure, and (c) bias in the exposure evolution between 2004-2014. The red line represents the reference dataset for population dynamics within the 2004 IOT limit.

By 2014, the reference dataset projected a decrease of 18% in the population living within the areas affected by the 2004 IOT, compared to 2004 (Fig. 2.18c). This aligns with the total population dynamics after the tsunami within the AOI. Although the global datasets also captured this decreasing trend (except for GHS-POP), they largely underestimated the decrease, with GlobPop projecting the lowest decrease of 2,481 people or 28,997 people less than the estimation provided by the reference dataset.

2.8 Summary of Chapter II

The reference dataset revealed that before the 2004 IOT, built-up areas were predominantly concentrated within zones that were later affected by the tsunami. A decade later, built-up areas expanded in both affected and non-affected areas. However, the distribution pattern in 2014 remained consistent with that in 2004, with affected areas still accounting for a larger share of built-up coverage. Notably, between 2004-2014, the greatest increase occurred in the non-affected areas.

The analysis showed that the affected areas increased by 211.25 ha (Fig. 2.19a), while non-affected areas expanded by 507.9 ha—more than twice as much (Fig. 2.19b). This finding suggested that greater increases in total built-up areas did not necessarily correspond to proportional increases in tsunami-exposed areas.

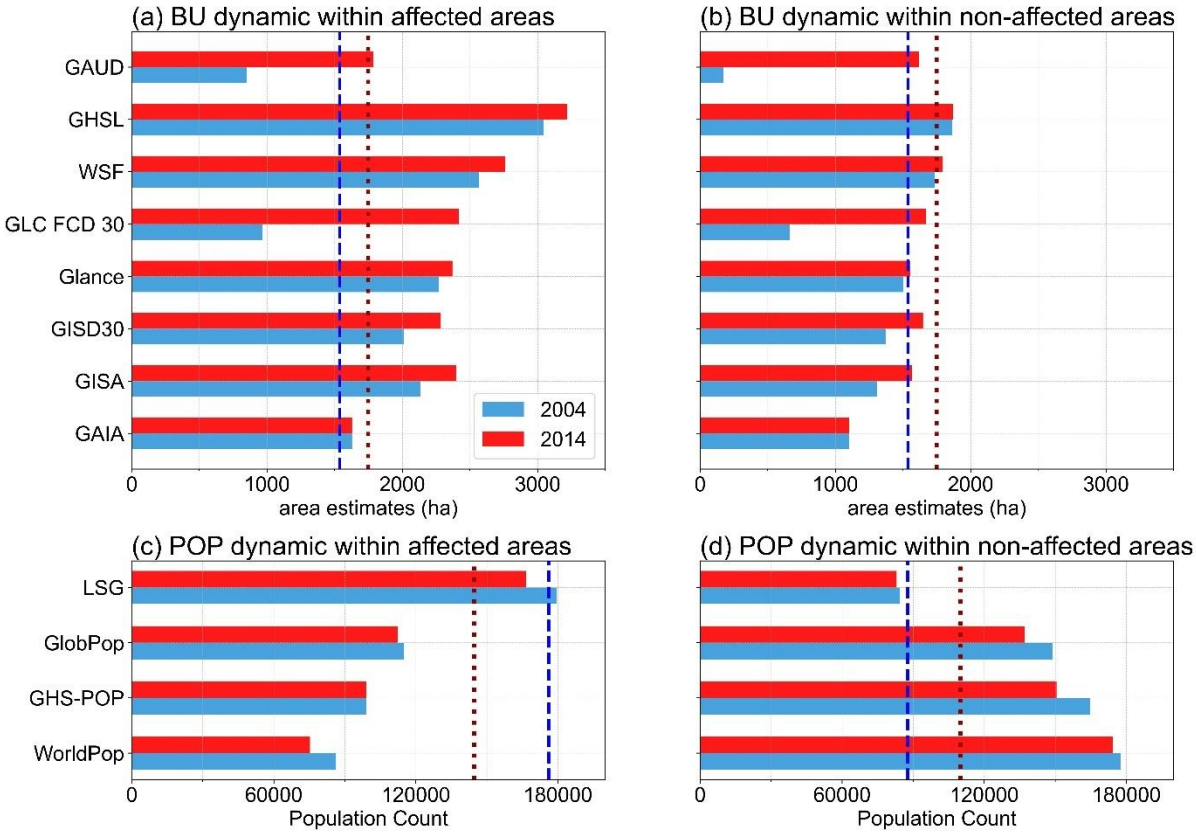


Fig. 2.19. The dynamics of built-up area and population estimates within: (a and c) affected areas of the 2004 IOT and (b and d) non-affected areas of the 2004 IOT. The blue and dark red dashed-lines represent the reference estimates for 2004 and 2014, respectively.

Dataset selection significantly influenced these observations. The magnitude of the built-up area increased considerably depending on the datasets used. Notably, except

for GISD30 and GAUD, global datasets indicated that the larger built-up increase between 2004-2014 occurred within affected areas. This pattern directly contradicts the evolution projected using the reference dataset.

The population distributions exhibited contrasting patterns. Both the local and global datasets demonstrated comparable trends within the affected areas, showing population decreases during 2004-2014 (Fig. 2.19c). However, the datasets diverged regarding non-affected areas: global datasets projected population decreases, while the local dataset indicated population increases (Fig. 2.19d).

These contrasting patterns have important implications for evaluating mitigation effectiveness. As mentioned in section 1.3, the AOI implemented spatial land use planning that designated non-affected areas as new development zones. Evaluation using global datasets like GHS-BUILT could have led to conclusions of ineffective planning, since built-up areas in new development zones appeared to grow by only 6 ha over the decade.

In contrast, population data from global datasets (WorldPop, GHS-POP, and GlobPop) could have suggested effective planning implementation, as these datasets indicated that most populations resided in new development zones rather than in tsunami-affected areas. In conclusion, these findings underscore the critical importance of understanding global dataset properties and selecting appropriate data for specific applications to minimize analytical errors that could lead to inaccurate tsunami risk assessments.

CHAPTER III:
SENSITIVITY OF DIGITAL ELEVATION AND LAND COVER ROUGHNESS MODELS

3.1. Digital Elevation Models (DEMs) Datasets

This study evaluated 11 digital elevation models (DEMs), consisting of original DEMs and their improved (error-reduced) versions. These DEMs varied in their creation methods, acquisition periods, and spatial resolutions, ranging from 30 to 90 m, as listed in Table 3.1. In addition to the improved versions from SRTM, COP30, and AW3D30, the TanDEM change elevation data (CRAWDEM) with a 30 m resolution was also incorporated into the analysis (Lachaise and Schweiβhelm, 2023). CRAWDEM was created by integrating new sensing data from 2017 to 2020 into the 30 m Edit TanDEM (EDEM) dataset (Gonzalez and Bueso-Bello, 2023). Additionally, the DEMNAS elevation model, a local DEM with an 8 m resolution introduced in 2018 (tanahair.indonesia.go.id), was also evaluated. DEMNAS employs multisource datasets, including IFSAR (5 m), TERRASAR-X (95 m), and ALOS PALSAR (11.25 m), along with additional mass point data derived from a stereo-plotting process. However, details about the acquisition period of DEMNAS's input datasets are limited.

Table 3.1. Characteristic of evaluated Digital Elevation Models (DEMs)

DEM (resolution)	Datum	Sensor	Acquisition and improvement
Original DEMs			
AW3D30 (30 m)	EGM96	Optical stereoscopic	2006 to 2011
COP30 (30 m)	EGM08	SAR interferometry	2010 to 2015
EDEM (30 m)	WGS84	SAR interferometry	2011 to 2015
SRTM (30 m)	EGM96	SAR interferometry	2000
Error-reduced DEMs			
CRAWDEM (30 m)	WGS84	Improved EDEM	Added new data from 2017-2021
DiluviumDEM (30 m)	EGM08	Improved COP30	Reduced vertical bias
FABDEM (30 m)	EGM08	Improved COP30	Removed building and vegetation biases
NASADEM (30 m)	EGM96	Improved SRTM	Void filling
CoastalDEM (90 m)	EGM96	Improved SRTM & NASADEM	Reduced vertical bias
MERIT (90 m)	EGM96	Improved SRTM & AW3D30	Reduced speckle, stripe, void, and vegetation bias
Local DEM			
DEMNAS (8 m)	EGM08	TERRASAR, IFSAR, and stereo-plotting mass points	

Some of these DEMs have been evaluated by numerous studies for various applications, including ecological terrain assessment (Bielski et al., 2023; Tran et al.,

2023; Li et al., 2022; Olajubu et al., 2021; Fassoni-Andrade et al., 2020), inundation modelling (Liu et al., 2021; McClean et al., 2020; Hawker et al., 2019; Hawker et al., 2018), sea level rise analysis (Seeger et al., 2023; Gesch, 2023; Kulp and Strauss, 2016), and terrain change detection (Chen et al., 2022; Brosens et al., 2022). Nevertheless, variations in terrain across different locations may lead to geographically specific error patterns (Hawker et al., 2018).

To address this, an elevation error analysis was performed within the research location before incorporating these data into the inundation models. Given the topographical setting of our area of interest (AOI), the error analysis focused on elevations lower than 10 m, a threshold to delimit the low-elevation coastal zone (Gesch, 2018). Prior to error analysis, all DEMs were harmonized to the same vertical datum of EGM2008, as the original datasets used two different datums (EGM1996 and EGM2008). Fig. 3.1 illustrates the spatial distribution of elevations across all evaluated DEMs within the low-elevation coastal zone, with blue areas representing water bodies (≤0 m) and the colour gradient from green to red indicating increasing elevation from 0 m to 10 m.

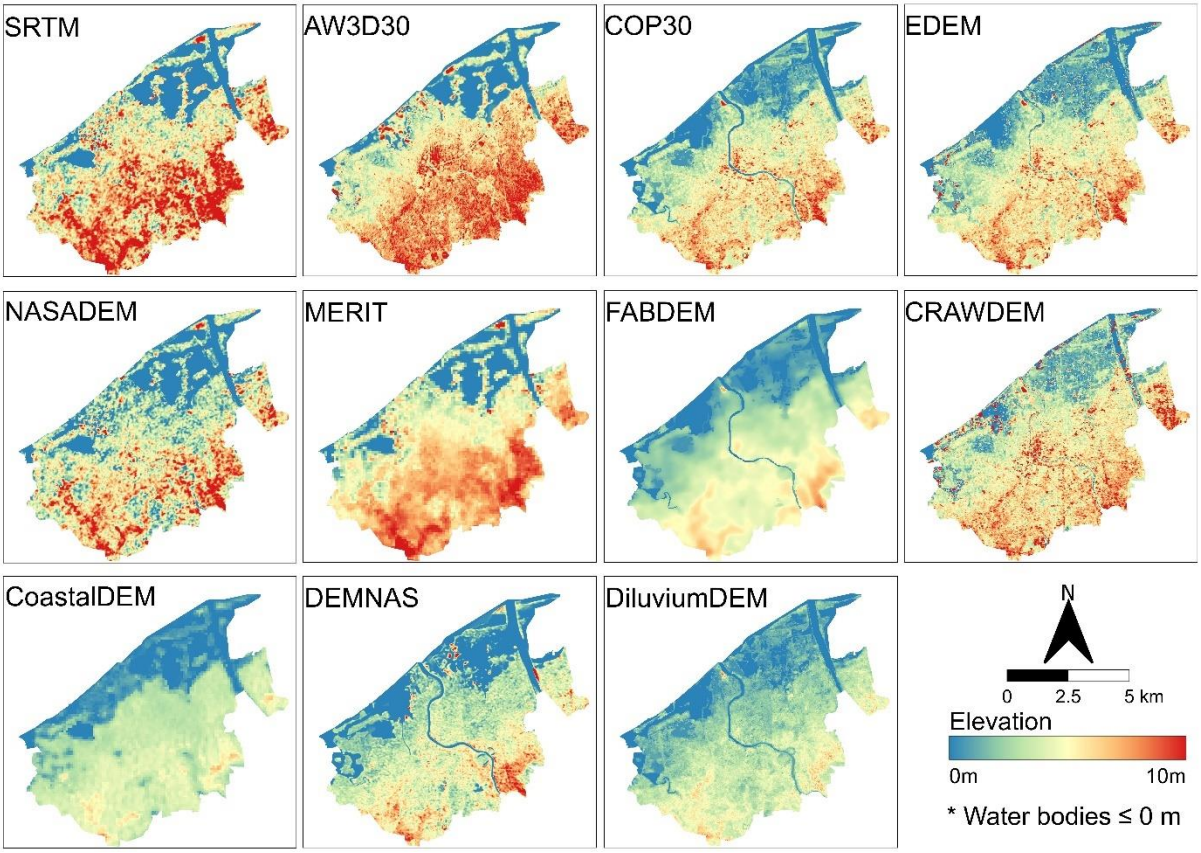


Fig. 3.1. Elevation distribution from evaluated DEMs within the AOI, vertical datum set to EGM2008

3.2. Elevation Error Analysis

Errors were defined as the discrepancies in elevation between the DEMs and the ground truth data (Dusseau et al., 2023). To mitigate bias in error interpretation, three statistical measures were employed: mean error (ME), mean absolute error (MAE), and root mean square error (RMSE) (equations 15-18). Due to the unavailability of local LiDAR data, elevation data from the global altimetry dataset ICESat-2 (version 006) were utilized as ground truth reference, sourced from the National Snow and Ice Data Center (https://nsidc.org).

The ICESat-2 satellite employs a photon-counting lidar system from Advanced Topographic Laser Altimeter System (ATLAS) instrument along with supporting technologies (GPS, star cameras, and ground processing systems) to measure photon travel time between the instrument and Earth's surface, enabling precise determination of reflected photon coordinates in terms of geodetic latitude and longitude (Neumann et al., 2023; Neuenschwander et al., 2023). Fig. 3.2 illustrates the ATLAS beam configuration, showing how laser pulses create three paired illumination spots on the surface, resulting in six ground tracks approximately 14 m in width as the satellite orbits Earth. These ground tracks are designated by their corresponding laser spot numbers, ranging from the leftmost track (GT1L) to the rightmost track (GT3R), with left and right tracks within each pair separated by approximately 90 m across-track and 2.5 km along-track. Additional details of ICESat-2 dataset are provided in Table 3.2.

The ICESat-2 data products are structured according to ground track organization, where tracks 1L and 1R constitute the first pair, tracks 2L and 2R form the second pair, and tracks 3L and 3R comprise the third pair. Each pair is associated with a theoretical Pair Track, representing the midpoint between the actual left and right beam positions, with these pair tracks spaced approximately 3 km apart in the across-track direction. ICESat-2 provides several products, including the geolocated photon data (ATL03) (Neumann et al., 2023) and a land and vegetation height product (ATL08) (Neuenschwander et al., 2023), with along-track resolutions of 20 m and 100 m, respectively.

$$\Delta h_i = h.DEM_i - \ h.ground\ reference_i \quad (15)$$

$$ME = \frac{\sum_i^N \Delta h_i}{N} \quad (16)$$

$$MAE = \frac{\sum_i^N |\Delta h_i|}{N} \quad (17)$$

$$RMSE = \sqrt{\frac{\sum_i^N (\Delta h_i)^2}{N}} \quad (18)$$

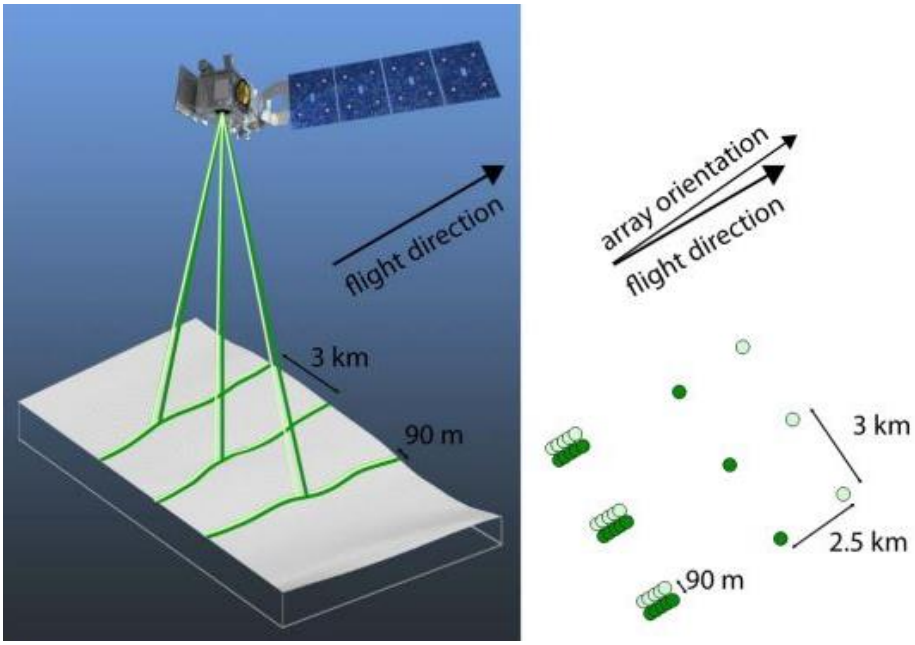


Fig. 3.2. ATLAS/ICESat-2 beam configuration. Adopted from Neumann et al. (2023)

Table 3.2. Characteristic of ICESat-2 dataset

Parameter	Characteristic
Mission	ICESat-2
Type	Discrete photon
Main objective	Cryosphere monitoring
Duration	2018-2024 (ongoing)
Orbit Inclination	92°
Beam footprint	11 m
# tracks	6 (in 3 strong/weak pairs)
Along track spacing	0.7 m (20 m for ATL08)
Across track spacing	3 km/90 m between pair
Swath width	6.6 km
Beam frequency	532 nm (green)
Vertical accuracy	0.91 cm MAE

The ICESat-2 provided 40 photon laser tracks encompassing the AOI from December 2018 to June 2023 (Fig. 3.3). Using the Python implementations of the Photon Research and Engineering Analysis Library (PhoReal) (Neuenschwander and Magruder, 2023), the photon signals were converted into elevation values. PhoReal can classify ATL03 photon signals into noise, canopy, and ground elevation data by utilising labelled information from ATL08 products.

The conversion of 40 photon laser tracks resulted in 460,000 ICESat-2 ground elevation points, which were used to analyse the elevation errors. Fig. 3.4 illustrates the detailed workflow of the elevation error analysis. Using information from the local land cover map in 2014 (LC-14; see Fig. 2.15), the analysis was conducted across various land cover types: all classes (excluding water), built-up areas, and highly vegetated regions (forest).

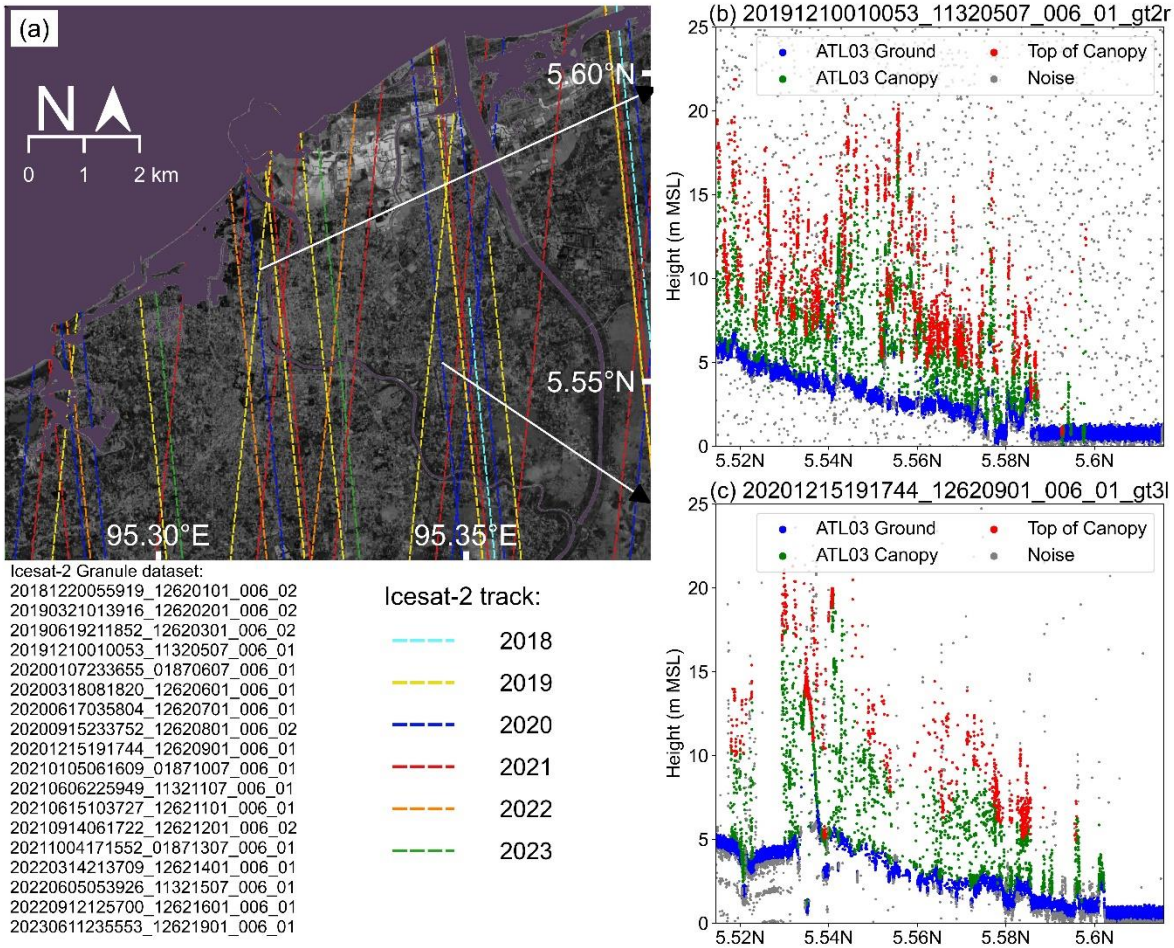


Fig. 3.3. (a) ICESat-2 track covered the AOI from 2018 to 2023. Sample of ATL03 photon classification for (b) Right ground track on 10 December 2019, and (c) Left ground track on 15 December 2020.

Table 3.3 summarises the error metrics for each DEMs at elevations below 10 m across various land cover types. The enhanced versions of COP30, namely FABDEM and DiluviumDEM, consistently demonstrated superior RMSE and MAE values compared to other DEM datasets. On average, FABDEM exhibited a slightly better performance than DiluviumDEM, with elevation errors being nearly normally distributed (ME = 2 cm) (Fig. 3.5). The error distributions further indicated that DiluviumDEM had a higher incidence of negative errors than FABDEM in built-up areas. However, when evaluated using an alternative error metric, DiluviumDEM surpassed FABDEM, achieving an RMSE of 1.02 m in built-up areas and 1.25 m in highly vegetated regions.

Furthermore, among the improved SRTM variants, CoastalDEM outperformed MERIT and NASADEM, particularly in built-up areas (RMSE: 1.27 m and MAE: 1 m) and highly vegetated regions (RMSE: 1.4 m and MAE: 1.07 m). However, CoastalDEM still exhibited significant negative biases, with an average of 1 m, resulting in an underestimation of the elevation of coastal low-lying areas when compared to ground truth data.

Table 3.3. Calculated statistical elevation error metrics

Metric (m):	RMSE			ME			MAE		
DEM [Resolution]	***AVG***	***BU***	***HV***	***AVG***	***BU***	***HV***	***AVG***	***BU***	***HV***
AW3D30 [30 m]	3.13	4.12	4.27	1.66	3.43	3.54	2.54	3.46	3.63
COP30 [30 m]	1.64	2.11	2.85	0.37	1.41	1.83	0.94	1.63	1.97
CRAWDEM [30 m]	2.06	2.39	3.14	0.83	1.73	2.16	1.26	1.90	2.29
CoastalDEM [90 m]	1.53	1.27	1.40	-0.98	-0.34	-0.73	1.30	1.00	1.07
DEMNAS [8 m]	2.35	1.94	3.16	-0.32	0.72	2.12	1.78	1.40	2.42
DiluviumDEM [30 m]	1.18	1.02	1.25	-0.36	-0.21	0.10	0.70	0.71	0.81
EDEM [30 m]	1.78	1.97	3.00	0.48	1.30	2.02	1.01	1.53	2.19
FABDEM [30 m]	1.09	1.11	1.38	0.02	0.43	0.65	0.59	0.81	0.97
MERIT [90 m]	2.74	3.30	3.57	1.58	2.80	3.00	2.29	2.86	3.13
NASADEM [30 m]	2.14	2.49	3.70	0.07	1.20	2.31	1.60	1.91	2.73
SRTM [30 m]	3.13	3.87	5.21	1.56	3.06	4.25	2.45	3.17	4.32

***AVG**: Average errors. All land cover classes, except water; **BU**: built-up area; **HV**: High-vegetation*

Notably, the average accuracy of the 8 m resolution local DEMNAS was lower than that of certain global DEMs with coarser resolutions, such as FABDEM, DiluviumDEM, CoastalDEM, COP30, and NASADEM. Nevertheless, DEMNAS demonstrated

superior performance compared to MERIT across various land cover types. It also exhibited greater accuracy than COP30 and NASADEM in built-up areas.

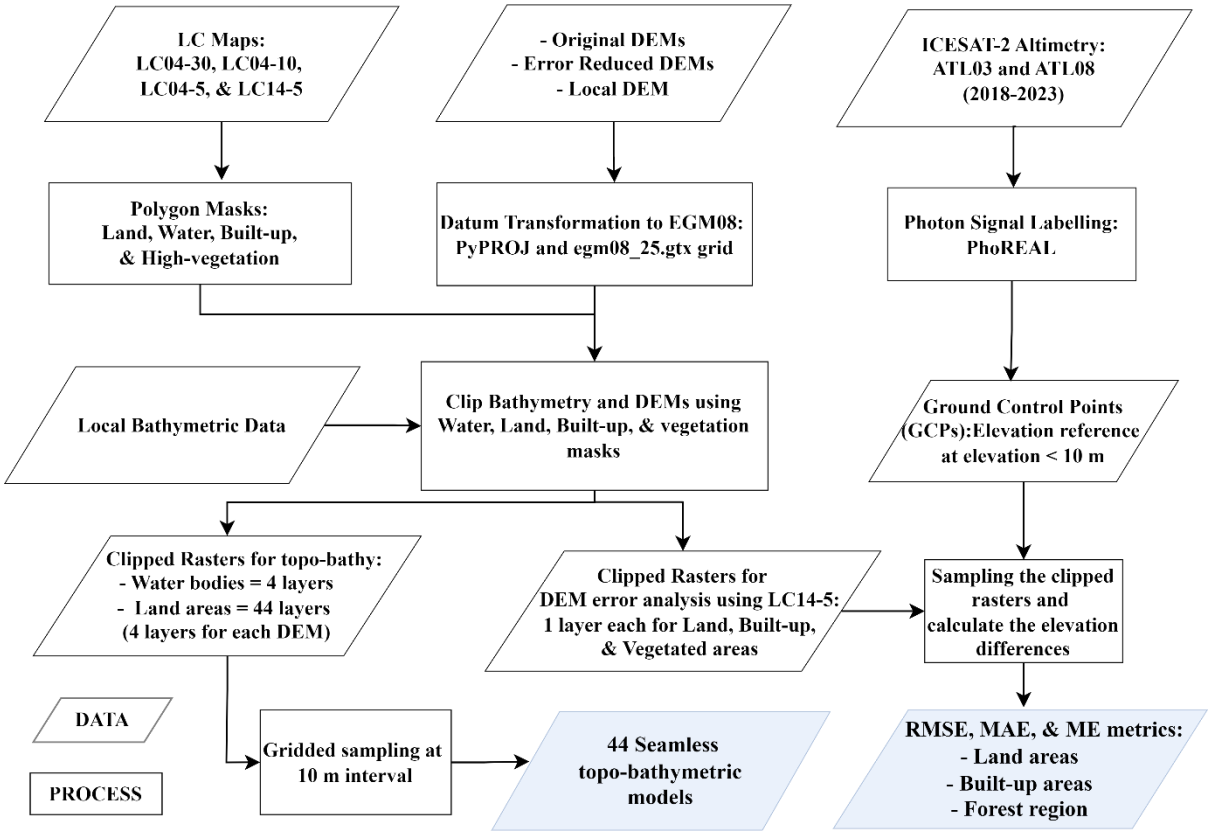


Fig. 3.4. Workflow for the DEMs error analysis and seamless topo-bathymetric model generation

In general, the accuracy of the error-reduced DEMs consistently surpassed the original DEMs, with the exception of CRAWDEM. This exception likely stems from insufficient calibration of newly added data (Lachaise and Schweiβhelm, 2023). Consequently, EDEM performed better than CRAWDEM in nine separate comparisons, despite CRAWDEM being designed as an improved version of EDEM.

Among the error-reduced DEMs, MERIT was observed to be the least accurate elevation model, with ME reaching 1.58 m. Similarly, its original datasets, AW3D30 and SRTM, showed the highest overestimation compared to other original datasets of EDEM and COP30. Specifically, positive errors within built-up areas reached 4.12 m for AW3D30 and 3.87 m for SRTM. For SRTM, these errors increased significantly up to 5 m in highly vegetated regions.

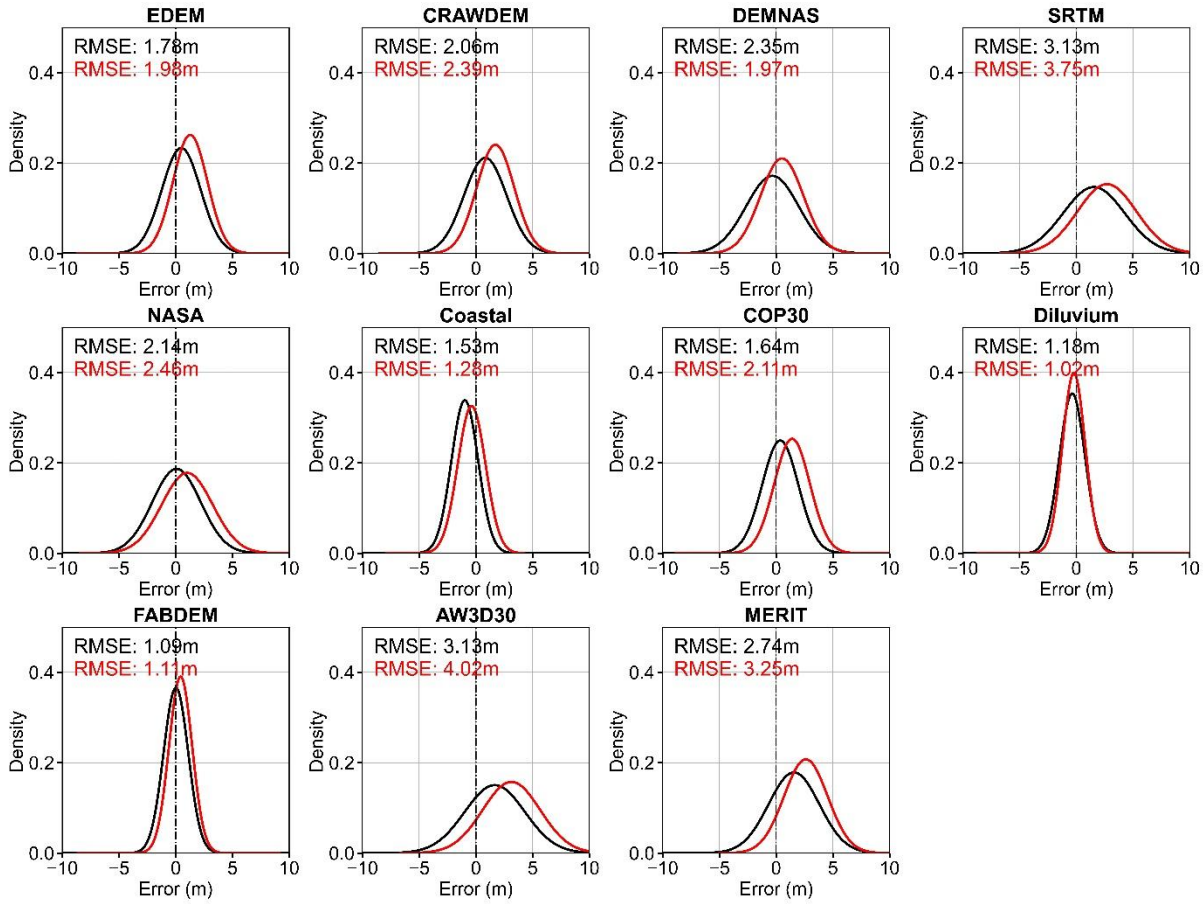


Fig. 3.5. The error distributions of evaluated DEMs, with red line indicates built-up areas and black line represents all land cover type (except water class).

3.3. Land Cover Roughness (LCR) Models

To develop the LCR model, a Manning's n value was assigned to each land-cover class, utilizing coefficients as listed in Table 3.4. Bricker et al. (2015) observed that commonly employed Manning coefficients tend to underestimate the dampening effect of dense vegetation and are inadequate for modelling tsunami flow in urban environments. Therefore, this study adopted Manning coefficients derived from extensive experiments and field measurements.

To address the impact of variations in roughness coefficients on tsunami estimates, two distinct sets of roughness coefficients were employed based on Bunya et al. (2010) and Koshimura et al. (2009). Both studies provide density-based Manning coefficients for built-up areas. The density of built-up areas was determined by calculating the ratio of built-up pixels to total pixels within a 1-hectare area. Subsequently, built-up area density was classified into three categories: low (ratio <30%), moderate (ratio 30-70%), and high (ratio >70%).

Table 3.4. Manning roughness coefficient values

Class	Manning's n value ($s/m^{1/3}$)	Manning's n value ($s/m^{1/3}$)
Water	0.013	0.013
Bare	0.025	0.020
Low vegetation	0.020	0.040
High vegetation	0.030	0.180
Low-density built-up (BU). Ratio: < 30%*	0.053	0.050
Moderate-density built-up (BU). Ratio: 30% -60%	0.094	0.120
High-density built-up (BU). Ratio: > 30%	0.172	0.150
Source	Koshimura et al. (2009)	Bunya et al. (2010)

***** Ratio of built-up area

To investigate how spatial resolution affects inundation modelling accuracy, multiresolution LCR models representing terrain features in the BNA region before the 2004 Indian Ocean Tsunami (IOT) were developed at three different resolutions: 30 m, 10 m, and 5 m. The development process involved:

1. Using the 2004 land cover map (LC-04) to generate a 5 m resolution LCR-04 model.
2. Resampling the LC-04 map to create a 10 m resolution LCR-04 model.
3. Developing a 30 m resolution LCR-04 model by following the land cover classification method described in section 2.5.4 (Chapter-II), using 30 m Landsat 5 TM imagery from 2004 as input (see Table 2.4)

To assess how land cover changes affect inundation properties, inundation models using the 2004 and 2014 LCR models were compared. In total, eight LCR models were developed, encompassing three spatial resolutions (30 m, 10 m, and 5 m), two temporal periods (2004 and 2014), and two different roughness coefficient sources, as shown in Fig. 3.6.

3.4. Relationship Between DEMs, LCR Models, and Inundation Accuracy

To evaluate how combinations of global DEMs and LCR models might influence inundation accuracy, the Mw 9.2 2004 Indian Ocean Tsunami (IOT) inundation was reproduced using various combinations of global DEMs and LCR models. The inversion fault model developed by Koshimura et al. (2009) was employed as the tsunami source model. This inversion model has been widely used to reproduce the 2004 IOT impact within BNA region (Jihad et al., 2023; Tursina et al., 2021; Syamsidik et al., 2019).

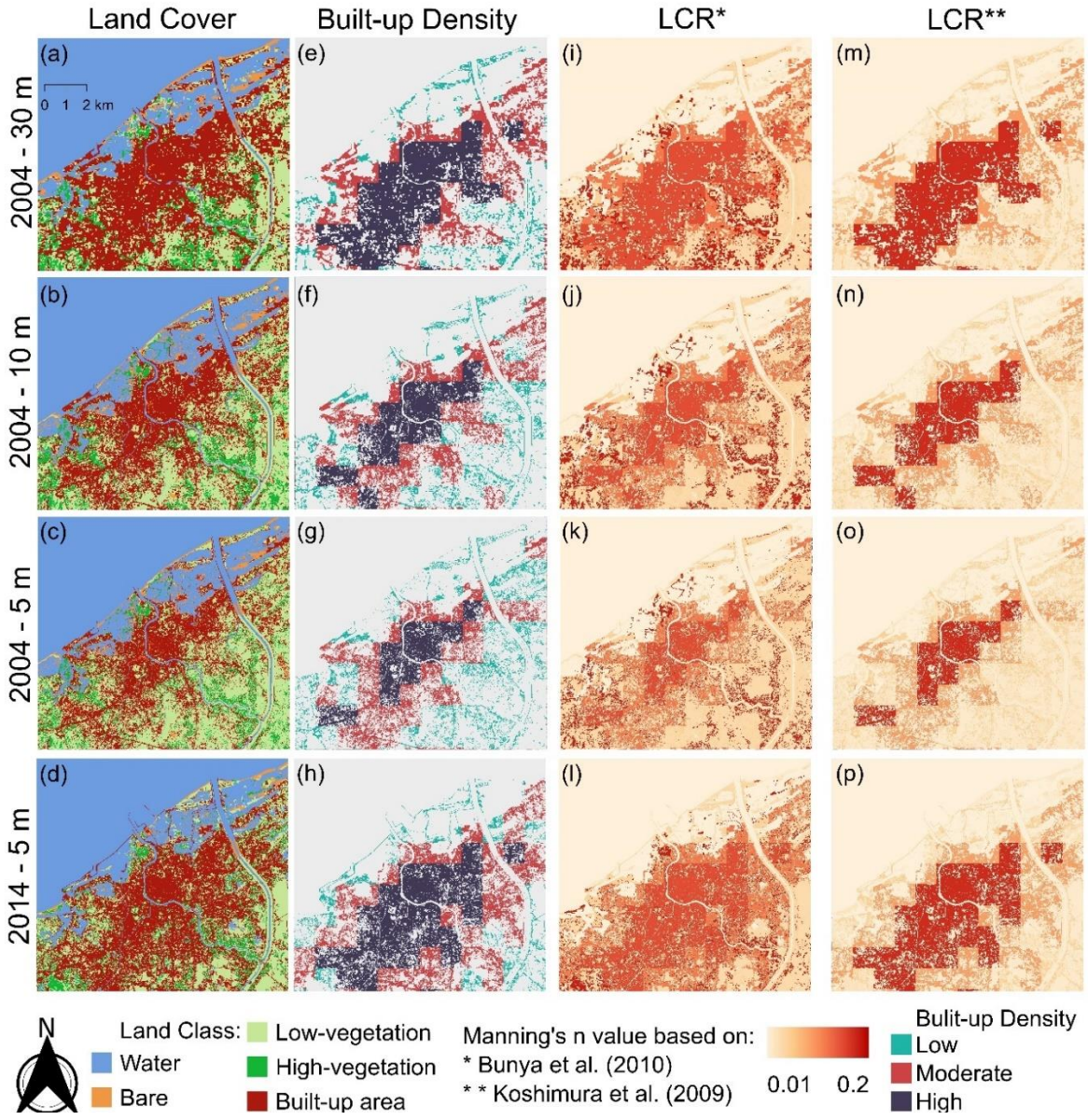


Fig. 3.6. Multiresolution LCR models, covering the computational region of Layer 4 (see Fig. 3.7).

