Doctoral Dissertation

A Framework for Forecasting High-Impact, Low-Probability Events: A Case Study of Volcanic Tsunami

Elmo Juanara

Supervisor: Prof. Chi Yung Lam

Graduate School of Advanced Science and Technology

Japan Advanced Institute of Science and Technology

Knowledge Science

September, 2025

JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY

Abstract

Creative Society Design Research Area

Graduate School of Advanced Science and Technology

Knowledge Science

A Framework for Forecasting High-Impact, Low-Probability Events: A Case Study of Volcanic Tsunami

By Elmo Juanara

High-Impact, Low-Probability (HILP) disasters are infrequent but extreme events with the capacity to cause substantial harm, requiring proactive measures to mitigate their unpredictable and devastating effects. The impacts of HILP events frequently cascade across sectors, resulting in secondary and tertiary effects. Recent events illustrate the profound challenges posed by HILP disasters. The 2018 Anak Krakatau tsunami struck without warning, highlighting the limitations of current early warning systems in mitigating sudden impacts. On a global scale, the 2020 COVID-19 pandemic showcased how a low-probability event could lead to catastrophic societal and economic disruptions. These examples highlight the devastating impact of HILP events, which is often exacerbated by a lack of preparedness compared to similar-strength events occurring in high-probability zones or predictable timings. Consequently, research on HILP events forecasting is both urgent and essential for advancing disaster preparedness and prevention. The aim of this study is to enhance disaster preparedness and prevention strategies for HILP events. Objectively, proposing a novel framework for forecasting methods and strengthening risk management models to minimize HILP events impact. The significance lies in advancing disaster research by providing the forecasting methods to mitigate the impacts of such unpredictable and devastating disasters.

The methodology for developing the forecasting framework is structured into three phases: Conceptualization, Simulation and Database Generation, and Analysis and Implementation.

Conceptualization. The conceptualization phase establishes the theoretical basis for the framework by critically examining the nature of HILP events. A comprehensive review of significant historical HILP disasters is conducted to identify their characteristics, impacts, and gaps in current preparedness and prevention strategies.

Simulation and Database Generation. Accurate prediction of HILP events necessitates a comprehensive dataset encompassing a wide range of possible conditions. However, due to the infrequent nature of such events, data collection through field surveys or historical records is often inadequate. To address this limitation, simulations are conducted to systematically generate diverse event scenarios by considering various parameters. The resulting simulation outputs are validated against empirical data from historical or real events to ensure their

reliability and accuracy. Once validated, these results are organized into a precomputed database, providing a robust foundation for further phase of analysis and implementation.

Analysis and Implementation. The analysis and implementation phase leverages the precomputed database to develop predictive and decision-making capabilities through following stages. Scenario categorization: Data mining and clustering techniques are applied to group the simulated scenarios based on shared characteristics, enabling efficient pattern recognition and rapid event matching. Predictive modeling: Advanced machine learning techniques are employed to predict complete hazards using limited input from observations. This approach enables the framework to generate full forecasts in real-time, significantly enhancing early warning capabilities. Classification of warning levels: Rule-based reasoning is applied to classify the warning level HILP. These thresholds are calibrated to align with local standards and practical evacuation applications. Performance evaluation: Established metrics from the machine learning domain were used to evaluate the performance of prediction capability, including validation loss, root mean square error (RMSE), F1 score, accuracy, and confusion matrices.

The proposed framework was implemented and evaluated through a case study of the 2018 Anak Krakatau volcanic tsunami, a quintessential example HILP event. Volcanic tsunamis are selected as the case study, owing to their significant threat to rapidly growing coastal populations, their sudden onset without seismic precursors, and the inherent challenges in issuing timely warnings. The designed framework was used to address the forecasting gaps by integrating scenario analysis, predictive modeling, and classification techniques. A total of 1,000 hypothetical tsunami scenarios were simulated by varying collapse parameters, such as collapse volume, dip angle, and direction on coastal and synthetic observation stations. These simulations formed the foundation of a precomputed database. Validation against historical data ensured the simulated scenarios accurately reflected real-world dynamics. The framework integrated the static nature of the pre-computed database with the dynamic demands of real-time disaster response. When new event data, such as the first 9 minutes of tsunami synthetic observations, were received, the system rapidly matched the data to the most similar scenario in the database. To provide actionable insights on coastal observation stations into predefined warning levels: “No Warning,” “Minor Tsunami,” “Tsunami,” and “Major Tsunami,” based on wave height thresholds.

The case study provides an evaluation of the proposed framework's performance. Various machine learning algorithms were applied to predict the maximum amplitude waves at coastal observation stations. Random Forest proved to be the most reliable model, performing best at three out of four coastal stations. Gradient Boosting also showed strong performance at one station but faced challenges in real-time applications due to its computational demands. Neural Networks, while flexible, were less suitable for real-time use because of their high resource requirements and frequent large errors. For waveform forecasting, the LSTM Test model, using full observation data, achieved the lowest error metrics (MSE: 0.0153, MAE: 0.0965), outperforming its more complex variant. However, with shorter input durations (9 minutes), all models experienced significant performance declines, underscoring the impact of data

constraints. Additionally, a decision tree model trained on simulated scenarios successfully classified warning levels with more than 90% accuracy, effectively categorizing events based on wave height thresholds.

The results highlight the potential of integrating precomputed databases with machine learning techniques to enhance forecasting and decision-making for HILP events. The proposed framework contributes to a structured methodology for managing tsunami risks, enabling accurate predictions, effective classification, and timely warning dissemination to mitigate impacts on coastal communities. By generating full forecasts from limited initial observations, the framework demonstrates adaptability to generate early warning with diverse scenarios and conditions.

The method has limitations in its current implementation, as it is specifically tailored to water-related HILP hazards, relying on specific types of data structures. Future work may focus on expanding the framework's scope to include multi-hazard scenarios by integrating diverse data sources. Further efforts could involve adapting the framework to address non-water-related HILP disasters, such as pandemics or nuclear events, which require different dynamics and data structures.

Keywords: high impact low probability, disaster preparedness, volcanic tsunami, simulation, predictive modelling, machine learning, early warning

Declaration of Authorship

I, Elmo Juanara, declare that this thesis titled, “Research on intelligence decision support system for high-consequence low-probability disaster early warning: vase of volcanic tsunamis” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at JAIST.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ELMO JUANARA :

Date: August 25 2025 :

Acknowledgements

I would like to express my heartfelt gratitude to everyone who has supported and guided me throughout my doctoral journey at the Japan Advanced Institute of Science and Technology (JAIST).

First and foremost, I extend my deepest appreciation to my supervisor, Prof. Lam Chi Yung, Ph.D., for his invaluable guidance, expertise, and encouragement, as well as for creating numerous learning opportunities and support that greatly enriched my doctoral studies. I am equally grateful to my second supervisor, Prof. Hideomi Gokon, Ph.D., for his constructive feedback and unwavering support regarding disaster research. I also wish to thank Dr. Mohammad Heidarzadeh from the University of Bath, UK, for his input and discussion on my external minor research project.

I am also indebted to the RIKEN Center for Computational Science, Dr. Iyan Eka Mulia and Dr. Takemasa Miyoshi as a leader of Data Assimilation Team for their support and guidance in my research. Also thank you to the National Institute of Advanced Industrial Science and Technology (AIST), especially Dr. Kawabata Daisaku for providing me with opportunities to enhance my research skills and gain exposure to advanced knowledge in simulation and machine learning applications for volcanic tsunami and slope disaster risk analysis.

My sincere gratitude extends to JAIST, the Iwatani Naoji Foundation, and BAZNAS Indonesia for their generous financial support through research scholarships. These resources have been instrumental in allowing me to conduct extensive research, engage in international collaborations and conferences, and share my findings at key conferences.

I am also deeply thankful to my family for accompanying me during my time in Japan and for their unwavering moral support. Their encouragement has been a constant source of strength and motivation throughout this journey. This dissertation would not have been possible without the collective support of each of you. Thank you for making this achievement a reality.

Finally, all praise and gratitude be to Allah SWT, whose mercy, guidance, and blessings have made this journey possible from beginning to end.

Related Publications

The following peer-reviewed publications were directly derived from the research conducted in this doctoral thesis. Several chapters are based on or adapted from these published works, with appropriate modifications to ensure coherence and continuity within the thesis structure.

1. **Juanara, Elmo,** C.Y., Lam (2025). *Deep learning-based waveform forecasting and physical simulations for volcano collapse tsunami early warning: The Anak Krakatau case*. Ocean Engineering, 335, 121683. https://doi.org/10.1016/j.oceaneng.2025.121683
→ Contributes to Chapter 6, particularly Section 6.2 on deep learning-based waveform prediction and real-time forecasting.

2. **Juanara, Elmo,** C.Y., Lam (2025). *Machine Learning Approaches for Early Warning of Tsunami Induced by Volcano Flank Collapse and Implication for Future Risk Management: Case of Anak Krakatau*. Ocean Modelling, 194, 102497. https://doi.org/10.1016/j.ocemod.2025.102497
→ Forms the basis for Chapter 6, specifically Section 6.1 focusing on tsunami maximum amplitude prediction frameworks using machine learning methods and initial test with short-term input waveform forecasting.

3. **Juanara, Elmo,** C.Y., Lam (2024). *Classification of Non-Seismic Tsunami Early Warning Level Using Decision Tree Algorithm*. Journal of Information Systems Engineering and Business Intelligence, 10(3), 378–391. https://doi.org/10.20473/jisebi.10.3.378-391
→ Provides content for Chapter 6, particularly Section 6.3 on classification of tsunami warning levels using supervised learning.

4. **Juanara, Elmo,** C.Y., Lam (2025). *Impact Assessment of Volcanic Tsunamis in Coastal Regions for Disaster Risk Reduction*. BIO Web of Conferences, 155, 02001. https://doi.org/10.1051/bioconf/202515502001
→ Supports the discussions and findings of impact analysis in Chapter 3 (methodology) and Chapter 4 (case study context and modelling basis).

Table of Contents

List of Figures

List of Tables

List of Equations

Chapter 1: Research Background

1.1 High-Impact Low-Probability (HILP) Events

High-Impact Low-Probability (HILP) events are rare occurrences with catastrophic consequences, posing significant challenges in risk and disaster management. These events defy conventional risk evaluation methods, as their low frequency makes them difficult to predict using historical data and statistical modelling (Lakshita & Nair, 2021). Despite their rarity, the severity of their consequences necessitates careful consideration in risk mitigation and decision-making processes.

HILP events are distinct from high-probability risks, which typically follow recognizable patterns and allow for structured preparedness measures. Conventional risk management frameworks prioritize frequent occurrences with well-documented probabilities, enabling the development of response strategies based on past experiences. However, HILP events operate outside these conventional expectations, requiring alternative approaches to assessment and planning (Aven, 2013; Taleb, 2007).

The increasing complexity of global systems, along with factors such as climate change, technological advancements, and urbanization, has heightened the relevance of HILP events in modern risk analysis. Their potential to cause widespread disruption across multiple sectors highlights the necessity of integrating advanced forecasting methodologies and resilience strategies into disaster preparedness efforts. Addressing the challenges associated with HILP events requires a shift from traditional probability-based models toward comprehensive frameworks that incorporate systemic risk evaluation and proactive planning (Krausmann & Necci, 2021).

Recognizing the significance of HILP events is crucial for improving risk management strategies and enhancing societal resilience. A deeper understanding of their implications can guide decision-makers in developing more adaptive and forward-looking approaches, ensuring that preparedness measures account for even the most unpredictable and severe threats.

1.2 Example of HILP Events

Several historical events exemplify the devastating consequences of HILP occurrences:

- 2004 Indian Ocean Tsunami: Triggered by a magnitude 9.1 earthquake, this tsunami caused over 230,000 fatalities (Latter, 1981) across 14 countries due to the absence of an early warning system (Nations, 2005).
- 1883 Krakatau Eruption: The explosive volcanic eruption generated massive tsunamis, killing over 36,000 people and affecting global atmospheric conditions (Winchester, 2003).

- Deepwater Horizon Oil Spill (2010): An industrial accident with catastrophic environmental and economic consequences, demonstrating the long-term impact of technological failures (Joye, 2015).
- 2011 Great East Japan Earthquake and Tsunami: The disaster led to cascading failures, including the Fukushima Daiichi nuclear meltdown, highlighting the systemic vulnerabilities of modern infrastructure (Ranghieri & Ishiwatari, 2014).
- 2018 Anak Krakatau Eruption and Tsunami: The collapse of a significant portion of Anak Krakatau's southwestern flank triggered a tsunami that struck coastal areas of Indonesia's Sunda Strait, causing over 400 deaths and injuring thousands. Unlike earthquake-induced tsunamis, this event occurred without significant seismic activity, making it more difficult to detect using conventional tsunami warning systems (Grilli et al., 2019).
- COVID-19 Pandemic: A biological HILP event that exposed weaknesses in global health systems, disrupted economies, and intensified social inequalities (Singh et al., 2021).

1.3 Problems Associated with HILP Events

High-Impact Low-Probability (HILP) events present unique problems in risk and disaster management. Due to their characteristics, conventional risk models often struggle to anticipate their full impact. Understanding these problems is essential for developing predictive methodologies and effective disaster response mechanisms.

- Cascading Failure

HILP events not only cause immediate destruction but also generate cascading failures across interconnected systems. Conventional risk models often underestimate their probability, leading to inadequate preparedness. For instance, the 2004 Indian Ocean Tsunami revealed vulnerabilities in global disaster management due to the absence of an early warning system (Nations, 2005). Triggered by a magnitude 9.1 undersea earthquake, the tsunami devastated coastal communities across 14 countries, causing over 230,000 fatalities (Latter, 1981). Similarly, the 1883 Krakatau eruption triggered a tsunami that killed over 36,000 people and caused global atmospheric changes (Winchester, 2003). The 2011 Great East Japan Earthquake and tsunami in 2011 revealed the cascading nature of HILP risks, where the natural disaster led to the Fukushima Daiichi nuclear meltdown, amplifying the scale of the crisis and its long-term effects on public health, the environment, and the economy (Ranghieri & Ishiwatari, 2014). These events were largely unexpected in their full magnitude, underscoring their classification as HILP phenomena.

Moreover, different from localized disasters, they propagate secondary and tertiary failures, worsening vulnerabilities. This phenomenon is particularly evident in technological and infrastructural networks (Lam & Shimizu, 2021), where disruptions in one sector can precipitate failures in multiple others. The 2011 Japan earthquake and tsunami overwhelmed coastal defenses, causing the Fukushima Daiichi nuclear disaster, leading to long-term environmental and economic consequences (Tsuboi et al., 2022). Similarly, cascading effects

can be observed in the Chernobyl nuclear disaster of 1986. The explosion at Reactor No. 4 released a massive amount of radioactive material into the atmosphere, forcing the evacuation of entire communities and rendering large swaths of land uninhabitable (Møller & Mousseau, 2006). However, the crisis did not stop at radiation exposure. The event exposed systemic failures in governance, crisis communication, and technological oversight within the Soviet Union. The mismanagement of information, combined with the government's delayed response, led to widespread public distrust, economic strain, and long-term health consequences.

- Occur Sporadically

HILP events occur sporadically, making them difficult to incorporate into standard risk models. Their infrequency results in limited historical data, hindering accurate probabilistic predictions (Masys et al., 2014). The unpredictability of HILP events is further exacerbated by evolving environmental, geological, and technological systems. Even if historical records exist, they may not provide sufficient insight into the next extreme occurrence due to changes in underlying conditions. For example, the 1883 Krakatau eruption was not merely a volcanic event but a complex geological interaction leading to widespread tsunamis and climatic disruptions (Izumi, 1981). At the time, the scientific understanding of volcanic activity and its secondary effects was very limited, and there were no such warning systems in place to mitigate the impact. The 2004 Indian Ocean Tsunami, despite historical precedents, lacked a dedicated warning system due to the misconception that absence of recent disasters implied low risk (Helbing, 2013).

- Global Consequences and Multi-Generational

Furthermore, the cascading nature of HILP events is not limited to natural or technological disasters but extends to biological crises as well. The COVID-19 pandemic illustrated how a single biological event could disrupt global systems on an unprecedented scale. The initial outbreak of the virus in Wuhan, China, quickly escalated into a worldwide health emergency, overwhelming healthcare systems and forcing governments to implement drastic containment measures. However, the secondary effects of the pandemic proved equally catastrophic. Global supply chains collapsed due to border closures and factory shutdowns, financial markets experienced severe volatility, and social inequalities were exacerbated as marginalized communities bore the brunt of economic downturns (Singh et al., 2021). The pandemic underscored the deeply interconnected nature of modern societies, where disruptions in one domain can reverberate across multiple sectors, compounding the overall crisis.

HILP events often exceed coping capacities, leaving multi-generational impacts. Different from recurrent disasters where preparedness measures mitigate harm, HILP events overwhelm response mechanisms. The 2004 Indian Ocean Tsunami killed over 230,000 people due to inadequate coastal preparedness. The 2011 Great East Japan Earthquake caused economic losses exceeding $235 billion, disrupting industries and leading to nuclear contamination (Lee et al., 2012). Similarly, the Chernobyl nuclear disaster of 1986 serves as a good example, where a single explosion resulted in radiation contamination over thousands of square kilometers,

rendering entire regions uninhabitable for centuries (International Atomic Energy Agency, 2005).

- Long-Term Economic and Environmental

Economic consequences of HILP events persist long-term, frequently amounting to hundreds of billions of dollars in direct and indirect losses. COVID-19 pandemic inflicted one of the worst global recessions in modern history as extreme consequences, halting international trade, triggering mass unemployment, and forcing governments to inject unprecedented financial stimulus to stabilize economies (Nicola et al., 2020). Contrary to localized disasters, the economic consequences of HILP events often persist for years or even decades, requiring structural adjustments at both national and global levels to recover from the devastation. Another example, the Deepwater Horizon oil spill in 2010, while an industrial accident, had HILP characteristics in its scale and impact, releasing 4.9 million barrels of oil into the Gulf of Mexico, devastating marine ecosystems, and disrupting coastal economies reliant on fishing and tourism (Joye, 2015). Environmental HILP events have long-term ecological consequences, often requiring multi-decade recovery efforts that are costly, uncertain, and sometimes ineffective.

- Decision-Making and Preparedness

Additionally, HILP events introduce high uncertainty, complicating decision-making. Policymakers often struggle with incomplete information, making risk assessment and response planning difficult. This was particularly evident during the COVID-19 pandemic, where the global spread of the virus overwhelmed healthcare systems, disrupted economies, and exposed critical weaknesses in global governance structures (WHO, 2025). The 2004 Indian Ocean Tsunami is a prime example—before the disaster, tsunami detection infrastructure in the region was minimal, as the probability of a large-scale event was perceived as low (Nations, 2005). Only after 230,000 deaths and massive destruction did significant investments in early warning systems materialize. A similar pattern occurred with the Fukushima nuclear disaster in 2011, where pre-existing safety standards were based on outdated seismic risk models, underestimating the potential for an earthquake-induced tsunami (Funabashi & Kitazawa, 2012). These cases highlight how traditional risk assessment methods struggle to incorporate low-probability, high-impact scenarios.

Coupled with this, investing in preparedness for rare but catastrophic events is challenging, as immediate risks tend to take priority over uncertain threats. Risk communication during HILP events is also critical—delayed or opaque messaging can exacerbate crises. As a result, investments in disaster mitigation for HILP events tend to be reactionary rather than proactive. For example, it was only after the devastation of the 2004 Indian Ocean Tsunami that significant investments were made in regional tsunami early warning systems (Angove et al., 2019). Furthermore, Chernobyl authorities initially downplayed the severity of the disaster, leading to panic and long-term mistrust in nuclear governance (International Atomic Energy Agency, 2005). During the COVID-19 pandemic, governance failures in crisis response, with

many governments facing backlash over delayed lockdown measures, vaccine distribution inequalities, and misinformation management (Sharifi & Khavarian-Garmsir, 2020).

- Forecasting and Early warning

Moreover, HILP events also challenge early warning efforts. Many extreme disasters exhibit precursor signals that could potentially be used for forecasting, but because these signals are rarely observed, they are often overlooked or misinterpreted. For example, while seismologists monitor fault lines for stress accumulation, accurately predicting the precise timing and magnitude of an earthquake remains elusive due to the irregular intervals at which major seismic activity occurs. Similarly, pandemics such as COVID-19 emerge through complex interactions between ecological, virological, and human behavioral factors, making their exact onset difficult to anticipate. Despite prior warnings about the potential for global pandemics, including the SARS and MERS outbreaks, the world was still largely unprepared for the scale and speed of COVID-19's spread (Filip et al., 2022).

1.4 Research question

Understanding HILP events requires a structured research approach that accounts for their complexity, rarity, and cascading effects. This study addresses three main research questions, each corresponding to a critical phase in the analytical process: contextual grounding, scenario-based simulation, and predictive forecasting for early warning.

1.4.1 Research question 1. How can historical case studies and contextual factors inform the modelling basis and scenario design for HILP events?

This question aims to connect the historical understanding of HILP events with the practical setup of simulations. It examines how key characteristics, triggers, and environmental contexts of past events can inform the selection of modelling parameters, initial conditions, and scenario configurations, forming a coherent basis for simulation-based studies.

1.4.2 Research question 2. What simulation strategies and parameterization methods are effective for generating representative scenario databases for HILP events?

This question explores the methodological design for simulating HILP events, including scenario generation, parameter selection, and model validation. It seeks to establish a reproducible approach for developing synthetic databases to support further analysis.

1.4.3 Research question 3. How can predictive modelling be utilized to improve real-time forecasting and early warning for HILP events?

This question focuses on the development and application of data-driven predictive models to anticipate the evolution of HILP events. It emphasizes the use of machine learning techniques and precomputed scenario databases to generate timely forecasts and strengthen early warning systems, particularly in situations with limited observational data or time-critical constraints.

1.5 Aim and objectives

The aim of this study is to develop an integrated research framework for forecasting high-impact, low-probability (HILP) events by combining contextual understanding, simulation-based scenario generation, and data-driven predictive modeling. This framework is designed to improve hazard detection and support early warning systems through structured scenario analysis and machine learning–based forecasting models.

The study seeks to achieve the following objectives:

- To establish the contextual foundation and modeling basis for HILP event simulation by analyzing relevant historical cases and identifying key environmental, physical, and hazard-specific parameters that inform scenario design.
- To generate and validate a synthetic database of HILP scenarios through systematic simulation techniques, incorporating variations in initial conditions and evaluating the representativeness of the modeled outputs against real-world observations.
- To develop and evaluate data-driven forecasting models using machine learning approaches trained on the precomputed database, with the goal of enabling rapid and reliable early warning based on short-term input signals.

1.6 Significance

This study offers several innovative contributions:

- Dataset Generation for Predictive Modelling: a novel dataset consisting of 1,000 simulated HILP scenarios will serve as the foundation for predictive modelling, enabling faster and more accurate early warnings.
- Integration of Physical Simulation with Machine Learning Techniques: unlike existing frameworks that rely solely on physical simulations and modelling, this research integrates AI-driven approaches to enhance predictive capabilities.
- Quick Prediction with Short-Term Input for Early Warning: by leveraging a well-trained predictive model, this approach enables rapid early warning generation using short-term observations or limited initial input, addressing critical challenges in disaster response.
- Application to Volcanic Collapse Tsunami Risk: while tsunami forecasting has traditionally focused on seismic sources, this study extends risk modelling to volcanic collapse-induced tsunamis, an area that remains less understood.

1.7 Contribution

This study provides significant contributions to both theoretical and practical aspects of High-Impact Low-Probability (HILP) event management.

Theoretically, it develops a conceptual framework based on historical analyses of HILP events, identifying their key characteristics, impacts, and gaps in risk assessment. By synthesizing insights from disaster science, predictive analytics, and decision-making frameworks, this research expands the understanding of HILP risks and their systemic implications.

Methodologically, the study proposes a structured approach to simulating HILP events, facilitating the development of a comprehensive dataset for risk assessment and scenario analysis. It also introduces a data-driven methodology that enhances forecasting accuracy by simulation models and integrating machine learning.

From a practical perspective, this research proposed early warning capabilities by incorporating predictive modelling and decision-support systems for disaster response. Furthermore, it provides a structured framework that can be utilized by policymakers, disaster management agencies, and researchers to strengthen preparedness and mitigation strategies for HILP events.

1.8 Structure of dissertation

This dissertation is organized into seven chapters, each addressing an important aspect of the study on HILP events.

Chapter 1: Research Background. This chapter provides an overview of the study, outlining the research problem, aim and objectives, significance, and contribution of the study. It also presents the research questions, expected contributions, and the overall structure of the dissertation.

Chapter 2: Literature Review. This chapter reviews relevant literature on HILP events, risk assessment methodologies, predictive modeling techniques, and trend of early warning systems. It identifies theoretical frameworks, historical case studies, and existing gaps that this research aims to address.

Chapter 3: Methodology. This chapter details the research design, data sources, and methodological approaches used in the study. It explains the rationale behind the selection of simulation techniques, machine learning models, and decision-support frameworks for forecasting HILP events.

Chapter 4: Case Study Context and Modelling Basis. This chapter develops a conceptual framework for understanding HILP events based on historical analyses. It defines their characteristics, impacts, and the challenges in risk assessment, serving as the foundation for the simulation and predictive modelling processes.

Chapter 5: Simulation and Database Generation. This chapter focuses on the development of simulation models to generate synthetic datasets representing various HILP scenarios. It discusses the modelling techniques, software tools, and validation processes used to ensure the accuracy and reliability of the generated data.

Chapter 6: Analysis and Implementation. This chapter presents the analysis of the simulation results and the implementation of predictive models. It evaluates the performance of different forecasting techniques and decision-support systems, assessing their effectiveness in improving early warning and disaster preparedness.

Chapter 7: Discussion and Conclusion. The final chapter synthesizes the findings of the research, discussing their theoretical and practical implications. It also addresses the limitations of the study and provides recommendations for future research and policy development in HILP risk management.

This structured approach ensures a logical progression of ideas, from problem identification and theoretical foundation to methodological execution and practical application, notably contributing to a novel framework for understanding and managing HILP events.

Chapter 2: Literature review

2.1 Introduction

High-Impact Low-Probability (HILP) events, such as large-scale tsunamis, volcanic eruptions, pandemics, and industrial disasters, pose significant challenges in disaster risk reduction. Understanding and modeling these events require a thorough review of existing methodologies, their limitations, and potential improvements. This chapter provides a structured review of literature on HILP events, focusing on conceptualization, simulation and database generation, and analysis and implementation approaches. It identifies research gaps and links existing studies to the current research on volcanic tsunami early warning systems.

2.2 Definition and Existing Research on HILP Events

Understanding HILP events requires a systematic review of existing definitions, research efforts, and analytical frameworks. This section explores how HILP events have been conceptualized in the literature, highlights key studies that address these events, and examines existing methodological frameworks used to analyze them. The objective is to identify gaps in current approaches and establish a framework for forecasting the HILP events.

2.2.1 Definition of HILP from the Literature

HILP represents events that are rare yet catastrophic, demanding significant attention in disaster risk reduction. These events, such as mega-tsunamis, volcanic eruptions triggering tsunamis, global pandemics, and catastrophic nuclear accidents, challenge conventional risk management frameworks due to their unpredictability and systemic consequences. Unlike high-probability events, which often follow established patterns and most people are prepared (Velasco et al., 2023), HILP events defy conventional risk assessment models. The 2011 Great East Japan Earthquake serves as a stark example, where a massive tsunami not only caused widespread destruction and loss of life but also triggered the Fukushima Daiichi nuclear disaster, leading to cascading failures across environmental, technological, and societal systems (Mimura et al., 2011).

What makes HILP events particularly urgent is their capacity to disrupt interconnected systems on a global scale, far exceeding the scope of more frequent risks. Despite their rarity, their potential for devastating impacts highlights the need for innovative approaches in forecasting, preparedness, and governance. As global interdependencies and environmental challenges grow, prioritizing the study and mitigation of HILP risks is essential to building resilient societies and safeguarding against future catastrophes. HILP events are defined by their rarity, unpredictability, and extreme consequences, making them fundamentally different from more frequent risks.

2.2.2 Existing Research on HILP

Several studies have examined different types of HILP events, their research problems, and methodologies:

- Tsunamis and Earthquakes: Research on events like the 2004 Indian Ocean tsunami and 2011 Tohoku earthquake has primarily focused on numerical simulations, GPS-based real-time tsunami warning systems, and probabilistic seismic hazard assessments (Angove et al., 2019; Titov et al., 2005). Despite these advancements, uncertainties remain in predicting tsunami amplification and near-shore impacts.
- Volcanic Eruptions and Tsunami Generation: Studies on Anak Krakatau's 2018 tsunami have employed hydrodynamic modeling and slope stability assessments to understand landslide-generated tsunami mechanisms (Giachetti et al., 2012; Selva et al., 2021). However, the complexity of underwater landslides continues to pose challenges in forecasting wave heights accurately.
- Pandemics and Other HILP Events: Advancements are being made in pathogen detection and source tracing (Kamalrathne et al., 2020), enhanced biosurveillance capabilities (Creppage et al., 2024), real-time genomic surveillance (Struelens et al., 2024), and nasal host response-based screening for identifying undiagnosed respiratory viruses following the COVID-19 pandemic. However, the pandemic has also revealed the limitations of existing epidemiological models, emphasizing the need for dynamic, AI-driven forecasting techniques for HILP scenarios (Ferguson et al., 2020).

While significant progress has been made across these domains, key challenges persist. Uncertainties in hazard prediction, the complexity of multi-hazard interactions, and limitations in real-time forecasting highlight the need for integrated approaches.

2.2.3 Existing Frameworks for Studying HILP

Various analytical frameworks have been developed to study HILP events, each with distinct methodologies and applications in disaster risk assessment, mitigation, and response. These frameworks aim to address the inherent unpredictability of HILP events by incorporating probabilistic modeling, scenario analysis, and machine learning-based predictions.

The table 2.1 below summarizes the limitations of the existing frameworks:

Table 2.1 The limitation of current framework

Study and Year	Framework	Limitation
Behrens et al., 2021; Grezio et al., 2017	Probabilistic Risk Assessment (PRA)	Limited by historical data, underestimates rare events
Geist & Parsons, 2006; Hong Nguyen et al., 2014	Scenario-Based Approaches	Computationally expensive, requires extensive scenarios
Grossi, 2005	Catastrophe Modeling (CAT Models)	Simplified hazard assumptions, may not capture full complexity

Krivorotko et al., 2022; Rodríguez et al., 2022	Agent-Based Modeling (ABM)	Requires calibration, high computational cost
Fauzi & Mizutani, 2020; Liu et al., 2021	Machine Learning-Based Prediction Models	Requires large labelled datasets, risk of model biases

- *Probabilistic Risk Assessment (PRA)*

Probabilistic Risk Assessment (PRA) is a widely used methodology in disaster risk management, particularly in nuclear safety, earthquake engineering, and extreme event modeling (Behrens et al., 2021; Grezio et al., 2017). PRA estimates the likelihood of different disaster scenarios by incorporating uncertainties into risk calculations. This approach is beneficial for evaluating system reliability and failure probabilities under extreme conditions.

Limitations: Relies heavily on historical data, which may be insufficient for HILP events, leading to underestimation of rare catastrophic occurrences.

- *Scenario-Based Approaches*

Scenario-based frameworks simulate potential disaster outcomes by modeling extreme event scenarios. This approach has been widely used in tsunami, earthquake, and climate change studies to estimate potential impacts (Geist & Parsons, 2006; Hong Nguyen et al., 2014). Scenario-based models help in disaster preparedness planning by generating multiple what-if scenarios to assess the consequences of different hazard conditions.

Limitations: Computationally intensive, requires extensive scenario generation, and does not always capture the full range of possible disaster events.

- *Catastrophe Modeling (CAT Models)*

Catastrophe modeling, commonly applied in insurance and financial risk assessment, uses statistical models to simulate extreme events and estimate their potential losses (Grossi, 2005). CAT models incorporate hazard, exposure, and vulnerability data to quantify financial risks associated with HILP events.

Limitations: Assumes simplifications in hazard dynamics, may not fully capture the complexity of real-world disasters.

- *Agent-Based Modeling (ABM)*

Agent-based modeling (ABM) is used to simulate the behavior of individuals, organizations, and infrastructure networks under extreme event conditions. This framework has been applied to study human evacuation dynamics, supply chain disruptions, and disease spread during pandemics (Krivorotko et al., 2022; Rodríguez et al., 2022).

Limitations: Requires extensive calibration, computationally expensive for large-scale simulations.

- *Machine Learning-Based Prediction Models*

Recent advancements in artificial intelligence (AI) and machine learning (ML) have introduced new frameworks for predicting HILP events. ML-based models utilize vast datasets to detect patterns and predict potential hazards, such as tsunami propagation, earthquake aftershocks, and climate-related disasters (Fauzi & Mizutani, 2020; Liu et al., 2021).

Limitations: Requires large amounts of labeled data, vulnerable to model biases and overfitting.

2.3 Integrated Framework for Forecasting HILP Events

While existing frameworks provide valuable insights into different aspects of HILP event forecasting, none comprehensively address the challenges of volcanic tsunamis. These events require a multi-faceted approach that includes conceptualization to understand key mechanisms and hazard dynamics, simulation and database generation to develop realistic scenarios using numerical and AI-driven approaches, and analysis and implementation to deploy decision-support systems for real-time forecasting and mitigation.

As outlined in Chapter 1, the aim and objectives of this study focus on developing a Framework for Forecasting High-Impact, Low-Probability Events: A Case Study of Volcanic Tsunamis, which integrates these three components to enhance early warning systems. The following sections will explore each stage of the proposed framework in detail, demonstrating how it addresses the gaps identified in existing methodologies.

2.3.1 Context and Modelling Basis of HILP Events

Context and Modelling Basis HILP events requires a structured approach to understanding their defining characteristics, underlying mechanisms, and implications for disaster management. These events pose challenges due to their unpredictability, complex interactions with social and environmental systems, and the limitations of traditional risk assessment methods. To systematically analyze HILP events, researchers have proposed method that incorporate hazard characterization, vulnerability assessment, and adaptive decision-making.

Several HILP events have been extensively studied, not as isolated incidents but as part of broader research into disaster risk reduction. These case studies highlight methodological approaches used and analyzing findings as well as method deficiencies (see Table 2.2).

Table 2.2 Previous studies that focus on characteristics, mechanism, and implication of HILP

Disaster	Study and Year	Methodology	Key Findings	Method Deficiencies
2004 Indian Ocean Tsunami	Syamsidik et al., (2021)	Qualitative case study, literature review, policy document analysis	Rapid early recovery but slow long-term progress; disaster education improved yet engagement declined	Lacks new empirical data; limited generalizability; lacks quantitative validation
2011 Great East Japan	Nakasu et al., (2018)	Mixed-method forensic	High mortality due to poor land-use planning, inadequate	Single-case focus; difficult to quantify cultural beliefs; does

Earthquake & Tsunami		analysis (quantitative disaster data, qualitative interviews)	tsunami defenses, and false risk perceptions	not assess policy improvements
2018 Anak Krakatau Eruption & Tsunami	Firdaus et al., (2023)	Mixed qualitative approach (field observations, interviews, literature review)	Communities relied on natural omens; need to integrate scientific and local knowledge for effective preparedness	Small sample size; mostly qualitative—no quantitative resilience assessment; limited generalizabilit
COVID-19 Pandemic	Chang et al., (2022)	Structured literature review, comparative analysis, conceptual framework development	Economic disruptions were demand-driven; small businesses and supply chains were highly vulnerable; traditional disaster relief methods were ineffective	Review-based; lacks empirical data; limited focus on developing nations

The studies analyzed reveal recurring patterns and challenges in disaster risk reduction across different types of disasters. The 2004 Indian Ocean Tsunami case underscores the importance of sustained long-term recovery efforts, highlighting the tendency for disaster education and engagement to decline over time. This suggests that disaster management should not only focus on immediate relief but also on maintaining resilience through continuous education and policy enforcement.

The 2011 Great East Japan Earthquake and Tsunami study emphasizes the role of land-use planning and infrastructure in mitigating disaster impacts. It also highlights the dangers of over-reliance on structural defenses without adequate public awareness and risk perception. The findings indicate that effective disaster risk reduction must integrate both engineering solutions and public education to ensure effective evacuation strategies.

For the 2018 Anak Krakatau eruption and tsunami, the study points to the need for combining scientific and indigenous knowledge in early warning systems. Many affected communities relied on traditional signs that were ineffective for this type of tsunami, illustrating the necessity of culturally adapted disaster communication strategies that incorporate modern scientific findings while respecting local perspectives.

The COVID-19 pandemic study showcases the distinct nature of economic disruptions in disasters. Unlike physical destruction, pandemic-induced disruptions are largely demand-driven, requiring different recovery strategies. Traditional disaster relief mechanisms such as

infrastructure rebuilding were less applicable, demonstrating the need for flexible, interdisciplinary approaches that can address both health and economic crises.

2.3.2 Simulation and Database Generation

HILP events such as tsunamis and pandemics require a deep understanding that cannot rely solely on historical data. Therefore, simulation and database generation become key elements in disaster mitigation and response. Table 2.3 shown in depth reviews various studies that have utilized simulations and database generation for four major HILP events: the 2004 Indian Ocean Tsunami, the 2011 Great East Japan Earthquake and Tsunami, the 2018 Anak Krakatau Eruption and Tsunami, and the COVID-19 Pandemic.

Table 2.3 Previous studies that focus on simulation and database generation of HILP

Disaster	Study and Year	Simulation Methodology	Database Generation	Deficiencies
2004 Indian Ocean Tsunami	Giles et al., (2021), Satake, (2014), Titov et al., (2005)	Nonlinear shallow water equations (NSWE) for tsunami propagation; Precomputed tsunami scenario databases	Post-event surveys; Remote sensing; Precomputed tsunami scenarios for future prediction	Lack of real-time tsunami data in 2004; Precomputed scenarios may not capture extreme outlier events; Model uncertainty due to limited historical data
2011 Great East Japan Earthquake and Tsunami	Koshimura & Shuto, (2015), Mulia et al., (2022), Suppasri et al., (2021)	NSWE with high-resolution urban-scale simulations; Full 3D hydrodynamic models for tsunami-structure interactions	Tide gauge and DART buoy data; Seismic network recordings; Post-disaster field surveys; Expanded tsunami scenario databases post-2011	Initial warning underestimated wave height; Precomputed databases failed to include larger-than-expected earthquakes; Challenges integrating multi-hazard impacts (nuclear accident, fires)
2018 Anak Krakatau Eruption and Tsunami	Grilli et al., (2019), Syamsidik et al., (2020)	Coupled landslide and tsunami simulations; 3D Navier-Stokes	Satellite and aerial imagery; Field surveys; Volcano monitoring	No real-time early warning for volcanic collapse tsunamis; High

		models for collapse-triggered tsunami	(eruption records)	uncertainty in landslide volume estimates; No precursory signals detected before collapse
COVID-19 Pandemic	Dong et al., (2020), Lutz & Giabbanelli, (2022)	Compartmental SEIR models; Agent-based models (ABM); Network and Monte Carlo simulations	JHU COVID-19 dataset (daily case tracking); Mobility and genomic databases; Public health reporting and surveillance	Limited real-time data in early pandemic phases; Data inconsistencies across regions; Difficulty modeling human behavior shifts; Model generalization issues with new variants

The 2004 tsunami, one of the deadliest in recorded history, prompted significant advancements in tsunami simulation and warning systems. Studies by Titov et al., (2005) and Satake (2014) utilized nonlinear shallow water equations (NSWE) to model tsunami propagation. One of the key developments following this disaster was the creation of precomputed tsunami scenario databases, which allow for rapid forecasting in future tsunami events. Data collection relied heavily on post-event field surveys, as well as remote sensing techniques. However, a major deficiency identified was the lack of real-time tsunami data and short-term prediction, as no deep-sea monitoring system was in place at the time. Additionally, while precomputed scenarios improved forecasting, they inherently could not capture extreme outlier events, leading to uncertainty in model predictions due to the limited availability of historical data.