To generate tsunami wave based on the inversion fault model input, the Cornell multigrid coupled tsunami model (COMCOT) was employed (Wang and Power, 2011). This model solves both the linear and nonlinear two-dimensional shallow water equations (2D-SWE). 2D-SWE consists of the continuity equation (equation 19) and momentum equation in both the x- (equation 20) and y-directions (equation 21). The governing equations are expressed as:

$$\frac{\partial \eta}{\partial t} + \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) = -\frac{\partial h}{\partial t}; H = \eta + h \qquad (19)$$

$$\frac{\partial P}{\partial t}+\frac{\partial}{\partial x}\left(\frac{P^2}{H}\right)+\frac{\partial}{\partial y}\left(\frac{PQ}{H}\right)+gH\frac{\partial \eta}{\partial x}+F_x=0 \quad (20)$$

$$\frac{\partial Q}{\partial t}+\frac{\partial}{\partial x}\left(\frac{PQ}{H}\right)+\frac{\partial}{\partial y}\left(\frac{Q^2}{H}\right)+gH\frac{\partial \eta}{\partial y}+F_y=0 \quad (21)$$

$$F_x=\frac{gn^2}{H^{\frac{7}{3}}}P\sqrt{P^2+Q^2} \quad (22)$$

$$F_y=\frac{gn^2}{H^{\frac{7}{3}}}Q\sqrt{P^2+Q^2} \quad (23)$$

$$P=\int_{-h}^{\eta}udz=uH \quad (24)$$

$$Q=\int_{-h}^{\eta}vdz=vH \quad (25)$$

where *h* is water depth or land elevation; *η* represents water surface fluctuation; *P* and *Q* represent flux in the x- and y- directions, respectively; *t* is time; *g* is gravity; F_x and F_y represent bottom friction in the x- and y- directions, respectively; and *n* denotes the Manning's coefficient values, which are represented by the LCR models.

COMCOT assumes that the seafloor displacement would instantaneously occur after an earthquake event (Wang and Power, 2011). Using the elastic finite fault theory proposed by Okada (1985), COMCOT computes the initial sea surface deformation. The elastic fault theory models earthquake-induced deformation by conceptualizing a rectangular fault surface embedded within an elastic half-space medium. The fault surface serves as an approximation of the fracture zone between converging tectonic plates, where sudden slip movement generates seismic events.

When displacement occurs along this fault surface, it generates stress distributions throughout the surrounding elastic medium, resulting in surface deformation that manifests as vertical and horizontal seafloor movements. As illustrated in Fig. 3.7, the COMCOT model requires both focal fault geometry parameters and focal mechanism to compute the initial sea surface deformation. The focal mechanisms include fault's azimuthal orientation (*strike angle, θ*), inclination from horizontal (*dip angle, δ*), and direction of slip movement (*rake angle, λ*). Meanwhile, the fault geometry parameters include depth to the fault's upper edge (*focal depth, h*), and the fault's physical dimensions (*length, L,* and *width, W*).

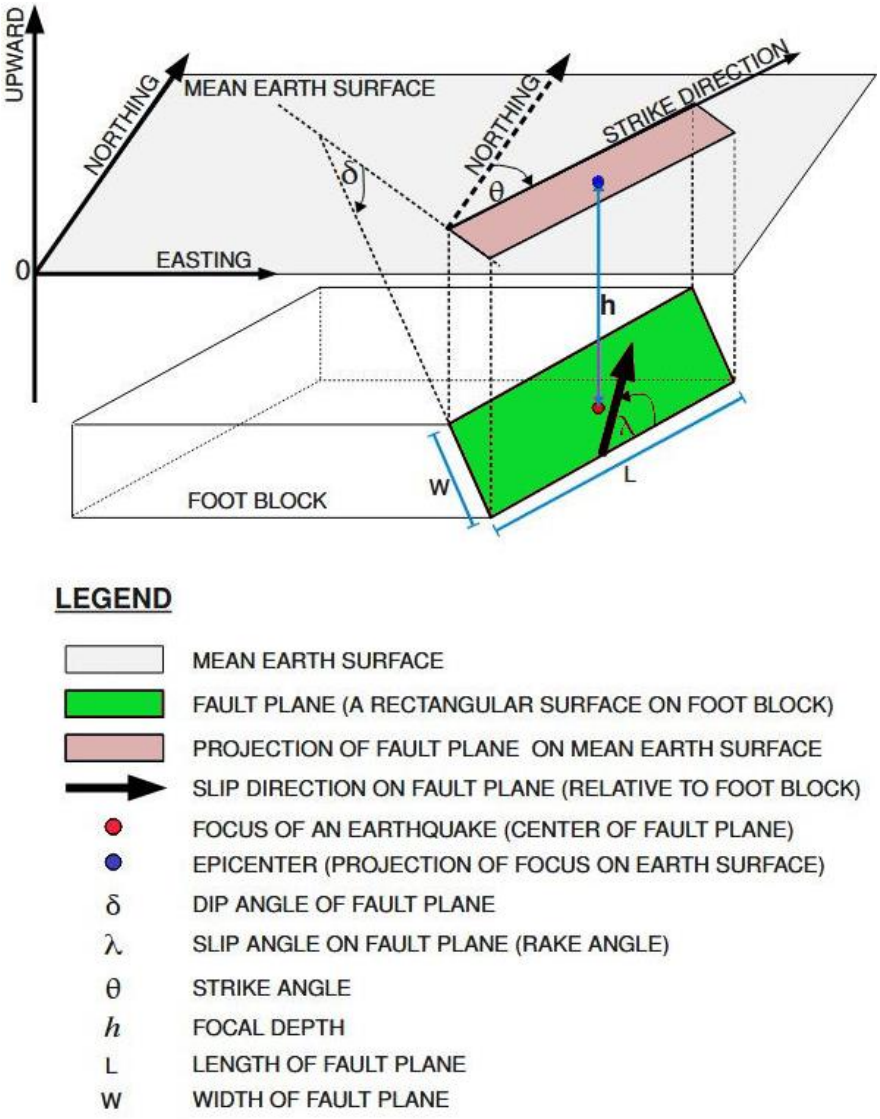


Fig. 3.7. Schematic illustration of fault plane geometry and associated parameters for initial sea surface deformation calculations. Colour-coded surfaces indicate the reference earth surface (light grey), the fault plane (green), and the projected fault area onto the earth's surface (light red). Modified from Wang and Power (2011).

For earthquakes involving complex rupture patterns, the source region can be discretized into multiple fault patches or segments. The cumulative deformation field is obtained through linear superposition of individual segment contributions, each computed using the elastic dislocation theory. This segmented approach allows for heterogeneous slip distribution across the rupture zone, producing more realistic spatial patterns of seafloor deformation.

The employed inversion fault model to generate the Mw 9.2 2004 IOT consisted of six segments, as illustrated in Fig. 3.8a. The detailed parameters of fault geometry and

focal mechanism are provided in Table 3.5. The initial sea surface deformation served as the initial condition for tsunami wave propagation in the numerical domains. To simulate the tsunami propagation from the Sunda-Andaman segment to the AOI, a four nested-grid configuration was employed as the numerical domain, as illustrated in Fig. 3.8b.

Table 3.5. Inversion fault model properties based on Koshimura et al. (2009)

Segment	Focal depth (km)	Length (km)	Width (km)	Strike (°)	Dip (°)	Rake (°)	Slip (m)
1	10	200	150	323	15	90	14.0
2	10	125	150	335	15	90	12.6
3	10	180	150	340	15	90	15.1
4	10	145	150	340	15	90	7.0
5	10	125	150	345	15	90	7.0
6	10	380	150	7	15	90	7.0

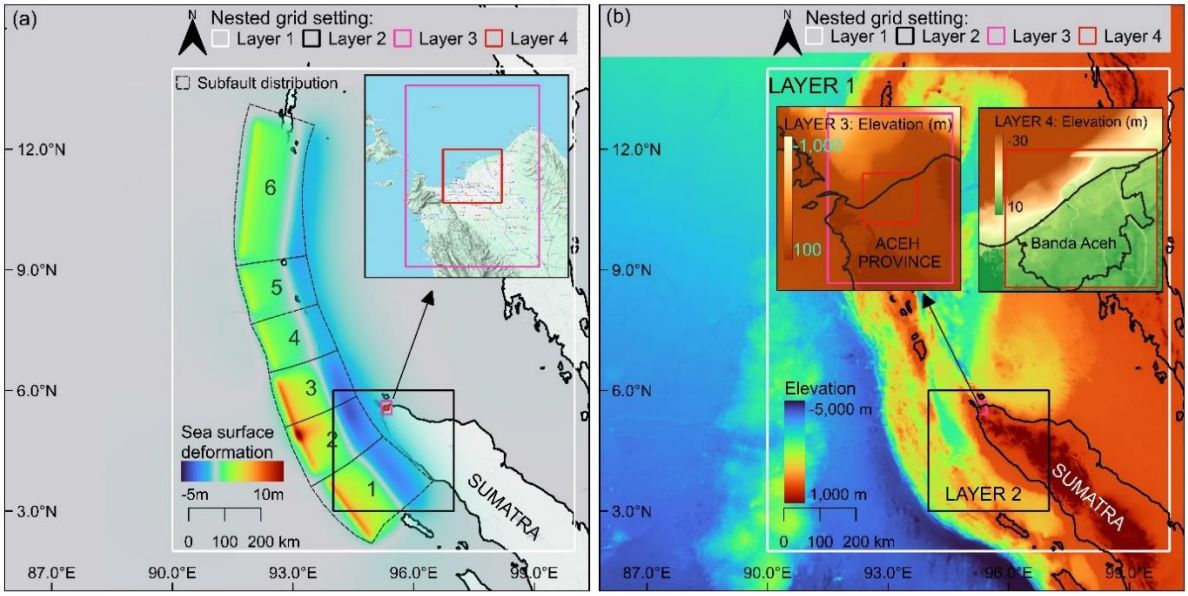


Fig. 3.8. (a) Subfault setting of inversion fault model and the initial sea surface deformation. Details of subfault properties are provided in Table 3.5. (b) Nested-grid layer setting and topo-bathymetric configuration for each layer.

Table 3.6 lists the details of the numerical domain configurations. The outermost domain (Layer 1), with a grid resolution of 1080 m, utilized the initial sea surface deformation as the initial condition, while the inner layers incorporated the wave heights and wave fluxes in the horizontal directions (x and y) along their nesting boundaries. The GEBCO 2023 Grid, with a resolution of 15 s (~450 m), served as the topo-bathymetric input for Layer 1 (GEBCO Bathymetric Compilation Group, 2023).

Layers 2 and 3 with grid resolutions of 216 m and 43.2 m, respectively, employed the Indonesian National Bathymetric Chart (BATNAS) with a resolution of 6 s (~180 m) (tanahair.indonesia.go.id). The topo-bathymetric input data for the innermost domain (Layer 4) integrated depth information from local bathymetric measurements and elevation data from the examined DEMs.

All topographic and bathymetric inputs were resampled to align them with the resolution of each numerical layer using a regular point-based sampling approach in GIS applications. To produce seamless topo-bathymetric input for Layer 4, multiresolution water masks derived from multiresolution land cover maps in Fig. 3.6 were used to identify land and water areas, including small water bodies such as ponds and rivers that may not be captured by local bathymetric data. As illustrated in Fig. 3.9, to adjust the water bodies detected by the masks but absent from the bathymetric data, a uniform depth of −0.5 m was assigned.

Table 3.6. Configuration of numerical domains

Domain	Layer 1	Layer 2	Layer 3	Layer 4
Ymin-Ymax (deg)	2 - 14	3 - 6	5.4 - 5.74	5.52 - 5.62
Xmin-Xmax (deg)	90 - 100	94 - 97	95.2 – 95.45	95.27 - 95.38
Grid size (m)	1080	216	43.2	10.8
Time step (s)	1.2	0.24	0.048	0.012
Layer ratio	1	5	5	4
Initial condition	Fault Model	Layer-1	Layer-2	Layer-3
Input depth	GEBCO_2023 Grid	BATNAS	BATNAS	Local bathymetry
Input elevation	GEBCO_2023 Grid	BATNAS	BATNAS	DEMs
Friction model	None	None	None	Uniform and 8 LCR
SWE type	Linear	Linear	Linear	Non-linear
Boundary condition	Radiation	Interpolated	Interpolated	Interpolated

The linear SWE scheme was implemented for Layers 1 to 3, while Layer 4 used the nonlinear SWE scheme. The outputs from Layer 4 were then used to develop inundation maps and evaluate inundation model performance. To evaluate the impact of DEMs on inundation model performance, DEMs were integrated with a uniform Manning's friction coefficient (Manning's n = 0.025). Concurrently, to assess the performance of DEMs coupled with variable Mannings, DEMs were incorporated with multiresolution LCR models as another scenario. In total, 99 inundation models were simulated, comprising 11 DEMs, 8 LCR models, and one uniform Manning model. All

computational topo-bathymetric grids were automatically updated at the start of each simulation to account for changes in seafloor and land deformation. Each scenario simulated tsunami propagation at a duration of 2 hours.

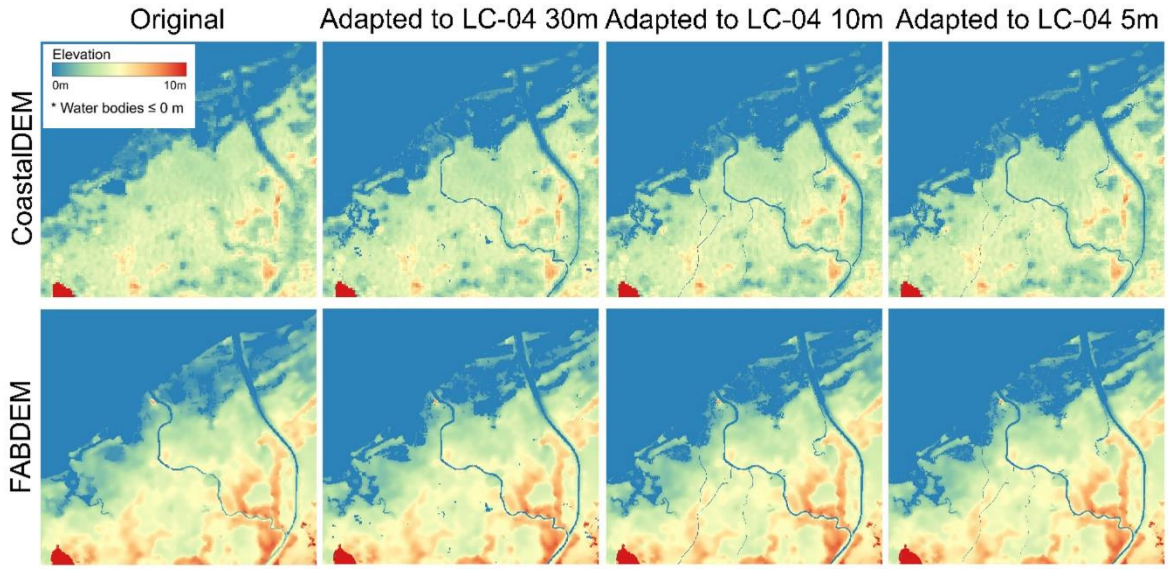


Fig. 3.9. Illustration of topo-bathymetric adjustment using water masks from land cover maps for CoastalDEM (top panel) and FABDEM (bottom panel).

3.5. Performance of Inundation Models

The performance of the inundation models was evaluated based on the simulated inundation maps. To evaluate the uncertainty of inundation performance, this study proposed three sequential assessments, incorporating historical data and the confidence level associated with inherent DEM error characteristics. First, the simulated inundation extents were compared against the 2004 IOT satellite-derived inundation limit (see section 2.5.3). A threshold of 25% difference was set as a benchmark for acceptable model performance. This accounts for the inherent challenges to accurately reproduce the 2004 IOT inundation due to limitations in pre-tsunami bathymetric dataset and constraints in the tsunami source model (Ocean Networks Canada et al., 2023; Sugawara, 2021).

Second, the simulated flow depth estimates were validated against field measurement data. The measured data were collected from various sources, including Tsuji et al. (2006), Borrero (2005), and Sugimoto et al., (2010) (Fig. 3.10a). The agreement between the simulated and observed flow depths was evaluated using Aida's functions, expressed as the geometric mean (K) and standard deviation (k) (Aida, 1978):

$$log\ K = \frac{1}{N}\sum_{i=1}^{N} log\ K_i \tag{26}$$

$$log\ k = \left[\frac{1}{N}\sum_{i=1}^{N}(log\ K_i)^2 - (log\ K)^2\right]^{1/2} \tag{27}$$

$$K_i = \frac{M_i}{S_i} \tag{28}$$

where M_i and S_i are the observed and simulated flow depths, respectively. The simulation results were considered to be in good agreement when Aida's parameters were: $0.8 \leq K \leq 1.2$ and $k \leq 1.60$ (Takeuchi et al., 2005).

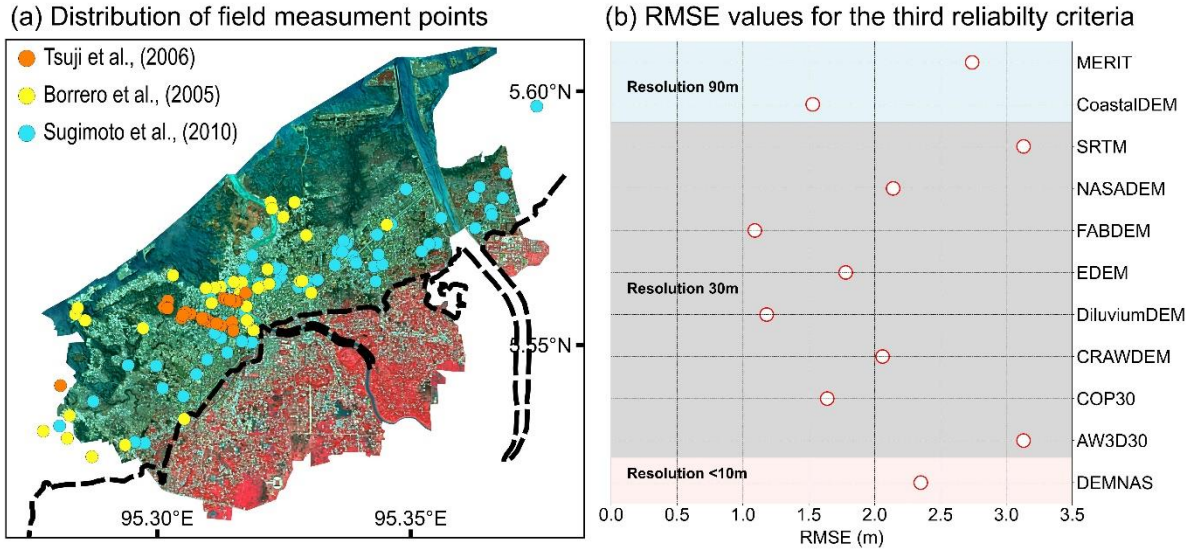


Fig. 3.10. (a) The 2004 IOT post-tsunami field measurement points. Details of observed points were provided in Appendix. Black dashed line indicated the inundation limit.(b). DEMs' RMSE values used for equations 29 and 30. These values were obtained from elevation errors analysis (see Table 3.3).

The uncertainties associated with mapping coastal flooding are affected by the characteristics of elevation data errors (Hinkel et al., 2021). To address this issue, a third criterion was introduced to assess the confidence level of any inundation model by quantifying the relationship between flood depth predictions and underlying elevation data uncertainties. Following Dusseau et al. (2023), an inundation model should achieve at least a 68% confidence level to be considered reliable for coastal flood exposure assessments, including tsunami inundation.

The confidence level can be determined by evaluating the relationship between the median of the maximum inundation height in terrestrial areas (ground elevation >0 m) derived from inundation simulations (*MIH*) and the root-mean-square error (RMSE) of the DEM used as the input for the inundation model (Fig. 3.10b) (Gesch, 2018). By assuming unbiased DEM errors (mean elevation errors = 0), the confidence level is derived from the standardized normal distribution of the ratio of MIH to DEM error. The standardized inundation height, *Z*, is computed as:

$$Z = \frac{MIH}{RMSE} \tag{29}$$

The confidence level, *CL*, is then calculated as the probability that the true inundation height lies within ± $Z/2$ standard deviations of the mean. This is expressed using the cumulative distribution function (*CDF*) of the standard normal distribution, Φ:

$$CL = \Phi\left(\frac{Z}{2}\right) - \Phi\left(-\frac{Z}{2}\right) \tag{30}$$

For example, an inundation model using an elevation dataset with an RMSE of 2 m achieved a minimum of 68% confidence level when the simulated MIH was 4 m (twice the RMSE value of the used DEM). Conversely, inundation models with lower MIH values yielded lower confidence levels. The minimum MIH values required to achieve 68% ($MIH_{68\%}$) and 95% ($MIH_{95\%}$) confidence level can be computed as:

$$MIH_{68\%} = RMSE\ x\ 2 \tag{31}$$

$$MIH_{95\%} = (RMSE\ x\ 1.96) x\ 2 \tag{32}$$

3.5.1 Performance of Inundation Models: Influence of DEMs

Fig. 3.11 presents the simulated inundation maps for the scenario that exclusively evaluated the influence of DEMs on inundation model performance. The inundation maps revealed varying degrees of agreement between the simulated and observed inundation extents across different DEMs. CoastalDEM demonstrates the closest agreement to the observed inundation extent, with an inundation extent difference of only 1.25%. FABDEM and DiluviumDEM also showed relatively good agreement, with

inundation extent differences of 4.89% and 5.90%, respectively. Based on the first reliability criterion (the difference between simulated and observed inundation extent must be less than 25%), six DEMs satisfied the given threshold: CoastalDEM, FABDEM, DiluviumDEM, COP30, EDEM, and local DEMNAS.

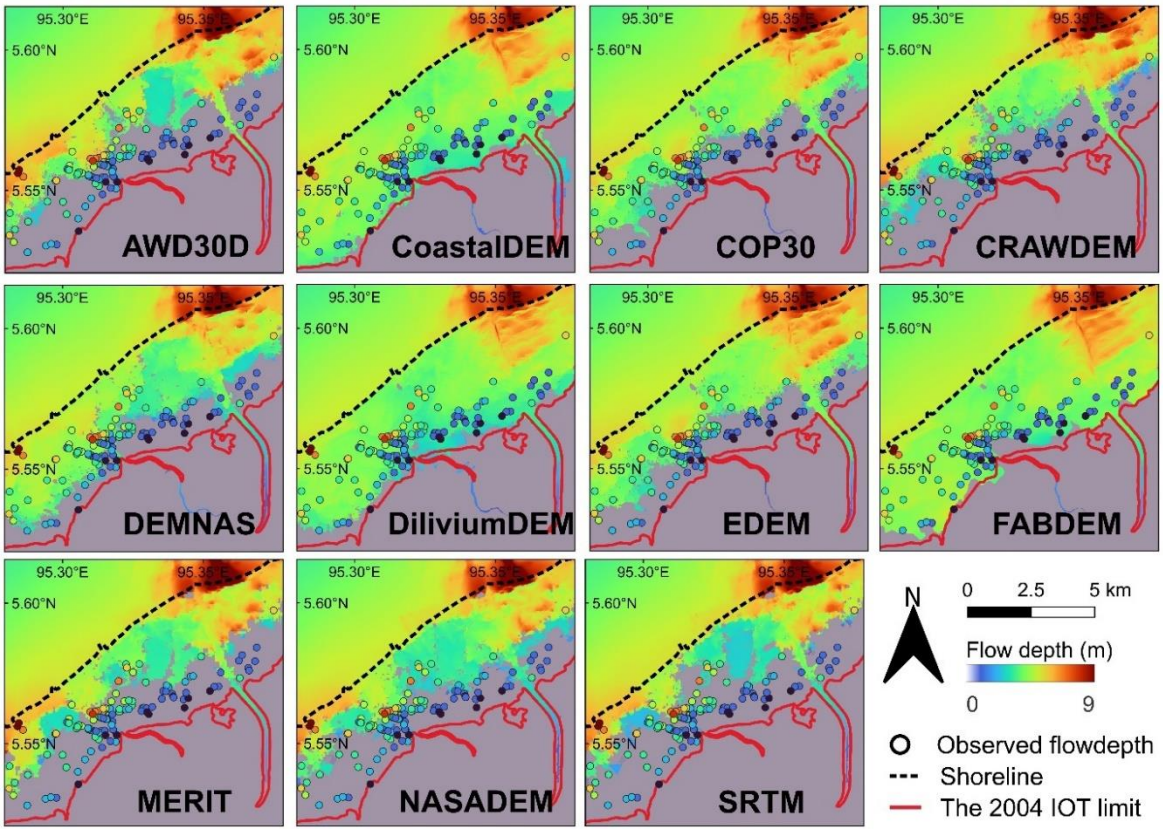


Fig. 3.11. The maximum flow depth distributions for DEMs coupled with uniform roughness Manning. The black dashed line indicates shoreline in 2004, and red line represents the 2004 IOT inundation limit.

The relationship between the DEM error characteristics and inundation bias was observed to be nonlinear (Fig. 3.12). Although DEMs with lower vertical elevation errors generally exhibited lower inundation bias, statistical testing revealed that this relationship is better represented by a quadratic function (Fig. 3.12a). Furthermore, MAE proved to be a superior indicator compared to RMSE for characterizing the relationship between DEMs' error characteristics and inundation accuracy (Fig. 3.12b).

Next, the inundation models that satisfied the first reliability criterion were then validated against observed inundation depth. As illustrated in Fig. 3.13, the results showed that none of these models met the second reliability criterion (i.e., Aida's parameters: $0.8 \leq K \leq 1.2$ and $k \leq 1.6$). Consequently, the third reliability criterion was

not examined, and it was concluded that inundation models coupled with uniform Manning coefficients were not reliable for reproducing the historical 2004 IOT event.

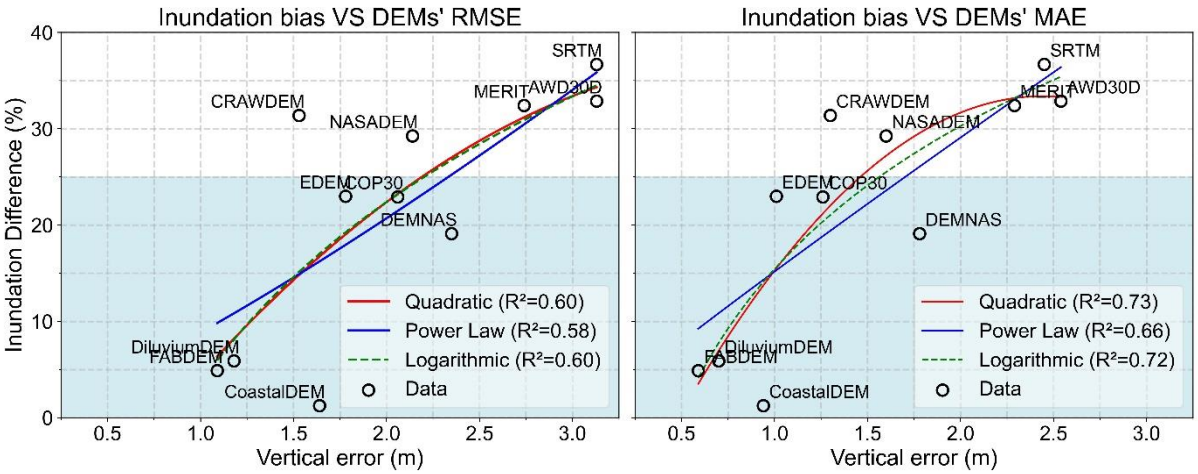


Fig. 3.12. Comparison of DEMs' errors and inundation extent differences between simulated and observed inundation. The left and right panels use the RMSE and MAE as elevation error indicators, respectively. The light blue shaded area indicates the acceptable 25% threshold for the inundation extent differences.

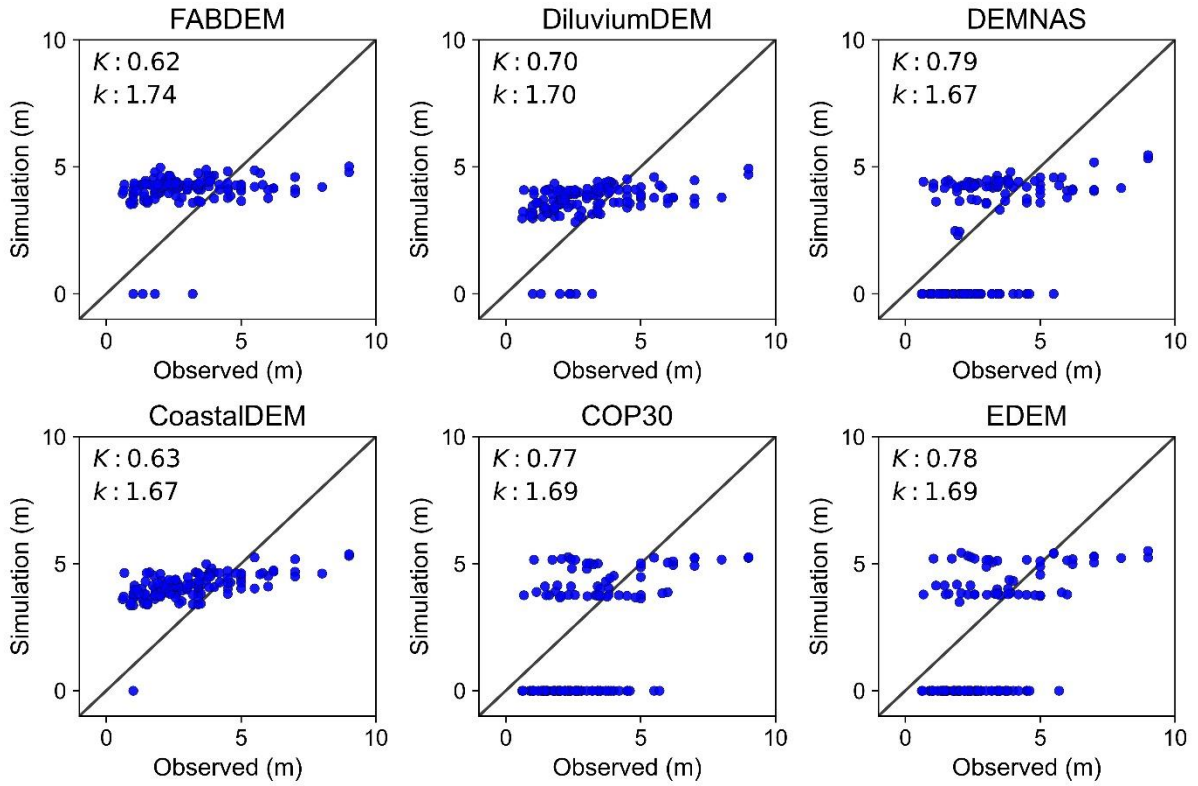


Fig. 3.13. Simulated vs. observed inundation depths for the scenario using DEMs coupled with uniform Manning coefficients, with *K* and *k* representing the geometric mean and standard deviation of Aida's parameters, respectively.

It is important to note that this conclusion stems from the proposed sequential evaluation approach, which requires models to satisfy all reliability criteria in sequence. If these criteria were applied independently, different conclusions might emerge. For instance, CRAWDEM could be considered reliable based solely on its confidence level estimate exceeding 68% (Table 3.7), while MERIT and SRTM might be assumed reliable based solely on Aida's parameters, despite their inundation extents covering less than 20% of validation points. These discrepancies underscore the importance of using sequential criteria when determining reliable elevation model inputs for tsunami simulation.

Table 3.7. Sequential validation assessment for DEMs coupled with uniform Manning

Input DEM	Validation-1	Validation-2			Validation-3			
	Inundation extent difference (%)	*N*	*K*	*k*	$MIH_{68\%}$ (m)	$MIH_{95\%}$ (m)	MIH_{sim} (m)	CL (%)
AWD30D	32.87	23	0.90	1.63	6.26	12.27	4.74	55.09
CoastalDEM	1.25	133	0.63	1.67	3.28	6.43	4.18	79.73
COP30	22.92	57	0.77	1.69	4.12	8.08	4.41	71.57
CRAWDEM	31.38	37	0.93	1.55	3.06	6.00	4.34	84.41
DEMNAS	19.11	71	0.79	1.67	4.70	9.21	4.16	62.39
DiluviumDEM	5.9	127	0.70	1.70	2.36	4.63	3.91	90.23
EDEM	22.99	56	0.78	1.69	3.56	6.98	4.43	78.64
FABDEM	4.89	130	0.62	1.74	2.18	4.27	4.32	95.24
MERIT	32.41	23	1.01	1.56	5.48	10.74	4.41	57.93
NASADEM	29.25	41	1.00	1.61	4.28	8.39	3.90	63.84
SRTM	36.68	17	1.16	1.58	6.26	12.27	4.25	50.32

N: number of observation points covered by inundation extent; CL: confidence level.

3.5.2 Performance of Inundation Models: Influence of DEMs and LCR Models

Figs. 3.14–3.16 present the simulated inundation maps for scenarios where DEMs were coupled with LCR models at resolutions of 30 m, 10 m, and 5 m, respectively. The first reliability assessment indicated that DEMs coupled with LCR models led to greater underestimation compared to when DEMs were combined with a uniform Manning's coefficient (Fig. 3.17). Given the acceptable threshold for the first reliability criterion, only four DEMs satisfied the given threshold: DEMNAS, FABDEM, CoastalDEM, and DiluviumDEM.

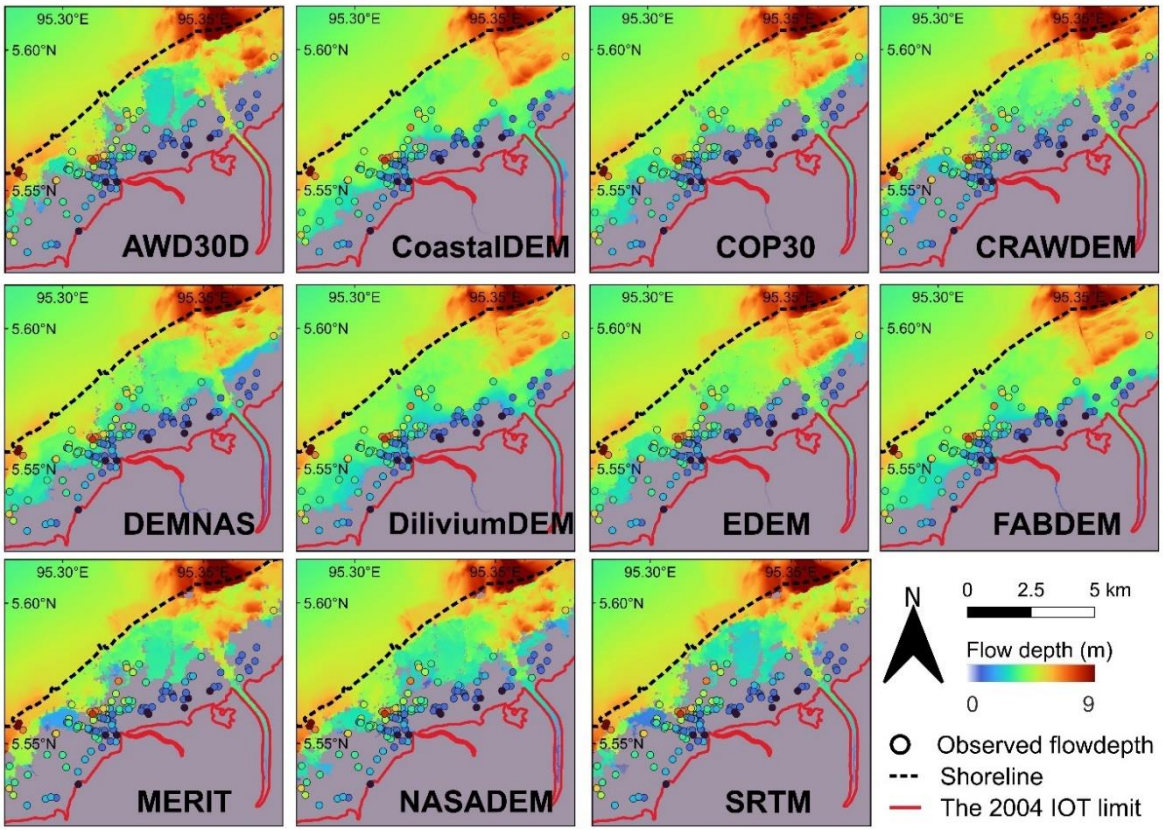


Fig. 3.14. The maximum flow depth distributions for DEMs coupled with LCR model at 30 m resolution. The black dashed line indicates shoreline in 2004, and red line represents the 2004 IOT inundation limit.