The 2011 tsunami event was one of the most well-documented disasters, allowing for highly refined tsunami simulations. Researchers such as Suppasri et al. (2021) and Koshimura et al. (2015) utilized high-resolution NSWE models with detailed urban-scale inundation mapping. Some studies incorporated 3D hydrodynamic models to evaluate the impact of tsunamis on infrastructure. Data collection for this event was extensive, with Japan's tide gauge and DART buoy networks recording detailed tsunami waveforms. Despite these advancements, deficiencies arose in the early warning system. The initial tsunami forecast underestimated wave heights, leading to inadequate evacuation measures. Furthermore, the precomputed scenario database approach failed to account for an earthquake of such magnitude, as the event exceeded previous assumptions about fault size and slip potential. Another significant challenge was the integration of cascading hazards, as the tsunami triggered fires, industrial

accidents, and the Fukushima nuclear crisis, which were beyond the scope of most tsunami-focused models.

The 2018 tsunami, triggered by a volcanic flank collapse rather than an earthquake, presented unique modeling challenges. Studies by Grilli et al. (2019) and Syamsidik et al. (2020) employed coupled landslide-tsunami simulations, incorporating 3D Navier-Stokes models to simulate the tsunami's unusual generation mechanism. Unlike earthquake-driven tsunamis, no precomputed database of volcanic collapse scenarios existed prior to this event, necessitating post-event scenario development. Data collection relied on satellite imagery, aerial surveys, and volcano monitoring records, yet a significant limitation was the lack of real-time early warning systems for such tsunamis. No precursory signals were detected before the collapse, and the high uncertainty in estimating the volume of the landslide made precise modeling difficult. This event highlighted the need for integrated volcano monitoring with tsunami warning systems to anticipate similar future disasters.

Unlike geophysical disasters, the COVID-19 pandemic required epidemiological simulations rather than physical wave modeling. Studies by Lutz & Giabbanelli (2022) and Barbano et al. (2023) employed compartmental SEIR models, agent-based models (ABM), and Monte Carlo network simulations to analyze virus spread. Data collection was vast and multifaceted, with the JHU COVID-19 dataset tracking daily cases in over 3,500 locations, alongside mobility data from tech companies, genomic sequencing databases, and public health reports. Despite these advancements, limitations emerged, particularly in real-time data quality during the early pandemic phases, leading to inconsistent reporting across regions. Furthermore, early epidemiological models struggled with generalization issues, as the emergence of new variants rendered previous assumptions outdated, necessitating frequent recalibrations.

Across these four HILP events, advancements in simulation methodologies and database generation have significantly enhanced disaster prediction and preparedness. Tsunami research has benefited from precomputed scenario databases, allowing for rapid forecasting, while pandemic modeling has leveraged real-time epidemiological data to adapt response strategies dynamically. However, key challenges remain, including uncertainty in outlier scenarios, failure to account for extreme or unforeseen events, and limitations in real-time data availability during early phases of disasters. Addressing these gaps will be crucial for improving database generation for HILP events.

2.3.3 Analysis and Implementation

High-Impact Low-Probability (HILP) events, such as massive tsunamis and pandemics, have driven research into machine learning (ML) for predictive modeling and early warning as part of the analysis and implementation of forecasting frameworks. As shown in Table 2.4, previous studies for each event are reviewed, focusing on ML methodologies and their limitations. Additionally, their deficiencies are highlighted to identify existing gaps in current approaches.

Table 2.4 Previous studies that focus on predictive modeling of HILP using machine learning

Disaster	Study and Year	Methodology	Deficiencies

2011 Great East Japan Earthquake and Tsunami	Mulia et al., (2022) & Suppasri et al., (2021b)	Deep learning models trained on offshore ocean-bottom sensor data to predict coastal tsunami heights in seconds; Probabilistic tsunami forecasting using Bayesian networks; AI-enhanced inundation mapping.	Underestimation of tsunami height in warnings due to reliance on earthquake magnitude; ML models less accurate for smaller tsunami events; model generalization issues and need for validation across real-world scenarios.
2018 Anak Krakatau Eruption and Tsunami	Ratnasari et al., (2023) & Ren et al., (2020)	Development of an early warning system integrating real-time monitoring and precomputed scenarios for tsunamis triggered by the collapse of Anak Krakatau volcano; Use of virtual observation stations around the source for rapid tsunami height	Limited observational data to fully validate real-time predictions; Dependence on precomputed scenarios that may not capture all real-world variations; Lack of collapse direction parameters.
COVID-19 Pandemic	Gao et al., (2020), Tiwari et al., (2023), Zhang et al., (2022)	Hybrid time-series forecasting (ARIMA + neural networks) for case prediction; Social media-based early warning classifiers using Twitter data (SVM, CNN, BERT); ML models predicting COVID-19 severity using clinical data.	Poor generalization of ML models across regions and pandemic phases; lack of adaptability to new virus variants and policy changes; biases in social media data; issues with ML interpretability in clinical applications.

The 2011 Great East Japan Earthquake and Tsunami revealed major deficiencies in early warning systems, particularly in accurately estimating tsunami heights. Mulia et al. (2022) and Suppasri et al. (2021) introduced deep learning models trained on offshore ocean-bottom sensor data, which enable real-time tsunami height predictions along coastlines within seconds. Probabilistic tsunami forecasting using Bayesian networks further enhanced the classification of tsunami severity levels. Additionally, predictive modelling-enhanced inundation mapping

was explored to improve evacuation strategies. However, despite these advancements, early warnings underestimated tsunami heights due to reliance on seismic magnitude alone, leading to delayed or insufficient responses. ML models also exhibited lower accuracy for smaller tsunami events, raising concerns about generalization. Furthermore, the need for extensive real-world validation remains a challenge to ensure reliability in practical scenarios.

For the 2018 Anak Krakatau Eruption and Tsunami, the event posed unique challenges as it was triggered by a volcanic flank collapse, meaning traditional earthquake-based tsunami warning systems failed. Rinda et al. (2023) and Ren et al. (2023) introduced predictive modelling early warning systems integrating real-time monitoring and precomputed scenarios. These models incorporated virtual observation stations around the volcano, enabling rapid tsunami height predictions. However, limitations persist, including limited scenarios, dependence on precomputed scenarios that may not cover all real-world variations, and lack of collapse direction parameters.

The COVID-19 Pandemic presented a different type of HILP event, where ML was extensively used for epidemiological forecasting and healthcare decision-making. Gao et al. (2020), Weng et al. (2022), and Tiwari et al. (2022) developed hybrid time-series forecasting models (ARIMA + neural networks) for case predictions and social media-based early warning classifiers leveraging platforms like Twitter. ML models also played a role in predicting COVID-19 severity based on clinical data, helping allocate healthcare resources more effectively. Despite these advancements, several deficiencies were noted, including poor generalization of ML models across different regions and pandemic phases, difficulties in adapting to new virus variants and policy changes, bias in social media data, and concerns about ML interpretability in clinical settings.

Therefore, while ML-based predictive modeling and classification systems have significantly improved early warning capabilities for various HILP events, persistent challenges remain. These include data limitations, generalization issues, lack of parameters setting, and challenges in integrating real-time observations with warning classification. Addressing these deficiencies requires a multi-disciplinary approach that combines advanced ML techniques with robust placement of observation stations to enhance the accuracy and reliability of early warning systems.

2.4 Conclusion and Linking to Current Research

Although previous studies have been conducted on High-Impact, Low-Probability (HILP) events, most remain fragmented and lack integration. This study builds upon previous research by introducing an integrated framework that systematically enhances forecasting capabilities. The framework consists of three key components: conceptualization, which accounts for the unique characteristics of HILP events; simulation and database generation, which involve various scenarios with different parameter settings; and analysis implementation, which integrates machine learning techniques for rapid scenario selection to support early warning systems.

Chapter 3: Methodology

To address those challenges, this study presents a framework, illustrated in Figure 3.1, based on three fundamental stages: case study context & modelling basis, simulation and database generation, and analysis and implementation. These stages ensure that forecasting models are grounded in previous studies and historical cases to understand the characteristics and impacts of HILP events.

The next sections provide an overview of each phase, covering the key components and methods used in this study. A detailed explanation of the methodology will be presented in specific chapters: Chapter 4 will discuss the conceptualization of characteristics used and impacts for this study in more detail; Chapter 5 will cover the technical aspects, including simulations, database generation, software, and parameter settings, while Chapter 6 will focus on the method used on machine learning algorithms, hyperparameter, and evaluation metrics. This structure ensures a clear and more detailed explanation of the methodology related to each chapter.

At the end of this subchapter, a detailed information of the case study used in this research, the 2018 Anak Krakatau Volcano Collapse Tsunami, will be presented. This section will provide background information on the case study, an overview of the affected location, a review of previous studies conducted on the 2018 Anak Krakatau tsunami, and highlighting of the significance of choosing this event for the study.

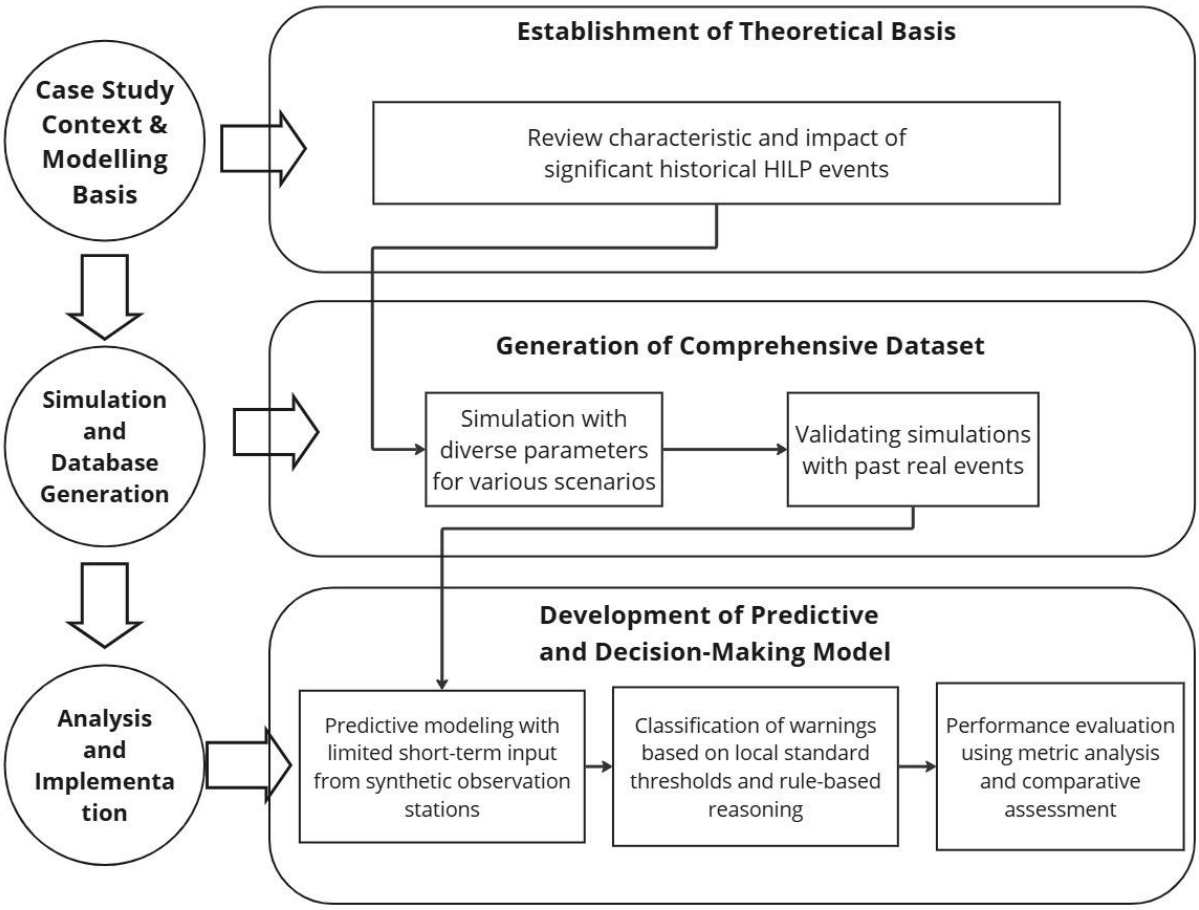


Figure 3.1 A Framework for Forecasting HILP events

3.1 Case Study Context and Modelling Basis

To establish a theoretical basis for the forecasting model, a comprehensive review was conducted on the several significant HILP events. This identified key characteristics and impacts. Historical cases of HILP events were analyzed to understand their underlying mechanisms.

Data were collected from journal papers and government reports published in the Scopus and ScienceDirect databases using HILP event-related keywords, including:

- "1883 Krakatau Eruption"
- "1986 Chernobyl Nuclear Disaster"
- "2004 Indian Ocean Tsunami"
- "2008-2009 Global Financial Crisis"
- "2010 Deepwater Horizon Oil Spill"
- "2011 Great East Japan Earthquake and Tsunami"
- "2018 Anak Krakatau Volcano Eruption Tsunami"
- "2019 COVID-19 Pandemic"

Additionally, general HILP disaster keywords were used, namely:

- "High-Impact Low-Probability Events"
- "Risk Assessment"

- "Risk Analysis"

This search yielded approximately final selected 85 papers and reports published between 1921 and 2025. Each document was reviewed individually to identify key findings from past HILP events. Based on this review, the characteristics, mechanisms, and parameters associated with each event were sorted out (Table 3.1).

Table 3.1 List of key findings from past HILP events and theory

HILP event and theory	**Key findings (characteristics, mechanisms, impacts, parameters)**
1883 Krakatau Eruption	• Occur sporadically • Evolving environmental and geological systems • Lack scientific understanding • No warning system to mitigate impact • Sheer magnitude explosion and subsequent tsunamis were not anticipated
1986 Chernobyl Nuclear Disaster	• Release massive amount of radioactive material into atmosphere • Forcing evacuation of entire communities • Rendering large swaths of land uninhabitable • System failures in government and public distrust
2001 9/11 Terrorist Attack	• Neglecting multiple warning signs and precursors indicates the possibility of an airline-based attack. • Failure of imagination decision-makers did not adequately assess the feasibility of an extreme yet plausible attack scenario.
2004 Indian Ocean Tsunami	• Absence of early warning • Unexpected
2008-2009 Global Financial Crisis	• Overreliance on stable market assumptions • Important to distinguish between temporary market fluctuations and systemic financial instability
2010 Deepwater Horizon Oil Spill	• Industrial accident releasing 4.9 million barrels of oil • Devastating marine ecosystems • Disrupting coastal economies
2011 Great East Japan Earthquake and Tsunami	• Cascading event • Long-term effect on environmental and economic consequences • Nuclear melting • Pre-existing safety standards outdated

	• Underestimate potential of cascading disasters
2018 Anak Krakatau Volcano Eruption Tsunami	• No early warning • Absence of tide gauges near the volcano • Volcano collapse (volume, slope angle, direction) cannot be estimated • Tsunami travel time is approximately 30 minutes to reach coastline • Evacuation cannot be made • The volume, thickness, and velocity of a landslide/collapse are key variables in tsunami formation
2019 COVID-19 Pandemic	• Complex interaction between ecological, virological, and human behavior • Global recessions • Pathogen detection and source tracing for forecasting and early warning • Real-time genomic surveillance • Nasal host response-based screening for identifying undiagnosed respiratory viruses
Risk Assessment & Analysis	• Rely on historical data
High-Impact Low-Probability Events	• Beyond current knowledge • People unprepared • Cascading failures • Preparation tends to reactionary rather than proactive

From the key findings list above, several impacts, characteristics, and parameters were identified for each historical HILP event. This analysis also took into account key factors essential for conceptualizing HILP events, providing a structured understanding of their underlying mechanisms. Furthermore, the findings help in identifying relevant parameters and characteristics that can be utilized as input for developing simulations and constructing a comprehensive database.

3.2 Simulation and Database Generation

The simulation and database generation phase were designed to overcome data limitations associated with HILP events by doing numerical simulations and using real-event data to validate the simulated result. To address this, this phase focused on creating a comprehensive dataset that captures a wide range of possible disaster scenarios based on the parameters and characteristics identified in the conceptualization phase. This dataset serves as a database for predictive modeling.

Simulation

The simulation process was essential in generating HILP event scenarios beyond the limitations of historical records. This study utilized the Cornell Multi-grid Coupled Tsunami Model (COMCOT), developed by Wang (2009), to conduct numerical simulations of landslide-generated tsunamis.

Numerical simulation is a computational technique that uses mathematical models to replicate complex physical processes. This approach is cost-effective as it eliminates the need for large-scale laboratory experiments (see Table 3.2). In tsunami modeling, numerical simulations allow researchers to simulate wave generation and propagation under various conditions, providing valuable insights into potential disaster scenarios that cannot be directly observed or tested in real life. This approach is particularly well-suited for HILP events, where historical and real-time data is scarce or insufficient.

Table 3.2 Comparison physical model with numerical simulation approaches

Approach	Method	Advantages	Disadvantages
Physical model experiments	Scale models in laboratory	Capture nonlinear effects, real data	Expensive, not always generalizable
Numerical simulation	Computer-based numerical methods	Handles complex cases, flexible	Depends on model accuracy, requires validation

A diverse set of parameters from past events and previous studies was incorporated, with additional refinements introduced by the authors to enhance the comprehensiveness of the simulations. The objective was to simulate 1,000 tsunami event scenarios triggered by volcanic collapses.

To ensure a broad representation of potential tsunami events, the simulations covered multiple variations in volcanic collapse characteristics, including landslide volume, slope failure angle, collapse direction, and initial wave dynamics. This approach provided a comprehensive dataset for predictive analysis and facilitated the identification of general, extreme and worst-case tsunami scenarios.

Validation

Validation is a crucial step to ensure that simulated tsunami scenarios accurately reflect real-events, allowing the generated dataset to be effectively utilized for predictive modeling (Table 3.3). This process involves comparing simulation results with observed historical data, particularly from tide gauge records that captured real tsunami waveforms (wave generation and propagation).

Table 3.3 Validation approach description

Aspect	Description
Purpose	Ensure simulated tsunami scenarios accurately reflect real events for predictive modeling
Method	Compare simulation results with historical data, particularly tide gauge records

Data used	2018 Anak Krakatau volcanic collapse tsunami; waveforms recorded at four tide gauge stations
Validation approach	Comparing recorded waveform plots with numerical simulation results
Expected outcome	Assess consistency between simulated and observed tsunami to enhance simulation and database reliability

In this study, validation was conducted using data from the 2018 Anak Krakatau volcanic collapse tsunami. During this event, tsunami waveforms were recorded at four tide gauge stations along the coastline. These recorded waveforms were systematically compared with numerical simulation results to assess the consistency between simulated and observed tsunami characteristics, ensuring the reliability of the model for future forecasting applications.

3.3 Analysis and Implementation

The analysis and implementation phase focuses on utilizing simulated tsunami scenarios as input for a predictive model to forecast tsunami waveforms and serve as a decision-making tool for issuing early warning levels. This phase involves developing predictive modelling, classifying warning levels, and evaluating system performance.

Predictive Modelling

Machine learning techniques are applied to develop a predictive model for estimating maximum tsunami amplitude and forecasting full tsunami waveforms.

The dataset is processed to extract and train the model using the maximum recorded tsunami heights from both tide gauges and synthetic observation stations, artificial observation points in oceanography used to monitor and simulate ocean wave movements for improved analysis and forecasting. The model then utilizes the recorded maximum tsunami heights from synthetic observation stations to predict the maximum tsunami height at coastal tide gauges.

For forecasting the full tsunami waveforms, time-series analysis models are trained using limited short-term input from synthetic observation stations. The model is optimized to function with minimal input data, allowing for rapid forecasting even with early observations. This capability makes it highly effective for early warning systems, providing sufficient lead time for response and evacuation.

Classification of Warning Levels

To translate predictive outputs into actionable intelligence, a classification system is implemented to determine warning levels. This system incorporates region-specific risk parameters (local thresholds) and rule-based reasoning to decide the appropriate warning level for the affected areas.

- Local thresholds refer to predefined, region-specific criteria or limits that determine when a warning should be issued. These thresholds vary based on geographical conditions and the type of HILP event, ensuring that the framework is adaptable to different scenarios. In this study, local thresholds are based on regulations set by the

Meteorology, Climatology, and Geophysical Agency of Indonesia (BMKG), which defines the tsunami wave height limits and the criteria for issuing early warnings in the study area.

- Rule-based reasoning utilizes a set of logical rules to evaluate the situation based on available data. It processes inputs and applies predefined criteria to determine the appropriate action, such as issuing a warning.

The classification framework is designed to balance the trade-off between false alarms and missed warnings, optimizing its effectiveness in early warning applications.

Performance Evaluation

To assess the accuracy and efficiency of the forecasting framework, performance metrics namely Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and F1-score are employed. The predictive model is benchmarked against traditional hydrodynamic simulations to quantify improvements in computational speed and forecasting accuracy.

By integrating predictive modelling, warning classification, and performance evaluation, this phase ensures that the forecasting framework is capable of supporting early warning decision-making for volcanic tsunamis. The refined system enhances the accuracy, adaptability, and applicability of disaster prediction, strengthening preparedness and response strategies for HILP events.

3.4 Case study: 2018 Anak Krakatau Volcanic Tsunami, Indonesia

3.4.1 Background

The December 22, 2018, Anak Krakatau Volcano (AKV) tsunami serves as an important case study for validating a proposed forecasting framework for HILP event in this study. At approximately 20:55 local time, AKV experienced a major lateral collapse following six months of eruptive activity that began in June (Grilli et al., 2019). The tsunami's catastrophic impact—resulting in 437 fatalities and severe coastal damage across Java and Sumatra—underscored the unpredictable yet highly destructive nature of volcanic tsunamis (Grilli et al., 2019).

Following this event, frequent small eruptions at AKV have continued to raise concerns about the possibility of another tsunami. Given that AKV is expected to remain highly active and a persistent threat, enhancing forecasting capabilities for tsunami early warning becomes imperative.

3.4.2 Study location

Anak Krakatau Volcano (Figure 3.2) is an active stratovolcano located in the Sunda Strait, which separates the islands of Java and Sumatra and lies between Banten and Lampung Provinces, Indonesia. These provinces play a significant role in connecting Java and Sumatra, two of the largest and most populous islands in the country. Additionally, Banten Province is located near Jakarta, Indonesia's capital and economic center, making it a region of high strategic importance for national business and governance.

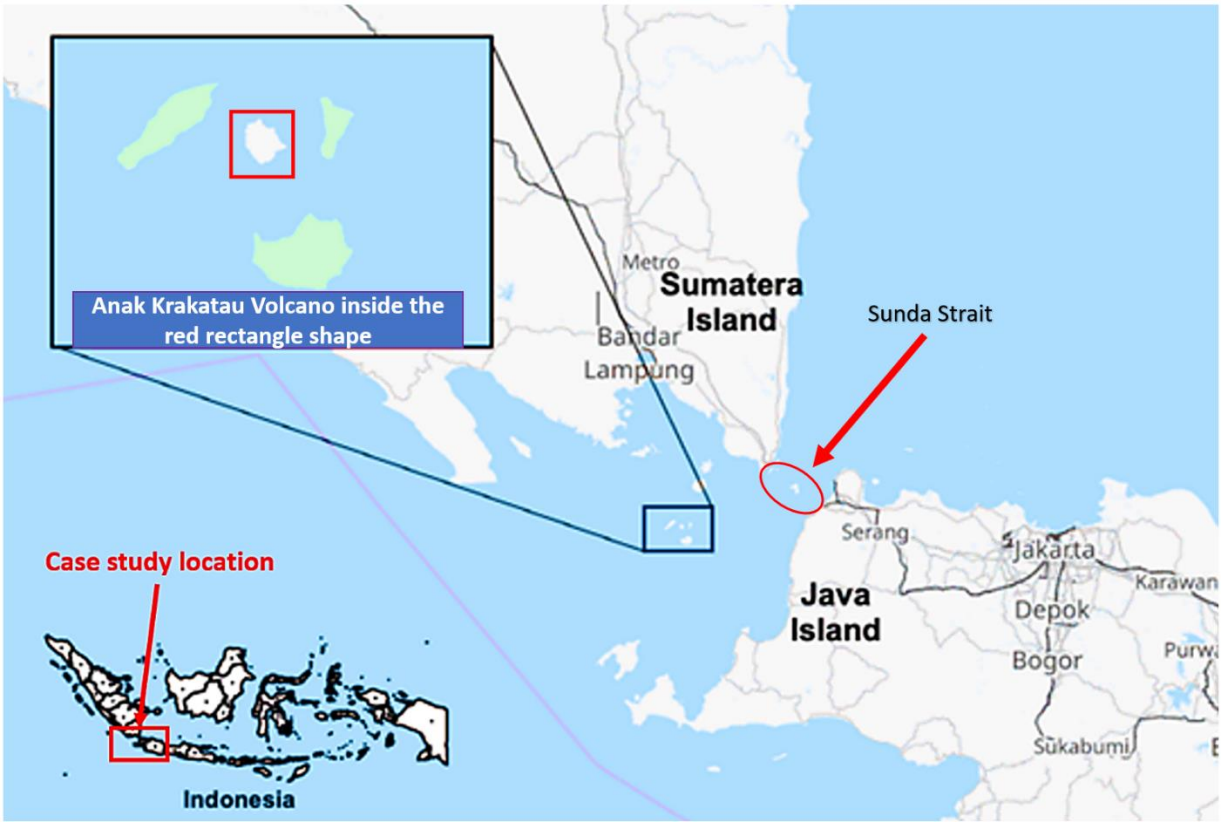


Figure 3.2 Map of Anak Krakatau Volcano, Indonesia. Created using an open street map

One of the key districts in Banten is Pandeglang, positioned at 6.3092°S and 106.1047°E, making it the westernmost point of Java. Covering an area of approximately 2,746 square kilometers with 1,413,897 population, and located 210 kilometers from Jakarta, Pandeglang is widely known for its natural beauty and cultural richness. The district features Tanjung Lesung, a popular tourist destination renowned for its pristine beaches and diverse marine biodiversity. Additionally, Peucang Island, part of the Ujung Kulon National Park, is a favored attraction among both domestic and international visitors due to its untouched landscapes and unique wildlife.

The Sunda Strait between Java and Sumatera islands, itself is a major maritime corridor essential for Indonesia's economy. Every day, approximately 145 ships pass through this busy waterway, totaling around 53,068 ship movements annually (hubla.dephub.go.id). The Merak Port in Banten and Bakauheni Port in Lampung serve as key ferry terminals connecting the two islands, facilitating the transport of passengers and vehicles. Hundreds of ferry crossings take place daily between these ports, supporting regional commerce and mobility.

However, Pandeglang's proximity to Anak Krakatau poses significant geological hazards, particularly tsunamis. To safeguard communities in the region, it is essential to establish continuous monitoring, implement early warning systems, and disaster preparedness strategies. Given Anak Krakatau's history of volcanic activity, proactive risk mitigation measures, supported by forecasting methodologies, are crucial to protecting both the population and critical infrastructure in the area.

Chapter 4: Case Study Context and Modelling Basis

4.1 Context and relevance of case study

Volcanic tsunamis represent a rare but devastating natural hazard, triggered by volcanic activity that can initiate a series of cascading events, culminating in massive ocean waves. These tsunamis have caused significant destruction in coastal regions worldwide, with impacts ranging from infrastructure damage to considerable loss of life. The cascading nature of volcanic collapses beginning with volcanic eruptions, followed by flank collapse, and ending in ocean displacement creates complex scenarios for disaster management and prediction (Paris *et al.*, 2014). As coastal populations continue to grow, there is an urgent need to enhance our understanding of the risks associated with volcanic tsunamis to develop more effective disaster reduction strategies (Grezio *et al.*, 2017).

Historically, volcanic tsunamis have led to some of the deadliest natural disasters, including the infamous Krakatoa eruption in 1883, which generated waves over 30 meters high and caused over 36,000 fatalities (Nazli and Kenji, 1995). Similar events, though smaller in scale, have occurred across the Sunda Strait regions, 2018 Anak Krakatau Volcano, highlighting the need for more detailed risk assessments of these rare but high-impact hazards (Grilli *et al.*, 2019). These tsunamis follow volcanic collapses entering the ocean, creating waves that propagate over long distances, affecting coastal communities far from the eruption site (Brown *et al.*, 2017). However, the unpredictability and the multifaceted nature of these tsunamis make their risks difficult to assess using traditional seismic-based early warning systems (Behrens *et al.*, 2021).

This chapter aims to address the existing gap by conceptualizing and revealing the characteristics of the volcanic tsunami that occurred in 2018. This chapter sets the context for simulation-based scenario development and risk assessment presented in the following chapters. This assessment is crucial for estimating the potential consequences within the defined study area.

By unveiling both the characteristics and the associated impacts, this chapter provides a clearer understanding of the unique risks posed by volcanic tsunamis and offers essential insights for developing more resilient coastal areas.

4.2 Result and Discussion

This section presents a detailed discussion of the key characteristics and impact analysis of the 2018 Anak Krakatau volcanic tsunami event.

4.2.1 Key characteristics and collapse parameters

To support the discussion of key characteristics and input parameters relevant to the 2018 Anak Krakatau volcanic tsunami, Table 4.1 presents a structured summary of significant component that define the event's physical and modelling. These components were compiled based on

findings from previous studies from other researchers (modelling, simulation, post-event investigations, and field survey), and serve as critical parameters for understanding and simulating high-impact, low-probability (HILP) disasters such as volcanic tsunamis.

Each component listed in the table captures different dimensions of the disaster's complexity, from the geological trigger to the hydrodynamic consequences. The mechanism of the landslide—described as a sudden lateral flank collapse with single massive block (en masse failure) and limited fragmentation— explain how an impulsive and localized geological event can instantaneously generate a basin-scale tsunami. This physical understanding was backed by remote sensing data such as SAR imagery and post-event marine surveys from other researchers in the previous studies.

Table 4.1 Detail of characteristics and parameters of the case study on the 2018 Anak Krakatau volcanic tsunami

Component	Key Characteristics
Mechanism of Landslide	Sudden lateral flank collapse with single massive block (en masse failure) and limited fragmentation, which rapidly displaced seawater and generated a tsunami. This interpretation is based on post-event marine surveys and SAR imagery analysis
Collapse Model	Scenario-based two-layer (water–slide) coupled model. Efficient for representing tsunami generation from variable collapse scenarios (volume, angle, direction)
Tsunami Model	Nonlinear Shallow Water Equations (NSWE). Suitable for shallow/coastal propagation and short-period landslide-induced waves
Rheology Assumption	Assumes rigid-body motion (solid block approximation) without internal deformation—ideal for rapid, high-momentum collapses
Collapse Volume	0.175–0.313 km^3 (likely ~0.224 km^3), constrained from marine geophysical data and satellite-based geometry analysis
Collapse Thickness	Up to ~250 m vertical failure depth from summit (~335 m) to sea floor
Collapse Length	Approximately 1.0–4.0 km downslope
Collapse Width	Estimated 0.5–2.0 km along the lateral extent of the failure scar
Collapse Direction	Southwest (approx. 225°), aligned with structural weakness and vent location
Bathymetry/Topography Source	Combined from SRTM30+, GEBCO, and DEMNAS with refinement using BIG tide data and post-event mapping near the source region
Simulation Duration	120 minutes. Tsunami travel time to coastline ~30 minutes the fastest reach coastal; simulation captures full wave evolution and near/far-field impacts

Distinctive Features	Multiple collapse scenarios generated to capture uncertainty and support predictive modelling-based tsunami early warning
Validation Dataset	Tide gauge waveform data from Marina Jambu (MJ) and Kota Agung (KA) during the 22 December 2018 event

To represent the physics of wave generation, a scenario-based two-layer coupled model was adopted. This approach integrates the interaction between the landslide body and surrounding water, enabling the simulation of diverse collapse scenarios by varying parameters such as volume, direction, and geometry. The propagation of the tsunami was simulated using Nonlinear Shallow Water Equations (NSWE), which are well suited for capturing short-period waves and coastal inundation, particularly in shallow marine environments.

The simulation assumed rigid-body motion without internal deformation—an idealized but appropriate approximation for rapid, high-momentum landslide collapses. Key physical parameters included collapse volume (0.175–0.313 km^3, likely ~0.224 km^3), vertical failure thickness (~250 m), downslope length (1.0–4.0 km), and lateral width (0.5–2.0 km). The direction of the collapse was estimated at approximately 225°, oriented southwest, in line with observed structural weaknesses and the location of the vent breach.

The bathymetry and topography data used in the simulation were derived from a combination of global and local sources, including SRTM30+, GEBCO, and DEMNAS, further refined using BIG tidal station data and post-event elevation mapping near the source area.

Simulation durations were set at 120 minutes, with coastal arrival times ranging between 30 and 60 minutes. The earliest wave arrival, observed at Marina Jambu (MJ), occurred approximately 33 minutes after the collapse. This window reflects the time-critical nature of tsunami early warning for near-field events.

The inclusion of multiple collapse scenarios was essential to account for the inherent uncertainty of volcanic tsunami events. These scenarios captured a wide range of parameter combinations and allowed evaluation of potential hazard variability. Model validation was performed using tide gauge waveform data from both MJ and Kota Agung (KA), confirming that the simulations could reproduce observed tsunami characteristics with reasonable accuracy.

4.2.2 Impact analysis: physical and spatial damages

The 2018 Anak Krakatau volcanic tsunami exemplifies the destructive potential of high-impact, low-probability (HILP) events, particularly in coastal areas with high exposure and limited preparedness. Table 4.2 presents a detailed summary of findings of the observed impacts, capturing key physical parameters and spatial characteristics that shaped the tsunami's effects across the Sunda Strait region.

Table 4.2 Impact analysis of the 2018 Anak Krakatau volcanic tsunami

Impact Component	Detail Findings
Fatalities	437 deaths across affected coastal regions; more than 50 fatalities occurred during a beachfront concert
Runup Height	Ranged from 0.9 to 13.5 meters depending on location and topographic conditions. Runup height is the highest vertical

	point reached by the tsunami on land, relative to mean sea level
Tsunami Height	Measured between 1.4 and 6.3 meters above ground level. Tsunami height is the vertical height of the wave as it reaches the coast or tide gauge station, typically measured from sea level to crest
Inundation Distance	Varied between 18 and 330 meters inland. Inundation distance is the horizontal distance the tsunami travels inland from the shoreline
Arrival Time	Tsunami waves arrived 30 to 60 minutes after the collapse
Earliest Arrival Location	Marina Jambu (MJ) tide gauge recorded the fastest, with waves arriving approximately 33 minutes after the collapse
Wave Period	Short-period waves (~6.6–7.4 minutes), increasing frequency of wave impacts and reducing evacuation time
Energy Concentration	Tsunami energy was mainly directed toward the southwestern coast of Java, particularly around Tanjung Lesung (where Marina Jambu gauge located)
Damage Characteristics	Severe damage occurred within 100 m of the shoreline; debris and lack of escape routes contributed to fatalities
Protective Factors	Areas located more than 100 meters inland or on elevated ground experienced less damage and fewer casualties

The tsunami caused 437 confirmed fatalities, with over 50 deaths reported during a beachfront concert, underscoring the lack of preparedness and early warning in the region. Wave runup heights varied significantly, ranging from 0.9 to 13.5 meters depending on topographic conditions. This vertical reach of water inland contributed to severe local damage and high casualty rates.

Tsunami heights, measured from sea level to crest at tide gauge locations, ranged from 1.4 to 6.3 meters, while inundation distances extended from 18 to 330 meters inland. The horizontal reach of the tsunami was heavily influenced by shoreline morphology and terrain slope, contributing to spatial variability in impact severity.

The waves arrived 30 to 60 minutes after the collapse, with the earliest detection at Marina Jambu tide gauge station, approximately 33 minutes post-collapse. Despite this window being theoretically sufficient for evacuation, the lack of public awareness, limited communication, and the presence of short-period wave trains (~6.6–7.4 minutes) complicated effective response. These rapid successive waves increased the frequency of impacts and reduced recovery time between wave arrivals.

Tsunami energy was predominantly directed southwest, with the highest concentration of impact observed in areas such as Tanjung Lesung, where Marina Jambu is located. Here, damage was particularly severe within 100 meters of the shoreline due to high wave energy, dense coastal settlements, and limited escape infrastructure. In contrast, regions located more than 100 meters inland or at elevated terrain experienced significantly less damage and lower fatalities, demonstrating the critical role of protective topography.

4.3 Summary

This chapter presented a conceptual foundation to understand the nature, mechanisms, and consequences of volcanic tsunamis, with a specific focus on the 2018 Anak Krakatau event. Volcanic tsunamis are complex cascading hazards that typically begin with volcanic eruption, followed by partial flank collapse, and culminate in tsunami generation and coastal inundation.

The 2018 Anak Krakatau event represents a rare, high-impact, non-seismic tsunami, providing a critical empirical benchmark for generalization. As a naturally occurring phenomenon, this event exhibits several distinctive features that justify its selection as a reference model for scenario development. These include: (1) a sudden sectoral collapse without prior seismic trigger, highlighting the importance of non-seismic sources; (2) directionally concentrated wave energy due to the collapse orientation toward the southwest; (3) short wave arrival times of 30–60 minutes, underscoring the limited lead time for early warning; and (4) pronounced spatial variability of tsunami impact influenced by coastal morphology, despite moderate offshore wave heights.

The physical parameters of the collapse—including volume (~0.224 km³), thickness (~250 m), slope angle, direction, and failure geometry—were derived from post-event bathymetric and geophysical surveys. These were integrated into a scenario-based modelling framework utilizing a two-layer coupled collapse-tsunami model governed by the nonlinear shallow water equations (NSWE), which are well-suited for simulating the key dynamics of volcanic tsunami generation and propagation.

Further analysis of tsunami impact indicators, such as runup height and maximum wave height, revealed significant spatial heterogeneity. These variations were primarily controlled by the distance from the source and local bathymetric/topographic conditions. The earliest wave detection occurred at Marina Jambu tide gauge approximately 33 minutes after collapse initiation, illustrating the narrow operational window for effective response in similar contexts.

By grounding the scenario framework in these empirically observed and physically meaningful features, this chapter establishes the basis for developing a synthetic scenario database in Chapter 5. The generalization of these features aims to support future modelling and early warning system design for other volcanic islands with similar collapse-prone morphologies.

To consolidate these insights, Table 4.3 outlines the specific features of the 2018 Anak Krakatau event as natural phenomena that serve as a conceptual and empirical foundation for generalization and synthetic scenario construction.