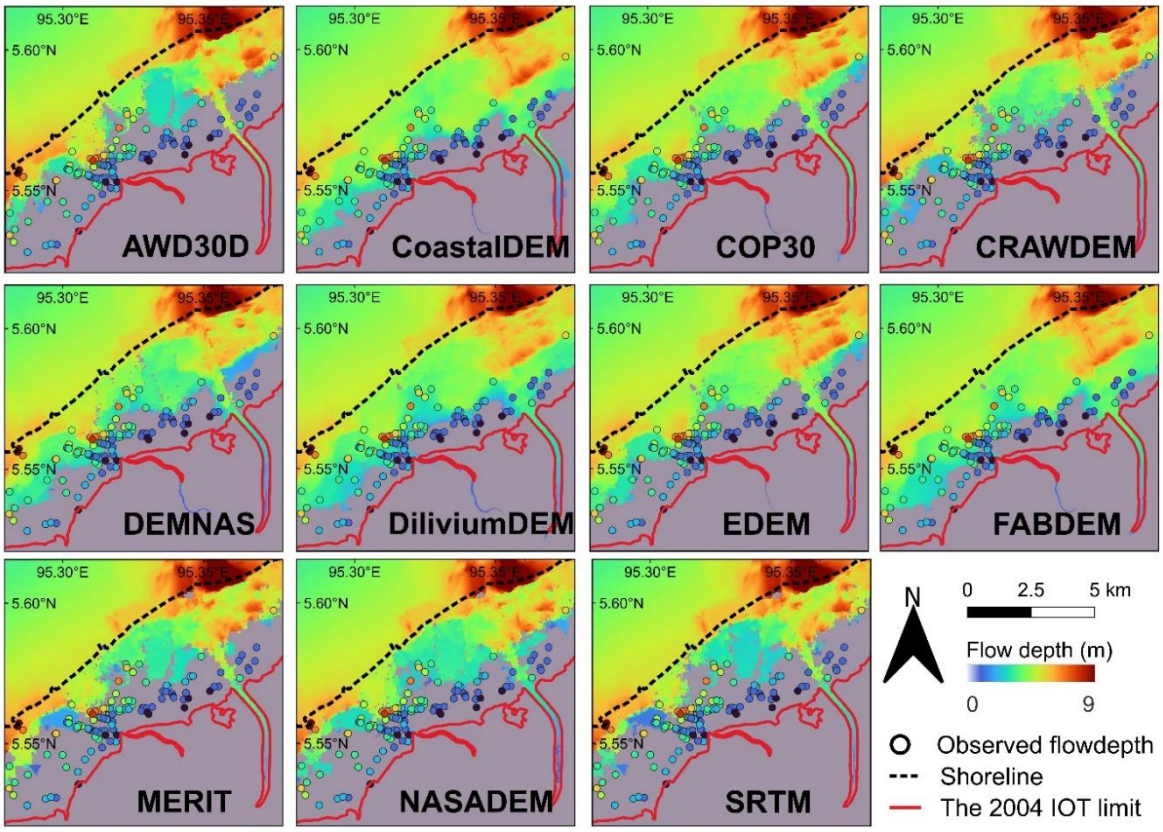


Fig. 3.15. The maximum flow depth distributions for DEMs coupled with LCR model at 10 m resolution. The black dashed line indicates shoreline in 2004, and red line represents the 2004 IOT inundation limit.

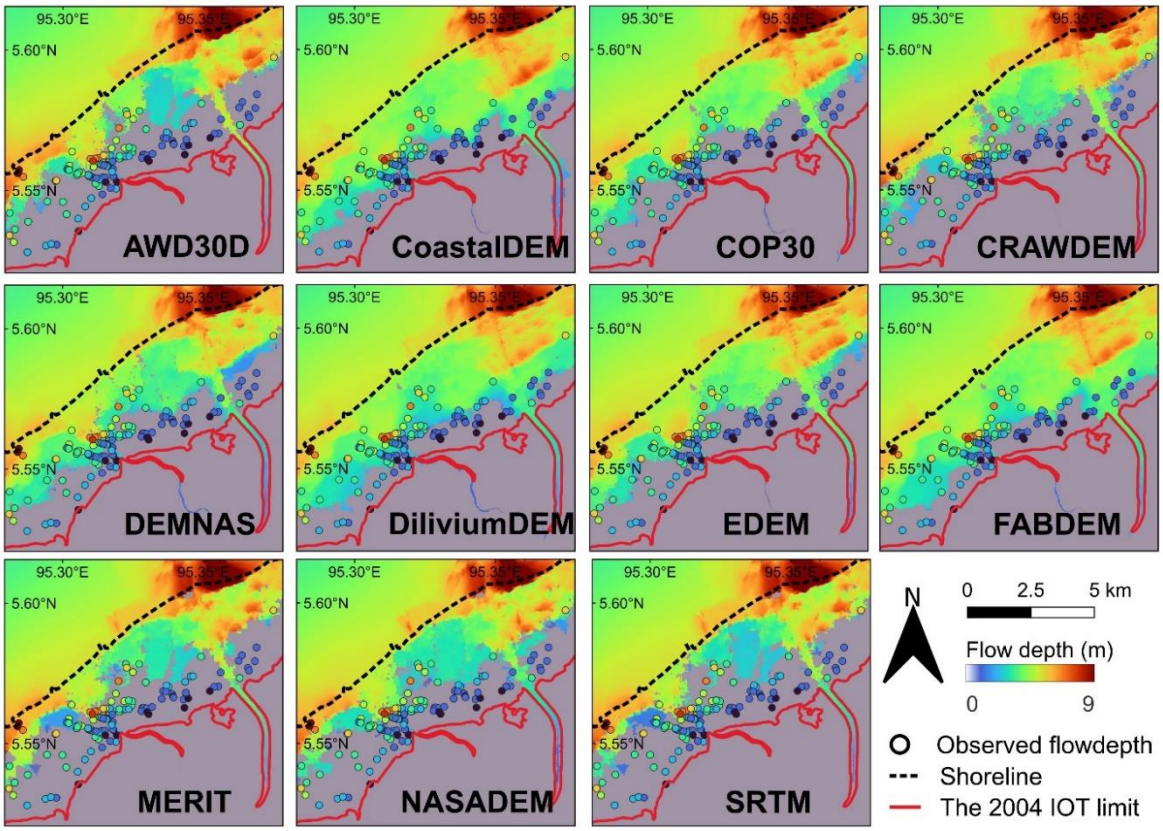


Fig. 3.16. The maximum flow depth distributions for DEMs coupled with LCR model at 5 m resolution. The black dashed line indicates shoreline in 2004, and red line represents the 2004 IOT inundation limit.

The effect of LCR model spatial resolution on inundation extent was observed to be minimal, with differences of less than 5%. Similar to the scenario using uniform Manning coefficients, the integration of DEMs and LCR models showed that DEMs with lower vertical errors tend to produce less inundation bias. Here, the relationship between DEM error characteristics and inundation bias was better represented by a logarithmic function rather than a quadratic function (see Fig. 3.17).

The inundation models that met the first reliability criterion were further evaluated using a second reliability assessment. The assessment revealed that combinations of DEMs with uniform Manning coefficients did not satisfy the AIDA parameters (Fig. 3.18). Interestingly, when coupled with LCR models, only CoastalDEM and DiluviumDEM satisfied the given thresholds. Again, the effect of the spatial resolution of the LCR model remained minimal.

For the combinations of CoastalDEM and DiluviumDEM with LCR models that satisfied both reliability criteria, confidence levels were then calculated. First, maximum flow depth distributions were examined to estimate the median inundation height (MIH).

The results indicated non-normal distributions with positive skewness (Fig. 3.19). The analysis showed that LCR models with finer resolution produced lower MIH estimates, likely because finer-resolution models better represent drainage systems (small rivers and ponds), which ultimately influence simulated flow depth (Koyama and Yamada, 2022). Additionally, CoastalDEM exhibited higher tsunami MIH estimates than DiluviumDEM, probably due to CoastalDEM's larger negative bias (averaging −1 m).

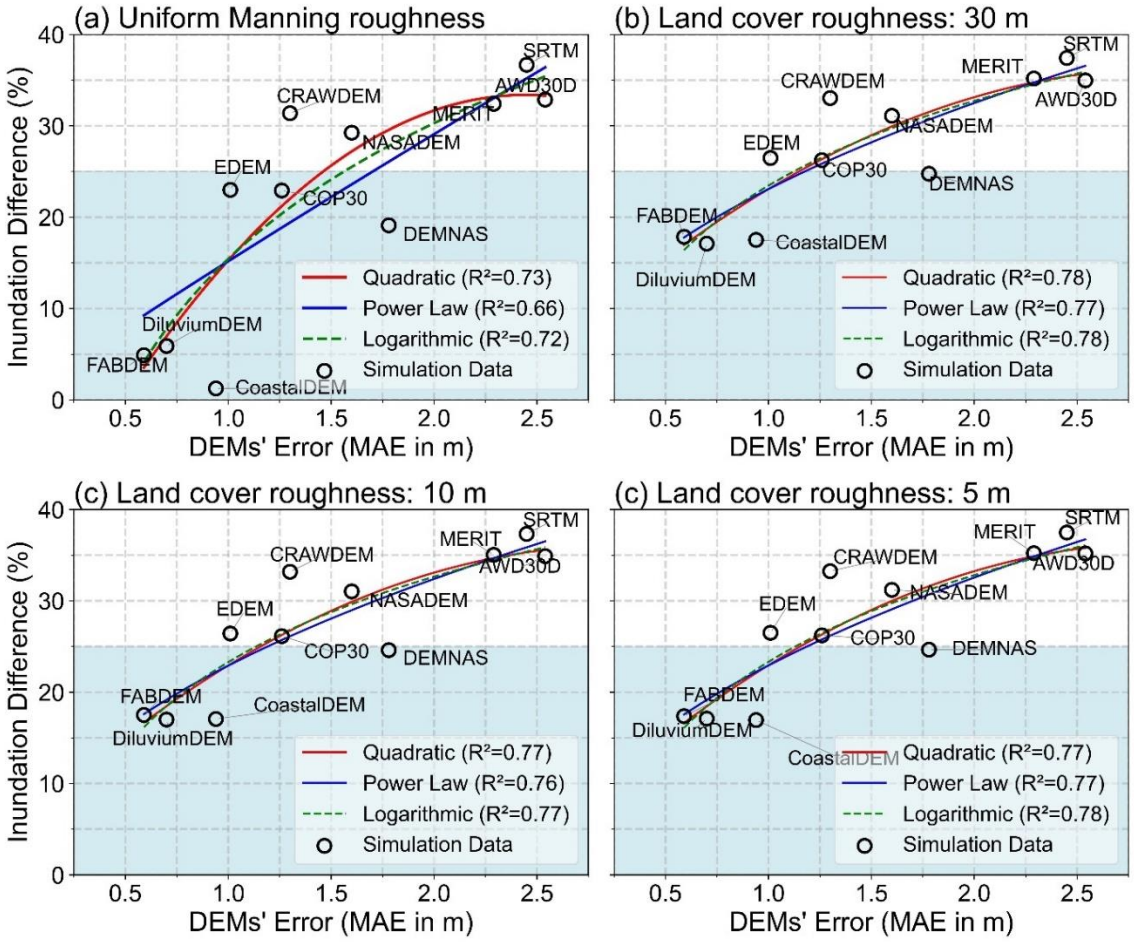


Fig. 3.17. First validation assessment for scenarios where DEMs coupled with: (a) a uniform Manning's coefficient, (b) LCR model at 30 m resolution, (c) LCR model at 10 m resolution, and (d) LCR model at 5 m resolution. The light blue shaded area indicate threshold region for acceptable model performance.

Given the skewed flow depth distributions, a log-normal distribution was employed to compute the fitted-median values, which served as MIH estimates for confidence level calculation (Glimsdal et al., 2019). The computed confidence levels revealed that both CoastalDEM and DiluviumDEM exceeded the minimum threshold (confidence level ≥68%), with DiluviumDEM achieving higher confidence level estimates (88.6% on average). This superior performance is related to DiluviumDEM's better vertical

accuracy compared to CoastalDEM (Table 3.8). The effect of LCR model resolution on confidence levels was minimal, with differences of less than 5%.

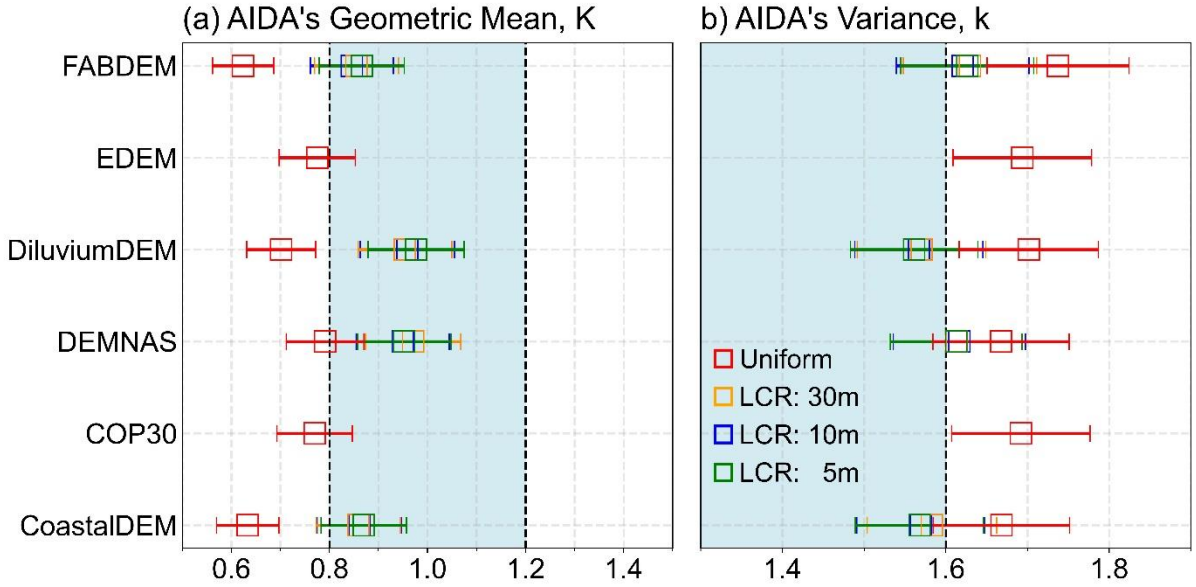


Fig. 3.18. Second validation assessment ─ Calculated Aida's parameters for inundation models that satisfy the first reliability criteria: (a) geometric mean (*K*), and (b) variance (*k*). The black dashed line indicates the acceptable threshold.

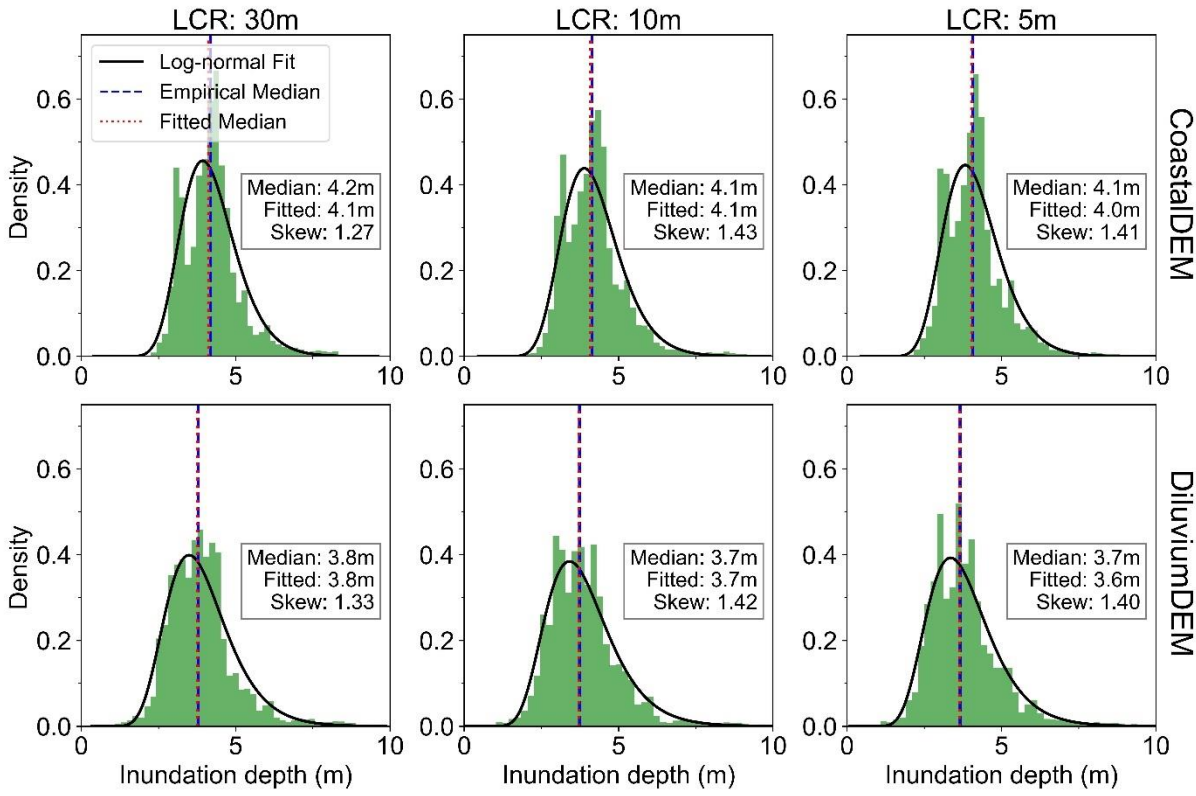


Fig. 3.19. Histograms of simulated flow depths for inundation models that satisfied the second reliability criterion.

Table 3.8. Confidence level (CL) assessment for models satisfied second reliability criterion

DEM	RMSE	Minimum MIH to reach:		Computed MIH relative to LCR models:		
		$CL_{68\%}$	$CL_{95\%}$	LCR: 30 m	LCR: 10 m	LCR: 5 m
CoastalDEM	1.53 m	3.06 m	6.00 m	81.97	81.97	80.89
DiluviumDEM	1.18 m	2.36 m	4.63 m	89.26	88.31	88.31

It is important to note that finer-resolution LCR models generally yield slightly lower confidence levels. This is associated with their MIH estimates, as the finer models produce lower inundation depths due to more precise mapping of drainage systems (see Fig. 3.19). Furthermore, coarser LCR models, which rely on lower-resolution land cover maps, identified more built-up areas (see Fig. 3.6). The larger proportions of the built-up areas likely enhance flow resistance, decrease flow infiltration capacity, and result in greater flow depths. In conclusion, after assessing all reliability criteria, inundation models that incorporate LCR models with either CoastalDEM or DiluviumDEM have proven to be more reliable choices for tsunami modelling.

3.5.3 Performance of Inundation Models: Influence of Manning Coefficients

It should be noted that, all the results discussed in sections 3.5.2 were derived from inundation models using Manning coefficients from Bunya et al. (2010). To evaluate how variations in Manning coefficients affect flood inundation performance, simulated inundation maps from models using the coefficients from Bunya et al. (2010) and Koshimura et al. (2009) were compared. For simplicity, the comparison was limited to inundation models using CoastalDEM and DiluviumDEM, as these were the only DEMs that met all reliability criteria in the previous assessment. Additionally, since spatial resolution showed minimal effect on outcomes, only the finest 5 m resolution LCR model was used for this comparison.

The analysis revealed that models using Manning coefficients from Koshimura et al. (2009) produced larger inundation extents, resulting in smaller differences between simulated areas and the observed inundation limit (Table 3.9). These models also generated higher inundation depths, as indicated by the higher MIH estimates compared with models using coefficients from Bunya et al. (2010). Furthermore, the simulated flow depths showed better agreement with the observational data, as represented by lower RMSE values. These patterns were exhibited by both CoastalDEM and DiluviumDEM and can be attributed to the lower Manning

coefficients for vegetated areas provided by Koshimura et al. (2009) compared with Bunya et al. (2010) (see Table 3.4). The lower Manning coefficients might reduce vegetation's damping effect and consequently lead to larger inundation extents and deeper inundation depths.

Despite these differences, the reliability parameters exhibited an average variation of less than 5% between the two coefficient sources, indicating a minimal overall impact on model performance. Nevertheless, these minor differences warrant careful interpretation, particularly when evaluating inundation extent. For instance, COP30 with LCR models utilizing Manning coefficients from Bunya et al. (2010) demonstrated inundation differences averaging 26% against the actual 2004 IOT event—1% above the acceptable threshold (see section 3.5.2). Employing Manning coefficients from Koshimura et al. (2009) may increase inundation extent and enable COP30 to meet the first reliability criterion. However, this adjustment does not ensure compliance with the other two reliability criteria.

Table 3.9. Validation assessments comparison from different source of Manning coefficients

Reliability Parameter	CoastalDEM & LCR 5 m. Manning coefficient source:		DiluviumDEM & LCR 5 m. Manning coefficient source:	
	Bunya et al. (2010)	Koshimura et al. (2009)	Bunya et al. (2010)	Koshimura et al. (2009)
First reliability criteria				
Inundation extent (Difference to 2004 IOT)	16.94%	11.05%	17.10%	12.94%
Second reliability criteria				
Flow depth agreement:				
RMSE	1.72 m	1.64 m	1.74 m	1.60 m
Aida's *K*	0.87	0.80	0.98	0.90
Aida's *k*	1.57	1.55	1.56	1.55
Third reliability criteria				
MIH	4.0 m	4.2 m	3.7 m	3.9 m
Confidence level	80.89%	83.01%	88.31%	90.16%

3.5.4 Performance of Inundation Models: Effect of Land Cover Change

To evaluate the sensitivity of inundation models to land cover change effects, simulation results from models using LCR derived from land cover maps of 2004 (LC-

04) and 2014 (LC-14) were compared, as shown in Fig. 3.20. Both LCR models were at 5 m resolution and employed Manning coefficients from Bunya et al. (2010). The results revealed that land cover change had minimal impact on variations in total inundated areas. Additionally, land cover changes also had a minimal impact on inundation depth, as shown in Fig. 3.21.

Despite this, the differences would still influence exposure assessments, particularly in high-density urban areas. Furthermore, contrasting patterns were observed for these two DEMs. CoastalDEM showed an increase in total inundated areas, from 2389.44 ha in 2004 to 2397.22 ha in 2014. Conversely, DiluviumDEM projected a decrease of approximately 3 ha over the same period. Additionally, the flow depth distributions exhibited slightly different skewness rates

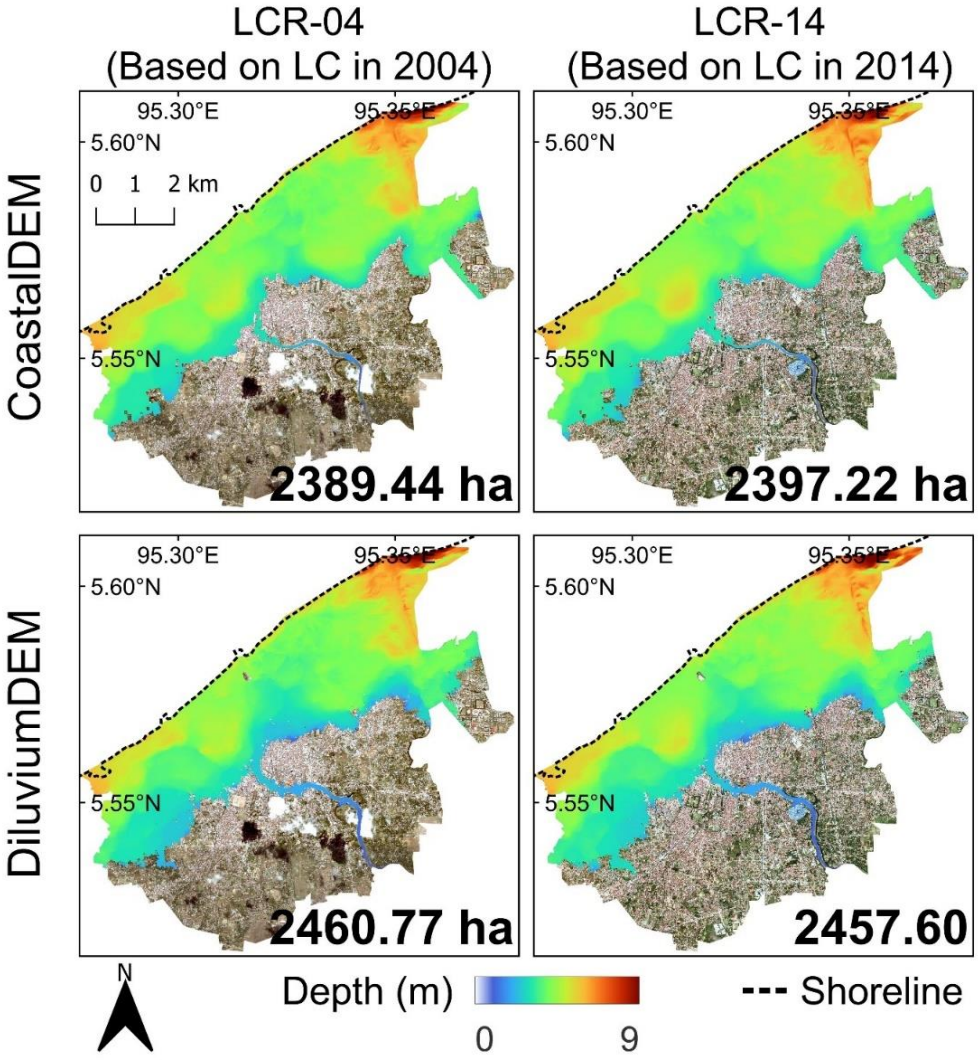


Fig. 3.20. Comparison of total inundated areas (land and inland water) due to land cover changes between 2004-2014 period.

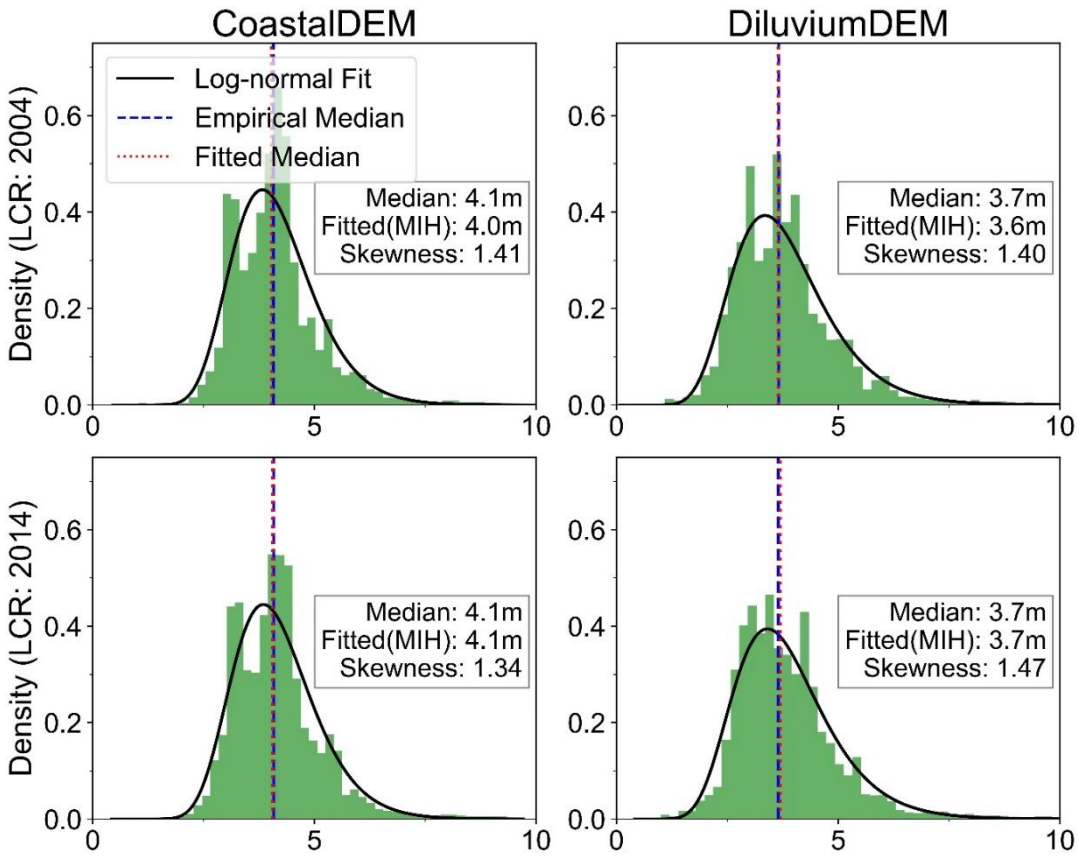


Fig. 3.21. Comparison of flow depth distributions (land and inland water) due to land cover changes.

3.6. Summary of Chapter III

In summary, the results indicate that two global error-reduced DEMs—CoastalDEM and DiluviumDEM—are more reliable elevation inputs for inundation modelling when coupled with LCR models, with DiluviumDEM demonstrating a higher confidence level than CoastalDEM. The effect of the LCR model spatial resolution on the inundation model accuracy was observed less significant.

Nevertheless, finer-resolution LCR models are recommended, as they provide a superior representation of terrain features. Given that most DEMs were developed after the 2004 IOT event, and land cover changes demonstrated minimal effects on tsunami inundation patterns, the findings suggest that DEMs exert a more dominant influence than LCR models in determining tsunami intensity measurements.

CHAPTER IV:
COMPOUNDING BIAS AND CROSS-UNCERTAINTY EXPOSURE ASSESSMENT

4.1. Compounding Bias Assessment

As discussed in the previous chapters, both DEMs and exposure datasets introduce bias into tsunami exposure assessments. To examine the effects of these compounding biases, exposure estimates derived from global datasets were compared against local reference datasets. The satellite-derived 2004 Indian Ocean Tsunami (IOT) inundation limit served as the hazard reference, while local datasets provided the exposure reference.

The simulated hazard was derived from the scenario using the 2004 LCR model at 5 m resolution coupled with DiluviumDEM. DiluviumDEM was selected based on its superior confidence level compared to CoastalDEM, as detailed in chapter 3. For the simulated exposure estimates, the least biased datasets were selected. Based on the relative bias assessment for 2004 period, the 30 m GAIA and 1 km GlobPop were chosen as built-up areas and gridded population datasets, respectively.

The results revealed that DEMs introduced larger biases than exposure datasets (Table 4.1). For built-up areas, GAIA caused a 6% overestimation, while DiluviumDEM introduced a 52% underestimation. Integrating both datasets reduced the underestimation to 34.7%, although this remained nearly six times larger than the bias introduced by GAIA alone.

Table 4.1. Estimation of built-up areas and populations exposed to the 2004 IOT.

Exposure dataset	**Tsunami inundation limit**	
	Observed	Simulation (DiluviumDEM)
Exposed built-up areas		
Reference (BU-04)	1535.71 ha	736.58 ha
Global dataset (GAIA)	1627.80 ha	1002.11 ha
Exposed population		
Reference (POP-04)	176,003	104,189
Global dataset (GlobPop)	114,882	65,416

A consistent bias trend was observed in population exposure estimates. GlobPop alone caused a 35% underestimation, whereas the DEM increased the underestimation to 40.8%. The combination of GlobPop and DiluviumDEM further increased the underestimation rate to 62.8%. This indicates that compounding biases are substantially larger than individual biases, particularly those from exposure datasets.

4.2. Modelling Expected Future Tsunamigenic Event

Previous research by Jihad et al. (2020) suggested that an Mw 8.7 earthquake could potentially occur within the Sunda-Andaman segment due to the seismic gap. Earthquakes with this typical magnitude can generate tsunamis within the area of interest (AOI), as reported by Horspool et al. (2014). To forecast the tsunami risk from future tsunamigenic events, the probabilistic tsunami models were employed (Grezio et al., 2017). To date, probabilistic models, also known as probabilistic tsunami hazard assessments (PTHA), are predominantly used for pre-disaster risk assessments. Briefly, PTHA quantifies the relationship between tsunami heights and their exceedance probabilities over a specified period (Behrens et al., 2021).

PTHA addresses the uncertainty caused by variations in earthquake properties (seismic source, fault segmentation, slip distribution, and maximum magnitude) by simulating a set of random hypothetical tsunamigenic scenarios that pose a threat to the site of interest. One of the outputs of PTHA is the hazard curve (Fig. 4.1), which depicts the annual rate of tsunamigenic events (i.e., events per year) affecting specific locations exceeding a certain wave height.

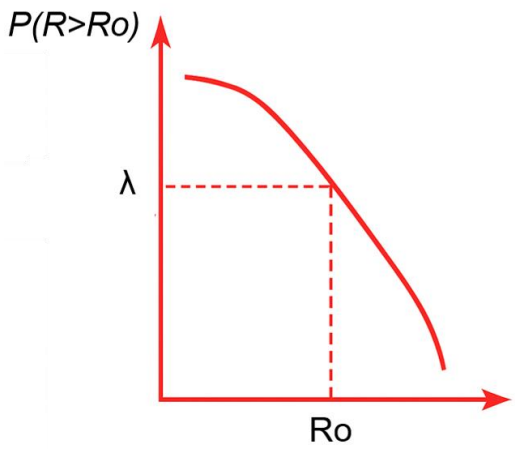


Fig. 4.1. Schematic illustration of the tsunami hazard curve. Modified from Geist and Lynett (2014)

An often-implicit assumption in numerous probabilistic analyses of tsunamis and other natural hazards is that these events occur randomly over time and are independent of one another, a concept known as the Poisson process. In this process, the interval between tsunami occurrences follows an exponential distribution. Consequently, the hazard curve illustrates the probability that one or more tsunamis will reach or surpass the specified wave height on the horizontal axis within a given exposure time (*T*), as expressed by $P(R \geq R0) = 1 - exp(-\lambda T)$, where λ represents the constant rate of occurrence of these tsunamis over time, as depicted in Fig. 4.1 (Geist and Lynett, 2014).

Recently, numerous PTHA have been conducted, including local, regional, and global scale assessments. These assessments have resulted in synthetic earthquake databases. One of them is the PTHA18 database, which is publicly accessible and consists of hundreds of thousands of synthetic earthquake scenarios, developed from both uniform and heterogeneous slip models (Davies and Griffin, 2018). Synthetic earthquake scenarios in the PTHA18 database have been validated against several historical tsunamis (Davies and Griffin, 2020; Davies, 2019).

PTHA18 modelled earthquake-tsunami scenarios and their frequencies at a global scale. These scenarios were modelled using several seismic sources (Fig. 4.2a), including the Sunda Subduction zone, which is the closest earthquake source zone to our area of interest (AOI). On each source zone, a large suite of hypothetical earthquake scenarios was created, with earthquake magnitudes ranging from Mw 7.2 to Mw 9.6. All earthquakes are represented as linear combinations of slip on the unit sources, with the dimension of each unit source being 50x50 km^2 (Fig. 4.2b). The moment magnitude values are related to the earthquake's slip and area by the following definitions:

$$M_0 = \sum_{j \in unit\ sources} S_j A_j \mu_j \tag{33}$$

$$M_w = \frac{2}{3}(log_{10}(M_0) - 9.05) \tag{34}$$

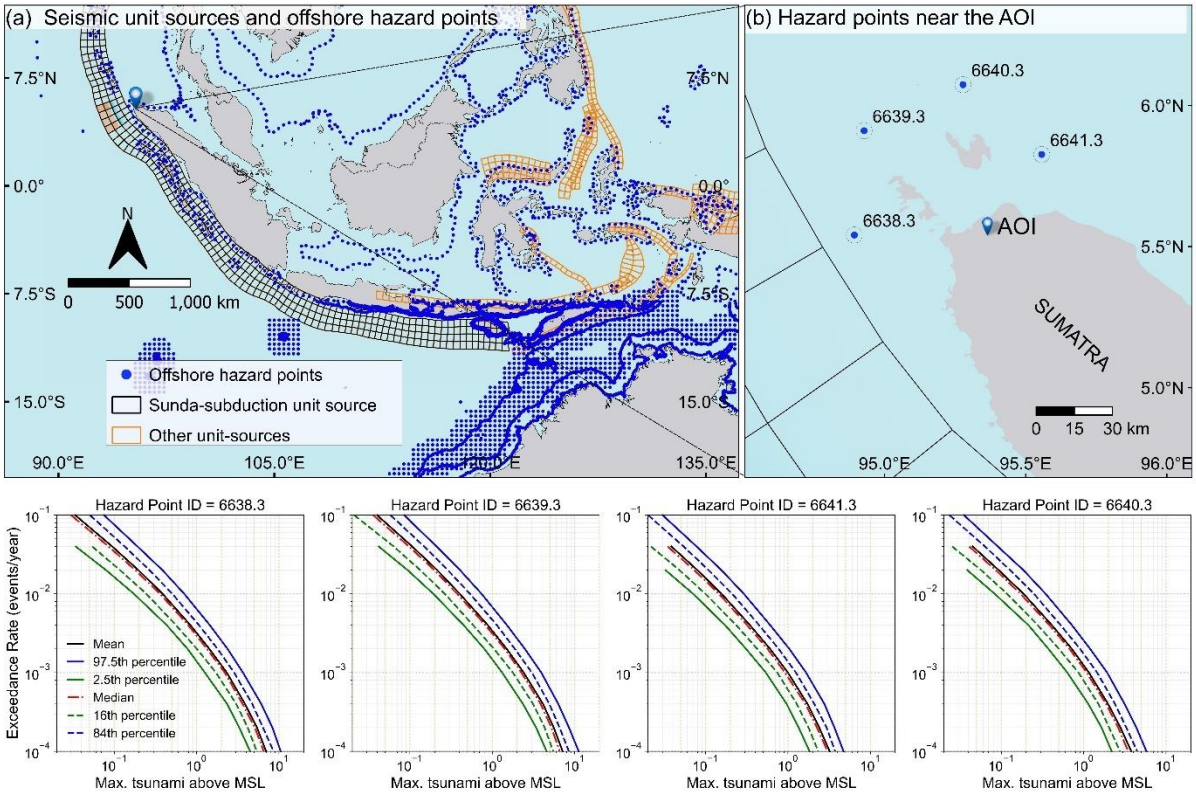
where M_0 is the seismic moment (kg m^2/s^2), M_w is the moment magnitude, S_j is the slip (m) on the *j'th* unit-source, A_j is the *j'th* unit-source area (m^2), and μ_j is the *j'th* unit-source shear modulus (kg/ms^2).