Table 4.3 Specific Feature of 2018 Event as Natural Phenomena for Generalization Purpose

Feature	Summary of Findings and Relevance for Generalization
Non-seismic Trigger Mechanism	The tsunami was generated without preceding seismic activity, highlighting the importance of modelling non-seismic triggers like volcanic flank collapse. This expands early warning scope beyond conventional earthquake-based systems
Sudden and Rapid Collapse	The collapse occurred within a short duration (~1–2 minutes), producing a concentrated wave energy release. This

	characteristic is crucial for calibrating the temporal resolution in scenario modelling
Directional Energy Propagation	Collapse directed southwest ward resulted in asymmetric wave propagation. This highlights the need to consider collapse directionality in future volcanic tsunami scenarios
Localized but High Impact	Despite moderate offshore wave heights, coastal runup and inundation were severe in specific regions, underscoring the influence of local bathymetry and amplification effects
Short Warning Time	Wave arrival at the nearest tide gauges occurred within ~30 minutes, emphasizing the narrow window for early detection and response in similar events
Topography-Sensitive Impact Distribution	The spatial variability of inundation and runup was strongly controlled by coastal morphology, a key factor for generalizing hazard zones in varied geographic settings
Observability and Empirical Data Availability	The event is one of the rare modern volcanic tsunamis with rich satellite, tide gauge, and post-event survey data, making it a robust reference case for model validation and scenario construction

Chapter 5: Simulation and Database Generation

5.1 Overview of simulation approach

The challenge of developing an effective early warning system for volcanic tsunamis lies in the complexity and unpredictability of these events. Unlike tectonic tsunamis, volcanic tsunamis often occur with little or no warning compared to tsunami generated by earthquakes (Maeda et al., 2015) and could be classified as a silent tsunami (Mulia, Watada, et al., 2020), making it difficult to provide early warning. This unpredictability, combined with the unique dynamics of volcanic activity, requires specialized data and modelling approaches to capture the full range of possible scenarios. The December 2018 Anak Krakatau tsunami event underscored the need for more advanced modelling frameworks by demonstrating how a relatively small-scale collapse can lead to significant tsunami impacts without typical precursor signals.

To address these challenges, the predictive modelling for volcanic tsunami prediction requires a comprehensive database of potential tsunami events. This begins with detailed numerical simulations that account for various volcanic collapse parameters, such as collapse direction, dip angle, thickness, length, width, and volume. However, capturing the complexity of volcanic tsunamis necessitates generating a broad array of scenarios to model diverse possible outcomes accurately.

In this chapter, a comprehensive multi-scenario tsunami simulation database comprising 1,000 distinct scenarios is developed to address the forecasting challenges posed by tsunamis triggered by volcanic flank collapse. The scenario construction is grounded in the key geophysical parameters identified in the preceding chapter, including variations in collapse volume, failure geometry, slope angle, and directional orientation. These parameters were selected based on their critical influence on tsunami generation and coastal impact, as demonstrated through numerical sensitivity analyses.

The choice to generate 1,000 simulation scenarios is supported by both methodological and strategic considerations. From a methodological perspective, this quantity is necessary to comprehensively cover the expanded parameter space, thereby capturing a wide range of physically plausible collapse configurations. By systematically varying each scenarios parameter across its relevant range, the dataset is capable of encompassing diverse initial conditions and potential outliers. This is particularly important given the inherently uncertain and site-specific nature of volcanic flank collapse mechanisms, which lack the standardized magnitude-depth relationships commonly associated with seismic sources.

At the initial stage of this study, a smaller number of simulation scenarios, 160 and 320, were tested for preliminary machine learning applications. The 160-scenario set was used in a classification model based on decision tree algorithms, while the 320-scenario set was

developed for predicting maximum tsunami amplitude using conventional machine learning methods. For those purposes, the results were acceptable. However, as the study progressed toward the use of deep learning models aimed at forecasting the full waveform based on limited initial observations, it became evident that a more extensive and diverse dataset was required to support mode training and testing.

The final decision to use 1,000 scenarios was made to address this need. The scenario set was expanded by enhancing the variation and resolution of the key input parameters, building upon previous studies (Heidarzadeh, 2020; Mulia, 2020; Ratnasari, 2023) while improving coverage across potential collapse configurations. Although generating more than 1,000 scenarios, such as the 1,300 cases used by Melgar et al. (2016) for earthquake-induced tsunami modelling, was technically possible, practical constraints such as computational time, data storage, and training efficiency informed the decision to limit the number. Furthermore, most recent studies employing deep learning for geophysical forecasting typically use datasets around of 1,000 samples or more to ensure model stability and generalization capability.

Strategically, the selection of 1,000 scenarios also reflects a balance between completeness and feasibility. While smaller datasets may be sufficient for simpler tasks such as maximum amplitude classification, deep learning models trained for sequence-to-sequence forecasting require significantly larger training sets to learn complex spatiotemporal patterns. The adopted scenario count thus represents an informed and justified compromise based on both practical and scientific considerations.

Finally, the generated dataset includes three principal data outputs for each scenario: (1) spatial distribution maps of tsunami propagation and maximum wave amplitude, (2) time-series waveforms recorded at predefined tide gauge locations, and (3) metadata describing the input parameters and derived outputs for each simulation. These outputs serve as the foundation for subsequent stages of model training and validation.

To ensure the physical credibility of the simulated scenarios, a validation phase was conducted by comparing selected outputs against empirical observations from the 22 December 2018 Anak Krakatau tsunami event. The alignment between simulated and observed waveforms, particularly in terms of arrival time, amplitude, and waveform shape, provides essential confirmation of the model's capacity to replicate real-world tsunami behaviour, thereby reinforcing the scientific validity and practical utility of the constructed database.

5.2 Tsunami mechanism induced by volcanic flank collapse or landslide

Tsunamis are long-wavelength ocean waves typically generated by large-scale disturbances of the seafloor or water surface. They can result from various geophysical processes, including undersea earthquakes, submarine landslides, and volcanic flank collapses. For effective tsunami hazard mitigation, it is essential to model tsunami waveforms to understand the wave heights that may impact coastal areas. These waveforms can be derived from both historical events and hypothetical scenarios. To produce realistic and reliable tsunami height estimates, numerical models are employed to simulate wave propagation from the source region to the coastline.

5.2.1 Governing equation: nonlinear shallow water equation (NSWE)

Since as the tsunami propagates and approaches a coastal area where the tide gauges located, linear shallow water equations are no longer valid. The wave length of the incident tsunami become shorter and amplitude become larger as the leading wave of a tsunami propagates into shallow water (approach the continental shelf and coastal). Therefore, the nonlinear convective inertia force and bottom friction term become important, while the dispersion term diminishes. The Nonlinear Shallow Water Equation (NSWE) including bottom friction effects are adequate to describe flow motion in coastal zone (Figure 5.1).

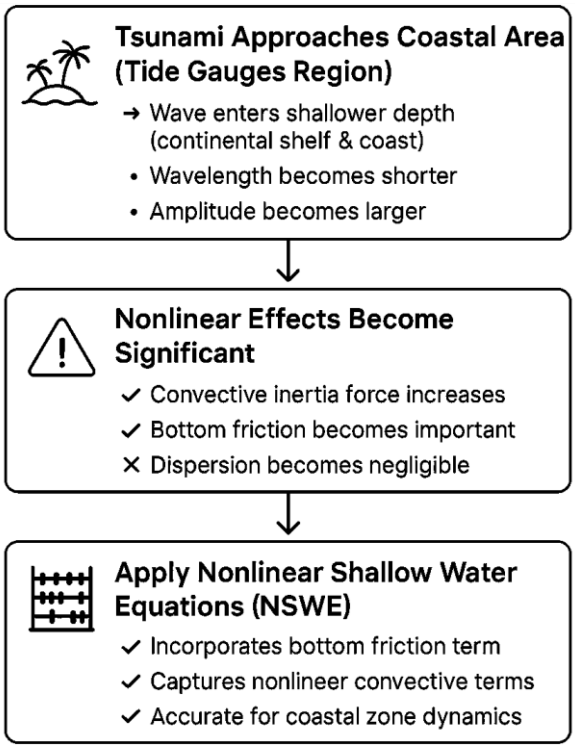


Figure 5. 1 NSWE explanation background

To discretize the nonlinear convective terms, the software uses a finite difference method formulated in conservative form. The solver incorporates numerical techniques that help maintain shock consistency, using a staggered grid and flux, where variables like surface elevation (η) and velocity components (u, v) are placed at different grid points to improve numerical stability, and an upwind flux scheme that considers the direction of flow to reduce oscillations near shocks. A more detailed explanation is provided in Subsection 5.2.2.

5.2.2 Numerical method: leap-frog finite difference method

The Finite Difference Method (FDM) was used in this study. The FDM is a numerical technique that approximates derivatives by calculating differences between values at discrete grid points. In the context of this dissertation, the FDM is crucial for simulating tsunami wave propagation resulting from volcanic flank collapse events, such as the 2018 Anak Krakatau case study.

Specifically, software used in this study implements a staggered grid using a leap-frog time-stepping scheme. The explicit leap-frog finite difference scheme for linear shallow water equations can be expressed as:

$$\frac{\partial_\eta}{\partial_t} = \frac{\eta_{i,j}^{n+1/2} - \eta_{i,j}^{n-1/2}}{\Delta t}$$

∂ (the partial derivative symbol) denotes the rate of change of a multivariable function with respect to one variable, in this case, time t, while keeping other variables such as spatial coordinates i and j constant. This is appropriate because the shallow water equations involve both spatial and temporal variables.

In this setup, the water surface elevation (η) is calculated at half times steps (e.g., $t = \left(n - \frac{1}{2}\right)\Delta t, (n + 1/2)\Delta t$), while the volume fluxes *P* and *Q* (representing the product of water depth and velocity in X and Y directions, respectively) are calculated at full time steps (e.g., $t = n\Delta t, (\, n + 1)\Delta t$). These calculations alternate in time to maintain numerical stability and computational efficiency.

Figure 5.2 shows the grid system used in the software, where the water surface displacement is evaluated at the center of each grid cell, while the fluxes (*P* and *Q*) are evaluated at the edges. This structure enables accurate resolution of nonlinear shallow water dynamics, which are critical for simulating the rapid wavefronts generated by landslide-induced tsunamis.

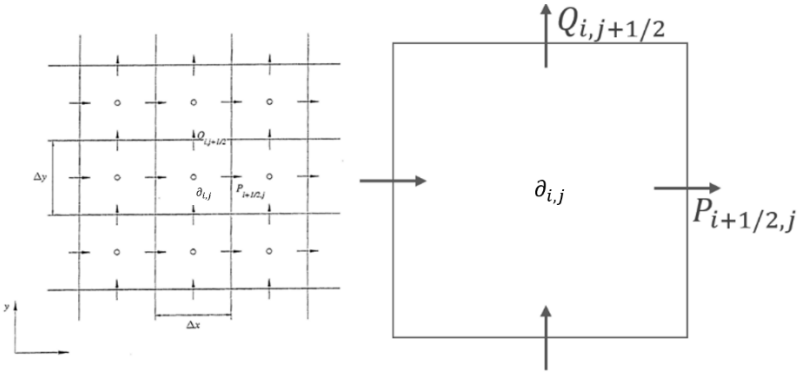


Figure 5. 2 Sketch of grid setup for numerical simulation

This method was selected for its computational efficiency, especially important for the large-scale generation of thousand synthetic tsunami scenarios in this study. It ensures the model can simulate near-field tsunami waveforms quickly and with sufficient accuracy to train deep learning models for early warning purposes.

5.2.3 Landslide mechanism and transient seafloor motion from volcanic landslides

In this study, the landslide is modelled as a time-evolving mass movement resulting from the flank collapse of Anak Krakatau. A kinematic modelling approach is adopted, whereby the landslide motion is defined through predetermined spatial and temporal parameters, such as collapse volume, direction, and duration, without explicitly simulating the internal forces, material strength, or rheological properties that drive the collapse. This approach emphasizes the specification of when, where, and how the mass moves, rather than why it moves.

To implement this, the software uses a sequence of bathymetric grids (snapshots, figure below) to represent the evolving shape and location of the sliding mass. These snapshots encode

vertical seafloor displacement at multiple time intervals and are used to calculate the rate of bathymetric change.

Moreover, unlike earthquake-generated tsunamis, which assume instantaneous seafloor displacement, tsunamis generated by volcanic flank collapses involve transient seafloor motion. The motion occurs over several seconds to minutes, during which a large volume of material moves downslope, displacing the overlying water and initiating tsunami waves. Software supports this scenario through direct inclusion of the time derivative of bathymetry $\frac{\partial \eta}{\partial t}$ in the continuity equation, allowing simulation of time-evolving seabed deformation.

The tsunami source is implemented via a sequence of bathymetric snapshots that capture the evolving geometry of the sliding mass. These snapshots provide vertical displacement data at successive time intervals, which are used to compute the rate of change of water depth (Figure 5.3). This dynamic displacement serves as the forcing mechanism in the model, generating realistic tsunami waveforms.

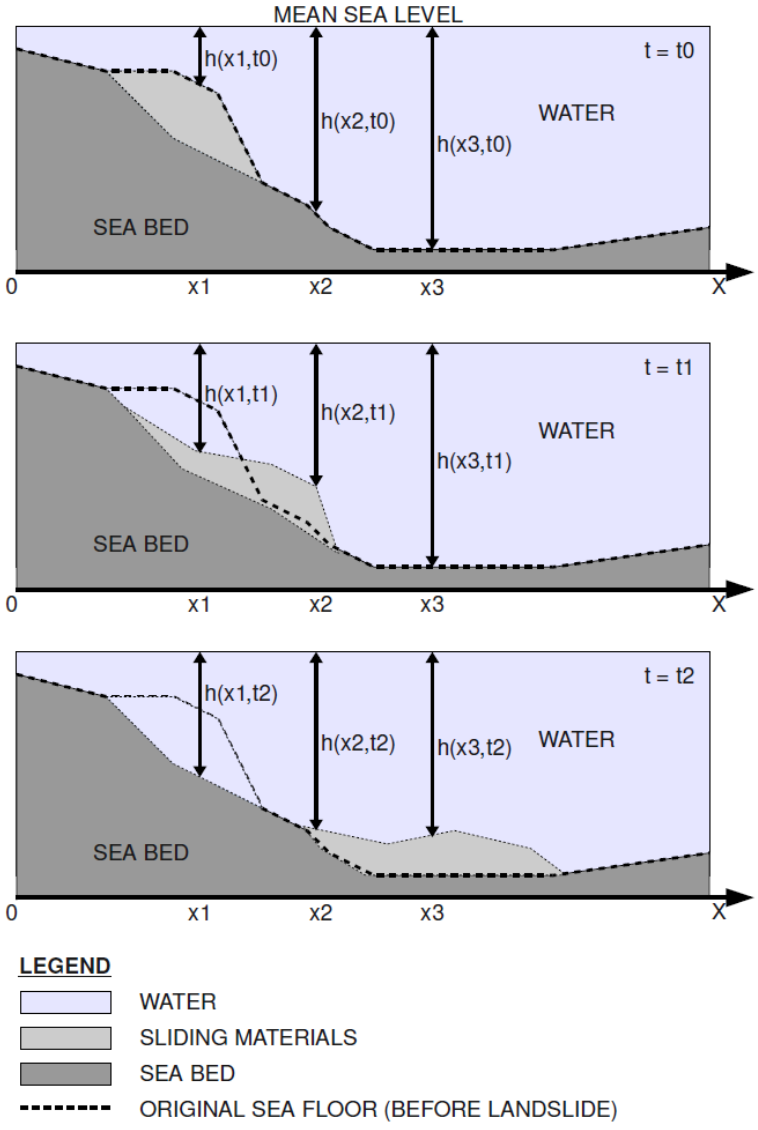


Figure 5. 3 Sketch of transient sea floor motion

5.2.4 Input configuration and source parameters

Each volcanic landslide scenario in this study is defined by a set of physical parameters that influence the tsunami generation process. These parameters include the estimated collapse volume, slope angle, failure geometry (e.g., length, width, depth), direction of movement, and the duration of the landslide. Based on these inputs, the deformation of the seafloor is represented through a series of time-stepped bathymetric grids that reflect the evolution of the sliding mass over time. These grids are used to compute the vertical seafloor velocity $\frac{\partial_\eta}{\partial_t}$ at each time step, which acts as the forcing term in the continuity equation for tsunami generation.

The model is designed to simulate a wide range of plausible collapse scenarios by systematically varying input parameters across hundreds of synthetic cases. This enables the construction of a comprehensive scenario database. The complete procedure for generating the landslide source data, including model domain selection, grid resolution, nesting configuration, and the parameter space used to define scenario variability, is described in detail in Section 5.3.2: Model Setup and Parameter Configuration.

5.3 Data generation

Due to the rarity of volcanic tsunami landslide data and the need for large datasets for predictive modelling, synthetic data is generated through simulations. Figure 5.4 presents a detailed flowchart outlining the synthetic data generation process for volcanic tsunami simulations. Given the scarcity of real-world tsunami events triggered by volcanic collapses, this synthetic approach ensures sufficient data volume for model development and testing. The process begins with data preparation, which includes selecting or designing a suitable landslide tsunami numerical model and configuring initial simulation parameters based on collapse volume, geometry, and location.

Following this step, tide gauge observation points are defined strategically around the source region to capture waveforms with high temporal resolution. These settings are followed by initial landslide scenario configuration, which determines the specific parameter combinations to be tested in each simulation run. The configured model then executes a tsunami numerical simulation, producing wave propagation results and time-series waveform data.

The validation step compares the simulated outputs against real observations, such as tide gauge data from the 2018 Anak Krakatau tsunami, to assess physical realism and numerical accuracy. If the accuracy is acceptable, the simulation outputs are incorporated into the volcanic tsunami synthetic dataset. If not, the workflow iterates by refining the landslide parameter set or simulation configuration until the desired fidelity is achieved.

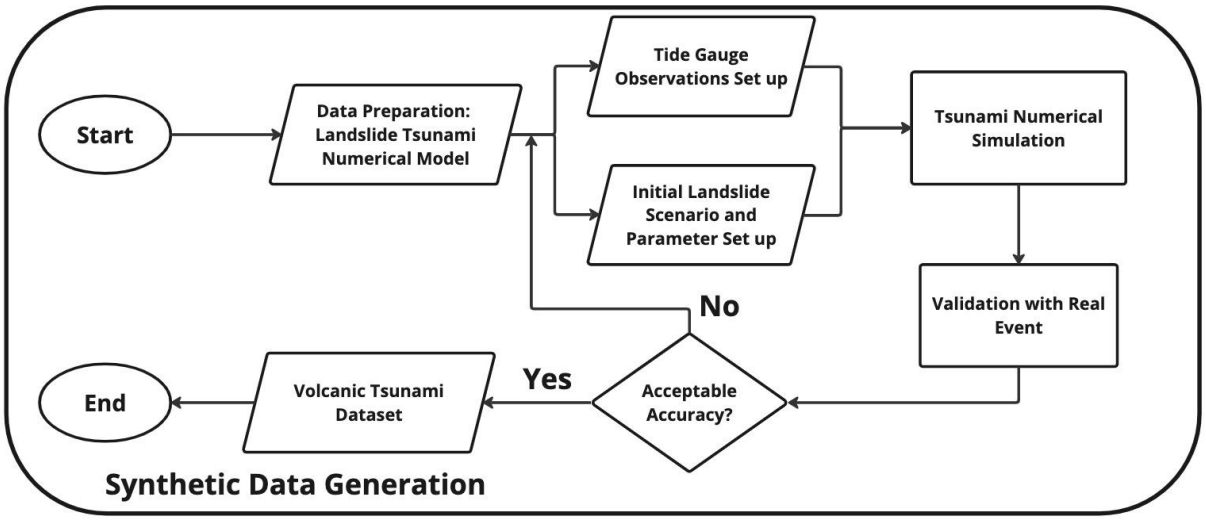


Figure 5.4 Volcanic tsunamis data generation flowchart

5.3.1 Landslide tsunami numerical simulation

When a tsunami propagates from deep water towards a continental shelf and coastal areas such as Java and Sumatra, the wave's behavior changes significantly. In the case of Anak Krakatau volcano, linear shallow water equations become less accurate due to the decreasing water depth, which causes the wavelength to shorten and the amplitude to increase as the leading wave enters shallower water. Nonlinear shallow water equations (NSWE) must be applied to account for these changes. The nonlinear convective inertia forces, bottom friction, and the effects of coastal boundaries become critical to accurately capture the tsunami dynamics (Woo & Cho, 1998). As the tsunami approaches the coastline, the significance of frequency dispersion and the Coriolis effect diminishes, while nonlinear factors such as bottom friction, coastline geometry, and the interaction with the seabed and land surface gain importance. Therefore, the nonlinear shallow water equations provide a more comprehensive framework for simulating tsunami propagation in coastal regions, where wave-breaking, inundation, and the interaction with complex topography must be considered.