The optimal method for translating an offshore PTHA scenario into an onshore hazard at a specific site involves simulating inundation for each scenario $e \in E$, where E represents the set of all offshore PTHA scenarios or a subset of interest, such as all scenarios within a particular source zone (Power et al., 2017; Davies & Griffin, 2020; Basili et al., 2021). For each individual earthquake scenario, PTHA18 estimates the frequency at which earthquake-tsunami waves exceeding a specified "size" occur at thousands of observation points, also referred to as hazard points. These observation points are distributed globally and are located offshore at a depth of approximately 100 m (Fig. 4.2a and 4.2b). Therefore, the estimated tsunami heights at these offshore hazard points cannot be directly used to represent onshore tsunami conditions.

Assuming that the offshore PTHA18 accounts for hazard uncertainties through multiple scenario-frequency models $i \in I$, where I denotes the set of all alternative scenario-frequency models, the exceedance rate (λi) for each scenario-frequency model can be quantified as

$$\lambda_i(Q > QT) = \sum_{e \in E} r_i(e)\, 1_{(Q(e)>QT)} \tag{35}$$

In this context, $\lambda i(Q>QT)$ represents the exceedance rate, defined as the average number of events per year, for which a specific quantity of interest Q, such as tsunami height, surpasses a given threshold QT, under the scenario-frequency model i. Each scenario e possesses its own quantity of interest $Q(e)$, which is determined through high-resolution inundation modelling. The offshore PTHA specifies the long-term occurrence rate of each scenario e, denoted as $ri(e)$ (events/year), contingent upon the scenario-frequency model i. The indicator function $1_{(Q(e) > QT)}$ assumes a value of 1 if $Q(e)>QT$ and 0 if otherwise. The uncertainty in the exceedance rate, characterized by variation with $i \in I$, can be summarized using the mean and percentiles, as illustrated in Fig. 4.2c.



(c) Maximum tsunami vs excedance-rate curve at four hazard points near the AOI

Fig. 4.2. (a) Distribution of offshore hazard points and seismic unit sources in PTHA18. (b) The relevant hazard points near the AOI. (c). The corresponding exceedance-rate for hazard points near the AOI.

4.3 Defining Relevant Offshore Hazard Points

The PTHA18 database has been used for tsunami hazard assessments in Samoa (Giblin and Damlamian, 2022) and Tonga (Davies et al., 2022). To select the relevant scenarios for design earthquakes in PTHA18, a relevant hazard point must first be defined (Giblin et al., 2022). Four hazard points surround the AOI (Fig. 4.2b), of which two were considered most relevant, given their proximity to the Sunda subduction zone and perpendicular orientation to the AOI (Fig. 4.3a). To select the most relevant hazard point, the following steps were applied:

1. The 2004 IOT was reproduced using PTHA18 hypothetical scenarios.
2. The best-fitting scenario was selected.
3. The simulated wave profile from the best-fitting scenario at the hazard points was compared with reference data.

To reproduce the 2004 IOT, the return-period method based on Giblin et al. (2022) was followed. Codes to apply this method are provided in the Appendix. The synthetic earthquake scenarios of PTHA18 were filtered at a certain magnitude and specified

return periods. This produced a small scenario subset and reduced computational time, although it cannot explicitly quantify uncertainty. Given the objective of selecting the most relevant hazard point, uncertainty was initially neglected. To define the best-fitting scenario, the simulated wave profiles at two onshore points (Fig. 4.3b) were compared with the simulated wave profile from the inversion fault model by Koshimura et al. (2009) (see Section 3.4). The inversion fault model was assumed as the reference.

Mw 9.2 was set as the defined magnitude to represent the 2004 IOT and 758 years as the annual recurrence interval (return period) for the 2004 IOT. Both values were obtained from Davies (2019). Based on Fujii et al. (2021), the slip rate for Mw 9.2 scenarios was defined as ranging from 5 to 20 m. Given these criteria, 36 hypothetical earthquake scenarios based on the uniform slip model and 93 scenarios based on the heterogeneous slip model were identified. A list of the selected scenarios is provided in the appendix.

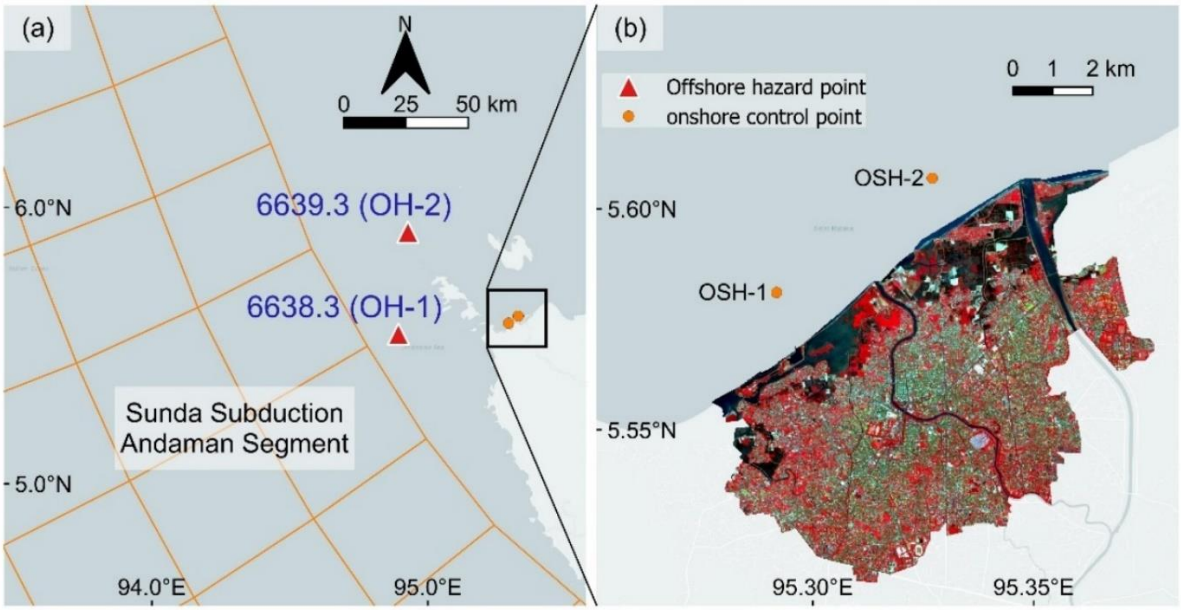


Fig. 4.3. The location of: (a) offshore hazard points (OH) and onshore control points (OSH)

Because the focus was on the Sunda-Andaman segment, the analysis was further limited to earthquake scenarios whose centroids (epicentres) were located within the Sunda-Andaman region, yielding 12 scenarios. The focal mechanism for all scenarios is thrust, which is also referred to as the reverse fault (Fig. 4.4). The details of the focal mechanism are listed in Table 4.2. PTHA18 provides the initial sea surface deformation files for each scenario in its database. Using the initial sea surface

information (Fig. 4.5), tsunami simulations were performed. The same numerical model configuration as in Section 3.4 was employed. However, the focus was only on tsunami wave profiles at offshore and onshore locations, whereas the inundation assessments were neglected.

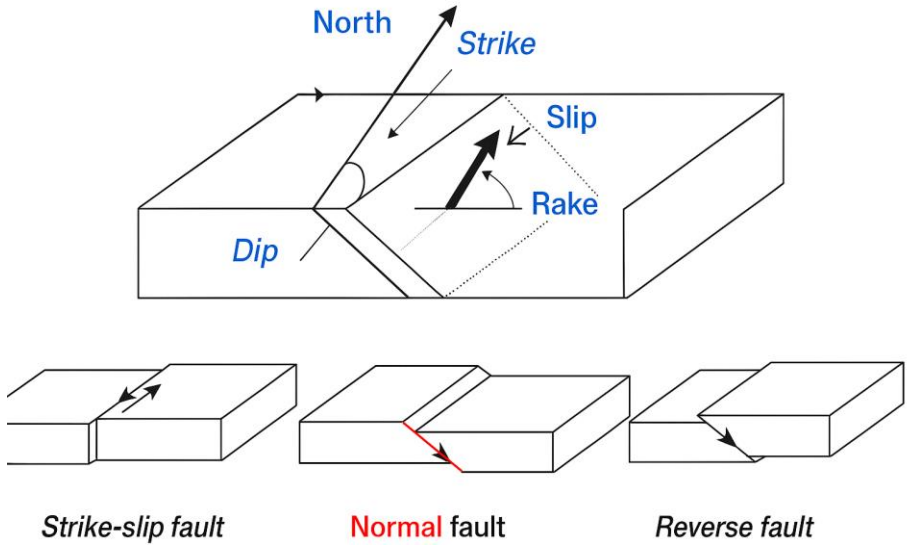


Fig. 4.4. Illustration of earthquake focal mechanisms and notation definition for parameters in Table 4.2

Table 4.2. Earthquake focal mechanism with slip model – HS: Heterogeneous; U: Uniform

No.	Scenario ID	Slip model	Depth of epicentre (km)	Dip (^{0})	Rake (^{0})	Strike (^{0})
1	104248	HS	10	10	90	273
2	104390	HS	40	21	90	270
3	104433	HS	9	9	90	275
4	107452	U	22	16	90	272
5	105933	HS	9	9	90	276
6	106012	HS	40	22	90	277
7	106036	HS	9	10	90	277
8	107580	U	39	24	90	283
9	107469	HS	9	10	90	277
10	107487	HS	23	19	90	275
11	107606	HS	10	11	90	277
12	107661	HS	24	20	90	275

The comparison of wave profiles at onshore points indicated that most synthetic scenarios had similar waveforms, consisting of a leading depression followed by an elevated wave, as shown in Fig. 4.6. However, compared to the reference model, the

wave peak positions of all scenarios were not perfectly aligned, particularly at OSH-1. Scenarios 104433 and 107469 showed better alignment, especially for the leading depression wave. Additionally, these scenario produced the wavelength which resembles that of the reference model. However, they exhibited a slimmer shape in OHS-2.

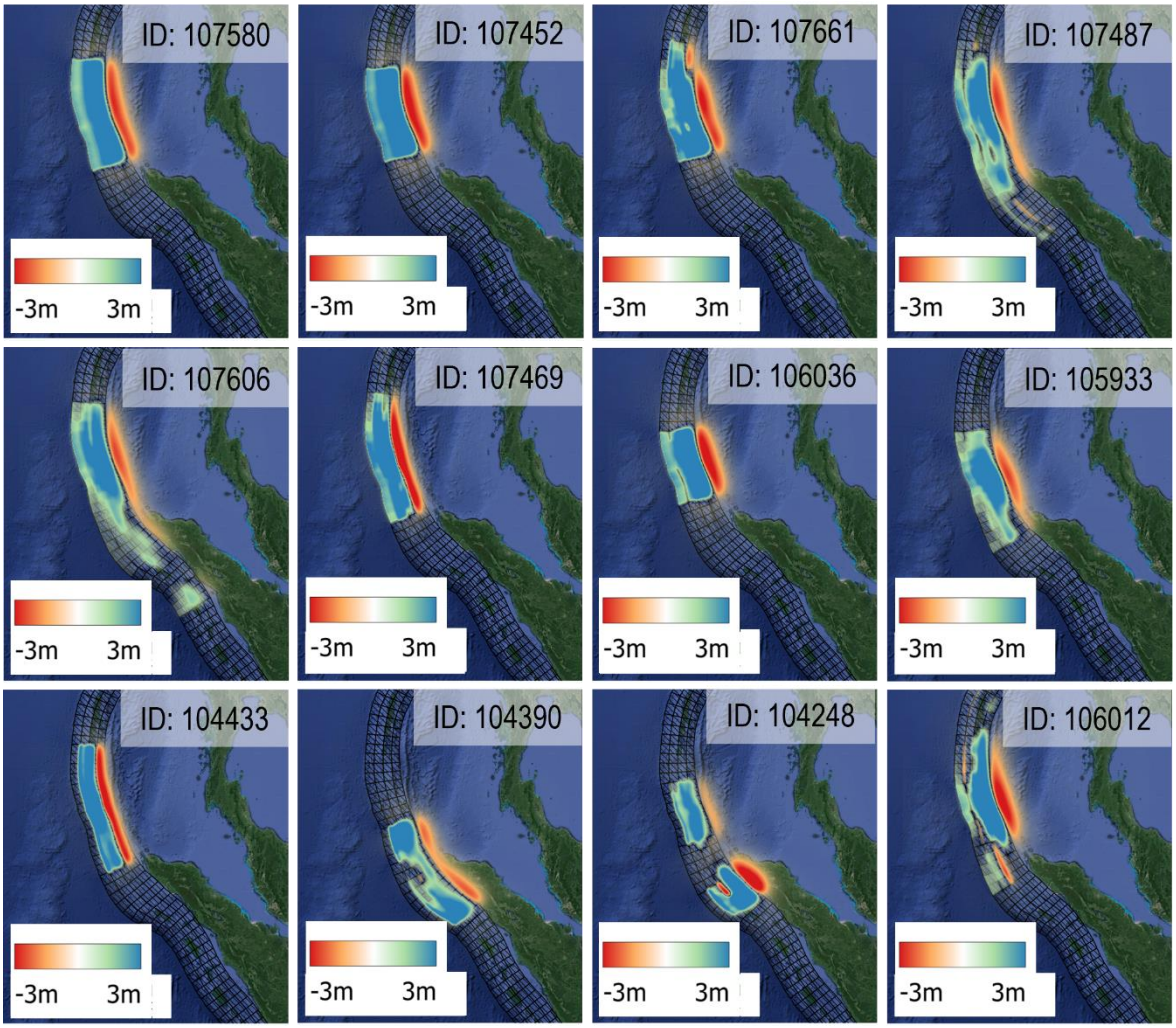


Fig. 4.5. Initial sea surface deformation for the Mw 9.2 selected scenarios. ID represents the scenario identification in the PTHA18 database. The black-gridded box shows Sunda subduction.

Although the reliability of PTHA18 has been validated by previous research (Davies and Griffin, 2020; Davies, 2019), the similarity in wave profiles from scenarios 104433 and 107469 also indicates that synthetic earthquake models from PTHA18 are capable of reproducing historical tsunamis within our AOI. Both scenarios were based on a heterogeneous slip model. This confirms that the actual slip distribution is heterogeneous rather than uniform, as observed by Fujita et al. (2024). Furthermore,

the onshore wave profile comparison indicated that scenario 104433 provided a better fit than scenario 107469. The evaluation of wave profiles at offshore hazard points was then based on the output from scenario 104433.

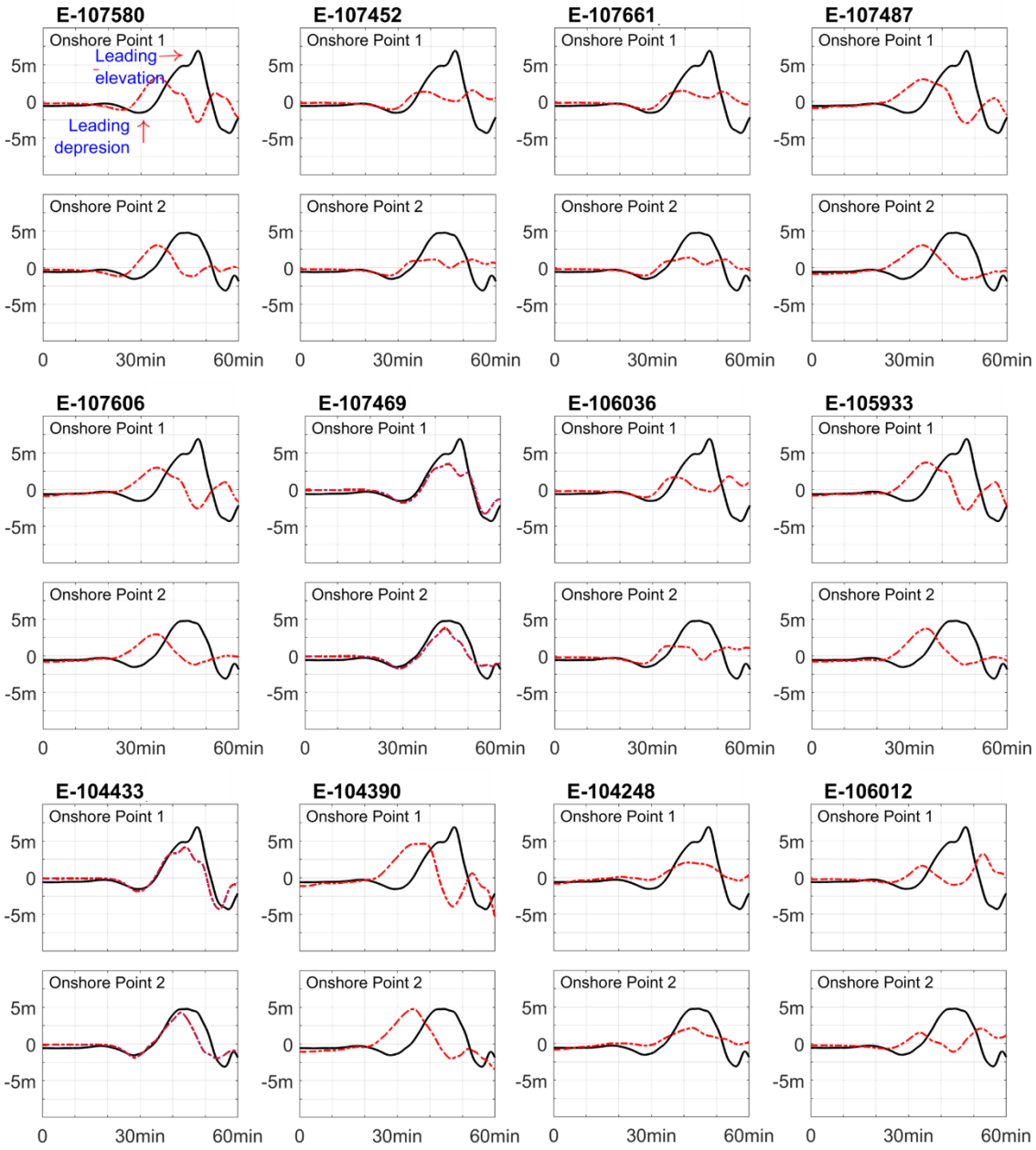


Fig. 4.6. Surface wave profiles at onshore points (OSH). The black and red lines represent the inversion fault model (reference) and the PTHA18 scenarios, respectively, while *E* indicates event or scenario.

Two key metrics were used to evaluate the accuracy of the tsunami offshore wave profile comparisons. The first is the normalized root-mean-square error (NRMSE). This metric measures the differences in water surface elevation between the reference and

scenario 104433, with lower values (0-1 scale) indicating better agreement in tsunami amplitude, expressed as:

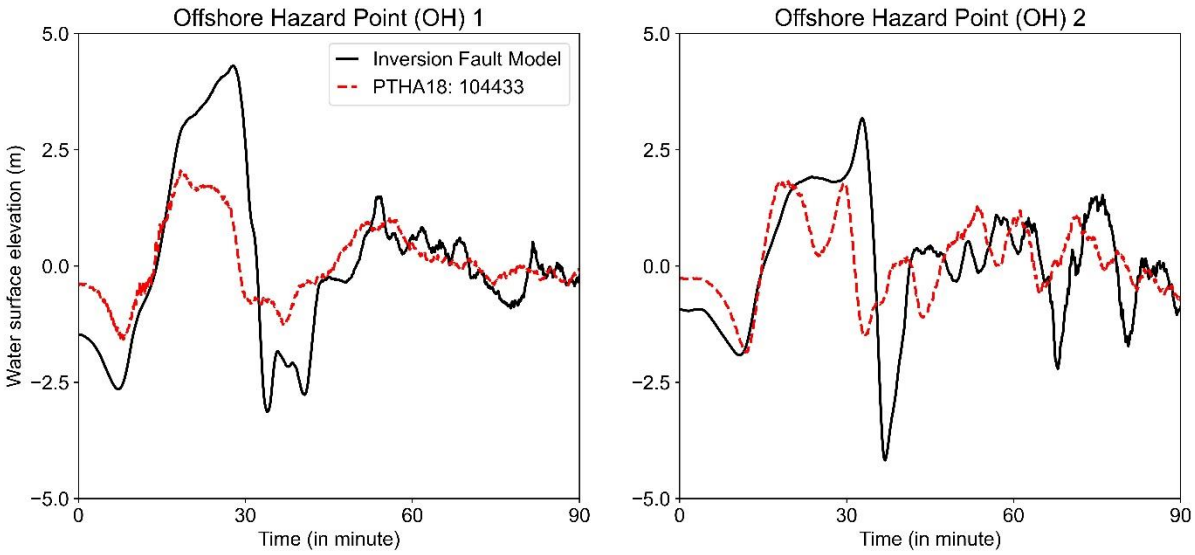
$$NRMSE = \frac{\frac{1}{N}\sum_{i=1}^{N}(\eta_{ref}(t_i) - \eta_{104433}(t_i))^2}{\max(\eta_{ref}) - \min(\eta_{ref})} \quad (36)$$

where $\eta_{ref}(t_i)$ and $\eta_{104433}(t_i)$ are water surface elevation at time t_i from the reference and scenario 104433, respectively.

Second, the cross-correlation maximum. The cross-correlation analysis was performed to assess waveform similarity and timing alignment between two time series (i.e., the reference and scenario 104433) as a function of time lag. The time lag, τ, reveals whether the hypothetical scenario predicted tsunami wave arrivals too early or too late, while the cross-correlation maximum represents the highest correlation value across all time lags and value closer to 1 signify nearly identical waveform patterns. The cross-correlation, $CC(\tau)$, can be expressed as:

$$CC(\tau) = \frac{\sum_{i=1}^{N}(\eta_{ref}(t_i) \cdot \eta_{104433}(t_i + \tau))}{\sqrt{\sum_{i=1}^{N}(\eta_{ref}^2(t_i) \cdot \sqrt{\sum_{i=1}^{N}(\eta_{104433}^2(t_i)}}} \quad (37)$$

The results showed that scenario 104433 showed a strong agreement in tsunami amplitude, with NRMSE values of 0.16 at offshore point 1 (OH-1) and 0.17 at point 2 (OH-2) (Fig. 4.7; see also Fig. 4.3). In terms of waveform shape, OH-1 exhibited better similarity than OH-2, with cross-correlation maximum values of 0.9 and 0.72, respectively. Considering all metrics, OHS-1 ─with ID 6638.3 in the PTHA18 database─ was selected as the most relevant hazard point and used as the reference for sampling Mw 8.7 scenarios from the database.



Accuracy Metric	OH-1	OH-2
NRMSE (0-1; 0~ perfect; <0.5~ good agreement)	0.16	0.17
CC maximum (0-1; 1~ perfect; >0.5~ good agreement)	0.90	0.72

Fig. 4.7. Wave profiles at two offshore hazard points (OH). The black line represents the inversion fault model (reference). The red lines indicate PTHA18 scenario 104433.

4.4 Sampling Mw 8.7 Earthquake Scenarios from PTHA18 Database

As mentioned in the previous sections, an ideal approach to solve the hazard uncertainty is by simulating all scenarios from the design earthquake magnitude. However, this approach requires an intensive number of inundation simulations and is ultimately computationally infeasible. Therefore, this study adopted the stratified random importance sampling approach developed by Davies et al. (2022). This approach focuses on a subset of scenarios that potentially generate large inundation. As with stratified sampling, the set of all scenarios *E* is split into multiple bins *Eb* corresponding to magnitude ranges *Mw,b*. Then, a fixed number of scenarios, *N(Mw,b)*, were randomly sampled from each bin based on their importance, *I(e)*, using weighted random sampling. The sampling weight is expressed as

$$w_{b,i}^{SIS}(e) = \frac{I(e)r*(e)}{\sum_{e \in E_b} I(e)r*(e)} \tag{38}$$

where $r*(e)$ should be non-zero for all scenarios that have $ri(e) > 0$. If only one scenario frequency model $i \in I$ is used, the natural choice is $r*(e) = ri(e)$. Using this weighting sampling method, then the 'all scenarios' exceedance-rate curve of all scenarios (equation 35) may be represented by a random scenario by:

$$\lambda_i^{SIS}(Q > Q^T) = \sum_{M_{w,b} \in magnitue\ bins} \lambda_i^{SIS}\left(Q > Q^T \mid M_{w,b}\right) \tag{39}$$

$$\lambda_i^{SIS}\left(Q > Q^T \mid M_{w,b}\right) = \lambda_i\left(M_{w,b}\right)\widehat{q_{b,i,T}} \tag{40}$$

$$\lambda_i\left(M_{w,b}\right) = \sum_{e \in E} r_i(e) \tag{41}$$

$$\widehat{q_{b,i,T}} = \frac{\left(\sum_{e \in E_{b,i}^{SIS}} \emptyset_{b,i}^{SIS}\ (e)\ 1_{(Q(e)>QT)}\right)}{N\left(M_{w,b}\right)} \tag{42}$$

$$\emptyset_{b,i}^{SIS}(\mathrm{e}) = \left(\frac{r_i(e)}{\lambda_i\left(M_{w,b}\right)}\right) /\ \ w_{b,i}^{SIS}(e) \tag{43}$$

The application of equation 39 results in the exceedance rate curve shown in Fig. 4.2c. The variance, denoted as σ^2, can be computed analytically as follows if the tsunami height *Q(e)* at hazard points is known:

$$\sigma^2\left(\ \lambda_i^{SIS}(Q > Q^T)\right)\ = \sum_{M_{w,b} \in magnitue\ bins} \sigma^2\left(\lambda_i^{SIS}\left(Q > Q^T \mid M_{w,b}\right)\right) \tag{44}$$

$$\sigma^2\left(\lambda_i^{SIS}\left(Q > Q^T \mid M_{w,b}\right)\right)\ \ = \frac{\left(\lambda_i\left(M_{w,b}\right)\right)^2}{N\left(M_{w,b}\right)} \sum_{e \in E} \left(\left[\ 1_{(Q(e)>QT)} \emptyset_{b,i}^{SIS}(e) - P_{b,i,T}\right]^2 w_{b,i}^{SIS}(e)\right) \tag{45}$$

$$P_{b,i,T}\ \ = \frac{\sum_{e \in E} r_i(e)\ 1_{(Q(e)>QT)}}{\sum_{e \in E} r_i(e)} \tag{46}$$

The central argument is that if numerous independent estimates of the Monte Carlo exceedance rate (equation 39) are generated through repeated sampling, the variance of these estimates will converge to that described by equation 44 as the number of repetitions increases. Given that equation 39 is unbiased, the variance of the Monte Carlo exceedance rates is equivalent to the variance of the Monte Carlo error, which has a mean of zero.

The variance of the Monte Carlo exceedance rate (equation 44) can be calculated to determine whether *E* represents the entire set of scenarios in the PTHA or a specific subset, such as a particular source zone. When the results are computed separately for distinct subsets, the variance of their combined results can be determined by summing the individual variances. It is relatively straightforward to compute equation 44 at a selected offshore hazard point, as exemplified by Davies and Griffin (2018).

To select the relevant subset scenarios for Mw 8.7, the 1 m wave height was set as the defined threshold, *QT*, following the minimum wave category set in the Indonesian Tsunami Early Warning System (InaTEWS) to trigger an evacuation warning (Jihad et al., 2023). Finally, to ensure that the selected samples sufficiently represent variability across all scenarios, the exceedance rate curves from equation 39 were compared with the median exceedance rate from all scenarios. To quantify the accuracy of this comparison, Davies et al. (2022) suggested that the confidence interval should be ~95%. The confidence interval can be approximated by:

$$\lambda_i^{SIS}(Q > Q^T) \pm 1.96\sqrt{\sigma^2(\lambda_i^{SS}(Q > Q^T))} \tag{47}$$

Initially, all magnitudes available in the PTHA18 database were sampled and trial and error was used to determine the optimal number of scenarios. Repetitive sampling was performed, with 10,000 repetitions for the sample sizes of 300, 600, and 1200 scenarios. The approximated confidence interval improved from 93.71% (300 samples) to 94.21% (600 samples) and further increased to 94.59% (1200 samples). Fig. 4.8 illustrates the comparison of exceedance rates computed by equation 39 against the median of all scenarios.

4.5 Optimal Sample for Mw 8.7 Hypothetical Scenarios

A total of 600 scenarios, *Ntot*, were selected for the sampling strategy because the computed confidence levels for the total scenarios of 600 and 1,200 were nearly similar. This decision also considered computational limitations. It should be noted that this total of 600 scenarios was to sample all magnitudes, ranging from Mw 7.2 to Mw 9.6. The main difference between stratification by magnitude and stratification by

importance is that the sample size for each magnitude bin is not identical (Davies et al., 2022).

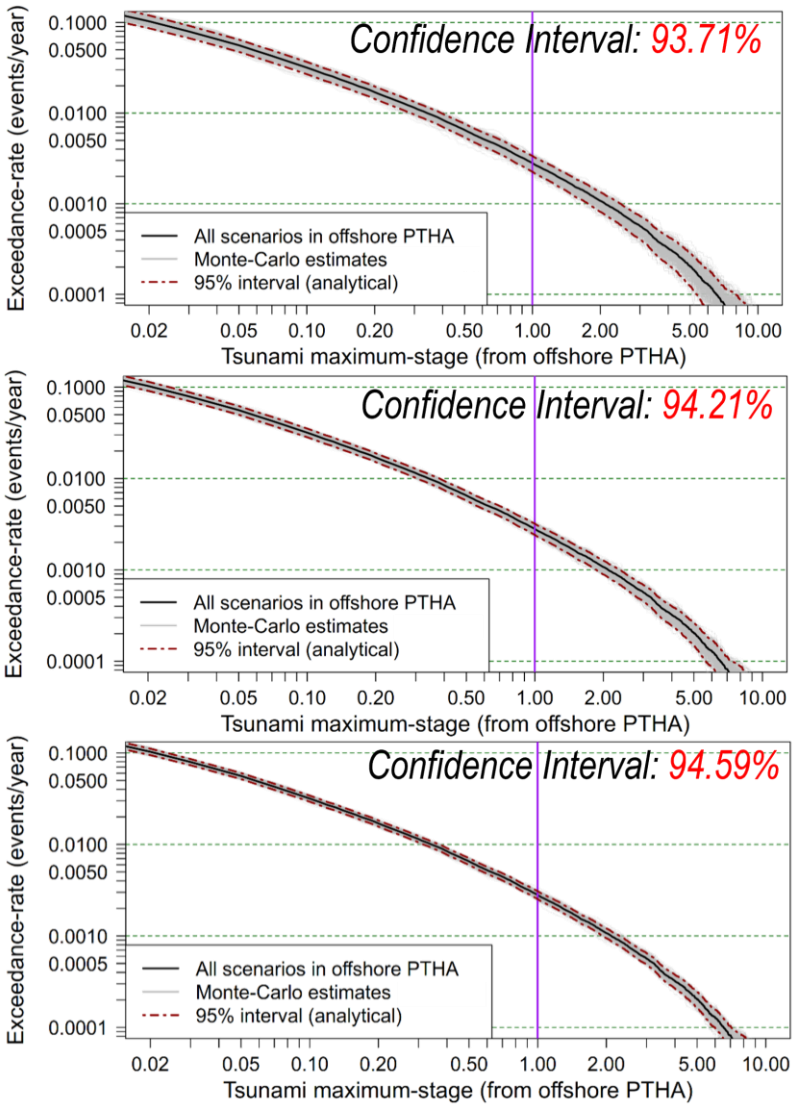


Fig. 4.8. Stratified importance sampling and their corresponding confidence intervals, with total scenarios of (top panel) 300, (middle panel) 600, and (bottom panel) 1200, respectively.

The optimal sample size for each magnitude bin, N_i, which can minimize the variance of the Monte Carlo errors (equation 39)–for any given N_{tot}, scenario-frequency model i, and threshold QT–can be computed by:

$$N_{tot} = \sum_{M_{w,b} \in magnitue\ bins} N(M_{w,b}) \tag{48}$$

$$N_i(M_{w,b} \mid Q^T) = \frac{N_{tot}\sqrt{\alpha_i(M_{w,b} \mid Q^T)}}{\sum_{M_{w,b} \in magnitue\ bins}\sqrt{\alpha_i(M_{w,b} \mid Q^T)}} \tag{49}$$

$$\alpha_i(M_{w,b} \mid Q^T) = \left(\lambda_i(M_{w,b})\right)^2 \left(\sum_{e \in E}\left(\left[1_{(Q(e)>QT)}\emptyset_{b,i}^{SIS}(e) - P_{b,i,T}\right]^2 w_{b,i}^{SIS}(e)\right)\right) \tag{50}$$

The calculated optimal sample for each magnitude bin is shown in Fig. 4.9, with 40 scenarios selected for the Mw 8.7. Fig. 4.10 shows the initial sea surface deformation for the selected scenarios. The stratified importance sampling and extraction of initial sea surface deformation from the selected scenarios were performed using codes provided by Davies et al. (2022). A link to access the codes is provided in the Appendix.

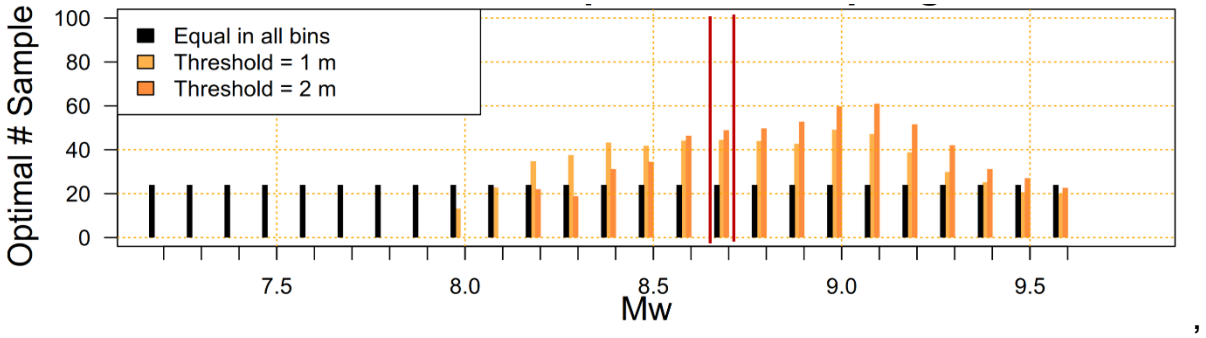


Fig. 4.9. Optimal sample for each magnitude bin. The black bars indicate equal sample sizes for each magnitude bin, given a total scenario of 600. The coloured bars show optimal sample sizes calculated for two different tsunami height thresholds, *QT*, demonstrating the variation in sampling strategy based on threshold selection. The analysis focused on a threshold of 1 m.

4.6 Tsunami Inundation Simulation for Mw 8.7 Scenarios

Inundation simulations were performed using the initial sea surface deformation information from the selected scenarios. The same numerical model configuration used to reproduce the 2004 IOT was employed (see Fig. 3.8). DiluviumDEM was set as the elevation model and the LCR model at 5 m resolution was used for the roughness model. To represent current terrain features, the LCR model was derived from the 2024 land cover map (see Fig. 2.15c), with Manning coefficients based on Bunya et al. (2010). In total, 40 scenarios were simulated, each representing two hours of tsunami conditions.

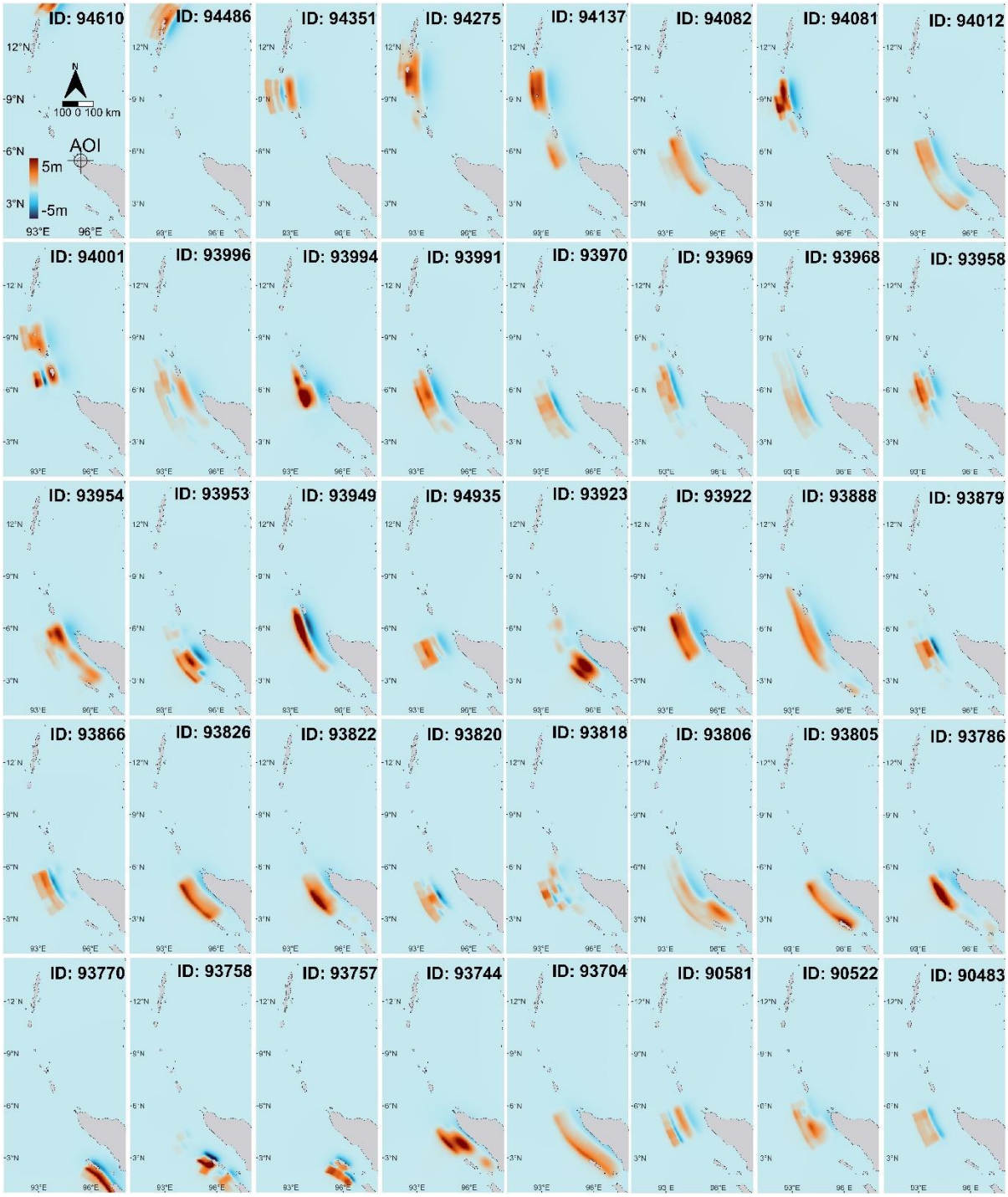


Fig. 4.10. Initial sea surface deformation for selected scenarios of Mw 8.7. The colour bar indicates the water surface elevation (in m).