In this study, we used COMCOT (Cornell Multi-grid Coupled Tsunami Model) developed by Wang (2009) for generating the landslide tsunami numerical simulations. COMCOT is a widely recognized model based on shallow water theory. The latest online version, COMCOT v1.7, has been successfully applied in various tsunami scenario simulations in previous research (e.g., Lynett *et al.*, 2012; Heo *et al.*, 2019; Dang *et al.*, 2023). Additionally, studies by Gusman *et al.* (2019) have also implemented this model for simulating landslide-induced tsunamis. Therefore, COMCOT v1.7 was employed in this study to simulate tsunami propagation.

COMCOT model solves the nonlinear shallow water equations (NSWE) in spherical coordinates, using a grid resolution of 0.2 arc minutes, equivalent to 12 arc seconds or approximately 370 meters (Table 5.1). This spatial resolution is used to capture the dynamics of tsunami propagation while balancing computational cost limitations. Shallow Water Equation (SWE) and Nonlinear Shallow Water Equation (NSWE) models are widely adopted by researchers due to their computational efficiency and reasonable accuracy in simulating

tsunami propagation and run-up. Previous studies (e.g., Apotsos, Jaffe and Gelfenbaum, 2011; Cheung *et al.*, 2011; Sonja van Leeuwen, Paul Tett, David Mills, 2015; Dang *et al.*, 2023) have successfully applied these models to reconstruct historical tsunami events.

Table 5. 1 COMCOT Grid and Bathymetry Configuration

Parameter	Value
Run this layer? (0: yes, 1: no)	0 (yes)
Coordinate system	0 (spherical)
Governing equations (0: linear, 1: linear)	1 (nonlinear)
Grid size (dx, sph:minute, cart: meter)	0.05
Time step (second)	1.0
Bottom friction switch? (0: Yes, 1: No)	0 (yes)
Manning' s roughness coef. (for fric.option=0)	0.013
Layer output option? (0: Z+Hu+Hv, 1: Z only)	0
X_start	104.489832
X_end	106.023233
Y_start	-6.842972
Y_end	-5.415588
File name of bathymetry data	180m.xyz
Data format option	3 (etopo)
Grid identification number	01
Grid level	1

While COMCOT may be considered simpler than more advanced models such as Boussinesq-type or Navier-Stokes-based solvers, its selection in this study was a deliberate and strategic decision driven by its computational efficiency. Since the objective was to generate a large-scale synthetic database of 1,000 landslide-induced tsunami scenarios for training deep learning models, a lightweight and fast simulation tool was essential. COMCOT implementation of the nonlinear shallow water equations (NSWE) offered a practical balance between simulation speed and physical fidelity, enabling the modelling of tsunami generation, propagation, and coastal amplification with sufficient accuracy. The Manning's roughness coefficient was set to 0.013, a value adopted from previous tsunami modelling studies that demonstrated good agreement with observed data and realistic wave propagation characteristics (e.g., Grilli et al., 2019; Mulia et al., 2020).

In addition, COMCOT simulates landslide-generated tsunamis by applying a time-dependent seafloor deformation as a dynamic boundary condition. The landslide is parameterized using an ellipsoidal body that moves downslope, with prescribed geometric and kinematic parameters such as initial volume, length, width, thickness, dip angle, and sliding velocity. This approach allows the model to approximate the mass failure process without explicitly simulating rheology, assuming a rigid-body behaviour consistent with previous tsunami generation studies. The governing equations (Equations 5.1-5.5) ensure that the simulation accurately represents tsunami behavior as it transitions from deep waters across continental shelves and into coastal region.

Equation 5.1 – 5.5 Nonlinear shallow water equations

$$\frac{\partial \eta}{\partial t} + \frac{1}{R\cos\varphi}\left\{\frac{\partial P}{\partial \psi} + \frac{\partial}{\partial \varphi}(\cos\varphi Q\right\} = -\frac{\partial h}{\partial t} \tag{5.1}$$

$$\frac{\partial \mathrm{P}}{\partial t} + \frac{1}{R\cos\varphi}\frac{\partial}{\partial \psi}\left\{\frac{P^2}{H}\right\} + \frac{1}{R}\frac{\partial}{\partial \psi}\left\{\frac{PQ}{H}\right\} + \frac{gH}{R\cos\varphi}\frac{\partial \eta}{\partial \psi} - fQ + F_x = 0 \quad (5.2)$$

$$\frac{\partial Q}{\partial t} + \frac{1}{R\cos\varphi}\frac{\partial}{\partial \psi}\left\{\frac{PQ}{H}\right\} + \frac{1}{R}\frac{\partial}{\partial \psi}\left\{\frac{Q^2}{H}\right\} + \frac{gH}{R}\frac{\partial \eta}{\partial \varphi} + fP + F_y = 0 \quad (5.3)$$

$$F_x = \frac{gn^2}{H^{\frac{7}{3}}} P(P^2 + Q^2)^{1/2} \quad (5.4)$$

$$F_x = \frac{gn^2}{H^{\frac{7}{3}}} Q(P^2 + Q^2)^{1/2} \quad (5.5)$$

The variables used in the Nonlinear Shallow Water Equations above (Equations 5.1–5.5) are defined as follows:

η: Free surface elevation (height of water above mean sea level)
h: Still water depth
$H = \eta + h$: Total water depth
P, Q: Depth-integrated momentum fluxes in longitudinal (ψ) and latitudinal (φ) directions (m^2/s)
R: Earth's radius
ψ, φ: Longitude and latitude, respectively
g: Gravitational acceleration
f: Coriolis parameter
F_x, F_y: Bottom friction components derived using Manning's equation, as shown in Equations (5.4–5.5)

These equations describe the interaction between surface elevation, mass flux, and external forces. The nonlinear terms (e.g., $\frac{P^2}{H}, \frac{Q^2}{H}, \frac{PQ}{H}$) represent convective acceleration, which becomes significant in shallow regions. Gravity terms reflect pressure gradients, while Coriolis terms account for Earth's rotation. Bottom friction terms introduce energy dissipation due to seabed roughness.

To clarify the structure, relationship, and role of the governing equations used in this study, Table 5.1 summarizes the function of each component in the Nonlinear Shallow Water Equations (NSWE) system. While Equations (5.1) to (5.3) form the core of the dynamic model governing tsunami propagation, Equations (5.4) and (5.5) serve as supporting expressions to quantify bottom friction forces essential for realistic simulation in nearshore environments.

Table 5.2 Role and Function of Equations in the NSWE Model

Equation(s)	Type & Function	Description
(5.1)	Continuity Equation	Represents mass conservation. Describes changes in free surface elevation η\eta due to divergence of flow.

Equation(s)	Type & Function	Description
(5.2) & (5.3)	Momentum Equations (Longitudinal and Latitudinal Directions)	Represent conservation of momentum in both horizontal directions. Include effects of advection, gravity, Coriolis force, and bottom friction.
(5.4) & (5.5)	Auxiliary Equations: Bottom Friction Terms using Manning's Formula	Provide expressions for calculating bottom friction components FxF_x and FyF_y, which appear in Equations (5.2) and (5.3). These depend on velocity, water depth, and seabed roughness.

The computational implementation in COMCOT uses a staggered-grid finite difference method. Scalar quantities like surface elevation are computed at grid cell centers, while momentum fluxes are computed at cell edges. A leap-frog time-stepping scheme is applied for temporal accuracy, and central difference schemes are used for spatial derivatives. This setup enables efficient and stable numerical solutions that capture key tsunami behaviors such as shoaling, refraction, and coastal inundation.

A comprehensive overview of the COMCOT software workflow is illustrated in Figure 5.5, which presents the sequential modeling steps and data processing procedures implemented in the simulation framework.

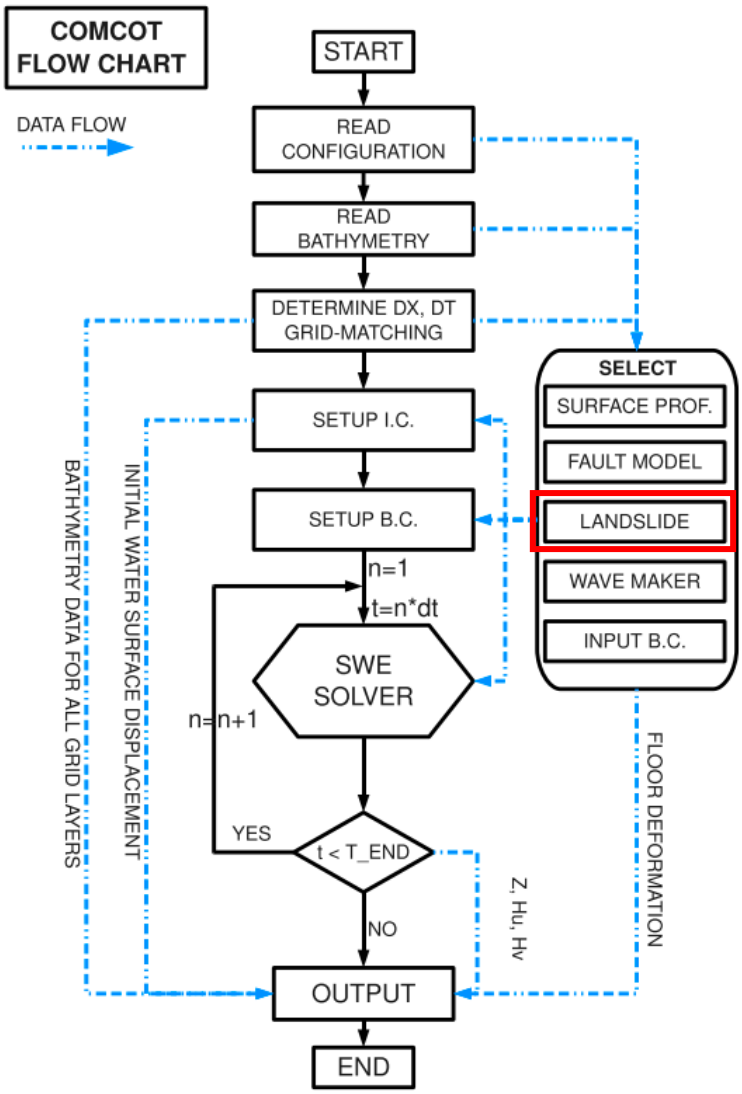


Figure 5.5 Flow chart of COMCOT. The study focused on landslides in the red box

Figure 5.2 illustrates the internal computational workflow of the COMCOT model utilized in this study to simulate tsunami propagation resulting from volcanic flank collapse. The sequence begins with the reading of the configuration file, which defines essential parameters for the simulation. Following this, high-resolution bathymetric and topographic data are loaded into the model. For this purpose, the study employed the DEMNAS dataset, an authoritative source developed by the Geospatial Information Agency of Indonesia (BIG), accessible via [https://tanahair.indonesia.go.id/demnas/#/]. These geospatial inputs are crucial for accurately representing the underwater terrain and coastal morphology, thereby enhancing the physical realism of the simulation.

Subsequently, the computational grid is generated to discretize the simulation domain, enabling numerical calculations over the area of interest. The landslide mechanism is then selected as the tsunami source. In this context, the landslide model is adopted because it offers a physically consistent and computationally tractable approach to representing volcanic collapse-induced tsunami generation. This choice is particularly relevant for simulating impulsive mass

movement phenomena, such as the 2018 Anak Krakatau event, where a rapid flank failure triggered significant wave generation.

By integrating landslide dynamics with shallow water hydrodynamics, the COMCOT model proceeds to execute the tsunami simulation, capturing wave generation, propagation, and coastal impact. Following this initialization phase, the model proceeds to configure the initial conditions (IC), encompassing the specification of the initial water surface displacement. Subsequent to the IC setup, boundary conditions (BC) are delineated to complete the model setup.

Subsequently, the simulation progresses to the computation of the NSWE solver, executing numerical algorithms to resolve the fluid dynamics within the defined domain. Post-NSWE solver computation, the model advances to address floor deformation phenomena, capturing alterations in the underwater terrain induced by dynamic forces. Finally, the simulation culminates in the generation of the tsunami wavefield, encapsulating parameters such as wave amplitude and temporal evolution, alongside the derivation of time-series records, which serve as comprehensive outputs delineating the spatiotemporal evolution of the simulated tsunami event.

5.3.2 Model setup and parameter configuration

The accurate setup of initial conditions is crucial for ensuring the physical validity and reliability of landslide tsunami simulations. These conditions define the starting state of the model and heavily influence the accuracy of waveform generation and spatial impact prediction. In the context of volcanic-collapse-induced tsunamis, such as the 2018 Anak Krakatau event, initial parameters must capture key physical characteristics of the landslide process (Figure 5.3).

This includes the geometry of the failure mass, the direction of collapse, and source material behaviour. Each of these elements plays a distinct role in shaping the resulting tsunami dynamics. The following subsection provides a detailed explanation of how these parameters are defined and integrated into the numerical model.

Table 5. 3 General COMCOT Numerical Simulation Configuration

Parameter	Value
Total run time (seconds)	7200.0
Time interval to save data (seconds)	600.0
Output option (Zmax & Timeseries)	2 (both)
Resuming time (if hot start) (seconds)	600.0
Minimum offshore depth (meter)	0.0
Initial condition type	3 (Landslide only)
Job description	NZ30sec bathymetry, spherical coordinates

5.3.2.1 Initial condition: landslide geometry, directions, material source parameters

Giachetti *et al* (2012) proposed a scenario that was comparable to what occurred on December 22, 2018, by modeling a tsunami generated by a flank collapse at Anak Krakatau. They used a 2-D depth-averaged hydrodynamic flow solver, a dynamic source model for a 0.28 m^3 flank

collapse, and a water wave propagation model in their simulation. The forecasting in this work was impressive, though not as accurate as it was in 2018. Their research yielded predictions for tsunami heights of about 45 meters in the immediate field and runup heights of 1.5 meters in Merak and 3.4 meters in Labuhan.

Following the December 2018 event, a number of scholars conducted simulation studies in an effort to compare the results of the simulation to the actual collapse event. Mulia *et al.* (2020) used a comparison of satellite pictures taken of the Anak Krakatau volcano before and after the event to estimate the collapse volume at 0.24 km^3. This collapse volume serves as the landslide tsunami's beginning condition. Other than volume, however, the COMCOT requires other parameter input for the landslide tsunami simulation, including thickness, length, and width (Figure 5.4). As a result, the study by Heidarzadeh *et al.* (2020) was modified for our investigation in order to ascertain the thickness and length parameters, and the volume was used to modify the width parameter. By taking into consideration the effect of underwater landslides on wave creation and propagation, this feature improves the realism of the original tsunami models.

Table 5. 4 Landslide Source Parameters (Example of One Scenario) in Software

Parameter	Value
Landslide start time (seconds)	0.0
Landslide end time (seconds)	420.0 (refer to previous studies 2018 event)
X (start) center of mass	105.430656
Y (start) center of mass	-6.102131
X (stop) center of mass	105.445474
Y (stop) center of mass	-6.102842
Typical slope angle (degree)	8.0
Length of sliding volume (m)	2913.0
Width of sliding volume (m)	833.2
Typical thickness of slide (m)	103.0

Finally, a studies by Mulia *et al.*, (2020) and Ratnasari *et al.* (2023) also mentioned the dip (degree) of 9° and 8°, respectively. These values are incorporated in this study, with an expanded parameter range of 5° to 9° to improve and broaden the scope of the analysis. The term dip angle represents the initial slope of the landslide body relative to the horizontal seafloor. This parameter controls the angle at which the mass begins to slide and is used consistently across all simulation scenarios to define the initiation dynamics of the landslide.

To support the development of a comprehensive synthetic tsunami waveform database, a total of 1000 landslide-induced tsunami simulation scenarios were executed. This number was selected to ensure adequate coverage of the parameter space, as informed by previous studies and further refined in the present research. The goal is to capture the inherent variability and uncertainty associated with volcanic collapse mechanisms. By systematically varying input parameters, such as landslide geometry, collapse direction, and collapse volume, the simulations aim to represent a wide range of physically plausible scenarios that are critical for the development of robust tsunami early warning applications.

Table 5.5 presents a representative subset of scenarios used as the basis for tsunami numerical simulations, illustrating the diversity of conditions explored in this study. The full detail information of all 1000 simulated scenarios, including detailed input configurations, is provided in the Supplementary Information to support reproducibility and enable further analysis.

Table 5.5 Landslide tsunami parameters
Note: The dip angle refers to the initial slope angle of the landslide along the failure plane, used to define the motion of the sliding mass

Scenario (sc)	Direction	Lon (start)	Lat (start)	Lon (stop)	Lat (stop)	Dip (°)	Thickness (A) in meter	Length (L) in meter	Width (W) in meter	Volume (V) in km^3
sc1	east	105.430656	-6.102131	105.445474	-6.102842	9	176	2906	78.2	0.04
sc61	northwest	105.417516	-6.09565	105.4096	-6.092904	9	86	2130	382.1	0.07
sc83	south	105.42366	-6.107302	105.423354	-6.119498	9	173	1404	617.5	0.15
sc110	southeast	105.43293	-6.10354	105.446	-6.1155	9	128	3460	519.3	0.23
sc157	west	105.42052	-6.0999	105.40475	-6.09999	9	51	1534	3834.6	0.30
sc403	south	105.42366	-6.107302	105.423354	-6.119498	7	176	2567	332	0.15
sc544	northwest	105.417516	-6.09565	105.4096	-6.092904	6	86	2975	664.5	0.17
sc682	northeast	105.427549	-6.095905	105.436475	-6.088797	5	132	3974	133.4	0.07
sc766	southwest	105.421887	-6.102357	105.412725	-6.115588	5	63	3541	851.7	0.19
…..										
sc1000	north	105.42448	-6.09382	105.42432	-6.07864	8	176	2972	630.9	0.33

In addition to defining landslide geometry, including collapse volume, it is essential to determine the direction of collapse, particularly because the tsunami scenarios in this study are induced by landslides. Historical records indicate that the 2018 Anak Krakatau flank collapse propagated primarily in a southwest direction. However, due to the inherently unpredictable nature of volcanic activity and the potential variability in collapse initiation points, it is critical to incorporate a broader directional strategy that accounts for future uncertainties.

Building on the initial approach proposed by Ratnasari *et al.* (2023), who considered four collapse directions, this study expands the directional range to eight compass-based orientations: north, northeast, east, southeast, south, southwest, west, and northwest. This expanded directional framework enables a more comprehensive assessment of possible collapse trajectories and their associated tsunami hazards. The spatial arrangement and distribution of these landslide directions are illustrated in Table 5.6 and Figure 5.6, providing an illustration data input in simulation software and visual overview of the directional assumptions adopted in the simulation scenarios.

Table 5. 6 Submarine Landslide Area and Input File Configuration (Example of One Scenario) in Software

Parameter	Value
X coord. of left/west edge of landslide area	105.369984
X coord. of right/east edge of landslide area	105.454957
Y coord. of bottom/south edge of landslide area	-6.201663
Y coord. of top/north edge of landslide area	-6.021925
Input data file name	landslide.ctl
Data format option	2 (function)

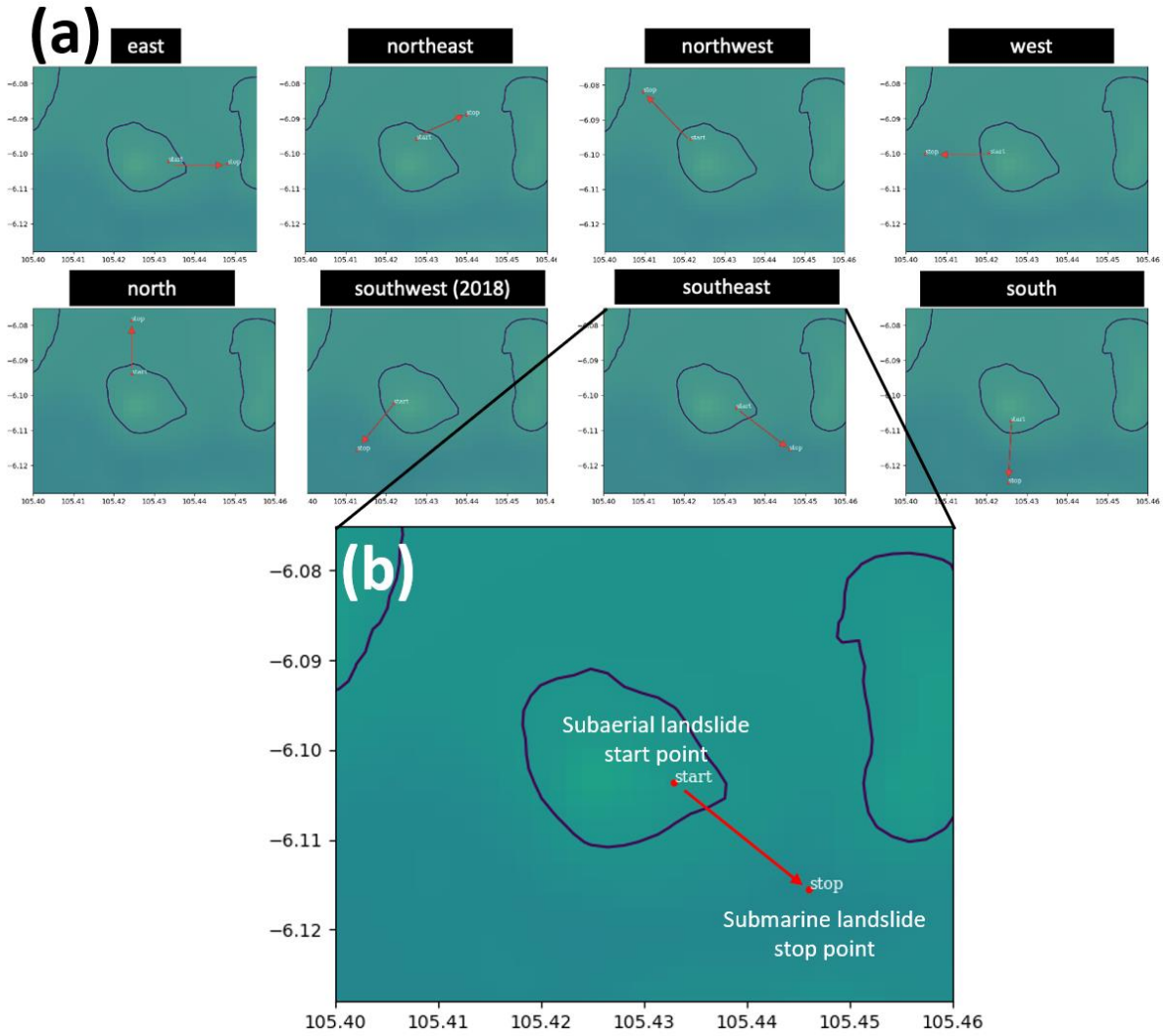


Figure 5.6 Landslide direction plan. (a) Eight directions landslide covers all future possibility tsunami: east, northeast, northwest, west, north, southwest, southeast, south (b) In detail of subaerial volcano landslide to submarine landslide stop point

5.3.3 Observation

In this study, we employed both original tide gauges and synthetic (or virtual) tide gauges as observation stations to monitor and analyze tsunami waveforms. In the context of tsunami modeling and early warning systems, observation refers to the systematic collection of tsunami wave characteristics, such as wave height, arrival time, period, and waveform evolution, at specific spatial and temporal points. In numerical simulations, these observation points serve not only to replicate physical measurements but also to track the propagation and transformation of tsunami waves under various scenario conditions. They are essential for evaluating model performance, enabling quantitative validation against real-world data, and supporting the development of predictive models for early warning applications.

The original tide gauges provide actual historical data recorded at established coastal monitoring sites, capturing key features of tsunami behavior during real events. In contrast, the synthetic tide gauges are computationally placed within the simulation domain, allowing the extraction of wave data from hypothetical or strategically important locations that may not be

physically instrumented. The integration of both types of gauges creates a comprehensive and flexible observational framework that enhances the robustness and spatial coverage of the tsunami analysis. Further details regarding the configuration and utilization of these observation stations are provided in the following subsection.

5.3.3.1 Original tide gauge

The original tide gauge data (Figure 5.7) were obtained from permanent coastal monitoring stations maintained by the Indonesian Geospatial Agency (Badan Informasi Geospasial/BIG) through their official tidal observation platform (http://tides.big.go.id). These tide gauges, physically installed at selected coastal locations, provide records of sea level variations, which are critical for both validating the outputs of tsunami simulations. By comparing the simulated waveforms with observed data, the accuracy, reliability, and physical realism of the numerical model can be thoroughly assessed.

In this study, four existing tide gauge stations were utilized to examine and analyze the tsunami waveform time series associated with the 2018 Anak Krakatau event. These stations, Marina Jambu, Ciwandan, Panjang, and Kota Agung, are geographically distributed along the coastlines of Java and Sumatra, and were selected based on their proximity to the tsunami source and the availability of continuous observational records. The inclusion of these gauges not only supports the technical validation of simulation results but also provides insight into spatial variability in wave arrival times and amplitudes different coastal regions.

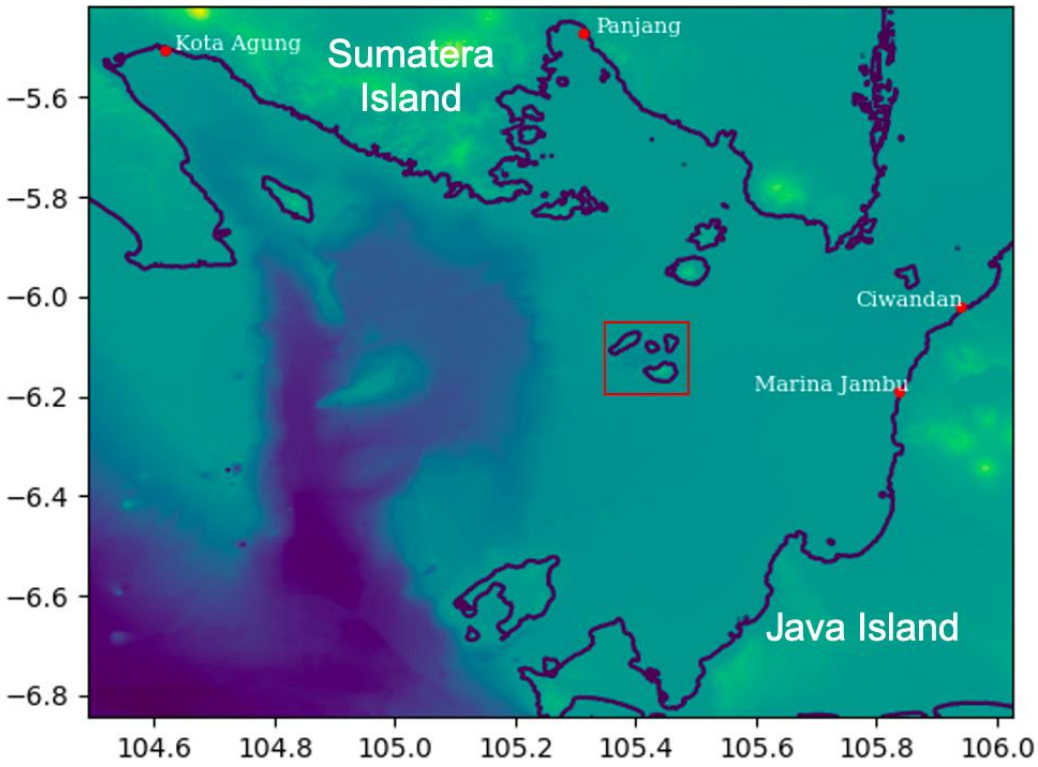


Figure 5.7 Four original tide gauges as observation stations located in Java and Sumatera Island. Anak Krakatau volcanic complex is indicated by the red rectangle

5.3.3.2 Synthetic tide gauge

Six synthetic tide gauge stations were recreated and strategically positioned around the islands near Anak Krakatau, as illustrated in Figure 5.8, following the configuration proposed by Ratnasari *et al.* (2023). Specifically, Stations 1 and 6 were located on Rakata Island, Stations 2 and 3 on Sertung Island, and Stations 4 and 5 on Panjang Island. These virtual stations served as designated input and output points for recording and analyzing tsunami waveform data within the simulation environment.

The decision to maintain all six synthetic stations in close proximity to Anak Krakatau was based on considerations of instrumental redundancy and operational resilience. Given the high-impact nature of volcanic tsunamis and the elevated risk of instrument damage due to extreme wave activity near the source, preserving multiple stations ensures continuity of data collection even if some gauges become non-functional. This approach aligns with the precautionary framework and recommendations outlined by Ratnasari *et al.* (2023), aiming to enhance the robustness of the early detection system under high-risk conditions.

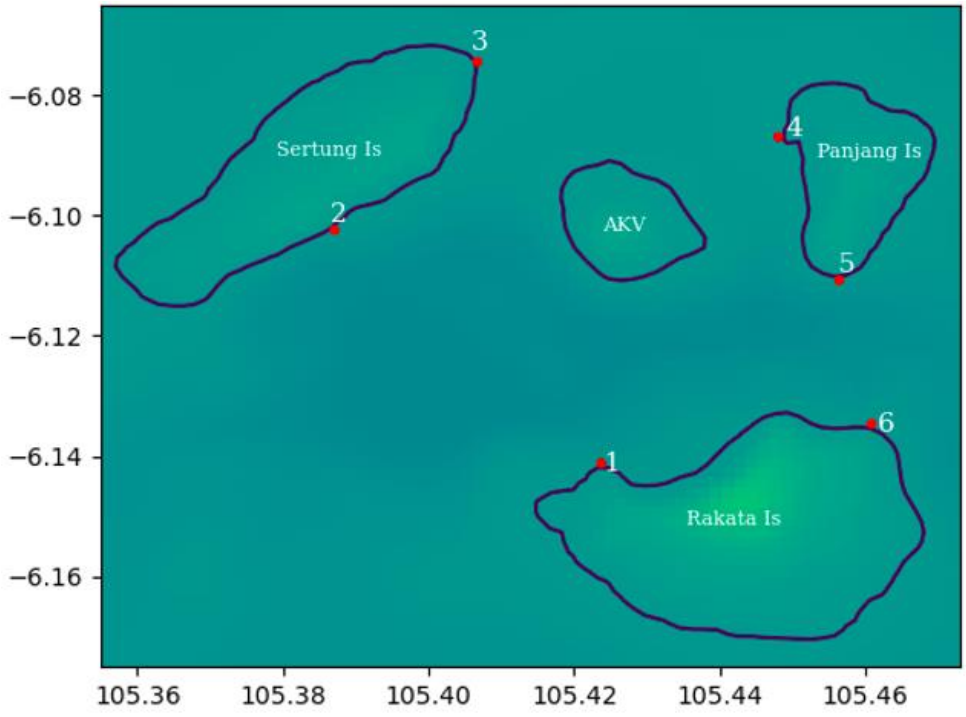


Figure 5.8 Six synthetic stations number placed on the islands around AKV: number 1 & 6 on Rakata island, number 2 & 3 on Sertung island, and number 4 & 5 on Panjang island

To support the simulation, analysis, and visualization processes in this study, a suite of Python-based tools was employed. COMCOT was utilized to generate 1000 synthetic tsunami scenarios triggered by landslide-induced collapses, based on the parameters defined in previous sections. For machine learning and deep learning implementation, the Keras library was adopted to execute various prediction models efficiently and flexibly.

In addition, Matplotlib, a comprehensive Python visualization library, was used to produce all waveform plots and graphical outputs, while NumPy was employed for numerical operations and array-based data processing. The integration of these Python libraries enabled the

development of clear, informative, and reproducible visualizations, which played a crucial role in interpreting the simulation results and communicating key findings throughout this study.

5.4 Validation: simulation result with real event

The database generated using the COMCOT software and parameters, requires validation against real event data to ensure the credibility and reliability of the simulation results. This validation process compares the simulated outputs with observed data from actual events. Additionally, this section presents the results of the simulations, including snapshots of wavefields, maximum tsunami amplitudes, and waveform time series. These outputs provide a comprehensive visualization of tsunami dynamics, capturing the propagation patterns, maximum wave heights, and temporal evolution of waveforms at observation points. The combination of validation and detailed simulation results strengthens the confidence in the generated database as a foundation for further modeling and predictive applications in intelligent decision support systems for early warning.

5.4.1 Numerical test with real event of 2018 Anak Krakatau tsunami

To assess the accuracy and reliability of the tsunami simulations in this study, a comparison (Figure 5.9) was conducted between the simulated waveforms and those recorded at four strategically located tide gauge stations in the event of December 2018. These tide gauge stations, namely Kota Agung and Panjang in Sumatera Island, and Ciwandan and Marina Jambu in Java Island, were strategically installed and maintained by the Indonesian government to monitor and record sea level variations.

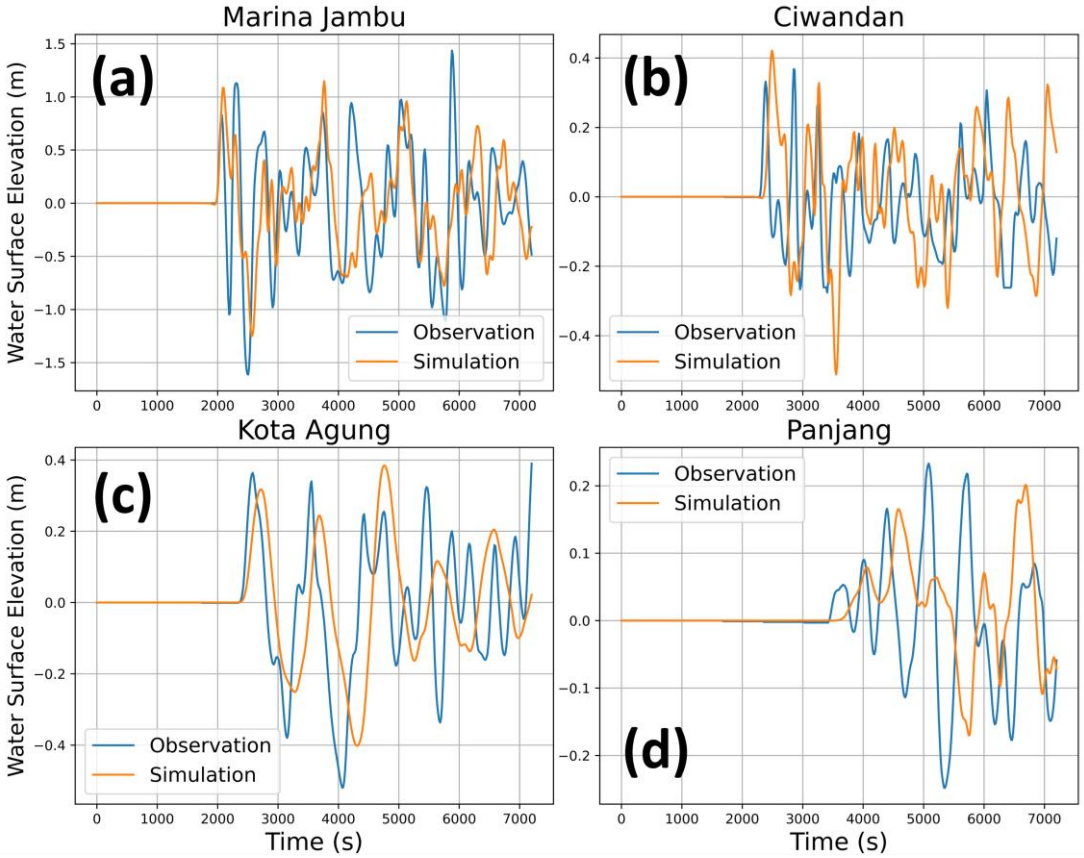


Figure 5.9 Comparison between simulated (orange) and observed (blue) tsunami waveforms from December 2018 Anak Krakatau volcano event at four tide gauges stations: (a) Marina Jambu, (b) Ciwandan, (C) Kota Agung, and (D) Panjang

The observed waveforms at Ciwandan, Marina Jambu, and Kota Agung demonstrated a good agreement with the simulated results. However, a significant discrepancy was identified at Panjang station (Figure 3d). Similar inconsistencies have also been reported in previous studies, including those by Grilli *et al.* (2019), Mulia *et al.* (2020), Paris *et al.* (2020) and Ratnasari *et al.* (2023). The differences in tsunami amplitude and arrival time may be attributed to the effects of shallow water conditions and energy dissipation, as the coastal environment near Panjang station involves complex interactions (Mulia *et al.*, 2020; (Zengaffinen *et al.*, 2020). Overall, the numerical simulation provides a reasonable explanation for the observed tsunami heights during the 2018 Anak Krakatau event, indicating that the selected software and parameters are suitable for simulating the generation and propagation of tsunamis triggered by the sector collapse of Anak Krakatau.

5.4.2 Snapshot of wavefield

This section presents a representative example of some wavefield simulation results, capturing snapshots of tsunami wave propagation from eight directions within the first 1 to 5 minutes following hypothetical landslide events. As shown in Figures 5.10 to 5.17, these snapshots illustrate the early-stage evolution of tsunami wavefields generated from landslide scenarios with eight different collapse directions: east, northeast, northwest, west, north, southwest, southeast, and south.

These visualization of snapshoots illustrate the dynamic evolution of wavefronts over time, beginning from the initial disturbance (0.000 minutes) to the formation and outward spread of tsunami waves up to 5.011 minutes. The amplitude distribution, represented in a red-blue color scale, captures both the crest (positive amplitude) and trough (negative amplitude) patterns relative to the still water level, providing insight into the asymmetrical energy dispersion caused by the directionality of the collapse.

These directional scenarios were developed to represent a broad spectrum of potential volcanic flank collapse configurations around Anak Krakatau. They offer a systematic basis for assessing how directional variation affects tsunami generation, wavefront geometry, and initial propagation dynamics. By analyzing these early waveforms, this chapter seeks to capture critical distinctions in amplitude, spatial spread, and directional energy distribution based on collapse orientation, insights that are essential for informing the subsequent predictive modeling phase. The resulting wavefield snapshots serve as both visual and quantitative references for the waveform analysis and modeling tasks discussed below.