The numerical simulation resulted in 40 inundation maps at 10 m resolution. To quantify the uncertainty in inundation extent generated by Mw 8.7 earthquakes, a percentile-based aggregation approach was applied. Each inundation map was stacked into a spatial ensemble, and the inundation values at each grid cell were statistically analysed. For every pixel location across the study area, the 16th, 50th

(median), and 84th percentiles of inundation depth were computed from the distribution of the 40 simulations. The 16th percentile provides a conservative estimate, indicating inundation depths that exceed 84% of the scenarios, which is useful for lower-bound hazard assessments. The 50th percentile (median) represents the central tendency and offers a typical expected inundation extent. The 84th percentile, in contrast, reflects an upper-bound estimate, indicating that inundation depths exceeded in only 16% of the scenarios, which helps identify worst-case conditions for risk planning.

This method accounts for variability across simulations while avoiding assumptions about underlying probability distributions, making it robust for decision-making in tsunami hazard mitigation. The resulting probabilistic maps allowed for spatially explicit hazard quantification, supporting differentiated risk assessments across coastal zones (Fig. 4.11).

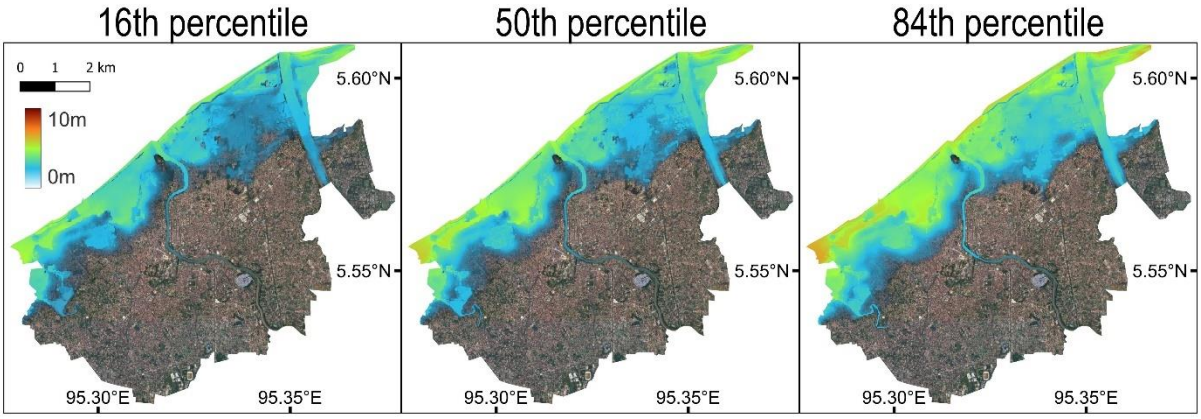


Fig. 4.11. Probabilistic tsunami inundation maps of Mw 8.7 scenarios showing the 16th, 50th, and 84th percentiles of maximum inundation depth.

4.7. Cross-uncertainty Population Exposure Estimates

The resulting probabilistic maps only account for uncertainty owing to variability in the tsunami sources (Grezio et al., 2017). However, the bias introduced by the elevation model has not been explicitly addressed. Given this, the probabilistic inundation maps were integrated with confidence levels derived from the elevation model used (see Section 3.5). First, the median inundation depth for each raster map was calculated (Fig. 4.12a-4.12c). Using the RMSE of DiluviumDEM (i.e., 1.18 m), the confidence level for each raster pixel was computed (Fig. 4.12d-4.12f). Subsequently, the raster

pixels with confidence levels higher than 68% and 95% were selected. Alternatively, using equations 31 and 32, the raster pixels with inundation depths less than 2.36 m for the 68% confidence level and less than 4.63 m for the 95% confidence level could be directly excluded. This process resulted in cross-uncertainty inundation maps, as illustrated in Fig. 4.12g to 4.12l.

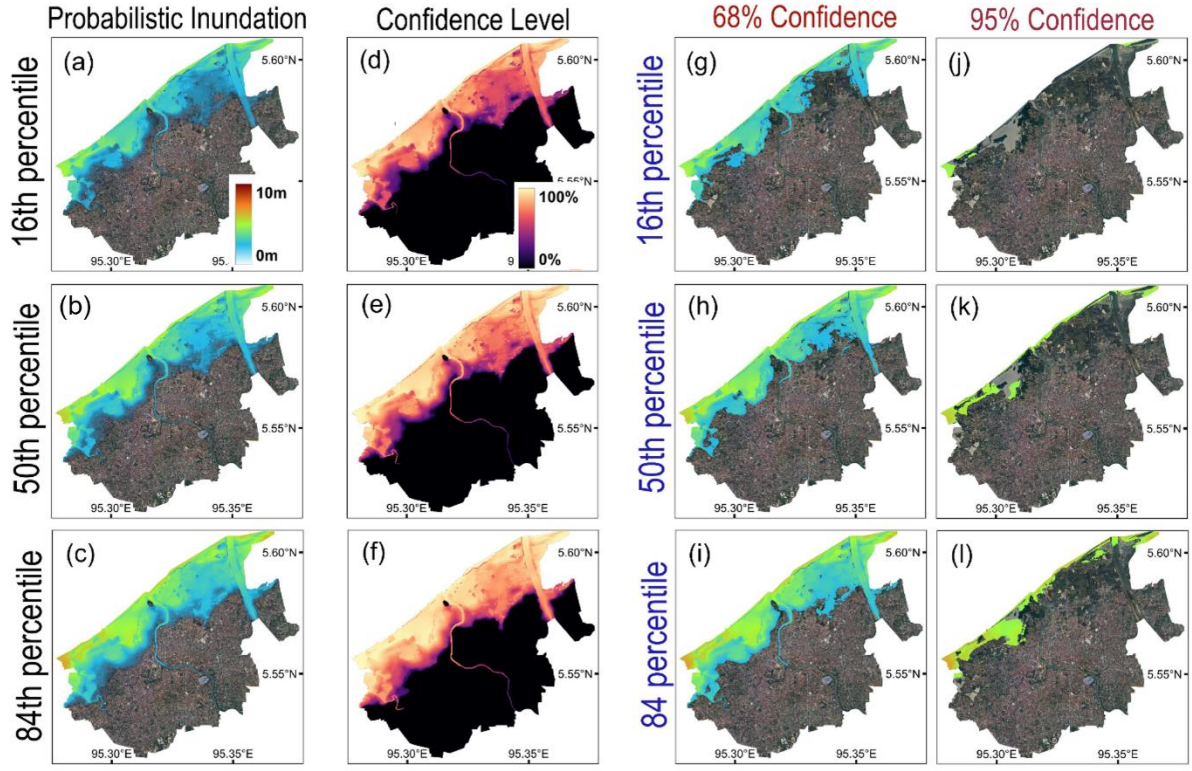


Fig. 4.12. (a-c) Probabilistic inundation maps. (d-f) Computed confidence levels for corresponding probabilistic inundation maps. (g-l). Cross-uncertainty inundation maps show the 16th, 50th, and 84th percentiles of the maximum inundation depth, with confidence levels of 68% and 95%, respectively.

Using the cross-uncertainty inundation maps and generated gridded population for 2024 (POP-24; see Fig. 2.16), the estimated number of Banda Aceh's population affected by the Mw 8.7 tsunami was calculated. The exposure assessment projected that the population potentially exposed to a tsunami of Mw 8.7 tsunami would range from a minimum of 563 people to a maximum of 36,306 people—14% of the total population (Fig. 4.13). Because the inundation maps with 95% confidence resulted in less inundation, they ultimately projected lower exposure estimates of 0.2% to 1.6% of the total population. In contrast, the inundation maps with 68% confidence projected much higher exposure estimates of 6–14% of the total population. For disaster risk preparedness, the median percentile (50th percentile) with a 95% confidence level

served as the lower-bound credible exposure estimates (i.e. 1,998 people) and the 68% confidence level as the upper-bound credible exposure estimates (i.e. 24,696 people).

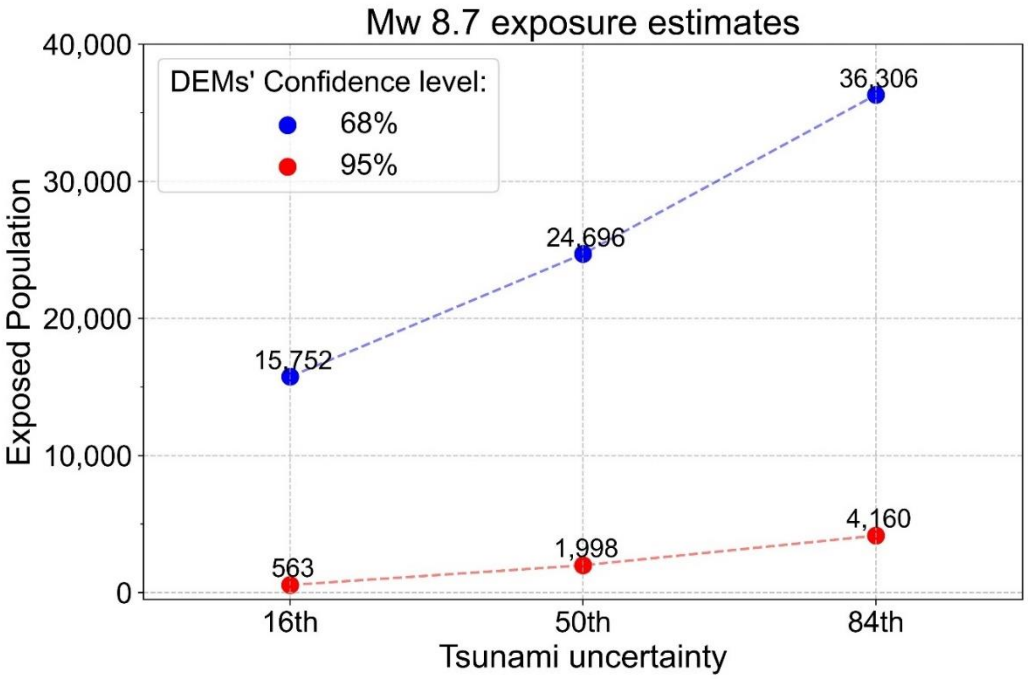


Fig. 4.13. The projection of Banda Aceh's population exposed to the Mw 8.7 tsunami. The red line serves as a lower-bound credible estimates and the blue line indicates an upper-bound credible estimates.

4.8. Summary of Chapter IV

In summary, this chapter's analysis revealed that combining global elevation models with exposure datasets led to larger bias to the exposure assessment. The results showed that elevation models introduced a larger bias than exposure datasets alone. Additionally, the cross-uncertainty inundation assessment for a Mw 8.7 tsunami scenario demonstrated that approximately 10% of the current population of Banda Aceh faces potential tsunami exposure.

CHAPTER V:
EXPOSURE AND MITIGATION SYSTEM

5.1. Evacuation System: Tsunami Shelter Capacity Vs. Exposed Population

The area of interest (AOI) is not equipped with tsunami coastal defense structure, such as sea wall (Syamsidik et al., 2019). As a result, this city relies on evacuation system to mitigate the tsunami risk. Five tsunami escape buildings were located in coastal-front region, as shown in Fig. 5.1. Murao et al. (2025) found that the effective capacity of these shelters are 5,560. This already included the use of space at roof.

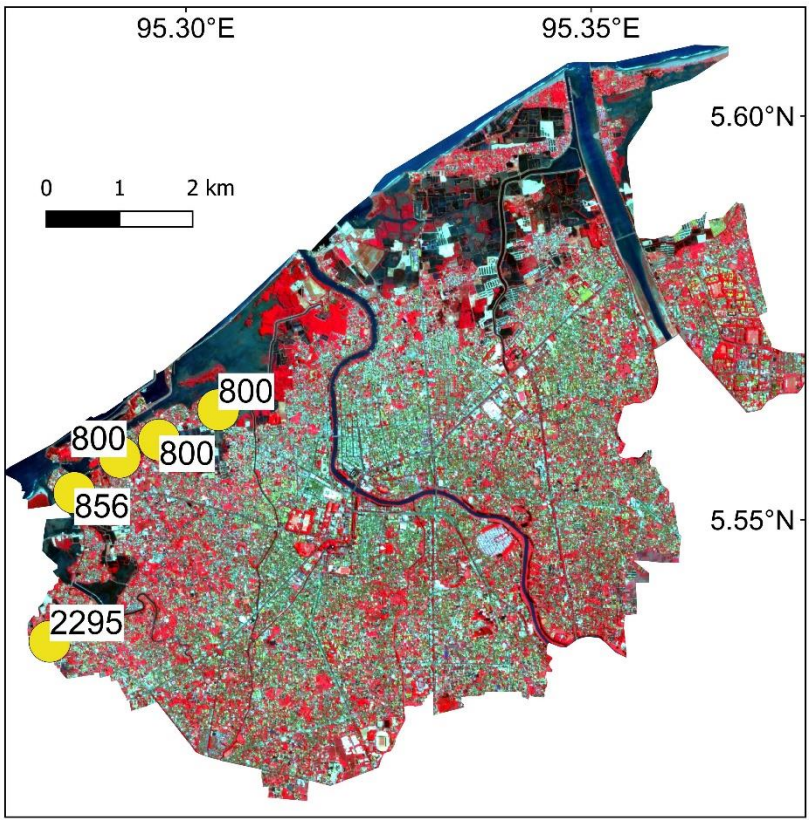


Fig. 5.1. Location of tsunami escape buildings (yellow circles) and their capacity.

Using the median lower-bound credible population exposure estimates, the analysis indicates that if the Mw 8.7 tsunami occurred in AOI, then the existing shelters could accommodate all 1,998 evacuees. The evacuees only occupied 36% of total shelters capacity. In contrast, when using the median upper-bound credible estimates, the existing shelters could only accommodate 23% of total evacuees, leaving 19,136 people unsheltered.

Nevertheless, it should be noted that these results merely compare projected exposed populations and shelter capacity. This study did not account for influence of evacuation behaviour, tsunami time arrival, route choice, and disaster knowledge. Despite these limitations, the analysis suggests that evacuation shelter capacity requires significant expansion to adequately serve the population under worst-case tsunami scenarios.

5.2. Spatial Planning: Spatial Land use Vs. Building Dynamics

As discussed in section 1.3, following the 2004 Indian Ocean Tsunami (IOT), local authority introduced a physical zoning system to regulate urban development in tsunami-prone areas (Fig. 5.2a). This zonation framework was subsequently integrated into spatial land use planning, incorporating historical hazard data while maintaining flexibility for controlled urban development (Fig. 5.2b). The land use planning itself was introduced in 2009 and later experienced some minor modification in 2017 (Banda Aceh Municipality, 2018). This study classified the land use planning into building and non-building areas.

The land use planning is also designed as a soft-mitigation strategy for tsunami disaster in the AOI. This such strategy is also highlighted in the Sendai Framework for Disaster Risk Reduction 2015-2030, especially for tsunami-prone regions with significant financial constraints, such as Indonesia and Chile (Takabatake, 2022; UNDDR, 2015). The integration of land-use planning with tsunami hazard mapping results in strategic urban spatial planning, assisting the planners in developing coastal cities while minimizing exposure to coastal hazards (Geiß et al., 2024; Rafliana et al., 2022; Vicuña et al., 2022).

In this context, monitoring urban evolution is essential to ensure the implementation of land-use planning, particularly where urban expansion cannot be regulated (León et al., 2022). Spatiotemporal building data are one of the key elements to characterize the evolution of built environments (Fuchs et al., 2015). To understand the degree to which the land use planning and temporal urban evolution have been consistent, the designated land use planning was compared with building evolution from 2018-2024, since the land use map was last modified in 2017.

However, the availability of authoritative spatiotemporal building data remains limited to high-income countries (Chamberlain et al., 2024). As an alternative, urban analysis in data-scarce regions largely relies on globally open building datasets, including the

volunteered geographic information (VGI) of OpenStreetMap (OSM) and published data from commercial companies, such as Microsoft's Global ML Building footprints (Microsoft, 2024) and Google Open Buildings Polygons (Sirko et al., 2021). While open building footprints could fill the gap where authoritative data are unavailable or incomplete, several issues regarding the quality of these datasets prevail (Herfort et al., 2023; Zhou et al., 2022).

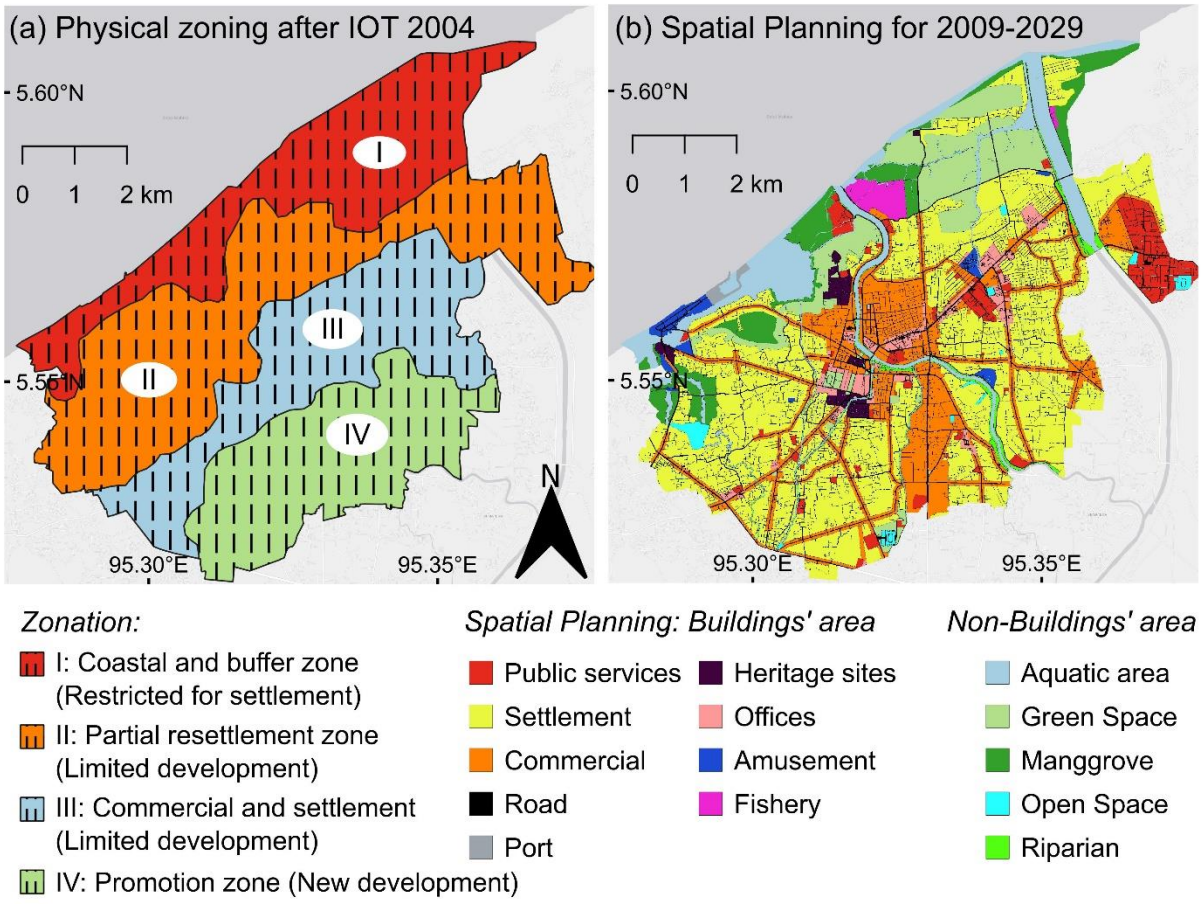


Fig. 5.2. (a) Physical zoning, and (b) Spatial land use planning map of 2009-2029 (Basemap: ESRI Light Gray).

OSM dataset is produced and edited through manual digitization performed by thousands of individual volunteered mappers and contributing organizations. Consequently, its accuracy and completeness are geographically varied (equation 53) (Chamberlain et al., 2024). Europe and North America regions have relatively higher levels of completeness (Herfort et al., 2023), while tsunami-prone regions, such as Chile and Indonesia, demonstrated lower levels of completeness by 32% and 26%, respectively (Zhou et al., 2022).

The temporal information of OSM building footprints can be derived using the OpenStreetMap History Database (OSHDB) framework (Raifer et al., 2019). OSHDB analysed the OSM's full database using object modification attribute information, e.g., added, removed, or modified. Given this, it may not completely record the building inventory for certain periods due to insufficient volunteered building mapping activity. For example, the addition of building footprints into OSM database stagnated in 2019 and resumed an increase in 2021 after the COVID-19 pandemic (Herfort et al., 2023). It should be noted that incomplete spatiotemporal datasets might mislead the risk evolution assessment (Fuchs et al., 2015). Meanwhile, both Microsoft and Google building datasets ─derived from automated feature identification based on machine-learning techniques─ are not accompanied by temporal attributes. This could limit their applicability to risk evolution analyses.

To develop spatiotemporal building data for the 2018-2024 period within the AOI, a simple backdating technique was applied. The 2024 building dataset served as the foundation for this analysis, compiled from OSM and Microsoft's Global ML Building footprints (Table 5.1), yielding 82,563 and 59,290 building footprints respectively. The Google Open Building dataset was excluded due to its limited coverage within the AOI region.

The backdating technique, also referred to as rapid backdating footprints generation (RBF), operates on the assumption that buildings are located within built-up areas. Built-up masks derived from satellite imagery were therefore used to identify which buildings from the 2024 dataset would have existed in 2018 and 2021, enabling the reconstruction of historical building distributions for these earlier time periods.

To evaluate the backdating performance, the backdating-generated footprints were compared against ground truth (GT) polygons. These GT polygons were derived from on-screen digitization of three historical images in Google Earth Pro (see Table 5.1). Several accuracy metrics were employed, including intersection over union (*IoU*), agreement rate, and completeness (equations 51-53) (see Fig. 5.3 for symbol denotations). Conversely, the building count was not included as an accuracy indicator because of inconsistent feature representations across datasets. For example, structures such as terraced housing can be mapped as either single or multiple polygon features (Herfort et al., 2023).

$$IoU = \frac{TP}{TP + FP + FN} \tag{51}$$

$$Aggreement = \frac{TP}{TP + FN}\ x\ 100\% \tag{52}$$

$$Completeness\ (\%) = \frac{Area\ Predicted}{Area\ Ground\ truth}\ x100\% \tag{53}$$

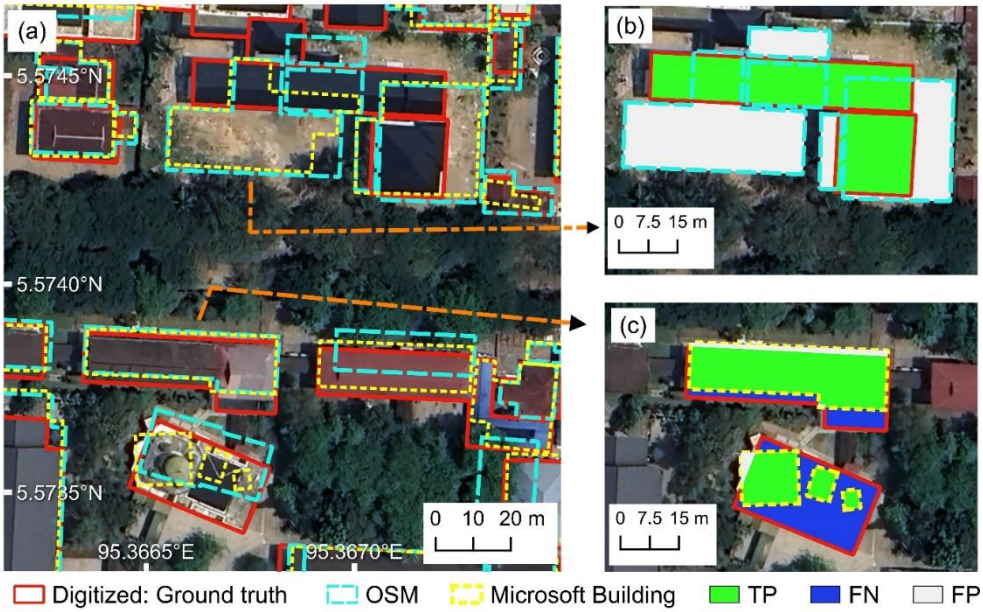


Fig. 5.3. (a) Overlayer of all building datasets. (b and c). Illustrations of true positive (TP), false positive (FP) and false negative (FN) relative to GT polygons. Background: Google Earth Pro Imagery.

Table 5.1. Input for built-up areas and backdating building footprint generation

Date	Type	Description
12 June 2018	Optical image	Dove Classic PS2, scene's ID: 20180612_033219_1015_3B 20180612_033220_1015_3B 20180612_033221_1015_3B
08 August 2021		SuperDove PSB.SD, scene's ID: 20210807_031345_46_2450_3B
15 June 2018	Building footprints	OSM-OSHDB, polygon counts: AOI: 66,951; ROI: 33,716
10 August 2021		AOI: 68,087; ROI: 33,880
01 July 2024		AOI: 82,563; ROI: 41,050
2016 – 2020		Microsoft's Global ML Buildings (MS) AOI: 59,290; ROI: 28,612
August 2018	Ground truth polygons (GT)	Google Earth Pro Imagery, counts: Building polygons: 44,144 (ROI)
September 2021		Building polygons: 48,402 (ROI)
March 2024		Building polygons: 50,939 (ROI)

5.2.1 Spatiotemporal Built-up Masks (2018-2024)

Given that built-up area information was already available for 2024 (see Fig. 2.16), built-up area maps were only developed for 2018 and 2021 periods. To obtain this information, the land cover classification was re-performed using the OBIA-RF method as described in section 2.5.1.d. Multitemporal four-band imagery of PlanetScope with 3 m geometric resolution was used (Table 5.1), retrieved via Planet Explorer (www.planet.com/explorer). All scenes were cloud-free and corresponded to orthorectified analytic surface reflectance products.

To increase the radiometric consistency of PlanetScope images from two different sensors (i.e., Dove Classic-PS2 and SuperDove-PSB.SD), all scenes were normalized to match Sentinel-2 spectral response. Spectral normalization was performed using the harmonize operation available in Planet Explorer (Planet, 2024). Fig. 5.4 shows the input image datasets and ground truth labelled data.

The spatial misalignment between PlanetScope multitemporal images was less than 10 m (Planet, 2023), with Leach et al. (2019) consistently finding that the spatial error was approximately 6 m. Given this accuracy, spatial co-registration was not performed because the geolocation accuracy was already considered sufficient for change detection analysis (Wegmueller et al., 2021).

The land cover classification for 2018 and 2021 achieved overall accuracies of 92.80% and 92.10%, respectively, with an average F1 score of 0.99 (Table 5.2). The lowest F1 score was observed in the bare class, with scores below 0.7 across all temporal variations. This lower accuracy might be attributed to misclassification between bare and built-up area pixels (Ettehadi Osgouei et al., 2019). Meanwhile, the classification accuracy of built-up areas was excellent, with F1 scores of 0.97 for 2018 and 0.96 for 2021. The 2018 and 2021 land cover maps were resampled to 5 m spatial resolution to ensure consistency with the 2024 land cover map (Fig. 5.5).

The land cover class-area estimates revealed that built-up areas were the dominant land cover type in BNA, covering approximately 62.89% (37.08 km²) of the BNA region. Built-up areas demonstrated a clear expansion trend between 2018-2024 (2018: 28.91 km²; 2021: 32.70 km²). This expansion subsequently caused a decline in other land classes, with a significant 44% decrease observed in the vegetation class. The

vegetation area estimate was 16.6 km² in 2018, decreased to 14.36 km² in 2021, and dropped to 9.30 km² by 2024.

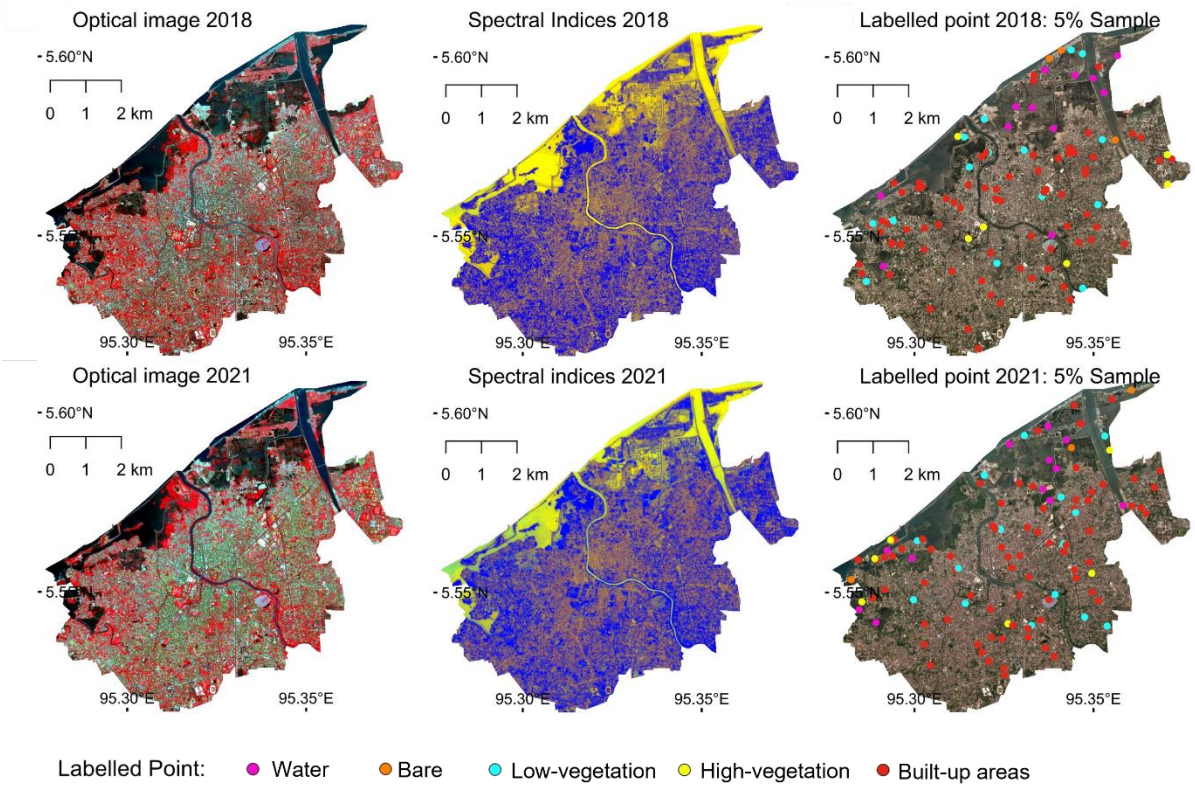


Fig. 5.4. Input datasets for land cover classification in 2018 and 2021.

Table 5.2. Land cover fraction, confusion matrix and classification accuracy

LC-18. Class-area estimates (in km²)—I: 9.46; II: 0.81; III: 16.60; IV: 3.17; V: 28.91											
		Truth Label					Accuracy Metrics				
*Land class**		*I*	*II*	*III*	*IV*	*V*	P	R	F1	Overall Accuracy	Kappa Index
Predicted Label	*I*	*128*	*0*	*1*	*0*	*1*	0.99	0.98	0.99	92.80%	0.88
	II	*0*	*11*	*1*	*0*	*8*	0.69	0.55	0.61		
	III	*1*	*1*	*129*	*20*	*10*	0.87	0.8	0.83		
	IV	*0*	*0*	*10*	*43*	*0*	0.68	0.81	0.74		
	V	*0*	*4*	*7*	*0*	*510*	0.96	0.98	0.97		
LC-21. Class-area estimates (in km²)—I: 9.07; II: 0.53; III: 14.36; IV: 2.29; V: 32.70											
		Truth Label					Accuracy Metrics				
Land class		*I*	*II*	*III*	*IV*	*V*	P	R	F1	Overall Accuracy	Kappa Index
Predicted Label	*I*	*121*	*1*	*2*	*0*	*0*	0.98	0.98	0.98	92.10%	0.87
	II	*0*	*22*	*3*	*0*	*15*	0.79	0.55	0.65		
	III	*3*	*2*	*113*	*9*	*17*	0.84	0.78	0.81		
	IV	*0*	*0*	*12*	*57*	*0*	0.86	0.83	0.84		
	V	*0*	*3*	*4*	*0*	*519*	0.94	0.99	0.96		

** I: water; II: bare; III: low-vegetation; IV: high-vegetation, and V: built-up areas*

Although less pronounced compared to vegetation, the inland water class also exhibited a declining pattern by approximately 1% every three years (2018: 9.46 km^2, 2021: 9.07 km^2, and 2024: 8.31 km^2). The decline may be associated with land reclamation for urban development, which simultaneously increasing the bare class estimates (see Fig. 5.5). By 2024, the bare areas rose to 128 ha, increased about 47 ha than in 2018. Meanwhile, the forestry areas (high-vegetation) exhibited a fluctuating trend: a decrease of 88 ha from 2018-2021, then succeeded by a 70 ha increase between 2021-2024. These temporal variations may be attributed to two factors. First, misclassification of forestry land cover in 2018, as indicated by a low F1 score of 0.74. Second, the observed expansion between 2021-2024 might result from mangrove development in wetland areas.

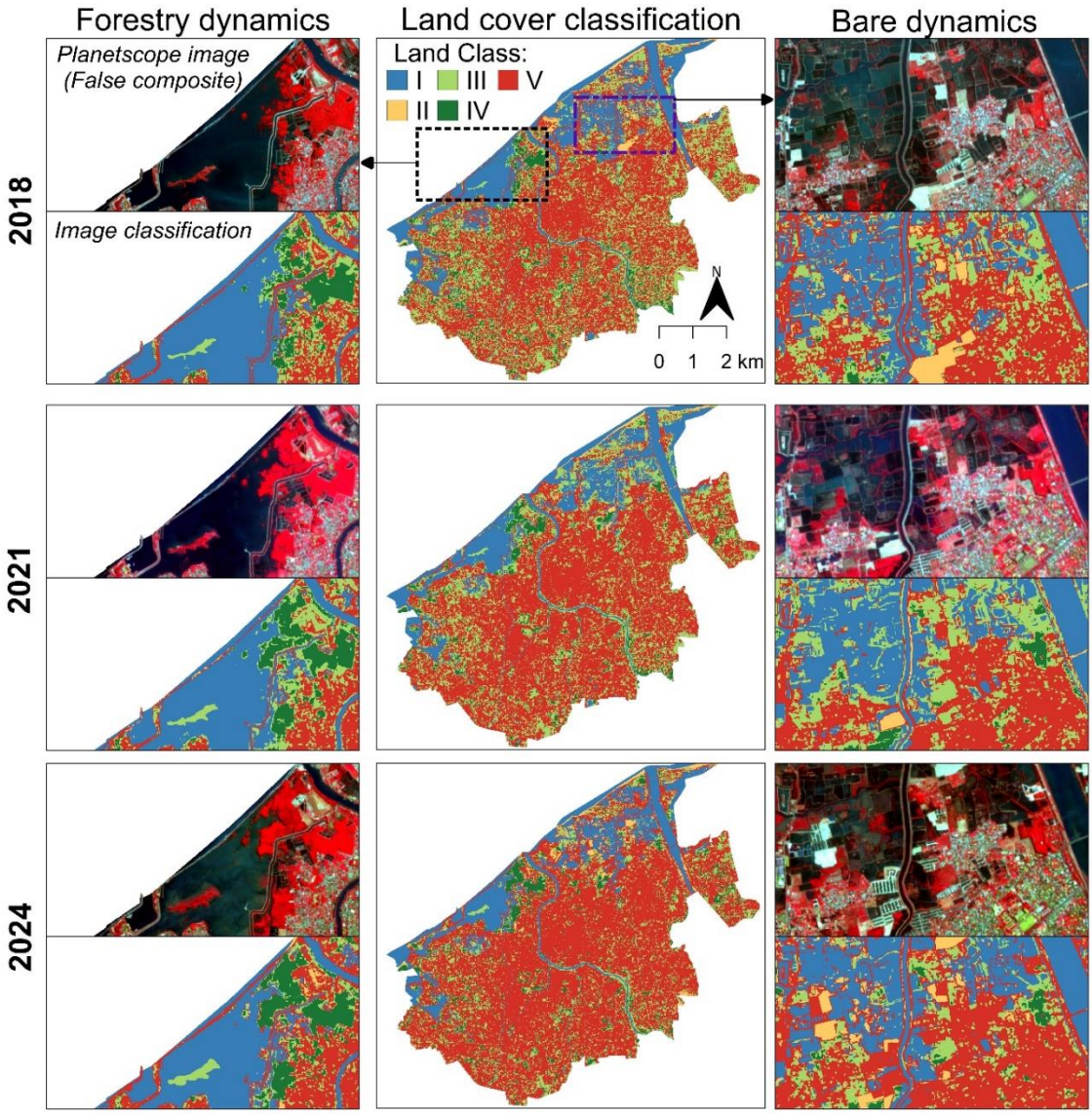


Fig. 5.5. Land cover maps of 2018-2024 and subsets showing dynamics of bare and forestry areas.

5.2.2 Spatiotemporal Building Footprints (2018-2024)

More than 50,000 GT polygons were digitized, covering an area of 5.54 km² for 2018, 5.99 km² for 2021, and 6.21 km² for 2024. Given the labor-intensive process of manual digitization, GT polygons only covered four subdistricts (subdistric ID: 1-4). These areas were designated as the region of interest (ROI). To ensure consistency with GT polygons, the backdating technique was initially performed exclusively within the ROI coverage. Three inputs served as contemporary building datasets: OSHDB 2024 (OSM), Microsoft's Global ML Buildings (MS), and a combination of both (OSMMS). To develop the OSMMS dataset, MS polygons that did not intersect with OSHDB 2024 were inspected. The non-intersecting polygons were then added to the OSHDB 2024 dataset.

All datasets were clipped to the region of interest (ROI) using administrative polygons retrieved from the Global Administrative Area Database (GADM) (https://gadm.org). To examine the impact of input data quality on the backdating method, GT 2024 was included as an additional input dataset. Based on accuracy assessment within the ROI, the most accurate dataset for each epoch was selected and the backdating approach was reapplied within the entire BNA region. The accuracy of backdating-generated buildings was also compared with spatiotemporal datasets derived from the OSHDB framework (see Table 5.1). The detailed workflow is illustrated in Fig. 5.6.

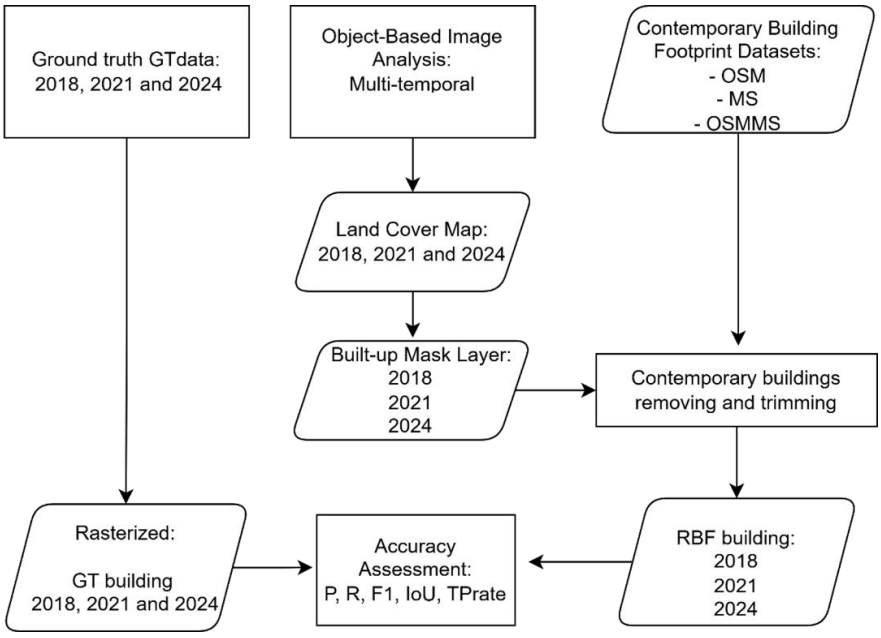


Fig. 5.6. Detailed pipeline of building backdating technique or also referred to as rapid backdating footprints generation (RBF)

The accuracy assessments showed that the agreement between validation polygons and OSM footprints was less than 60% for 2018 and 2021, with completeness levels of 80.12% and 75.24%, respectively (Table 5.3). However, completeness reached 90.19% by 2024, followed by an increase in agreement levels (74.85%) and *IoU* (0.65). The fluctuating level of completeness (i.e., decreasing in 2021 and a substantial rise in 2024) might reflect the volunteered mapping activity during and after the COVID-19 pandemic, as also reported by Herfort et al. (2023).

The increasing coverage of validation polygons aligned with built-up expansion within the ROI (Fig. 5.7). Overall, the 2018 and 2021 backdating generated buildings showed a better agreement with validation data than those in 2024. This suggests that the backdating approach is more effective in reconstructing historical data than the contemporary dataset.

Table 5.3. Accuracy assessment for generated building footprints. BU: Built-up area layer

Input	Mask	Output	Validation	IoU	Agreement	Completeness
OSHDB 2018	-	OSM 2018	GT 2018	0.46	57.16%	80.12%
OSHDB 2021	-	OSM 2021	GT 2021	0.46	55.80%	75.34%
OSHDB 2024	-	OSM 2024	GT 2024	0.65	74.85%	90.19%
GT 2024	BU-18	RBF-GT 2018	GT 2018	0.88	91.79%	95.98%
	BU-21	RBF-GT 2018	GT 2021	0.94	96.19%	98.16%
	BU-24	RBF-GT 2024	GT 2024	0.98	97.58%	97.52%
OSHDB (OSM) 2024	BU-18	RBF-OSM 2018	GT 2018	0.65	74.45%	89.69%
	BU-21	RBF-OSM 2021	GT 2021	0.66	74.86%	88.96%
	BU-24	RBF-OSM 2024	GT 2024	0.64	73.23%	87.09%
MS	BU-18	RBF-MS 2018	GT 2018	0.61	70.10%	84.44%
	BU-21	RBF-MS 2021	GT 2021	0.60	67.66%	81.15%
	BU-24	RBF-MS 2024	GT 2024	0.58	66.14%	79.26%
OSMMS	BU-18	RBF-OSMMS 2018	GT 2018	0.64	75.29%	92.49%
	BU-21	RBF-OSMMS 2021	GT 2021	0.65	75.27%	90.93%
	BU-24	RBF-OSMMS 2024	GT 2024	0.63	72.74%	88.96%

The completeness of backdating-generated buildings for 2018 and 2021 was higher than that of OSM footprints (i.e., OSHDB 2018 and 2021). Furthermore, Microsoft-based datasets (RBF-MS) had lower completeness than OSM-based data (RBF-OSM) (Fig. 5.8). This could be associated with the research location, as Microsoft's building data typically have better coverage for major cities (Chamberlain et al., 2024). The combination of OSM and Microsoft datasets (RBF-OSMMS) did not significantly

improve the completeness level of generated datasets (e.g., RBF-OSM 2018: 74.45%; RBF-OSMMS 2018: 75.29%). In addition, the performance of the backdating approach was observed to be influenced by the quality of the input dataset, as RBF-GT demonstrated more than 90% agreement across all periods. Given the accuracy assessment, OSMMS was selected as the input dataset. The backdating approach was then performed for the entire AOI region. For 2024, rather than using backdating-generated buildings, the original OSM 2024 data was used because of its high completeness level.

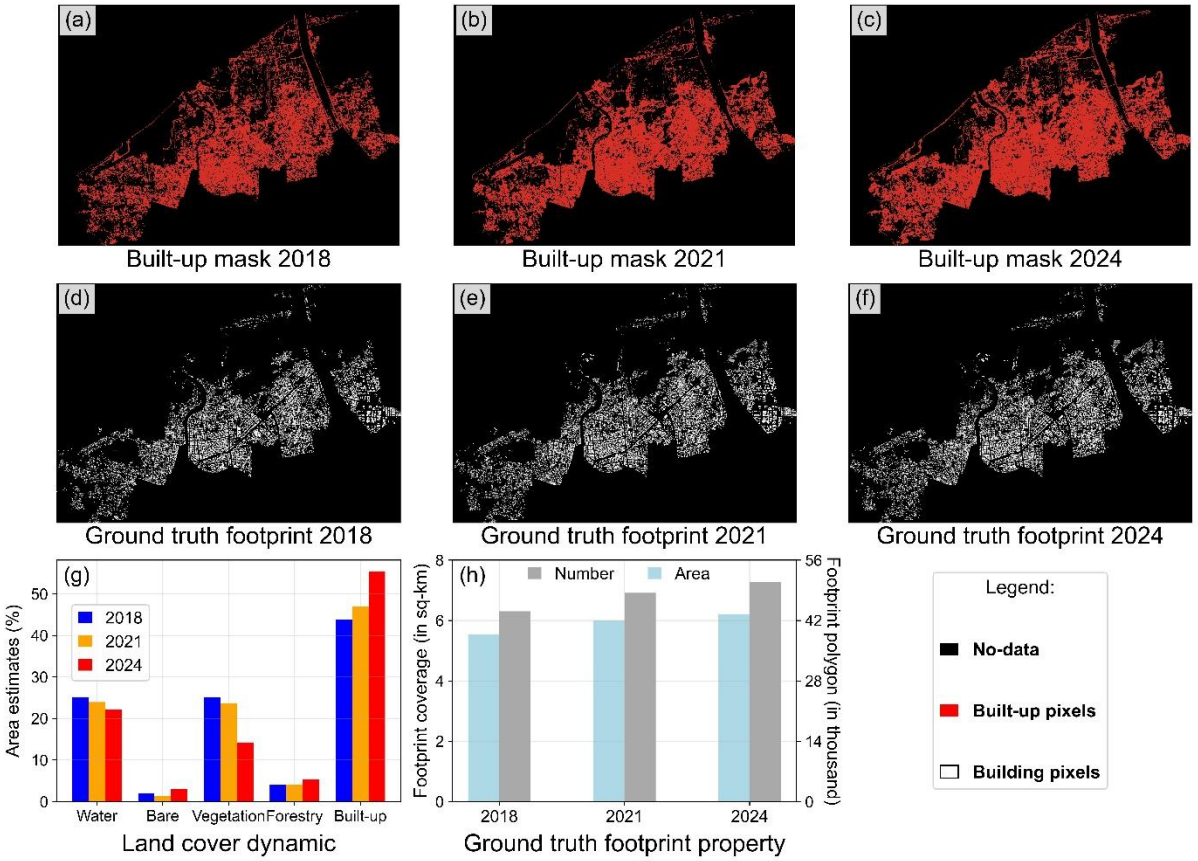


Fig. 5.7. (a-c) Multitemporal built-up masks. (d-f) Spatial distribution of GT polygons. (g) Land cover fraction within the ROI. (h) Properties of GT datasets.

Results showed that the building coverage reached 11.60 km^2 by 2024, increased from 11.40 km^2 in 2021 and 10.56 km^2 in 2018 (Fig. 5.9). This increase was consistent with the expansion of built-up areas. This highlights that built-up evolution could serve as an indicator for building dynamic behaviour. However, while built-up areas increased steadily at an average rate of 13% per three-year interval, building growth rates varied between intervals (2018-2021: 1.8%; 2021-2024: 8%).

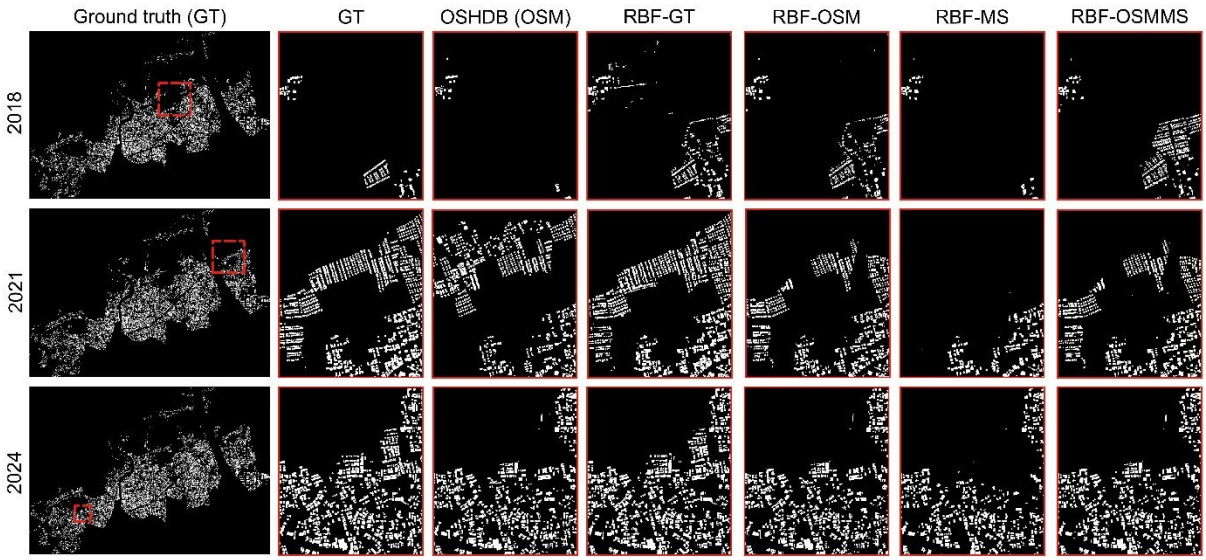


Fig. 5.8. Subset comparison of GT, OSM, and backdating building footprints within the ROI coverage.

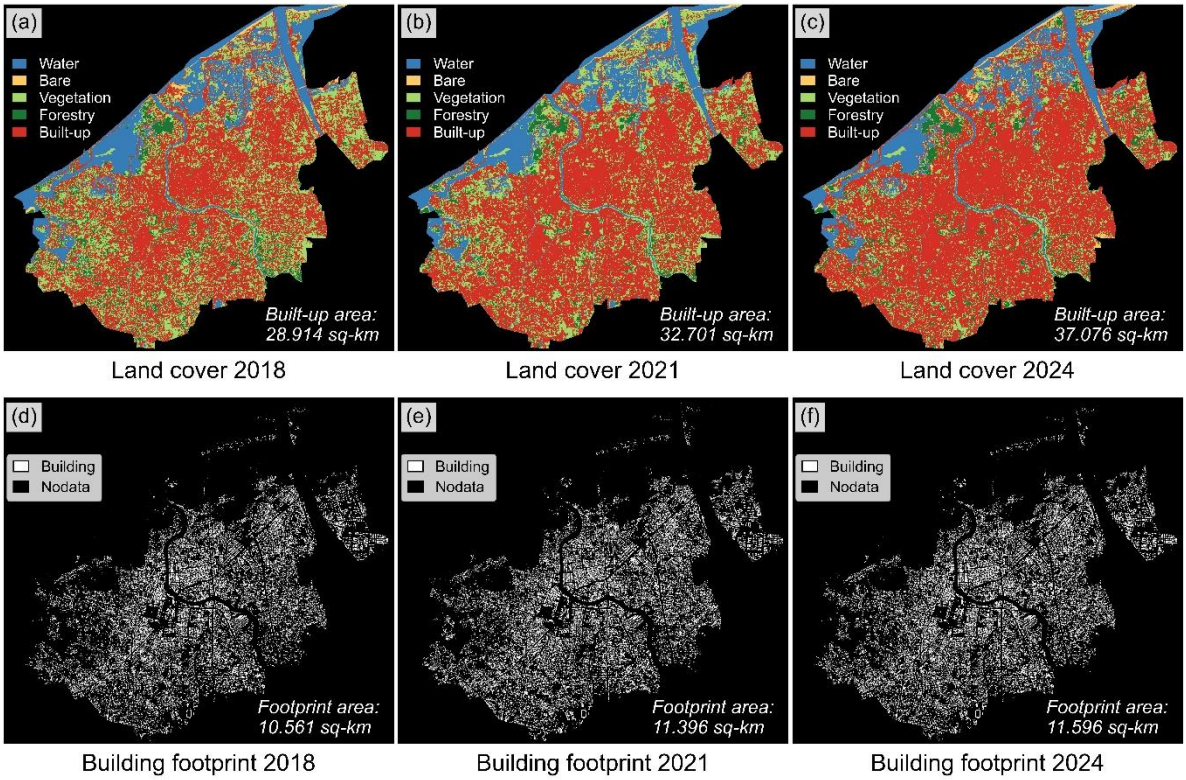


Fig. 5.9. The land cover maps and generated building footprints, respectively, within the AOI for periods: (a and d) 2018, (b and e) 2021, and (c and f) 2024.

5.2.3 Building Dynamics Assessment

Building dynamics were evaluated within three regions: the 2004 IOT inundation limit, physical zoning areas, and designated non-building zones based on land use planning. The Mw 8.7 tsunami inundation limit was not incorporated as an evaluation criterion

since the land use planning was designed for the worst-case tsunami scenario, which is the Mw 9.2 2004 IOT.

The analysis revealed that 56.77% of building stock in 2018 was located within the area affected by the 2004 IOT, with building coverage reaching nearly 6 km^2 or 600 ha (Fig. 5.10a). This estimate increased substantially by 44.72 ha in 2021 and followed by a modest expansion of 9.96 ha by 2024. A lower increase of 38.77 ha was observed in non-inundated areas by 2021. In contrast, by 2024, building development within this zone was slightly higher than that in former inundated areas (i.e., 10.03 ha). Despite this, building distribution from 2018 to 2024 exhibited a relatively similar pattern, where most buildings were concentrated within the 2004 IOT affected zone.

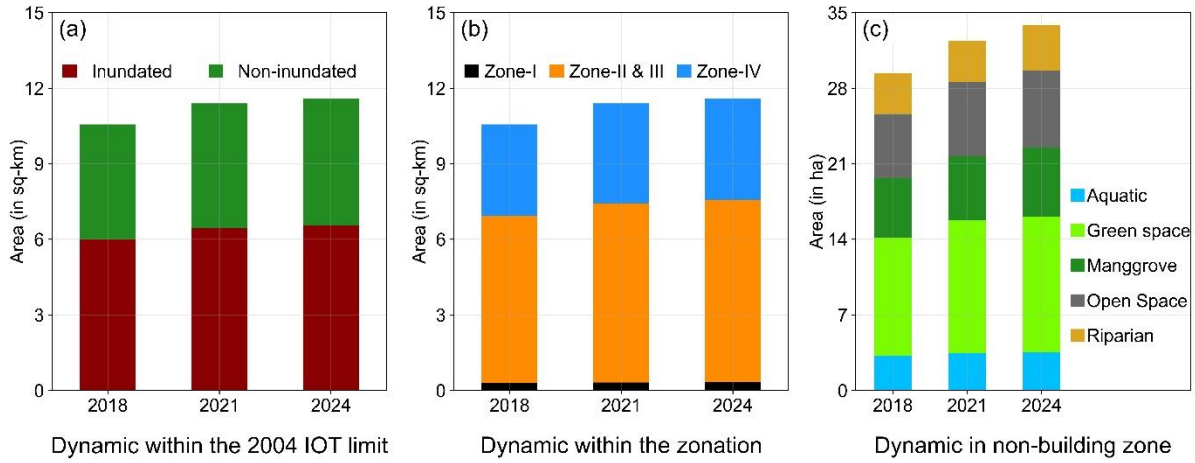


Fig. 5.10. Evolution assessment of building dynamics in the AOI within: (a) the 2004 IOT inundation limit, (b) physical zoning areas, and (c) non-building zones based on land use planning.

The 2004 IOT inundation extent encompassed entire restricted development area (Zone-I) and covered approximately 81% of limited development zone (Zone-II and III). Zone-I comprised a small number of buildings, approximately 3% of total buildings for each period. The building dynamics within this region were minimal, with building coverage increased by only 2.55 ha and 1.64 ha for 2021 and 2024, respectively. Meanwhile, Zone-II and III accounted for 63% of total buildings in 2018, or approximately 663.5 ha (Fig. 5.10b) . This number increased by 48.22 ha in 2021 and a slight of 12 ha in 2024. Likewise, a substantial building increase of 32.73 ha was observed by 2021 within the newly promoted zone (Zone-IV), followed by a limited expansion of 6 ha in 2024. However, Zone-IV only hosted approximately 35% of the total buildings in 2024.

A comparison between land-use planning and building dynamics revealed an upward trend of building encroachment into the non-building zone (2018: 29.39 ha; 2021: 32.42 ha, and 2024: 33.83 ha). A dominant encroachment was observed in green spaces with an incline of 1.65 ha between 2018 and 2024, accounting for an average of 37% of the total footprints in non-building zone (Fig. 5.10c).

Meanwhile, number of buildings in open spaces and mangrove areas by 2024 increased by 1.2 ha and 0.9 ha, respectively, compared to 2018. A relatively lower dynamics encroachment was observed in riverine (riparian) and aquatic areas, with 3.23 ha and 3.78 ha in 2018, respectively. A slight increase was observed by 3.55 ha and 4.2 ha by 2024 for both zones, respectively.

Given the results, the existing land use planning appears to have not fully incorporated the initial physical zoning scheme. This is confirmed by the newly promoted zone containing only an average of 35% of the building stock, with a cumulative increase of only 38.95 ha from 2018 to 2024. In contrast, the limited development zone experienced a nearly twofold increase, totalling 64.64 ha. This pattern resulted in building development being largely concentrated within the 2004 IOT inundation extent.

5.3. Summary of Chapter V

The analysis revealed that the existing capacity of tsunami shelter are not sufficient to accommodate the population which potentially exposed by Mw 8.7 tsunami. Additionally, building dynamic assessment showed that the city development does not fully comply with land-use planning. Furthermore, the rising trend in building encroachment, especially in green and open spaces, emphasizes deviations in urban planning implementation.

CHAPTER VI:
DISCUSSION AND CONCLUSION

6.1. Error-reduced DEMs, LCR model resolutions and Simulated Inundation

Given these results, the improved variants of global DEMs have fewer inherent errors in coastal low-lying areas than their original datasets. Some of these DEMs, with spatial resolutions ranging from 30 m to 90 m, outperformed the local DEMNAS at resolution of 8 m. This highlights that spatial resolution should not be used as a sole indicator to define the performance of elevation data, as it is influenced by various factors, including generation techniques and geographical settings (Liu et al., 2021; Hawker et al., 2018).

Despite having better accuracy, inundation models using error-reduced DEMs still underestimated the actual inundation extent of the 2004 Indian Ocean Tsunami (IOT). The underestimations likely correlate with several factors, such as the acquisition period of these DEMs (e.g., most data were collected between 2006 and 2020 or after the 2004 IOT) and limitations on the onshore bathymetric data used. As highlighted by Sugawara (2021), the evolution of coastal morphology and land cover changes after a tsunami disaster might present a significant challenge for accurately reconstructing historical tsunami estimates.

This underestimation may also correlate with the accuracy of the tsunami source model. It should be noted that the simulated inundations were constrained to the Mw 9.2 fault inversion model from Koshimura et al. (2009). Using alternative tsunami source models to reproduce the 2004 IOT inundation in the Banda Aceh region, such as the one proposed by Yanagisawa et al. (2010), may yield different simulation results and potentially lead to different agreement with the actual inundation.

Results from combining DEMs with roughness models demonstrated that uniform Manning coefficients consistently produced larger inundation extents compared to variable Manning or Land Cover Roughness (LCR) models. Consequently, the uniform Manning approach showed better agreement with the historical inundation limits. This phenomenon can be attributed to the enhanced flow resistance provided by the LCR models, which incorporate spatially variable damping effects from different land cover

types, particularly vegetation and built-up areas. The increased resistance reduces the flow velocities and limits tsunami propagation.

The sensitivity of the Manning coefficient values within the LCR models further illustrates this mechanism. The larger inundation extents produced when applying the coefficients from Koshimura et al. (2009) can be attributed to the lower resistance values assigned to vegetated areas compared with the coefficients from Bunya et al. (2010). Because vegetation covered approximately 30% of the area of interest (AOI) by 2004, this coefficient difference significantly influenced the overall flow dynamics and energy dissipation patterns across the simulation domain.

The sensitivity of the Manning coefficient values within the LCR models further illustrates this mechanism. The larger inundation extents produced when applying the coefficients from Koshimura et al. (2009) can be attributed to the lower resistance values assigned to vegetated areas compared with the coefficients from Bunya et al. (2010). Because vegetation covered approximately 30% of the area of interest (AOI) by 2004, this coefficient difference significantly influenced the overall flow dynamics and energy dissipation patterns across the simulation domain.

Furthermore, the analysis showed that CoastalDEM coupled with uniform Manning values resulted in the largest inundation extent. This is likely attributed to CoastalDEM having a larger negative bias than other error-reduced DEMs at nearly -1 m. This large negative bias underestimates ground elevation and eventually leads to greater inundation (Liu et al., 2021). However, when paired with LCR models at 10 m and 30 m resolutions, CoastalDEM resulted in lower inundation extents than DiluviumDEM. These findings suggest that the interaction between inundation and parameterization of surface friction is unique to each DEM. When focusing on overland inundation, CoastalDEM showed less overland inundation than FABDEM and DiluviumDEM did. This might be because CoastalDEM has more water pixels than the other two DEMs.

The analysis consistently showed that the spatial resolution of the LCR models did not significantly affect either inundation extent or flow depth. Although not statistically significant, a contrasting pattern was observed: LCR models with finer resolution produced larger inundation extents but smaller median maximum inundation heights (MIH). This contradiction can be explained by several interrelated mechanisms.

First, the relationship between the resolution and built-up area representation creates differential flow resistance patterns. Finer-resolution LCR models derived from higher-resolution land cover maps tend to classify fewer areas as built-up than coarser LCR models. Because built-up areas provide greater flow resistance than other land cover types, their lower representation in finer models decreases the overall flow impedance. This reduced resistance allows water to spread more widely across the landscape, resulting in larger inundation extents but correspondingly shallower depths. This typical phenomenon was also observed by Koyama & Yamada (2022).

Second, the inverse relationship between inundation extent and depth explains the lower MIH estimates. When floods spread over larger areas with reduced resistance, the same volume of water is distributed across a greater surface area, naturally resulting in smaller median depths. Third, finer-resolution LCR models provide a better representation of drainage systems (i.e., small rivers and ponds), which can influence the simulated flow depth. This was consistent with the inundation model setting, in which small water bodies were incorporated by adjusting the topo-bathymetric inputs based on the chosen land cover maps.

Finally, the proposed sequential validation assessment identified that the combination of CoastalDEM and DiluviumDEM with LCR models was the most reliable input for tsunami modelling. Based on the elevation error analysis, these two DEMs consistently exhibited lower errors, particularly within the built-up areas. This suggests that elevation error distributions within built-up areas might be an alternative indicator in selecting global DEMs for inundation modelling, especially in locations dominated by built-up areas. Additionally, considering the relationship between inundation bias and vertical errors of DEMs, MAE could be an alternative elevation error metric in selecting global DEMs rather than RMSE, especially for flood modelling applications.

6.2. Selecting Global Exposure Datasets for Local-scale Application

Evaluation of global exposure datasets showed an underestimation pattern for population and an overestimation pattern for built-up area estimates. The analysis revealed that intrinsic biases from each dataset will ultimately introduce bias into the exposure analysis. In addition, the calculated intrinsic biases showed temporal variation, with differences between 2004 and 2014 varying according to the dataset.

Therefore, data selection should be guided by specific analytical objectives. For instance, GAIA was selected to analyse the compounding bias given that its intrinsic bias in 2004 was lower compared to other datasets. It should be noted that the bias in 2004 was used as an indicator since the 2004 IOT inundation limit served as the hazard reference. While GAIA also showed a lower relative bias by 2014, it should not be used to evaluate the exposure evolution within the areas affected by the 2004 IOT. This is because GAIA showed no change in built-up area evolution, whereas the reference dataset suggested that exposed built-up areas increased by 15% compared to 2004. As for built-up exposure evolution between 2004-2014, GISD30 demonstrated a relatively better accuracy.

Beyond the bias assessment, understanding the properties of global exposure datasets is also essential before selecting data that fits the specific purposes. For instance, the generated local population datasets projected an increase of 8,805 people living within the area affected by the 2004 IOT between 2014-2024. While global datasets such as LandScan Global (LSG) do not provide population data for 2024, combining gridded population data from the 2024 local dataset and 2014 LSG to analyse the exposure evolution resulted in a decrease of 13,875 people. This discrepancy could lead to substantially different conclusions regarding risk evolution. Further analysis revealed that this discrepancy is potentially attributable to population overestimation by LSG in 2014.

Overall, this discrepancy underscores the importance of maintaining consistent datasets for tsunami exposure assessments. Where feasible, identical datasets should be employed to evaluate the temporal exposure evolution. When different datasets are utilized, their respective biases should be systematically characterized beforehand. Failure to account for these biases may introduce systematic errors, potentially resulting in misinterpretation of risk evolution patterns.

6.3. Tsunami Risk and Mitigation Strategy

The ratio of tsunami shelters and populations that are potentially exposed to the Mw 8.7 tsunami revealed that the existing shelter capacity is insufficient. Additionally, following the 2004 IOT, the newly promoted region was designed as a new city center. However, building dynamics assessment demonstrated that this region comprised only 35% of the total building inventory, suggesting that the majority of buildings were

concentrated in tsunami-prone areas. Moreover, the rising trend in building encroachment, especially in green and open spaces, emphasizes deviations in land use planning implementation in Banda Aceh.

The observed deviations in land use planning may be attributed to multiple factors, including inadequate enforcement of land use policies (Iuchi et al., 2023). The absence of periodic monitoring further exacerbates these deviations. While the analysis revealed poor implementation of land use planning based on building dynamics between 2014 and 2024, Takabatake (2022) found that these issues had already emerged during the reconstruction period beginning in 2005.

The initial master plan proposed relocation of populations from the most affected areas to safer zones. However, residential construction in new designated locations failed to achieve significant progress due to difficulties the local government encountered in acquiring land within the constrained timeframe.

Consequently, the urgent need for coastal residents to restore their livelihoods compelled the government to approve resettlement plans permitting their return to original residential areas. This finding is corroborated by Syamsidik et al. (2017), who identified rental costs and land prices as key determinants influencing Banda Aceh residents' continued coastal habitation despite their awareness of the devastating 2004 tsunami. Given these results, it is clear that the community remained as vulnerable as it was in 2004.

6.4. Limitations in Reported Results

Although the results of this study can provide valuable insights for tsunami exposure assessment in data-scarce regions, it should be noted that certain limitations still exist. These limitations are comprehensively discussed in the following sub-sections.

6.4.1 Limitations in Proposed Sequential Validation Approach

Although the proposed sequential validation approach could illuminate the sensitivity of input datasets to inundation model performance, several limitations warrant further consideration. First, the evaluation of DEMs and surface roughness parameterizations is highly sensitive to the established threshold. For example, given the agreement between the simulated and observed inundation extents, changing the acceptable

threshold for the inundation difference from 25% to 10% could suggest that the application of the LCR models results in unreliable inundation estimates. Likewise, reducing the threshold by 10% would make SRTM and AW3D30 classified as “reliable” despite their elevation error characteristics.

Second, while resulting in a less accurate prediction of inundation depths, the use of a uniform Manning’s coefficient has better agreement with the observed inundation extent. This suggests that the application of uniform Manning is still reliable to some extent. It is also crucial to highlight that uncertainties in reliability assessment may persist because of inaccurate records of tsunami measurements and potential geolocation offsets between the simulated and measured flow depth points.

Third, the reported results were constrained to a specific historical tsunami event and the flat coastal plain characteristics of the research location. For example, given its intrinsic elevation errors within the Banda Aceh region, DiluviumDEM and CoastalDEM would achieve the minimum confidence level when the simulated MIH were at 2.36 m and 3.06 m, respectively. This is strongly correlated to their elevation errors with RMSE of 1.18 m for DiluviumDEM and 1.53 m for CoastalDEM. It should be highlighted that those RMSE values are geographically specific elevation error values for the Banda Aceh area. Consequently, changing the research location might yield different elevation error values and yield different confidence level estimates. Thus, the results should not be used to provide an absolute indication of the accuracy of different input data choices, but rather to illuminate how different input data choices can modify the simulated inundation extent.

Finally, these limitations suggest that the relationships observed between model resolution, land cover representation, and inundation patterns may not be directly transferable to areas with different topographic characteristics or coastal configurations. The sensitivity patterns observed in the flat study area may differ significantly in regions with steeper topographies. Similarly, caution should be exercised when comparing the results of this study with other studies employing different hazard modelling approaches or different tsunami sources. Nevertheless, instead of independent individual assessment, these limitations should encourage the modeller to adopt the sequential validation approach to reduce uncertainty in evaluating inundation model performance.

6.4.2 Limitations in Exposed Population Projection

The analysis has estimated that nearly 2,000 up to 24,696 people in Banda Aceh by 2024 would be potentially exposed to an Mw 8.7 tsunami, depending on the exposure limit used. However, these exposure estimates should be interpreted cautiously, as they are constrained to the generated gridded population dataset (i.e., POP-24), which inherits uncertainties from the built-up area definitions used in land cover mapping.

As highlighted by Bonatz et al. (2024), variations in built-up area definitions can lead to discrepancies in population estimates. Because roads and all impervious surfaces were included as built-up areas, using other local datasets at a more granular building level may yield different exposure estimates. Additionally, the gridded population datasets may underestimate settlement-associated properties, including settlement density, number of building stories, and type of settlement (e.g., residential, schools, or offices).

Bias in the reported exposed population may also result from daily mobility patterns, particularly those related to work and school (Lloyd et al., 2019). This mobility potentially results in an unregistered or floating population (Wu and Zhang, 2021). For Banda Aceh, the floating population is reflected by local census data, which reported a population decline of 3000 people between 2018 and 2023 (BPS, 2024). In contrast, the same report revealed that the number of university students in Banda Aceh increased by 27,142 during the same period.

These contrasting records may also explain the contrasting pattern in Banda Aceh's evolution from 2004 to 2014: built-up areas increased by close to 720 ha, but the total population declined by nearly 15,000. Given its administrative role as the capital city of Aceh Province, Banda Aceh hosts many universities and government offices that could attract people from other districts and provinces (Meilianda et al., 2019). This influx potentially leads to the development of built-up areas (e.g., student apartments), even though the city's population may not be significantly increasing.

6.4.3 Limitations in Evacuation System Assessment

The analysis projected that more than 19,000 people, or 77% of the total exposed population, were potentially unsheltered from the Mw 8.7 tsunami. While these findings

can provide valuable insight for disaster preparedness in Banda Aceh, they should be interpreted with caution due to several limitations.

First, the population exposure estimates did not consider evacuation behaviour, evacuation time, routes, or tsunami time arrival. The exposed population could be minimized if unsheltered evacuees are able to evacuate to higher grounds away from the coastal areas. Obviously, this scenario would be possible if rigorous early warning systems are available, including alert information, evacuation routes, and disaster knowledge, such as where and how to evacuate. Future research should address these variables systematically.

Second, the exposure estimates assumed that existing shelters remained structurally intact during the earthquake and subsequent tsunami waves. However, the 2018 Palu tsunami showed that designated shelters may not survive earthquake damage (Koul and Mulchandani, 2021). If shelters are compromised by an earthquake, exposure estimates would increase accordingly. Therefore, future research should incorporate shelter damage assessments to address this limitation.

Third, the exposure assessment did not include alternative shelters to accommodate unsheltered evacuees. Jihad et al. (2023) and Murao et al. (2025) proposed using public buildings such as schools, government offices, and mosques as co-benefit structures for tsunami evacuation sites. Including these alternative shelters would ultimately reduce exposure projections. However, the analysis assumed that because these co-benefit structures are not officially designated for evacuation purposes, evacuees may not recognize them as evacuation sites. Therefore, future research should assess the public knowledge of these alternative shelters. Alternatively, the Banda Aceh municipality should officially designate these structures by installing shelter signage.

6.4.4 Limitations in Building Dynamics Assessment

The evaluation of building dynamics was conducted using spatiotemporal building data generated using a simple backdating method. However, this method has several limitations that warrant further investigation. First, the reported accuracies were strongly influenced by the quality of the contemporary building footprints and validation datasets. Because the validation polygons did not fully cover the Banda Aceh region, employing more complete validation data may yield different accuracy results.

Similarly, the performance of the backdating approach in other regions may differ due to variations in OSM and Microsoft's Global ML Buildings data quality. Hence, assessing the quality of the building input data prior to applying this method is recommended.

Second, the properties of built-up masks, including their resolution, definition of built-up areas, and collection techniques, determine the number of filtered building features. Built-up masks with a high resolution of 5 m were employed. Using built-up masks from global datasets with resolutions ranging from 30 m to 90 m potentially leads to different results. Considering that the effect of the built-up mask resolution was not discussed in this study, further research is strongly recommended to comprehensively address this issue.

Third, while the analysis revealed a substantial increase in building expansion within the inundation extent of the historical 2004 IOT, further research is required to evaluate its impact on human exposure. This is because the analysis did not classify the building categories (e.g., residential, public buildings, permanent buildings, or informal and temporary structures).

6.5. Recommendation

Given the obtained results and their limitations, this study offers some recommendations to address methodological gaps and improve tsunami disaster preparedness strategies:

1. It is highly recommended that future studies reproduce other historical tsunami events and apply sequential validation assessments to evaluate the global DEMs and LCR model performance. By doing so, the robustness of the sequential validation approach could be evaluated. This could additionally characterize the performance of error-reduced DEMs and LCR models across varying terrain characteristics (e.g., slope gradients and land cover configurations).
2. To evaluate population exposure assessment, future research should integrate evacuation modelling that accounts for tsunami arrival times and human behavioural responses during evacuation.
3. To evaluate building exposure assessment, future research should integrate official building cadastral data to identify building structures (e.g., timber-framed

buildings, confined masonry infill brick walls, or reinforced concrete structures). By doing so, the modeller could integrate the inundation depth with the building fragility function and finally result in building damage probability (Lahcene et al., 2021). Using building category information, modellers can develop more detailed and dynamic population exposure models by incorporating temporal population distribution patterns. For instance, populations can be allocated exclusively to residential buildings to model night-time exposure scenarios. Using appropriate weighting coefficients, the modeller could also proportionally distribute populations across all building types to develop exposure models for specific temporal scenarios, including office hours, lunch periods, and commuting times (i.e., travel to and from workplaces) (Dabbeek et al., 2025).

4. In case measured ground elevation data are not available, it is recommended that modellers conducting coastal flood risk assessments within the Banda Aceh region should employ error-reduced DEMs–either DiluviumDEM or CoastalDEM–rather than the local DEMNAS model, which introduces larger elevation errors.
5. For tsunami disaster risk reduction and preparedness, it is recommended that Banda Aceh's municipal government implement incentive programs targeting young residents and first-time homebuyers to promote residential development in designated low-risk areas or promoted zones. Additionally, constructing more evacuation shelters could facilitate evacuation processes and accommodate more residents in coastal areas.
6. It is strongly recommended that modellers conducting coastal flood risk assessments in data-scarce regions should perform sequential bias assessments prior to data utilization. The inclusion of cross-uncertainty assessments in inundation map products based on the elevation models employed is also encouraged. This approach would enable disaster managers and decision-makers in low- and middle-income countries to understand the accuracy of risk information provided, facilitating the formulation of appropriate risk reduction strategies within budget constraints, without overestimating or underestimating the tsunami risk levels
7. Finally, the transferability of the results and methods from this study to other locations requires careful consideration. Future research employing or adapting

the methodological framework of this study should carefully evaluate the above limitations within their specific contexts.

6.6. Summary and Conclusion

This research evaluated the applicability of global datasets–including 11 elevation models and 12 exposure datasets–for tsunami exposure assessment at the local scale. To evaluate these datasets, this study proposed a sequential validation framework to reduce the bias caused by individual reliability assessments. This study also proposed a cross-uncertainty inundation map concept, integrating probabilistic inundation with the confidence level from elevation data.

The results of this study indicated that the error-reduced variant of global elevation models has better elevation accuracy and is relevant for utilization in coastal flood risk assessment, especially CoastalDEM and DiluviumDEM. Additionally, this study proved that integrating land cover roughness (LCR) models with elevation models could improve inundation model accuracy to a certain extent. While the effect of spatial resolution of LCR models seems minimal, it is advised that using finer resolution LCR models can represent better terrain features.

This research has developed high-resolution local exposure datasets as a reference for reviewing applications of global exposure datasets at the local scale. Overall, global exposure datasets exhibit biases by overestimating the built-up environment and underestimating population estimates. Notably, elevation models introduced a larger bias in tsunami exposure estimates than exposure datasets. When both elevation models and exposure datasets were integrated, compounding bias occurred, which nearly doubled the error in exposure estimation.

In conclusion, global elevation models and exposure datasets represent viable resources for local tsunami exposure assessments when their limitations are appropriately acknowledged. A comprehensive understanding of the inherent biases within these datasets is essential to ensure the credibility of tsunami risk assessments.

6.7. Data Availability

Link to access all materials and codes used in this study are provided in the Appendix section.

BIBLIOGRAPHY

Aida, I. (1978). Reliability of a tsunami source model derived from fault parameters. *Journal of Physics of the Earth*, *26*(1), 57–73. https://doi.org/10.4294/jpe1952.26.57

Banda Aceh Municipality. (2017). *Report on Spatial Planning of Banda Aceh 2009-2029 (Revised Version)*. Banda Aceh Municipality. https://bappeda.bandaacehkota.go.id/galeri/rtrw/

Basili, R., Brizuela, B., Herrero, A., Iqbal, S., Lorito, S., Maesano, F. E., Murphy, S., Perfetti, P., Romano, F., Scala, A., Selva, J., Taroni, M., Tiberti, M. M., Thio, H. K., Tonini, R., Volpe, M., Glimsdal, S., Harbitz, C. B., Løvholt, F., … Zaytsev, A. (2021). The Making of the NEAM Tsunami Hazard Model 2018 (NEAMTHM18). *Frontiers in Earth Science*, *8*, 616594. https://doi.org/10.3389/feart.2020.616594

Behling, R., Roessner, S., Segl, K., Kleinschmit, B., & Kaufmann, H. (2014). Robust Automated Image Co-Registration of Optical Multi-Sensor Time Series Data: Database Generation for Multi-Temporal Landslide Detection. *Remote Sensing*, *6*(3), 2572–2600. https://doi.org/10.3390/rs6032572

Behrens, J., Løvholt, F., Jalayer, F., Lorito, S., Salgado-Gálvez, M. A., Sørensen, M., Abadie, S., Aguirre-Ayerbe, I., Aniel-Quiroga, I., Babeyko, A., Baiguera, M., Basili, R., Belliazzi, S., Grezio, A., Johnson, K., Murphy, S., Paris, R., Rafliana, I., De Risi, R., … Vyhmeister, E. (2021). Probabilistic Tsunami Hazard and Risk Analysis: A Review of Research Gaps. *Frontiers in Earth Science*, *9*, 628772. https://doi.org/10.3389/feart.2021.628772

Bernard, E., & Titov, V. (2015). Evolution of tsunami warning systems and products. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*(2053), 20140371. https://doi.org/10.1098/rsta.2014.0371

Bernhofen, M. V., Cooper, S., Trigg, M., Mdee, A., Carr, A., Bhave, A., Solano-Correa, Y. T., Pencue-Fierro, E. L., Teferi, E., Haile, A. T., Yusop, Z., Alias, N. E., Sa'adi, Z., Bin Ramzan, M. A., Dhanya, C. T., & Shukla, P. (2022). The Role of Global Data Sets for Riverine Flood Risk Management at National Scales. *Water Resources Research*, *58*(4), e2021WR031555. https://doi.org/10.1029/2021WR031555

Bielski, C., López-Vázquez, C., Grohmann, C. H., Guth, P. L., & Group, T. T. D. W. (2023). *Novel approach for ranking DEMs: Copernicus DEMs improve open global topography*. https://doi.org/10.48550/ARXIV.2302.08425

Bonatz, H., Reimann, L., & Vafeidis, A. T. (2024). Comparing built-up area datasets to assess urban exposure to coastal hazards in Europe. *Scientific Data*, *11*(1), 499. https://doi.org/10.1038/s41597-024-03339-4

Borrero, J. C. (2005). Field Survey of Northern Sumatra and Banda Aceh, Indonesia after the Tsunami and Earthquake of 26 December 2004. *Seismological Research Letters*, *76*(3), 312–320. https://doi.org/10.1785/gssrl.76.3.312

BPS Kota Banda Aceh. (2024). *Kota Banda Aceh Dalam Angka 2024*. https://drive.google.com/file/d/1j1-bsAxoDSjrU0P6obBD6SELMzC3oQ5d/view?usp=drive_link

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bricker, J. D., Gibson, S., Takagi, H., & Imamura, F. (2015). On the Need for Larger Manning's Roughness Coefficients in Depth-Integrated Tsunami Inundation Models. *Coastal Engineering Journal*, *57*(2), 1550005-1-1550005–1550013. https://doi.org/10.1142/S0578563415500059

Brosens, L., Campforts, B., Govers, G., Aldana-Jague, E., Razanamahandry, V. F., Razafimbelo, T., Rafolisy, T., & Jacobs, L. (2022). Comparative analysis of the Copernicus, TanDEM-X, and UAV-SfM digital elevation models to estimate lavaka (gully) volumes and mobilization rates in the Lake Alaotra region (Madagascar). *Earth Surface Dynamics*, *10*(2), 209–227. https://doi.org/10.5194/esurf-10-209-2022

Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., & Tait, A. M. (2022). Dynamic World, Near real-time global 10 m land use land cover mapping. *Scientific Data*, *9*(1), 251. https://doi.org/10.1038/s41597-022-01307-4

Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Herold, M., & Fritz, S. (2020). *Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe* (Version V3.0.1) [Dataset]. Zenodo. https://doi.org/10.5281/ZENODO.3939050

Bunya, S., Dietrich, J. C., Westerink, J. J., Ebersole, B. A., Smith, J. M., Atkinson, J. H., Jensen, R., Resio, D. T., Luettich, R. A., Dawson, C., Cardone, V. J., Cox, A. T., Powell, M. D., Westerink, H. J., & Roberts, H. J. (2010). A High-Resolution Coupled Riverine Flow, Tide, Wind, Wind Wave, and Storm Surge Model for Southern Louisiana and Mississippi. Part I: Model Development and Validation. *Monthly Weather Review*, *138*(2), 345–377. https://doi.org/10.1175/2009MWR2906.1

Cecioni, C., Iorio, V., Bellotti, G., & Grilli, S. T. (2023). Probabilistic landslide tsunami modeling of the 2018 Palu Bay event. *Coastal Engineering*, *183*, 104332. https://doi.org/10.1016/j.coastaleng.2023.104332

Chamberlain, H. R., Darin, E., Adewole, W. A., Jochem, W. C., Lazar, A. N., & Tatem, A. J. (2024). Building footprint data for countries in Africa: To what extent are existing data products comparable? *Computers, Environment and Urban Systems*, *110*, 102104. https://doi.org/10.1016/j.compenvurbsys.2024.102104

Chen, W., Yao, T., Zhang, G., Li, F., Zheng, G., Zhou, Y., & Xu, F. (2022). Towards ice-thickness inversion: An evaluation of global digital elevation models (DEMs) in the

glacierized Tibetan Plateau. *The Cryosphere*, *16*(1), 197–218. https://doi.org/10.5194/tc-16-197-2022

Chowdhury, Md. S. (2024). Comparison of accuracy and reliability of random forest, support vector machine, artificial neural network and maximum likelihood method in land use/cover classification of urban setting. *Environmental Challenges*, *14*, 100800. https://doi.org/10.1016/j.envc.2023.100800

Crawford, C. J., Roy, D. P., Arab, S., Barnes, C., Vermote, E., Hulley, G., Gerace, A., Choate, M., Engebretson, C., Micijevic, E., Schmidt, G., Anderson, C., Anderson, M., Bouchard, M., Cook, B., Dittmeier, R., Howard, D., Jenkerson, C., Kim, M., … Zahn, S. (2023). The 50-year Landsat collection 2 archive. *Science of Remote Sensing*, *8*, 100103. https://doi.org/10.1016/j.srs.2023.100103

Dabbeek, J., Crowley, H., Silva, V., & Ozcebe, S. (2025). Impact of population spatiotemporal patterns on earthquake human losses. *International Journal of Disaster Risk Reduction*, *122*, 105455. https://doi.org/10.1016/j.ijdrr.2025.105455

Davies, G. (2019). Tsunami variability from uncalibrated stochastic earthquake models: Tests against deep ocean observations 2006–2016. *Geophysical Journal International*, *218*(3), 1939–1960. https://doi.org/10.1093/gji/ggz260

Davies, G., & Griffin, J. (2018). *The 2018 Australian probabilistic tsunami hazard assessment: Hazard from earthquake generated tsunamis*. Geoscience Australia. https://doi.org/10.11636/Record.2018.041

Davies, G., & Griffin, J. (2020). Sensitivity of Probabilistic Tsunami Hazard Assessment to Far-Field Earthquake Slip Complexity and Rigidity Depth-Dependence: Case Study of Australia. *Pure and Applied Geophysics*, *177*(3), 1521–1548. https://doi.org/10.1007/s00024-019-02299-w

Davies, G., Griffin, J., Løvholt, F., Glimsdal, S., Harbitz, C., Thio, H. K., Lorito, S., Basili, R., Selva, J., Geist, E., & Baptista, M. A. (2018). A global probabilistic tsunami hazard assessment from earthquake sources. *Geological Society, London, Special Publications*, *456*(1), 219–244. https://doi.org/10.1144/SP456.5

Davies, G., Weber, R., Wilson, K., & Cummins, P. (2022). From offshore to onshore probabilistic tsunami hazard assessment via efficient Monte Carlo sampling. *Geophysical Journal International*, *230*(3), 1630–1651. https://doi.org/10.1093/gji/ggac140

De Risi, R., & Goda, K. (2017). Simulation-Based Probabilistic Tsunami Hazard Analysis: Empirical and Robust Hazard Predictions. *Pure and Applied Geophysics*, *174*(8), 3083–3106. https://doi.org/10.1007/s00024-017-1588-9

Dusseau, D., Zobel, Z., & Schwalm, C. R. (2023). DiluviumDEM: Enhanced accuracy in global coastal digital elevation models. *Remote Sensing of Environment*, *298*, 113812. https://doi.org/10.1016/j.rse.2023.113812

Ehrlich, D., Kemper, T., Pesaresi, M., & Corbane, C. (2018). Built-up area and population density: Two Essential Societal Variables to address climate hazard impact. *Environmental Science & Policy*, *90*, 73–82. https://doi.org/10.1016/j.envsci.2018.10.001

Esch, T., Heldens, W., Hirner, A., Keil, M., Marconcini, M., Roth, A., Zeidler, J., Dech, S., & Strano, E. (2017a). Breaking new ground in mapping human settlements from space – The Global Urban Footprint. *ISPRS Journal of Photogrammetry and Remote Sensing*, *134*, 30–42. https://doi.org/10.1016/j.isprsjprs.2017.10.012

Esch, T., Heldens, W., Hirner, A., Keil, M., Marconcini, M., Roth, A., Zeidler, J., Dech, S., & Strano, E. (2017b). Breaking new ground in mapping human settlements from space – The Global Urban Footprint. *ISPRS Journal of Photogrammetry and Remote Sensing*, *134*, 30–42. https://doi.org/10.1016/j.isprsjprs.2017.10.012

Ettehadi Osgouei, P., Kaya, S., Sertel, E., & Alganci, U. (2019). Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote Sensing*, *11*(3), 345. https://doi.org/10.3390/rs11030345

Fahrland, E., Paschko, H., Jacob, P., & Kahabka, H. (2022). *Copernicus DEM Product Handbook*. https://spacedata.copernicus.eu/documents/20123/121239/GEO1988-CopernicusDEM-SPE-002_ProductHandbook_I4.0.pdf

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, *45*(2), RG2004. https://doi.org/10.1029/2005RG000183

Fassoni-Andrade, A. C., Paiva, R. C. D. D., Rudorff, C. D. M., Barbosa, C. C. F., & Novo, E. M. L. D. M. (2020). High-resolution mapping of floodplain topography from space: A case study in the Amazon. *Remote Sensing of Environment*, *251*, 112065. https://doi.org/10.1016/j.rse.2020.112065

Fernández-Pacheco, V. M., López-Sánchez, C. A., Álvarez-Álvarez, E., López, M. J. S., García-Expósito, L., Yudego, E. A., & Carús-Candás, J. L. (2018). Estimation of PM10 Distribution using Landsat5 and Landsat8 Remote Sensing. *The 2nd International Research Conference on Sustainable Energy, Engineering, Materials and Environment*, 1430. https://doi.org/10.3390/proceedings2231430

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381–395. https://doi.org/10.1145/358669.358692

Frazier, A. E., & Hemingway, B. L. (2021). A Technical Review of Planet Smallsat Data: Practical Considerations for Processing and Using PlanetScope Imagery. *Remote Sensing*, *13*(19), 3930. https://doi.org/10.3390/rs13193930

Friedl, M. A., Woodcock, C. E., Olofsson, P., Zhu, Z., Loveland, T., Stanimirova, R., Arevalo, P., Bullock, E., Hu, K.-T., Zhang, Y., Turlej, K., Tarrio, K., McAvoy, K., Gorelick, N., Wang, J. A., Barber, C. P., & Souza, C. (2022). Medium Spatial Resolution Mapping of Global Land Cover and Land Cover Change Across Multiple Decades From Landsat. *Frontiers in Remote Sensing*, *3*, 894571. https://doi.org/10.3389/frsen.2022.894571

Fuchs, S., Keiler, M., & Zischg, A. (2015). A spatiotemporal multi-hazard exposure assessment based on property data. *Natural Hazards and Earth System Sciences*, *15*(9), 2127–2142. https://doi.org/10.5194/nhess-15-2127-2015

Fujii, Y., Satake, K., Watada, S., & Ho, T.-C. (2021). Re-examination of Slip Distribution of the 2004 Sumatra–Andaman Earthquake (Mw 9.2) by the Inversion of Tsunami Data Using Green's Functions Corrected for Compressible Seawater Over the Elastic Earth. *Pure and Applied Geophysics*, *178*(12), 4777–4796. https://doi.org/10.1007/s00024-021-02909-6

Fujita, S., Nomura, R., Moriguchi, S., Otake, Y., LeVeque, R. J., & Terada, K. (2024). Scenario Superposition Method for Real-Time Tsunami Prediction Using a Bayesian Approach. *Journal of Geophysical Research: Oceans*, *129*(12), e2024JC021565. https://doi.org/10.1029/2024JC021565

Fukui, N., Mori, N., Miyashita, T., Shimura, T., & Goda, K. (2022). Subgrid-scale modeling of tsunami inundation in coastal urban areas. *Coastal Engineering*, *177*, 104175. https://doi.org/10.1016/j.coastaleng.2022.104175

Gašparović, M., Medak, D., Pilaš, I., Jurjević, L., & Balenović, I. (2018). FUSION OF SENTINEL-2 AND PLANETSCOPE IMAGERY FOR VEGETATION DETECTION AND MONITORING. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII–1*, 155–160. https://doi.org/10.5194/isprs-archives-XLII-1-155-2018

Gayer, G., Leschka, S., Nöhren, I., Larsen, O., & Günther, H. (2010). Tsunami inundation modelling based on detailed roughness maps of densely populated areas. *Natural Hazards and Earth System Sciences*, *10*(8), 1679–1687. https://doi.org/10.5194/nhess-10-1679-2010

GEBCO Bathymetric Compilation Group 2023. (2023). *The GEBCO_2023 Grid—A continuous terrain model of the global oceans and land.* (Version 1) [Documents,Network Common Data Form]. NERC EDS British Oceanographic Data Centre NOC. https://doi.org/10.5285/F98B053B-0CBC-6C23-E053-6C86ABC0AF7B

Geiß, C., Maier, J., So, E., Schoepfer, E., Harig, S., Gómez Zapata, J. C., & Zhu, Y. (2024). Anticipating a risky future: Long short-term memory (LSTM) models for spatiotemporal extrapolation of population data in areas prone to earthquakes and tsunamis in Lima, Peru. *Natural Hazards and Earth System Sciences*, *24*(3), 1051–1064. https://doi.org/10.5194/nhess-24-1051-2024

Geiß, C., Schauß, A., Riedlinger, T., Dech, S., Zelaya, C., Guzmán, N., Hube, M. A., Arsanjani, J. J., & Taubenböck, H. (2017). Joint use of remote sensing data and volunteered geographic information for exposure estimation: Evidence from Valparaíso, Chile. *Natural Hazards*, *86*(S1), 81–105. https://doi.org/10.1007/s11069-016-2663-8

Geist, E., & Lynett, P. (2014). Source Processes for the Probabilistic Assessment of Tsunami Hazards. *Oceanography*, *27*(2), 86–93. https://doi.org/10.5670/oceanog.2014.43

Gesch, D. (2023). *Assessing Global Elevation Models for Mapping the Low Elevation Coastal Zone*. https://doi.org/10.5281/ZENODO.8011577

Gesch, D. B. (2018). Best Practices for Elevation-Based Assessments of Sea-Level Rise and Coastal Flooding Exposure. *Frontiers in Earth Science*, *6*, 230. https://doi.org/10.3389/feart.2018.00230

Gholizadeh, A., Žižala, D., Saberioon, M., & Borůvka, L. (2018). Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. *Remote Sensing of Environment*, *218*, 89–103. https://doi.org/10.1016/j.rse.2018.09.015

Gibbons, S. J., Lorito, S., De La Asunción, M., Volpe, M., Selva, J., Macías, J., Sánchez-Linares, C., Brizuela, B., Vöge, M., Tonini, R., Lanucara, P., Glimsdal, S., Romano, F., Meyer, J. C., & Løvholt, F. (2022). The Sensitivity of Tsunami Impact to Earthquake Source Parameters and Manning Friction in High-Resolution Inundation Simulations. *Frontiers in Earth Science*, *9*, 757618. https://doi.org/10.3389/feart.2021.757618

Giblin, J., & Damlamian, H. (2022). *Tsunami Hazard Assessment: Samoa Case Study* [SPC Technical Report SPC00069]. Pacific Community. https://pacificdata.org/data/dataset/oai-www-spc-int-3d002317-af3b-429e-a9cc-311b7c98ef7f

Giblin, J., Damlamian, H., Davies, G., Weber, R., & Wilson, K. (2022). *Earthquake Scenario Selection for Tsunami Inundation Hazard Assessment: Guidelines on using the 2018 Probabilistic Tsunami Hazard Assessment in the Pacific* [78pp]. https://doi.org/10.25607/OBP-1836

Glimsdal, S., Løvholt, F., Harbitz, C. B., Romano, F., Lorito, S., Orefice, S., Brizuela, B., Selva, J., Hoechner, A., Volpe, M., Babeyko, A., Tonini, R., Wronna, M., & Omira, R. (2019). A New Approximate Method for Quantifying Tsunami Maximum Inundation Height Probability. *Pure and Applied Geophysics*, *176*(7), 3227–3246. https://doi.org/10.1007/s00024-019-02091-w

Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., Zhang, W., & Zhou, Y. (2020). Annual maps of global artificial impervious area (GAIA) between

1985 and 2018. *Remote Sensing of Environment*, *236*, 111510. https://doi.org/10.1016/j.rse.2019.111510

Gonzalez, C., & Bueso-Bello, J. L. (2023). *TanDEM-X 30m Edited DEM Product Description*. https://geoservice.dlr.de/web/dataguide/tdm30/pdfs/TD_GS_PS_0215_TanDEM_X_30m_Edited_DEM_Product_Description_v1_1.pdf

Grezio, A., Babeyko, A., Baptista, M. A., Behrens, J., Costa, A., Davies, G., Geist, E. L., Glimsdal, S., González, F. I., Griffin, J., Harbitz, C. B., LeVeque, R. J., Lorito, S., Løvholt, F., Omira, R., Mueller, C., Paris, R., Parsons, T., Polet, J., … Thio, H. K. (2017). Probabilistic Tsunami Hazard Analysis: Multiple Sources and Global Applications: PROBABILISTIC TSUNAMI HAZARD ANALYSIS. *Reviews of Geophysics*, *55*(4), 1158–1198. https://doi.org/10.1002/2017RG000579

Griffin, J., Latief, H., Kongko, W., Harig, S., Horspool, N., Hanung, R., Rojali, A., Maher, N., Fuchs, A., Hossen, J., Upi, S., Edi Dewanto, S., Rakowsky, N., & Cummins, P. (2015). An evaluation of onshore digital elevation models for modeling tsunami inundation zones. *Frontiers in Earth Science*, *3*. https://doi.org/10.3389/feart.2015.00032

Grizonnet, M., Michel, J., Poughon, V., Inglada, J., Savinaud, M., & Cresson, R. (2017). Orfeo ToolBox: Open source processing of remote sensing images. *Open Geospatial Data, Software and Standards*, *2*(1), 15. https://doi.org/10.1186/s40965-017-0031-6

Harig, S., Zamora, N., Gubler, A., & Rakowsky, N. (2022). Systematic Comparison of Tsunami Simulations on the Chilean Coast Based on Different Numerical Approaches. *GeoHazards*, *3*(2), 345–370. https://doi.org/10.3390/geohazards3020018

Hawker, L., Bates, P., Neal, J., & Rougier, J. (2018). Perspectives on Digital Elevation Model (DEM) Simulation for Flood Modeling in the Absence of a High-Accuracy Open Access Global DEM. *Frontiers in Earth Science*, *6*, 233. https://doi.org/10.3389/feart.2018.00233

Hawker, L., Neal, J., & Bates, P. (2019). Accuracy assessment of the TanDEM-X 90 Digital Elevation Model for selected floodplain sites. *Remote Sensing of Environment*, *232*, 111319. https://doi.org/10.1016/j.rse.2019.111319

Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., & Neal, J. (2022). A 30 m global map of elevation with forests and buildings removed. *Environmental Research Letters*, *17*(2), 024016. https://doi.org/10.1088/1748-9326/ac4d4f

Heidarzadeh, M., & Mulia, I. E. (2023). A new dual earthquake and submarine landslide source model for the 28 September 2018 Palu (Sulawesi), Indonesia tsunami. *Coastal Engineering Journal*, *65*(1), 97–109. https://doi.org/10.1080/21664250.2022.2122293

Herfort, B., Lautenbach, S., Porto De Albuquerque, J., Anderson, J., & Zipf, A. (2023). A spatio-temporal analysis investigating completeness and inequalities of global urban

building data in OpenStreetMap. *Nature Communications*, *14*(1), 3985. https://doi.org/10.1038/s41467-023-39698-6

Hermosilla, T., Wulder, M. A., White, J. C., & Coops, N. C. (2022). Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes. *Remote Sensing of Environment*, *268*, 112780. https://doi.org/10.1016/j.rse.2021.112780

Hese, S., & Heyer, T. (2016). Earth observation data based rapid flood-extent modelling for tsunami-devastated coastal areas. *International Journal of Applied Earth Observation and Geoinformation*, *46*, 63–83. https://doi.org/10.1016/j.jag.2015.11.005

Hettiarachchi, S. (2018a). Establishing the Indian Ocean Tsunami Warning and Mitigation System for human and environmental security. *Procedia Engineering*, *212*, 1339–1346. https://doi.org/10.1016/j.proeng.2018.01.173

Hettiarachchi, S. (2018b). Establishing the Indian Ocean Tsunami Warning and Mitigation System for human and environmental security. *Procedia Engineering*, *212*, 1339–1346. https://doi.org/10.1016/j.proeng.2018.01.173

Hinkel, J., Feyen, L., Hemer, M., Le Cozannet, G., Lincke, D., Marcos, M., Mentaschi, L., Merkens, J. L., de Moel, H., Muis, S., Nicholls, R. J., Vafeidis, A. T., van de Wal, R. S. W., Vousdoukas, M. I., Wahl, T., Ward, P. J., & Wolff, C. (2021). Uncertainty and Bias in Global to Regional Scale Assessments of Current and Future Coastal Flood Risk. *Earth's Future*, *9*(7). https://doi.org/10.1029/2020EF001882

Ho, T., Satake, K., Watada, S., Hsieh, M., Chuang, R. Y., Aoki, Y., Mulia, I. E., Gusman, A. R., & Lu, C. (2021). Tsunami Induced by the Strike-Slip Fault of the 2018 Palu Earthquake (M_w = 7.5), Sulawesi Island, Indonesia. *Earth and Space Science*, *8*(6), e2020EA001400. https://doi.org/10.1029/2020EA001400

Horspool, N., Pranantyo, I., Griffin, J., Latief, H., Natawidjaja, D. H., Kongko, W., Cipta, A., Bustaman, B., Anugrah, S. D., & Thio, H. K. (2014). A probabilistic tsunami hazard assessment for Indonesia. *Natural Hazards and Earth System Sciences*, *14*(11), 3105–3122. https://doi.org/10.5194/nhess-14-3105-2014

Houborg, R., & McCabe, M. F. (2018). A Cubesat enabled Spatio-Temporal Enhancement Method (CESTEM) utilizing Planet, Landsat and MODIS data. *Remote Sensing of Environment*, *209*, 211–226. https://doi.org/10.1016/j.rse.2018.02.067

Huang, X., Song, Y., Yang, J., Wang, W., Ren, H., Dong, M., Feng, Y., Yin, H., & Li, J. (2022). Toward accurate mapping of 30-m time-series global impervious surface area (GISA). *International Journal of Applied Earth Observation and Geoinformation*, *109*, 102787. https://doi.org/10.1016/j.jag.2022.102787

Imamura, F., Boret, S. P., Suppasri, A., & Muhari, A. (2019). Recent occurrences of serious tsunami damage and the future challenges of tsunami disaster risk reduction. *Progress in Disaster Science*, *1*, 100009. https://doi.org/10.1016/j.pdisas.2019.100009

Iuchi, K., Takagi, H., Jibiki, Y., Kondo, T., Kusunoki, A., Hanifa, N. R., Pelupessy, D., Gayathri, R. T., & Olshansky, R. (2023). Questioning the hazard map-based rebuilding process: Learning from the 2018 Sulawesi earthquake in Indonesia. *Coastal Engineering Journal*, *65*(1), 126–148. https://doi.org/10.1080/21664250.2023.2165430

JICA. (2005). *The Study on The Urgent Rehabilitation And Reconstruction Support Program For Aceh Province And Affected Areas In North Sumatra* [FINAL REPORT]. https://openjicareport.jica.go.jp/pdf/11802741_02.pdf

Jihad, A., Muksin, U., Syamsidik, Ramli, M., Banyunegoro, V. H., Simanjuntak, A. V. H., & Rusdin, A. A. (2023). Tsunami evacuation sites in the northern Sumatra (Indonesia) determined based on the updated tsunami numerical simulations. *Progress in Disaster Science*, *18*, 100286. https://doi.org/10.1016/j.pdisas.2023.100286

Jihad, A., Muksin, U., Syamsidik, Suppasri, A., Ramli, M., & Banyunegoro, V. H. (2020). Coastal and settlement typologies-based tsunami modeling along the northern Sumatra seismic gap zone for disaster risk reduction action plans. *International Journal of Disaster Risk Reduction*, *51*, 101800. https://doi.org/10.1016/j.ijdrr.2020.101800

Ju, J., Zhou, Q., Freitag, B., Roy, D. P., Zhang, H. K., Sridhar, M., Mandel, J., Arab, S., Schmidt, G., Crawford, C. J., Gascon, F., Strobl, P. A., Masek, J. G., & Neigh, C. S. R. (2025). The Harmonized Landsat and Sentinel-2 version 2.0 surface reflectance dataset. *Remote Sensing of Environment*, *324*, 114723. https://doi.org/10.1016/j.rse.2025.114723

Kaiser, G., Scheele, L., Kortenhaus, A., Løvholt, F., Römer, H., & Leschka, S. (2011). The influence of land cover roughness on the results of high resolution tsunami inundation modeling. *Natural Hazards and Earth System Sciences*, *11*(9), 2521–2540. https://doi.org/10.5194/nhess-11-2521-2011

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., & Brumby, S. P. (2021). Global land use / land cover with Sentinel 2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 4704–4707. https://doi.org/10.1109/IGARSS47720.2021.9553499

Koshimura, S., Moya, L., Mas, E., & Bai, Y. (2020). Tsunami Damage Detection with Remote Sensing: A Review. *Geosciences*, *10*(5), 177. https://doi.org/10.3390/geosciences10050177

Koshimura, S., Oie, T., Yanagisawa, H., & Imamura, F. (2009). Developing Fragility Functions for Tsunami Damage Estimation Using Numerical Model and Post-Tsunami Data from Banda Aceh, Indonesia. *Coastal Engineering Journal*, *51*(3), 243–273. https://doi.org/10.1142/S0578563409002004

Kotaridis, I., & Lazaridou, M. (2020). OBJECT-BASED IMAGE ANALYSIS OF DIFFERENT SPATIAL RESOLUTION SATELLITE IMAGERIES IN URBAN AND SUBURBAN ENVIRONMENT. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B3-2020*, 105–112. https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-105-2020

Koul, S., & Mulchandani, H. (2021). Building Performance and Geotechnical Failures in 7.5M Palu Earthquake and Tsunami 28 September 2018. In T. G. Sitharam, R. Jakka, & L. Govindaraju (Eds.), *Local Site Effects and Ground Failures* (Vol. 117, pp. 335–343). Springer Singapore. https://doi.org/10.1007/978-981-15-9984-2_29

Koyama, N., & Yamada, T. (2022, March 28). *Analysis of inundation characteristics under various computational conditions for large-scale flood forecasting*. https://doi.org/10.5194/egusphere-egu22-10930

Kuffer, M., Owusu, M., Oliveira, L., Sliuzas, R., & Van Rijn, F. (2022). The Missing Millions in Maps: Exploring Causes of Uncertainties in Global Gridded Population Datasets. *ISPRS International Journal of Geo-Information*, *11*(7), 403. https://doi.org/10.3390/ijgi11070403

Kulp, S. A., & Strauss, B. H. (2018). CoastalDEM: A global coastal digital elevation model improved from SRTM using a neural network. *Remote Sensing of Environment*, *206*, 231–239. https://doi.org/10.1016/j.rse.2017.12.026

Kulp, S., & Strauss, B. H. (2016). Global DEM Errors Underpredict Coastal Vulnerability to Sea Level Rise and Flooding. *Frontiers in Earth Science*, *4*. https://doi.org/10.3389/feart.2016.00036

Lachaise, M., & Schweißhelm, B. (2023). *TD-GS-PS-0216 TanDEM-X 30m DEM Change Maps Product Description*. https://geoservice.dlr.de/web/dataguide/tdm30/pdfs/TD-GS-PS-0216_TanDEM-X_30m_DEM_Change_Maps_Product_Description_1.0.pdf

Lahcene, E., Ioannou, I., Suppasri, A., Pakoksung, K., Paulik, R., Syamsidik, S., Bouchette, F., & Imamura, F. (2021). Characteristics of building fragility curves for seismic and non-seismic tsunamis: Case studies of the 2018 Sunda Strait, 2018 Sulawesi–Palu, and 2004 Indian Ocean tsunamis. *Natural Hazards and Earth System Sciences*, *21*(8), 2313–2344. https://doi.org/10.5194/nhess-21-2313-2021

Láng-Ritter, J., Keskinen, M., & Tenkanen, H. (2025). Global gridded population datasets systematically underrepresent rural population. *Nature Communications*, *16*(1), 2170. https://doi.org/10.1038/s41467-025-56906-7

Laso Bayas, J. C., Ekadinata, A., Widayati, A., Marohn, C., & Cadisch, G. (2015). Resolution vs. Image quality in pre-tsunami imagery used for tsunami impact models in Aceh, Indonesia. *International Journal of Applied Earth Observation and Geoinformation*, *42*, 38–48. https://doi.org/10.1016/j.jag.2015.05.007

Leach, N., Coops, N. C., & Obrknezev, N. (2019a). Normalization method for multi-sensor high spatial and temporal resolution satellite imagery with radiometric inconsistencies. *Computers and Electronics in Agriculture*, *164*, 104893. https://doi.org/10.1016/j.compag.2019.104893

Leach, N., Coops, N. C., & Obrknezev, N. (2019b). Normalization method for multi-sensor high spatial and temporal resolution satellite imagery with radiometric inconsistencies. *Computers and Electronics in Agriculture*, *164*, 104893. https://doi.org/10.1016/j.compag.2019.104893

Lebakula, V., Sims, K., Reith, A., Rose, A., McKee, J., Coleman, P., Kaufman, J., Urban, M., Jochem, C., Whitlock, C., Ogden, M., Pyle, J., Roddy, D., Epting, J., & Bright, E. (2025). LandScan Global 30 Arcsecond Annual Global Gridded Population Datasets from 2000 to 2022. *Scientific Data*, *12*(1), 495. https://doi.org/10.1038/s41597-025-04817-z

León, J., Gubler, A., & Ogueda, A. (2022). Modelling geographical and built-environment attributes as predictors of human vulnerability during tsunami evacuations: A multi-case-study and paths to improvement. *Natural Hazards and Earth System Sciences*, *22*(9), 2857–2878. https://doi.org/10.5194/nhess-22-2857-2022

León, J., Mokrani, C., Catalán, P., Cienfuegos, R., & Femenías, C. (2019). The Role of Built Environment's Physical Urban Form in Supporting Rapid Tsunami Evacuations: Using Computer-Based Models and Real-World Data as Examination Tools. *Frontiers in Built Environment*, *4*, 89. https://doi.org/10.3389/fbuil.2018.00089

Leyk, S., Gaughan, A. E., Adamo, S. B., De Sherbinin, A., Balk, D., Freire, S., Rose, A., Stevens, F. R., Blankespoor, B., Frye, C., Comenetz, J., Sorichetta, A., MacManus, K., Pistolesi, L., Levy, M., Tatem, A. J., & Pesaresi, M. (2019). The spatial allocation of population: A review of large-scale gridded population data products and their fitness for use. *Earth System Science Data*, *11*(3), 1385–1409. https://doi.org/10.5194/essd-11-1385-2019

Li, B., Xu, X., Liu, X., Shi, Q., Zhuang, H., Cai, Y., & He, D. (2023). An improved global land cover mapping in 2015 with 30 m resolution (GLC-2015) based on a multisource product-fusion approach. *Earth System Science Data*, *15*(6), 2347–2373. https://doi.org/10.5194/essd-15-2347-2023

Li, H., Zhao, J., Yan, B., Yue, L., & Wang, L. (2022). Global DEMs vary from one to another: An evaluation of newly released Copernicus, NASA and AW3D30 DEM on selected terrains of China using ICESat-2 altimetry data. *International Journal of Digital Earth*, *15*(1), 1149–1168. https://doi.org/10.1080/17538947.2022.2094002

Li, Z., Scheffler, D., Coops, N. C., Leach, N., & Sachs, T. (2021). Towards analysis ready data of optical CubeSat images: Demonstrating a hierarchical normalization framework at a wetland site. *International Journal of Applied Earth Observation and Geoinformation*, *103*, 102502. https://doi.org/10.1016/j.jag.2021.102502

Liao, M., Wen, H., Yang, L., Wang, G., Xiang, X., & Liang, X. (2024). Improving the model robustness of flood hazard mapping based on hyperparameter optimization of random forest. *Expert Systems with Applications*, *241*, 122682. https://doi.org/10.1016/j.eswa.2023.122682

Liu, L., Cao, X., Li, S., & Jie, N. (2024). A 31-year (1990–2020) global gridded population dataset generated by cluster analysis and statistical learning. *Scientific Data*, *11*(1), 124. https://doi.org/10.1038/s41597-024-02913-0

Liu, Y., Bates, P. D., Neal, J. C., & Yamazaki, D. (2021). Bare-Earth DEM Generation in Urban Areas for Flood Inundation Simulation Using Global Digital Elevation Models. *Water Resources Research*, *57*(4). https://doi.org/10.1029/2020WR028516

Lloyd, C. T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F. R., Gaughan, A. E., Nieves, J. J., Hornby, G., MacManus, K., Sinha, P., Bondarenko, M., Sorichetta, A., & Tatem, A. J. (2019). Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data*, *3*(2), 108–139. https://doi.org/10.1080/20964471.2019.1625151

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 277–293. https://doi.org/10.1016/j.isprsjprs.2017.06.001

Mansinha, L., & Smylie, D. E. (1971). The displacement fields of inclined faults. *Bulletin of the Seismological Society of America*, *61*(5), 1433–1440. https://doi.org/10.1785/BSSA0610051433

Marconcini, M., Metz- Marconcini, A., Esch, T., & Gorelick, N. (2021). Understanding Current Trends in Global Urbanisation—The World Settlement Footprint Suite. *GI_Forum*, *1*, 33–38. https://doi.org/10.1553/giscience2021_01_s33

Marconcini, M., Metz-Marconcini, A., Üreyen, S., Palacios-Lopez, D., Hanke, W., Bachofer, F., Zeidler, J., Esch, T., Gorelick, N., Kakarla, A., Paganini, M., & Strano, E. (2020). Outlining where humans live, the World Settlement Footprint 2015. *Scientific Data*, *7*(1), 242. https://doi.org/10.1038/s41597-020-00580-5

Mas, E., Paulik, R., Pakoksung, K., Adriano, B., Moya, L., Suppasri, A., Muhari, A., Khomarudin, R., Yokoya, N., Matsuoka, M., & Koshimura, S. (2020). Characteristics of Tsunami Fragility Functions Developed Using Different Sources of Damage Data from the 2018 Sulawesi Earthquake and Tsunami. *Pure and Applied Geophysics*, *177*(6), 2437–2455. https://doi.org/10.1007/s00024-020-02501-4

McClean, F., Dawson, R., & Kilsby, C. (2020). Implications of Using Global Digital Elevation Models for Flood Risk Analysis in Cities. *Water Resources Research*, *56*(10), e2020WR028241. https://doi.org/10.1029/2020WR028241

Meilianda, E., Pradhan, B., Syamsidik, Comfort, L. K., Alfian, D., Juanda, R., Syahreza, S., & Munadi, K. (2019). Assessment of post-tsunami disaster land use/land cover change and potential impact of future sea-level rise to low-lying coastal areas: A case

study of Banda Aceh coast of Indonesia. *International Journal of Disaster Risk Reduction*, *41*, 101292. https://doi.org/10.1016/j.ijdrr.2019.101292

Montero, D., Aybar, C., Mahecha, M. D., Martinuzzi, F., Söchting, M., & Wieneke, S. (2023). A standardized catalogue of spectral indices to advance the use of remote sensing in Earth system research. *Scientific Data*, *10*(1), 197. https://doi.org/10.1038/s41597-023-02096-0

Murao, O., Sato, M., Sugiyasu, K., Miura, H., Khoiriyah, M., Saito, R., & Affan, M. (2025). Tsunami evacuation risk change associated with urban recovery in Banda Aceh after 2004 Aceh tsunami. *International Journal of Disaster Risk Reduction*, *121*, 105400. https://doi.org/10.1016/j.ijdrr.2025.105400

Nakata, K., Katsumata, A., & Muhari, A. (2020). Submarine landslide source models consistent with multiple tsunami records of the 2018 Palu tsunami, Sulawesi, Indonesia. *Earth, Planets and Space*, *72*(1), 44. https://doi.org/10.1186/s40623-020-01169-3

NASA JPL. (2020). *NASADEM Merged DEM Global 1 arc second nc V001* [Dataset]. NASA EOSDIS Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MEASURES/NASADEM/NASADEM_NC.001

Neuenschwander, A. L., K. L. Pitts, B. P. Jelley, J. Robbins, B. Klotz, S. C. Popescu, R. F. Nelson, D. Harding, D. Pederson, & R. Sheridan. (2023). *ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 6* [Dataset]. [object Object]. https://doi.org/10.5067/ATLAS/ATL08.006

Neuenschwander, A. L., & Magruder, L. A. (2023). PhoREAL: Increasing spatial and temporal data accessibly of ICESat-2 through open-source Python repository. In M. D. Turner, G. W. Kamerman, & L. A. Magruder (Eds.), *Laser Radar Technology and Applications XXVIII* (p. 18). SPIE. https://doi.org/10.1117/12.2667098

Neumann, T. A., A. Brenner, D. Hancock, J. Robbins, J. Saba, K. Harbeck, A. Gibbons, J. Lee, S. B. Luthcke, & T. Rebold. (2023). *ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 6* [Dataset]. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/ATLAS/ATL03.006

Nosavan, J., Moreau, A., & Hosford, S. (2020). *SPOT World Heritage catalogue: 30 years of SPOT 1-to-5 observation* [Other]. oral. https://doi.org/10.5194/egusphere-egu2020-8275

Ocean Networks Canada, Bosma, C., Shumlich, A., Rankin, M., Kouhi, S., & Amouzgar, R. (2023). Integrating Topographic and Bathymetric Data for High-Resolution Digital Elevation Modeling to Support Tsunami Hazard Mapping. *Oceanography*. https://doi.org/10.5670/oceanog.2023.s1.23

OECD. (2016). *OECD Regions at a Glance 2016*. OECD. https://doi.org/10.1787/reg_glance-2016-en

Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America*, *75*(4), 1135–1154. https://doi.org/10.1785/BSSA0750041135

Oktari, R. S., Sofyan, H., Hidayati, A., Rolanda, R., Syamsidik, Puturuhu, F., & Sasaki, D. (2025). Enhancing coastal community preparedness: The role of education, experience, and exposure to information. *Progress in Disaster Science*, 100429. https://doi.org/10.1016/j.pdisas.2025.100429

Olajubu, V., Trigg, M. A., Berretta, C., Sleigh, A., Chini, M., Matgen, P., Mojere, S., & Mulligan, J. (2021). Urban correction of global DEMs using building density for Nairobi, Kenya. *Earth Science Informatics*, *14*(3), 1383–1398. https://doi.org/10.1007/s12145-021-00647-w

Opdyke, A., & Fatima, K. (2024). Comparing the suitability of global gridded population datasets for local landslide risk assessments. *Natural Hazards*, *120*(3), 2415–2432. https://doi.org/10.1007/s11069-023-06283-5

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J. H., Carioli, A., Corbane, C., Dijkstra, L., Florio, P., Friedrich, H. K., Gao, J., Leyk, S., Lu, L., Maffenini, L., Mari-Rivero, I., Melchiorri, M., Syrris, V., Van Den Hoek, J., & Kemper, T. (2024). Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data. *International Journal of Digital Earth*, *17*(1), 2390454. https://doi.org/10.1080/17538947.2024.2390454

Phiri, D., Morgenroth, J., Xu, C., & Hermosilla, T. (2018). Effects of pre-processing methods on Landsat OLI-8 land cover classification using OBIA and random forests classifier. *International Journal of Applied Earth Observation and Geoinformation*, *73*, 170–178. https://doi.org/10.1016/j.jag.2018.06.014

Planet. (2023). *PlanetScope Product Specifications*. Planet. https://assets.planet.com/docs/Planet_PSScene_Imagery_Product_Spec_letter_screen.pdf

Planet. (2024). *Scene Level Normalization and Harmonization of Planet Dove Imagery*. Planet. https://assets.planet.com/docs/scene_level_normalization_of_planet_dove_imagery.pdf

Power, W., Wang, X., Lane, E., & Gillibrand, P. (2013). A Probabilistic Tsunami Hazard Study of the Auckland Region, Part I: Propagation Modelling and Tsunami Hazard Assessment at the Shoreline. *Pure and Applied Geophysics*, *170*(9–10), 1621–1634. https://doi.org/10.1007/s00024-012-0543-z

Pronk, M., Hooijer, A., Eilander, D., Haag, A., De Jong, T., Vousdoukas, M., Vernimmen, R., Ledoux, H., & Eleveld, M. (2024). DeltaDTM: A global coastal digital terrain model. *Scientific Data*, *11*(1), 273. https://doi.org/10.1038/s41597-024-03091-9

Raduszynski, T., & Numada, M. (2023). Measure and spatial identification of social vulnerability, exposure and risk to natural hazards in Japan using open data. *Scientific Reports*, *13*(1), 664. https://doi.org/10.1038/s41598-023-27831-w

Rafliana, I., Jalayer, F., Cerase, A., Cugliari, L., Baiguera, M., Salmanidou, D., Necmioğlu, Ö., Ayerbe, I. A., Lorito, S., Fraser, S., Løvholt, F., Babeyko, A., Salgado-Gálvez, M. A., Selva, J., De Risi, R., Sørensen, M. B., Behrens, J., Aniel-Quiroga, I., Del Zoppo, M., … Hancilar, U. (2022). Tsunami risk communication and management: Contemporary gaps and challenges. *International Journal of Disaster Risk Reduction*, *70*, 102771. https://doi.org/10.1016/j.ijdrr.2021.102771

Raifer, M., Troilo, R., Kowatsch, F., Auer, M., Loos, L., Marx, S., Przybill, K., Fendrich, S., Mocnik, F.-B., & Zipf, A. (2019). OSHDB: A framework for spatio-temporal analysis of OpenStreetMap history data. *Open Geospatial Data, Software and Standards*, *4*(1), 3. https://doi.org/10.1186/s40965-019-0061-3

Rogass, C., Segl, K., Kuester, T., & Kaufmann, H. (2013). Performance of correlation approaches for the evaluation of spatial distortion reductions. *Remote Sensing Letters*, *4*(12), 1214–1223. https://doi.org/10.1080/2150704X.2013.860565

Roy, D. P., Kovalskyy, V., Zhang, H. K., Vermote, E. F., Yan, L., Kumar, S. S., & Egorov, A. (2016). Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sensing of Environment*, *185*, 57–70. https://doi.org/10.1016/j.rse.2015.12.024

Sadashiva, V. K., Wang, X., Lin, S.-L., Lukovic, B., Heron, D. W., & Suppasri, A. (2022). Quantifying effects of explicit representation of buildings in tsunami inundation simulations. *International Journal of Disaster Risk Reduction*, *81*, 103277. https://doi.org/10.1016/j.ijdrr.2022.103277

Samela, C., Coluzzi, R., Imbrenda, V., Manfreda, S., & Lanfredi, M. (2022). Satellite flood detection integrating hydrogeomorphic and spectral indices. *GIScience & Remote Sensing*, *59*(1), 1997–2018. https://doi.org/10.1080/15481603.2022.2143670

Sarkawi, G. M., Feng, L., McCaughey, J. W., Meltzner, A. J., Susilo, S., Muksin, U., Socquet, A., Oktari, R. S., Adi, S. P., Bürgmann, R., & Hill, E. M. (2024). Insights into tectonic hazards since the 2004 Indian Ocean earthquake and tsunami. *Nature Reviews Earth & Environment*. https://doi.org/10.1038/s43017-024-00613-2

Satake, K. (2014). Advances in earthquake and tsunami sciences and disaster risk reduction since the 2004 Indian ocean tsunami. *Geoscience Letters*, *1*(1), 15. https://doi.org/10.1186/s40562-014-0015-7

Scheffler, D., Hollstein, A., Diedrich, H., Segl, K., & Hostert, P. (2017). AROSICS: An Automated and Robust Open-Source Image Co-Registration Software for Multi-Sensor Satellite Data. *Remote Sensing*, *9*(7), 676. https://doi.org/10.3390/rs9070676

Seeger, K., Minderhoud, P. S. J., Peffeköver, A., Vogel, A., Brückner, H., Kraas, F., Nay Win Oo, & Brill, D. (2023). Assessing land elevation in the Ayeyarwady Delta

(Myanmar) and its relevance for studying sea level rise and delta flooding. *Hydrology and Earth System Sciences*, *27*(11), 2257–2281. https://doi.org/10.5194/hess-27-2257-2023

Shaw, R., Izumi, T., Djalante, R., & Imamura, F. (2025). Introduction and Overview of Disaster Risk Landscape: Perspectives of Twenty Years from the Indian Ocean Tsunami of 2004. In R. Shaw, T. Izumi, R. Djalante, & F. Imamura (Eds.), *Two Decades from the Indian Ocean Tsunami* (pp. 1–10). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-2669-4_1

Shean, D. E., Alexandrov, O., Moratto, Z. M., Smith, B. E., Joughin, I. R., Porter, C., & Morin, P. (2016). An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *116*, 101–117. https://doi.org/10.1016/j.isprsjprs.2016.03.012

Shim, D. (2014). Remote sensing place: Satellite images as visual spatial imaginaries. *Geoforum*, *51*, 152–160. https://doi.org/10.1016/j.geoforum.2013.11.002

Sideris, K., Colson, D., Lightfoot, P., Heeley, L., & Robinson, P. (2020). *Review of image segmentation algorithms for analysing Sentinel-2 data over large geographical areas* (No. 655; JNCC Report). JNCC. https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

Sirko, W., Kashubin, S., Ritter, M., Annkah, A., Bouchareb, Y. S. E., Dauphin, Y., Keysers, D., Neumann, M., Cisse, M., & Quinn, J. (2021). *Continental-Scale Building Detection from High Resolution Satellite Imagery* (No. arXiv:2107.12283). arXiv. http://arxiv.org/abs/2107.12283

Sleeter, B. M., Wood, N. J., Soulard, C. E., & Wilson, T. S. (2017). Projecting community changes in hazard exposure to support long-term risk reduction: A case study of tsunami hazards in the U.S. Pacific Northwest. *International Journal of Disaster Risk Reduction*, *22*, 10–22. https://doi.org/10.1016/j.ijdrr.2017.02.015

Smith, A., Bates, P. D., Wing, O., Sampson, C., Quinn, N., & Neal, J. (2019). New estimates of flood exposure in developing countries using high-resolution population data. *Nature Communications*, *10*(1), 1814. https://doi.org/10.1038/s41467-019-09282-y

Somphong, C., Suppasri, A., Pakoksung, K., Nagasawa, T., Narita, Y., Tawatari, R., Iwai, S., Mabuchi, Y., Fujita, S., Moriguchi, S., Terada, K., Athanasius, C., & Imamura, F. (2022). Submarine landslide source modeling using the 3D slope stability analysis method for the 2018 Palu, Sulawesi, tsunami. *Natural Hazards and Earth System Sciences*, *22*(3), 891–907. https://doi.org/10.5194/nhess-22-891-2022

Sugawara, D. (2021). Numerical modeling of tsunami: Advances and future challenges after the 2011 Tohoku earthquake and tsunami. *Earth-Science Reviews*, *214*, 103498. https://doi.org/10.1016/j.earscirev.2020.103498

Sugimoto, M., Iemura, H., & Shaw, R. (2010). Tsunami height poles and disaster awareness: Memory, education and awareness of disaster on the reconstruction for resilient city in Banda Aceh, Indonesia. *Disaster Prevention and Management: An International Journal*, *19*(5), 527–540. https://doi.org/10.1108/09653561011091869

Sundaresan, A., Varshney, P. K., & Arora, M. K. (2007). Robustness of Change Detection Algorithms in the Presence of Registration Errors. *Photogrammetric Engineering & Remote Sensing*, *73*(4), 375–383. https://doi.org/10.14358/PERS.73.4.375

Suppasri, A., Goto, K., Muhari, A., Ranasinghe, P., Riyaz, M., Affan, M., Mas, E., Yasuda, M., & Imamura, F. (2015). A Decade After the 2004 Indian Ocean Tsunami: The Progress in Disaster Preparedness and Future Challenges in Indonesia, Sri Lanka, Thailand and the Maldives. *Pure and Applied Geophysics*, *172*(12), 3313–3341. https://doi.org/10.1007/s00024-015-1134-6

Swanwick, R. H., Read, Q. D., Guinn, S. M., Williamson, M. A., Hondula, K. L., & Elmore, A. J. (2022). Dasymetric population mapping based on US census data and 30-m gridded estimates of impervious surface. *Scientific Data*, *9*(1), 523. https://doi.org/10.1038/s41597-022-01603-z

Syamsidik, Al Farizi, M. D., Tursina, Yulianur, A., Rusydy, I., & Suppasri, A. (2023). Assessing probability of building damages due to tsunami hazards coupled with characteristics of buildings in Banda Aceh, Indonesia: A way to increase understanding of tsunami risks. *International Journal of Disaster Risk Reduction*, *90*, 103652. https://doi.org/10.1016/j.ijdrr.2023.103652

Syamsidik, Tursina, Suppasri, A., Al'ala, M., Luthfi, M., & Comfort, L. K. (2019). Assessing the tsunami mitigation effectiveness of the planned Banda Aceh Outer Ring Road (BORR), Indonesia. *Natural Hazards and Earth System Sciences*, *19*(1), 299–312. https://doi.org/10.5194/nhess-19-299-2019

Tadono, T., Nagai, H., Ishida, H., Oda, F., Naito, S., Minakawa, K., & Iwamoto, H. (2016). GENERATION OF THE 30 M-MESH GLOBAL DIGITAL SURFACE MODEL BY ALOS PRISM. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLI-B4*, 157–162. https://doi.org/10.5194/isprs-archives-XLI-B4-157-2016

Takabatake, T. (2022). Disaster mitigation and the protection of residents. In T. Shibayama & M. Esteban, *Coastal Disaster Surveys and Assessment for Risk Mitigation* (1st ed., pp. 203–228). CRC Press. https://doi.org/10.1201/9781003156161-8

TAKEUCHI, H., Murahsima, Y., Imamura, F., Shuto, N., & Yoshida, K. (2005). *Verification of tsunami run-up height records of Meiji Sanriku Tsunami and Showa Sanriku Tsunami on the coast of Iwate Prefecture using numerical simulation*. https://www.histeq.jp/kaishi_20/26-Takeuchi.pdf

Teluguntla, P., Thenkabail, P. S., Oliphant, A., Xiong, J., Gumma, M. K., Congalton, R. G., Yadav, K., & Huete, A. (2018). A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*, *144*, 325–340. https://doi.org/10.1016/j.isprsjprs.2018.07.017

Tiecke, T. G., Liu, X., Zhang, A., Gros, A., Li, N., Yetman, G., Kilic, T., Murray, S., Blankespoor, B., Prydz, E. B., & Dang, H.-A. H. (2017). *Mapping the world population one building at a time*. https://doi.org/10.48550/ARXIV.1712.05839

Tonini, R., Di Manna, P., Lorito, S., Selva, J., Volpe, M., Romano, F., Basili, R., Brizuela, B., Castro, M. J., de la Asunción, M., Di Bucci, D., Dolce, M., Garcia, A., Gibbons, S. J., Glimsdal, S., González-Vida, J. M., Løvholt, F., Macías, J., Piatanesi, A., … Vittori, E. (2021). Testing Tsunami Inundation Maps for Evacuation Planning in Italy. *Frontiers in Earth Science*, *9*, 628061. https://doi.org/10.3389/feart.2021.628061

Tran, T.-N.-D., Nguyen, B. Q., Vo, N. D., Le, M.-H., Nguyen, Q.-D., Lakshmi, V., & Bolten, J. D. (2023). Quantification of global Digital Elevation Model (DEM) – A case study of the newly released NASADEM for a river basin in Central Vietnam. *Journal of Hydrology: Regional Studies*, *45*, 101282. https://doi.org/10.1016/j.ejrh.2022.101282

Tsuji, Y., Tanioka, Y., Matsutomi, H., Nishimura, Y., Kamataki, T., Murakami, Y., Sakakiyama, T., Moore, A., Gelfenbaum, G., Nugroho, S., Waluyo, B., Sukanta, I., Triyono, R., Namegaya, Y., Earthquake Research Institute, University of Tokyo, Faculty of Natural Sciences, Hokkaido University, Faculty of Engineering and Resource Science, University of Akita, Akita City, Japan, Active Fault Research Center, National Institute of Advanced Industrial Science and Technology, Tsukuba Science City, Ibaragi Ken, Japan, Kansai Electric Power Co. Ltd., Osaka, Japan, … Meteorological and Geophysical Agency, Indonesia. (2006). Damage and Height Distribution of Sumatra Earthquake-Tsunami of December 26, 2004, in Banda Aceh City and its Environs. *Journal of Disaster Research*, *1*(1), 103–115. https://doi.org/10.20965/jdr.2006.p0103

Tsushima, H., Hino, R., Ohta, Y., Iinuma, T., & Miura, S. (2014). tFISH/RAPiD: Rapid improvement of near-field tsunami forecasting based on offshore tsunami data by incorporating onshore GNSS data. *Geophysical Research Letters*, *41*(10), 3390–3397. https://doi.org/10.1002/2014GL059863

Tu, Y.-H., Johansen, K., Aragon, B., El Hajj, M. M., & McCabe, M. F. (2022). The radiometric accuracy of the 8-band multi-spectral surface reflectance from the planet SuperDove constellation. *International Journal of Applied Earth Observation and Geoinformation*, *114*, 103035. https://doi.org/10.1016/j.jag.2022.103035

Tursina, Syamsidik, Kato, S., & Afifuddin, M. (2021). Coupling sea-level rise with tsunamis: Projected adverse impact of future tsunamis on Banda Aceh city, Indonesia. *International Journal of Disaster Risk Reduction*, *55*, 102084. https://doi.org/10.1016/j.ijdrr.2021.102084

US Geological Survey. Advanced National Seismic System (ANSS) comprehensive catalog (US Geological Survey, 2017)

United Nations International Strategy for Disaster Reduction. (2015). *Sendai Framework for Disaster Risk Reduction 2015-2030*. United Nations. https://www.undrr.org/media/16176/download

United Nations International Strategy for Disaster Reduction. (2016). *Tsunami Disaster Risk: Past impacts and projections*. United Nations.

United Nations International Strategy for Disaster Reduction. (2019). *Global Assessment Report on Disaster Risk Reduction 2019*. United Nations.

Vicuña, M., León, J., & Guzmán, S. (2022). Urban form planning and tsunami risk vulnerability: Analysis of 12 Chilean coastal cities. *Environment and Planning B: Urban Analytics and City Science*, *49*(7), 1967–1979. https://doi.org/10.1177/23998083221075635

Wang, X., & Power, W. L. (2011). *COMCOT: A tsunami generation propagation and run-up model*. GNS Science.

Wegmueller, S. A., Leach, N. R., & Townsend, P. A. (2021). LOESS radiometric correction for contiguous scenes (LORACCS): Improving the consistency of radiometry in high-resolution satellite image mosaics. *International Journal of Applied Earth Observation and Geoinformation*, *97*, 102290. https://doi.org/10.1016/j.jag.2020.102290

World Bank. (2024). *World Development Report 2024: The Middle-Income Trap*. Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-2078-6

Wu, L., & Zhang, J. (2021). The effects of human movements on urban climate over Eastern China. *Npj Urban Sustainability*, *1*(1), 36. https://doi.org/10.1038/s42949-021-00038-6

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations: Accurate Global Terrain Elevation map. *Geophysical Research Letters*, *44*(11), 5844–5853. https://doi.org/10.1002/2017GL072874

Yanagisawa, H., Koshimura, S., Miyagi, T., & Imamura, F. (2010). Tsunami damage reduction performance of a mangrove forest in Banda Aceh, Indonesia inferred from field data and a numerical model. *Journal of Geophysical Research: Oceans*, *115*(C6), 2009JC005587. https://doi.org/10.1029/2009JC005587

Yang, M., Hu, Y., Tian, H., Khan, F. A., Liu, Q., Goes, J. I., Gomes, H. D. R., & Kim, W. (2021). Atmospheric Correction of Airborne Hyperspectral CASI Data Using Polymer, 6S and FLAASH. *Remote Sensing*, *13*(24), 5062. https://doi.org/10.3390/rs13245062

Young, N. E., Anderson, R. S., Chignell, S. M., Vorster, A. G., Lawrence, R., & Evangelista, P. H. (2017). A survival guide to Landsat preprocessing. *Ecology*, *98*(4), 920–932. https://doi.org/10.1002/ecy.1730

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N.-E., … Arino, O. (2021). *ESA WorldCover 10 m 2020 v100* (Version v100) [Dataset]. Zenodo. https://doi.org/10.5281/ZENODO.5571936

Zeferino, L. B., Souza, L. F. T. D., Amaral, C. H. D., Fernandes Filho, E. I., & Oliveira, T. S. D. (2020). Does environmental data increase the accuracy of land use and land cover classification? *International Journal of Applied Earth Observation and Geoinformation*, *91*, 102128. https://doi.org/10.1016/j.jag.2020.102128

Zhang, X., Liu, L., Zhao, T., Gao, Y., Chen, X., & Mi, J. (2022). GISD30: Global 30 m impervious-surface dynamic dataset from 1985 to 2020 using time-series Landsat imagery on the Google Earth Engine platform. *Earth System Science Data*, *14*(4), 1831–1856. https://doi.org/10.5194/essd-14-1831-2022

Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X., & Liu, L. (2024). GLC_FCS30D: The first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. *Earth System Science Data*, *16*(3), 1353–1381. https://doi.org/10.5194/essd-16-1353-2024

Zhao, M., Cheng, C., Zhou, Y., Li, X., Shen, S., & Song, C. (2022). A global dataset of annual urban extents (1992–2020) from harmonized nighttime lights. *Earth System Science Data*, *14*(2), 517–534. https://doi.org/10.5194/essd-14-517-2022

Zhou, Q., Zhang, Y., Chang, K., & Brovelli, M. A. (2022). Assessing OSM building completeness for almost 13,000 cities globally. *International Journal of Digital Earth*, *15*(1), 2400–2421. https://doi.org/10.1080/17538947.2022.2159550

Zhou Wang, Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

APPENDIX

APPENDIX 1: Observed inundation depths from the 2004 Indian Ocean Tsunami

No	Latitude	Longitude	H (m)
1	5.5529	95.3190	0.60
2	5.5529	95.3189	1.00
3	5.5603	95.3304	1.00
4	5.5354	95.3053	1.00
5	5.5547	95.3178	1.35
6	5.5625	95.3282	1.50
7	5.5626	95.3286	1.50
8	5.5549	95.3176	1.80
9	5.5576	95.3176	2.00
10	5.5737	95.3453	2.40
11	5.5303	95.2936	2.40
12	5.5615	95.3204	2.60
13	5.5620	95.3221	2.80
14	5.5280	95.2871	2.80
15	5.5603	95.3162	3.00
16	5.5623	95.3166	3.00
17	5.5612	95.3199	3.40
18	5.5781	95.3224	3.40
19	5.5716	95.3294	3.50
20	5.5781	95.3271	4.00
21	5.5361	95.2827	4.00
22	5.5316	95.2822	4.20
23	5.5597	95.3161	4.50
24	5.5649	95.3217	4.50
25	5.5330	95.2775	4.50
26	5.5624	95.3151	5.00
27	5.5637	95.3031	5.00
28	5.5583	95.3107	5.00
29	5.5533	95.2972	5.00
30	5.5752	95.3248	5.00
31	5.5590	95.3154	5.50
32	5.5769	95.3226	6.00
33	5.5615	95.3125	6.20
34	5.5613	95.3113	6.20
35	5.5612	95.3096	7.00
36	5.5548	95.2857	7.00
37	5.5561	95.2838	9.00
38	5.5573	95.2843	9.00
39	5.5624	95.3296	1.40
40	5.5685	95.3355	1.95

No	Latitude	Longitude	H (m)
41	5.5657	95.3350	1.39
42	5.5675	95.3387	1.84
43	5.5738	95.3461	2.60
44	5.5717	95.3433	2.60
45	5.5655	95.3428	1.55
46	5.5634	95.3244	3.20
47	5.5621	95.3233	2.70
48	5.5621	95.3200	3.50
49	5.5505	95.3186	2.57
50	5.5614	95.3278	1.52
41	5.5657	95.3350	1.39
42	5.5675	95.3387	1.84
43	5.5738	95.3461	2.60
44	5.5717	95.3433	2.60
45	5.5655	95.3428	1.55
46	5.5634	95.3244	3.20
47	5.5621	95.3233	2.70
48	5.5621	95.3200	3.50
49	5.5505	95.3186	2.57
50	5.5614	95.3278	1.52
51	5.5610	95.3298	0.89
52	5.5628	95.3316	1.52
53	5.5750	95.3560	1.91
54	5.5701	95.3551	1.80
55	5.5699	95.3536	1.40
56	5.5684	95.3521	1.00
57	5.5766	95.3657	1.30
58	5.5792	95.3624	1.75
59	5.5790	95.3682	1.45
60	5.5812	95.3636	2.00
61	5.5659	95.3444	1.00
62	5.5679	95.3436	1.80
63	5.5625	95.3431	0.90
64	5.5700	95.3374	2.45
65	5.5705	95.3367	2.65
66	5.5666	95.3390	1.00
67	5.5662	95.3392	1.80
68	5.5683	95.3365	2.00
69	5.5839	95.3687	1.80
70	5.5757	95.3660	1.20

No	Latitude	Longitude	H (m)
71	5.5649	95.3245	3.40
72	5.5652	95.3233	3.50
73	5.5676	95.3187	4.60
74	5.5649	95.3171	4.50
75	5.5721	95.3197	7.00
76	5.5624	95.3143	6.00
77	5.5604	95.3105	8.00
78	5.5545	95.3136	2.70
79	5.5565	95.3172	2.30
80	5.5507	95.3167	2.00
81	5.5514	95.3125	3.40
82	5.5485	95.3138	1.80
83	5.5523	95.3113	3.80
84	5.5535	95.2974	5.80
85	5.5460	95.2942	3.90
86	5.5459	95.2997	3.70
87	5.5471	95.3098	2.20
88	5.5442	95.3077	2.20
89	5.5415	95.3010	3.40
90	5.5399	95.3051	2.30
91	5.5306	95.2975	2.00
92	5.5309	95.2955	2.35
93	5.5389	95.2874	3.80
94	5.5807	95.3491	3.70
95	5.5730	95.3627	3.20
96	5.5340	95.2807	5.70
97	5.5971	95.3748	5.50
98	5.5587	95.3021	4.80
99	5.5591	95.3134	3.41
100	5.5592	95.3131	3.24
101	5.5562	95.3065	4.49
102	5.5587	95.3148	2.31
103	5.5589	95.3147	2.42
104	5.5590	95.3138	3.11
105	5.5591	95.3136	3.00
106	5.5587	95.3158	2.03
107	5.5586	95.3155	2.07
108	5.5585	95.3154	2.18
109	5.5547	95.3096	3.43
110	5.5547	95.3096	2.06

No	Latitude	Longitude	H (m)
111	5.5553	95.3048	3.82
112	5.5575	95.3013	3.31
113	5.5575	95.3013	1.60
114	5.5563	95.3051	3.75
115	5.5573	95.3021	4.19
116	5.5573	95.3021	1.49
117	5.5573	95.3021	0.67
118	5.5565	95.3056	3.02
119	5.5572	95.3020	3.39
120	5.5589	95.3143	2.57
121	5.5589	95.3143	1.72
122	5.5589	95.3143	1.04
123	5.5559	95.3050	3.48
124	5.5545	95.3104	2.41
125	5.5545	95.3104	1.45
126	5.5539	95.3126	2.18
127	5.5552	95.3093	2.57
128	5.5545	95.3147	1.03
129	5.5543	95.3117	1.91
130	5.5529	95.3150	0.64
131	5.5550	95.3089	2.32
132	5.5603	95.3174	1.13
133	5.5554	95.3088	2.56
134	5.5420	95.2809	3.85

APPENDIX 2: List of Mw 9.2 hypothetical scenarios in PTHA18, obtained from return-period based method.

No.	Scenario ID in PHTA18 database	Slip model
1	71078	Heterogeneous
2	76250	Heterogeneous
3	76256	Heterogeneous
4	76281	Heterogeneous
5	81293	Heterogeneous
6	81324	Heterogeneous
7	81379	Heterogeneous
8	81471	Heterogeneous
9	81475	Heterogeneous
10	81515	Heterogeneous
11	81564	Heterogeneous
12	86451	Heterogeneous
13	86541	Heterogeneous
14	86604	Heterogeneous
15	86701	Heterogeneous
16	86719	Heterogeneous
17	90392	Heterogeneous
18	90512	Uniform
19	90520	Uniform
20	90561	Heterogeneous
21	90567	Heterogeneous
22	90589	Heterogeneous
23	90605	Heterogeneous
24	90617	Heterogeneous
25	90663	Heterogeneous
26	93888	Heterogeneous
27	94119	Heterogeneous
28	94127	Heterogeneous
29	97142	Uniform
30	97244	Heterogeneous
31	97284	Heterogeneous
32	97291	Uniform
33	97356	Uniform
34	97360	Uniform
35	97380	Uniform
36	97397	Uniform
37	97466	Heterogeneous
38	97492	Heterogeneous
39	100559	Heterogeneous
40	100589	Uniform

No.	Scenario ID in PHTA18 database	Slip model
41	100612	Heterogeneous
42	100626	Heterogeneous
43	100666	Heterogeneous
44	100772	Heterogeneous
45	100791	Uniform
46	100792	Heterogeneous
47	102661	Heterogeneous
48	102799	Uniform
49	102802	Heterogeneous
50	102840	Uniform
51	102853	Heterogeneous
52	102903	Heterogeneous
53	102906	Heterogeneous
54	102946	Uniform
55	104297	Uniform
56	104390	Heterogeneous
57	104403	Heterogeneous
58	104428	Uniform
59	104433	Heterogeneous
60	104512	Heterogeneous
61	104536	Heterogeneous
62	105828	Heterogeneous
63	105933	Heterogeneous
64	105953	Uniform
65	105968	Heterogeneous
66	105974	Uniform
67	105978	Uniform
68	106012	Heterogeneous
69	106021	Heterogeneous
70	106033	Uniform
71	106036	Heterogeneous
72	106143	Heterogeneous
73	106238	Heterogeneous
74	107272	Heterogeneous
75	107441	Uniform
76	107452	Uniform
77	107469	Heterogeneous
78	107487	Heterogeneous
79	107491	Heterogeneous
80	107563	Heterogeneous
81	107580	Uniform
82	107606	Heterogeneous
83	107661	Heterogeneous

No.	Scenario ID in PHTA18 database	Slip model
84	107718	Heterogeneous
85	108490	Uniform
86	108641	Heterogeneous
87	108742	Heterogeneous
88	108744	Uniform
89	108764	Uniform
90	108878	Heterogeneous
91	108921	Heterogeneous
92	108925	Heterogeneous
93	108963	Heterogeneous
94	109015	Uniform
95	109015	Heterogeneous
96	109029	Uniform
97	109041	Heterogeneous
98	109063	Heterogeneous
99	109130	Uniform
100	109153	Heterogeneous
101	109214	Heterogeneous
102	109224	Uniform
103	109256	Heterogeneous
104	109280	Heterogeneous
105	109313	Heterogeneous
106	109990	Heterogeneous
107	110180	Uniform
108	110219	Heterogeneous
109	110326	Heterogeneous
110	110346	Heterogeneous
111	110357	Uniform
112	110410	Heterogeneous
113	110417	Heterogeneous
114	110479	Heterogeneous
115	110514	Heterogeneous
116	110539	Heterogeneous
117	110564	Uniform
118	110590	Uniform
119	110596	Heterogeneous
120	110627	Uniform
121	110808	Heterogeneous
122	111044	Heterogeneous
123	111293	Heterogeneous
124	111508	Heterogeneous
125	111559	Uniform
126	111594	Heterogeneous

APPENDIX 3: Link to access codes applied in this study

1. PTHA18 scenario selection – Return-period based method: https://github.com/GeoscienceAustralia/ptha/blob/master/ptha_access/example_event_access_scripts/multi_site_scenario_selection/example_usage.md
2. PTHA18 scenario selection – Centroid-based filter: https://github.com/GeoscienceAustralia/ptha/tree/nearshore_testing_ptha/misc/nearshore_testing_ptha_2025/ptha18_scenarios_random/set_range_of_mw_and_centroid
3. PTHA18 scenario selection – Stratified-importance sampling: https://github.com/GeoscienceAustralia/ptha/blob/nearshore_testing_ptha/ptha_access/example_event_access_scripts/random_scenarios_non_uniform_and_importance_sampling/random_scenario_sampling.md
4. Orfeo Toolbox (OTB): https://www.orfeo-toolbox.org/ https://zenodo.org/doi/10.5281/zenodo.3522154
5. Python implementation of Orfeo Toolbox (PyOTB): https://pyotb.readthedocs.io/en/stable/ https://github.com/orfeotoolbox/pyotb/tree/develop/pyotb
6. PhoREAL: https://github.com/icesat-2UT/PhoREAL
7. AROSICS: https://zenodo.org/doi/10.5281/zenodo.3742909
8. IR-MAD: https://github.com/latmperkmol/ts-norm?tab=readme-ov-file
9. OBIA-RF: https://github.com/yahcut/OBIA-RF
10. Backdating Building Approach: https://github.com/yahcut/RBF

APPENDIX 4: Link to access the elevation models

1. SRTM v.3.0 1 Arc-Second Global: https://doi.org/10.5066/F7PR7TFT
2. NASADEM Merged DEM Global 1 arc-second V001: https://doi.org/10.5069/G93T9FD9
3. CoastalDEM 3 arc-second v2.1 non-commercial: https://go.climatecentral.org/coastaldem/
4. Copernicus GLO-30: https://doi.org/10.5069/G9028PQB
5. FABDEM V1-2: https://doi.org/10.5523/bris.s5hqmjcdj8yo2ibzi9b4ew3sn
6. DiluviumDEM V.1.0: https://zenodo.org/doi/10.5281/zenodo.8329293
7. EDEM: https://download.geoservice.dlr.de/TDM30_EDEM/
8. CRAWDEM: https://download.geoservice.dlr.de/TDM30_DCM/
9. AW3D30 V.4.1: https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30_e.htm
10. MERIT 3 arc second: https://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/
11. GEBCO 2023 Grid 15 arc-second: https://download.gebco.net/
12. DEMNAS 0.27 arc-second and BATNAS 6 arc-second: https://tanahair.indonesia.go.id/demnas/
13. ATL03 and ATL08 V.006 IceSAT-2: https://nsidc.org/data/icesat-2/data

APPENDIX 5: Link to access input satellite images

1. Landsat 5 TM and Landsat 8 OLI, Collection 2 Level-2: https://www.usgs.gov/ee
2. SPOT 5 images were acquired from Spot World Heritage Data Center (under CNES's Spot World Heritage Programme. Access: https://regards.cnes.fr/user/swh/modules/58
3. Orthorectification of SPOT 5 images: https://swh-2a-carto.fr/
4. The PlanetScope images were accessed from Planet Explorer under Planet's Program for Education and Research (E&R). Access: https://www.planet.com/industries/education-and-research/

APPENDIX 6: Link to access spatiotemporal building datasets

1. Microsoft Building: https://github.com/microsoft/GlobalMLBuildingFootprint
2. OSM building polygon were retrieved from OSHDB and processed using Python's Ohsome-py. Access: https://github.com/GIScience/ohsome-py/tree/master

APPENDIX 7: Link to access built-up area datasets

1. GHS-BUILT: https://human-settlement.emergency.copernicus.eu/download.php
2. GISD30: https://doi.org/10.5281/zenodo.5220816
3. GAIA: https://data-starcloud.pcl.ac.cn/data-nav/remote-sensing
4. GAUD: https://doi.org/10.6084/m9.figshare.16602224.v1
5. GLC FCD30: https://doi.org/10.5281/zenodo.8239305
6. Glance: https://lpdaac.usgs.gov/products/glance30v001/
7. WSF Evolution: https://geoservice.dlr.de/web/maps/eoc:wsfevolution
8. GISA v.2: https://doi.org/10.5281/zenodo.6476661
9. Our generated local dataset: https://github.com/yahcut/OBIA-RF/tree/main/Land%20cover%20map

APPENDIX 8: Link to access gridded population datasets

1. GHS-POP: https://human-settlement.emergency.copernicus.eu/download.php
2. WorldPop: https://hub.worldpop.org/geodata/listing?id=16
3. LSG: https://landscan.ornl.gov/
4. GlobPop: https://doi.org/10.5281/zenodo.10088105
5. Our generated local dataset: https://github.com/yahcut/OBIA-RF/tree/main/Gridded%20Population