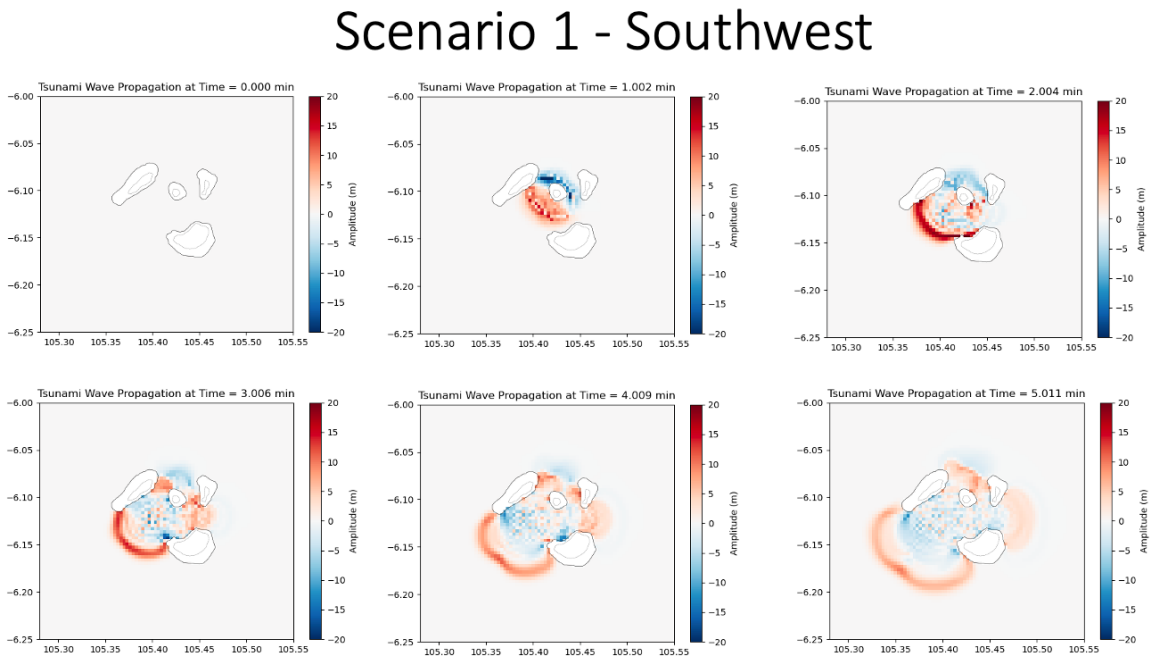


Figure 5.10 Snapshots of wavefield in 1–5 minutes in the southwest direction

Scenario 2 - South

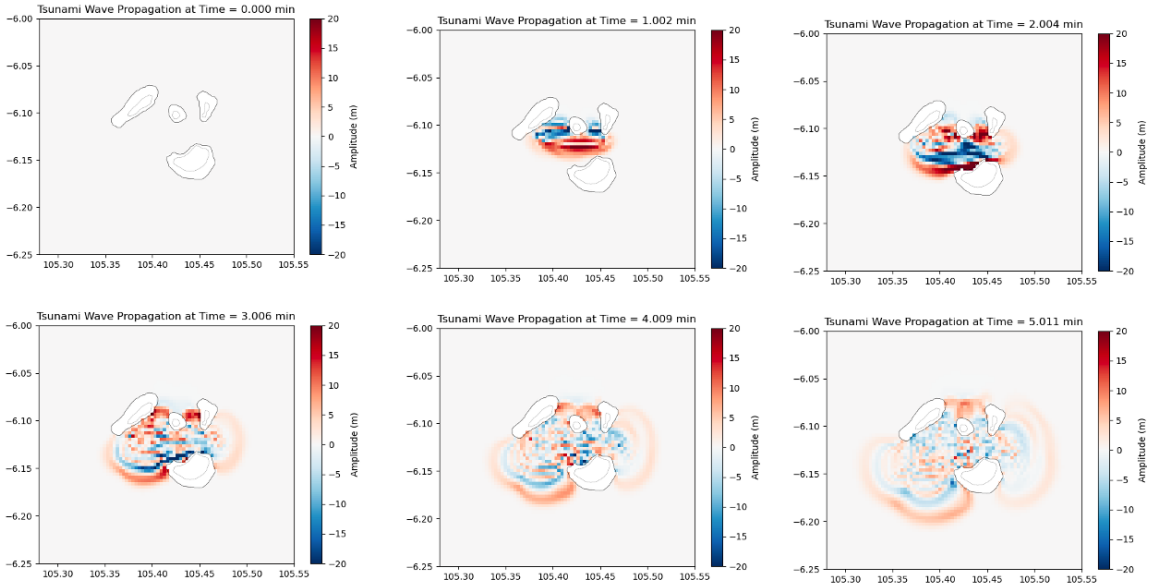


Figure 5.11 Snapshots of wavefield in 1–5 minutes in the south direction

Scenario 3 - Southeast

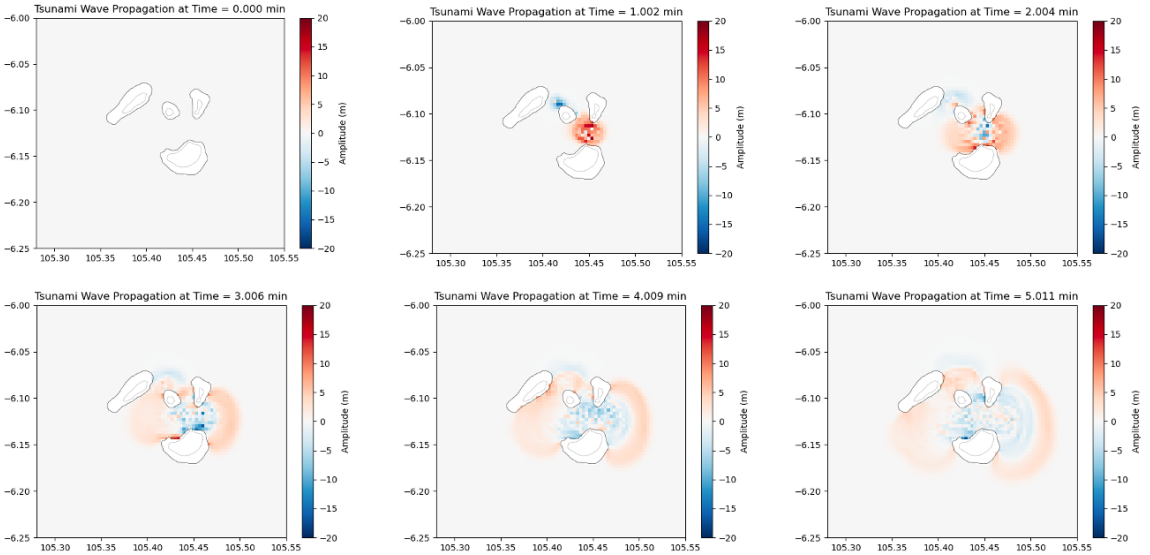


Figure 5.12 Snapshots of wavefield in 1–5 minutes in the southeast direction

Scenario 4 - East

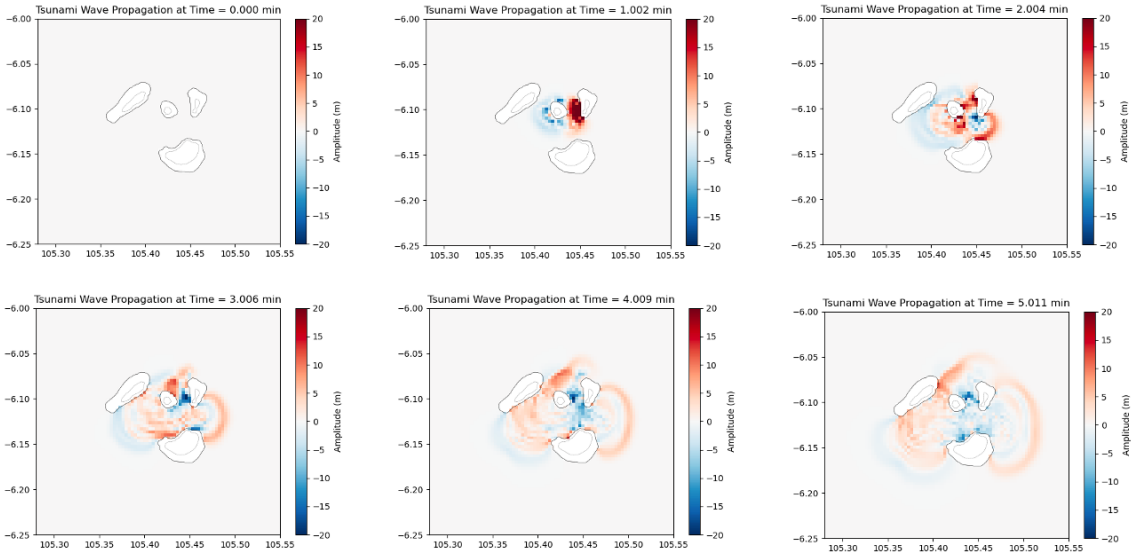


Figure 5.13 Snapshots of wavefield in 1–5 minutes in the east direction

Scenario 5 - Northeast

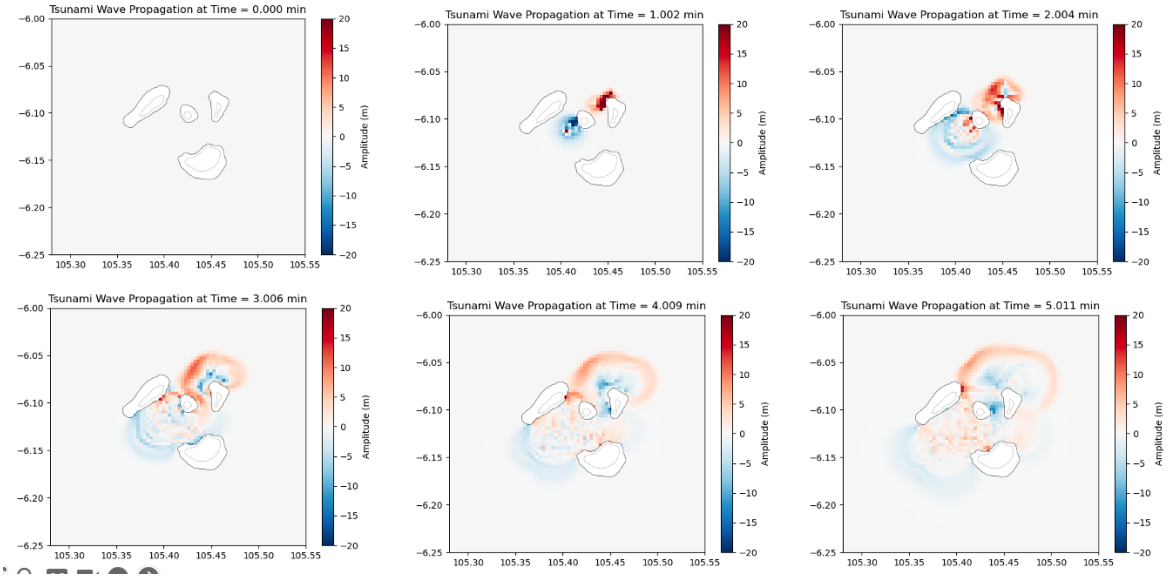


Figure 5.14 Snapshots of wavefield in 1–5 minutes in the northeast direction

Scenario 6 - North

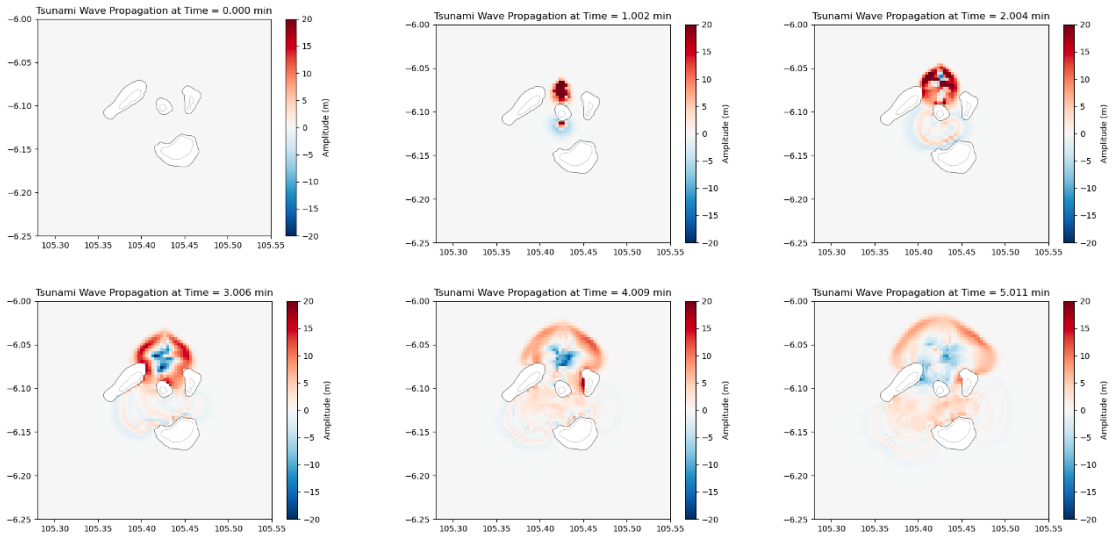


Figure 5.15 Snapshots of wavefield in 1–5 minutes in the north direction

Scenario 7 - Northwest

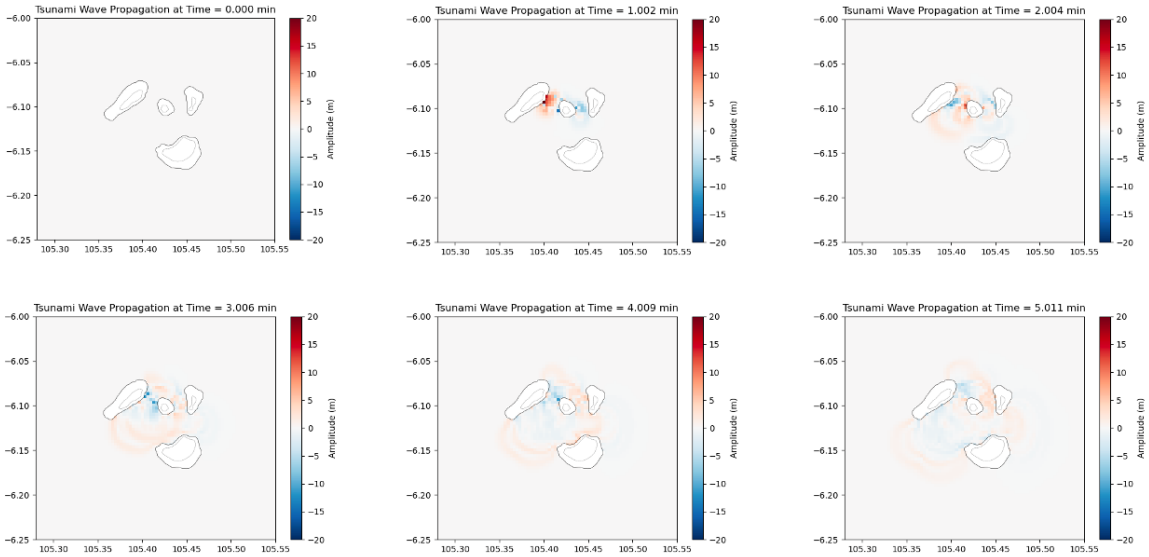


Figure 5.16 Snapshots of wavefield in 1–5 minutes in the northwest direction

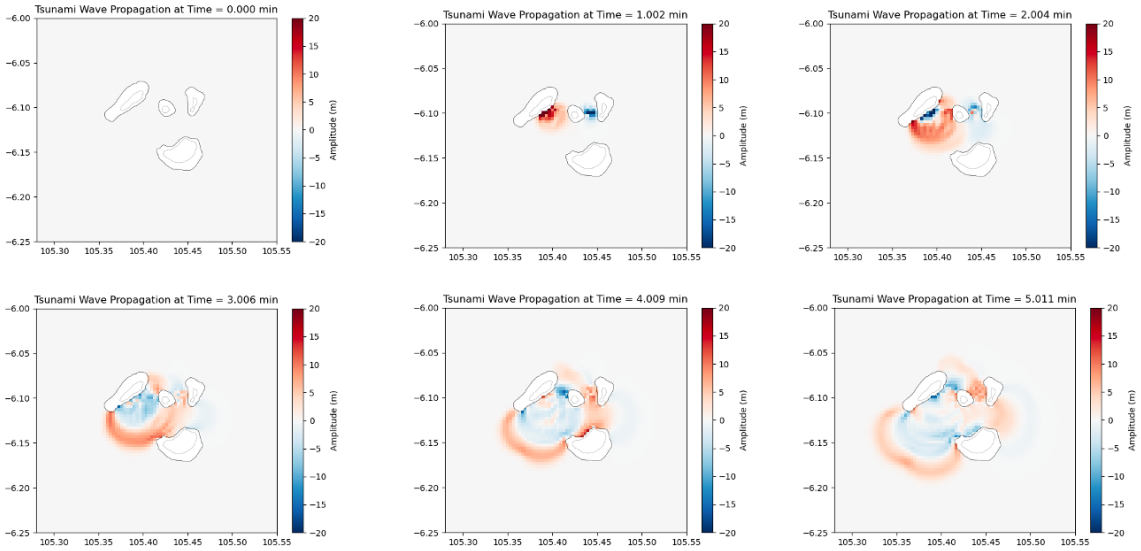


Figure 5.17 Example of snapshots of wavefield in 1–5 minutes in the west direction

5.4.3 Tsunami maximum amplitude

This section presents the simulation results of maximum tsunami amplitudes generated from a range of potential landslide directions, as depicted in Figures 5.18 to 5.19. The analysis encompasses eight distinct collapse directions: east, northeast, northwest, west, north, southwest, southeast, and south. These directional scenarios offer a comprehensive perspective on how the orientation of a volcanic landslide influences tsunami amplitude and energy distribution patterns.

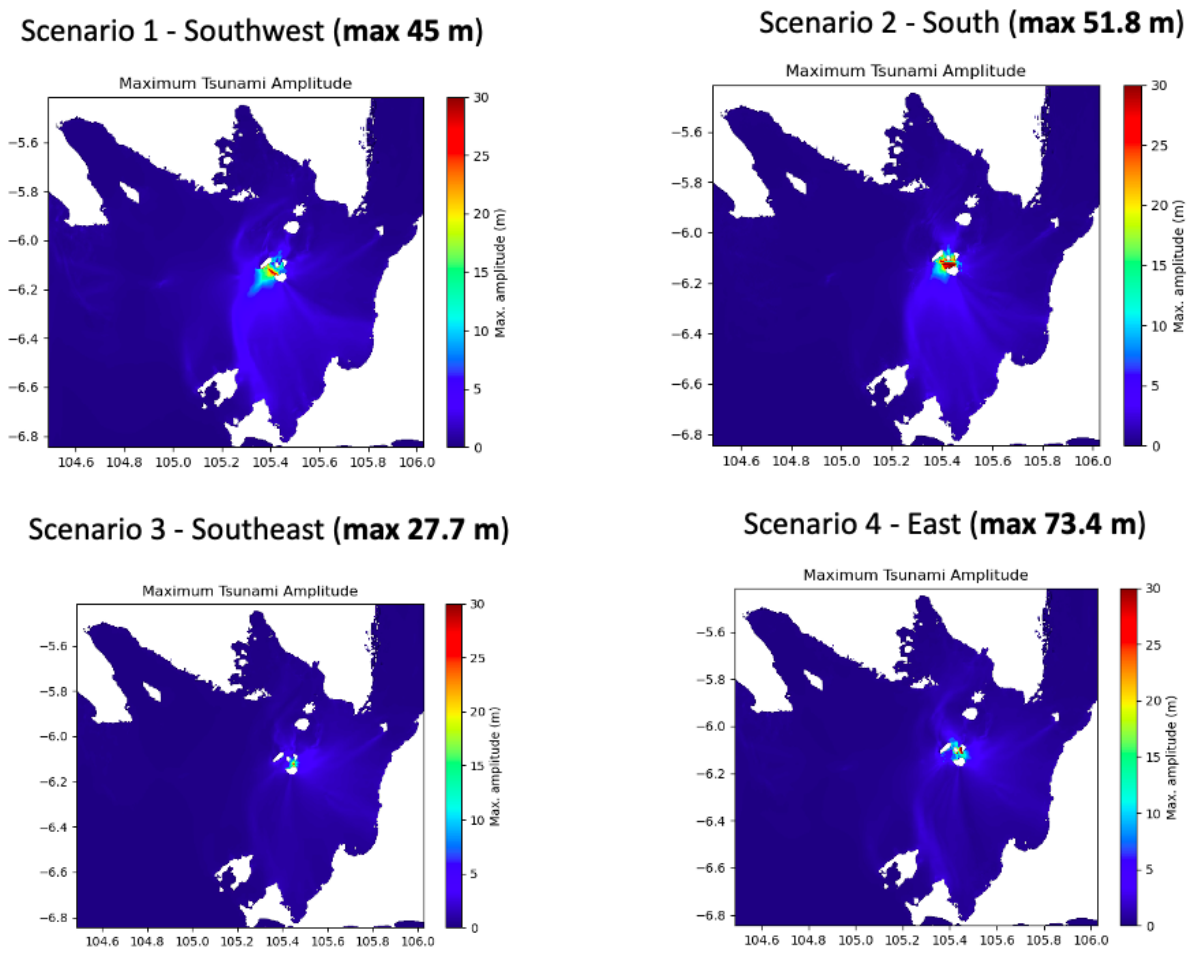


Figure 5.18 Maximum tsunami amplitude in the southwest, south, southeast, and east directions

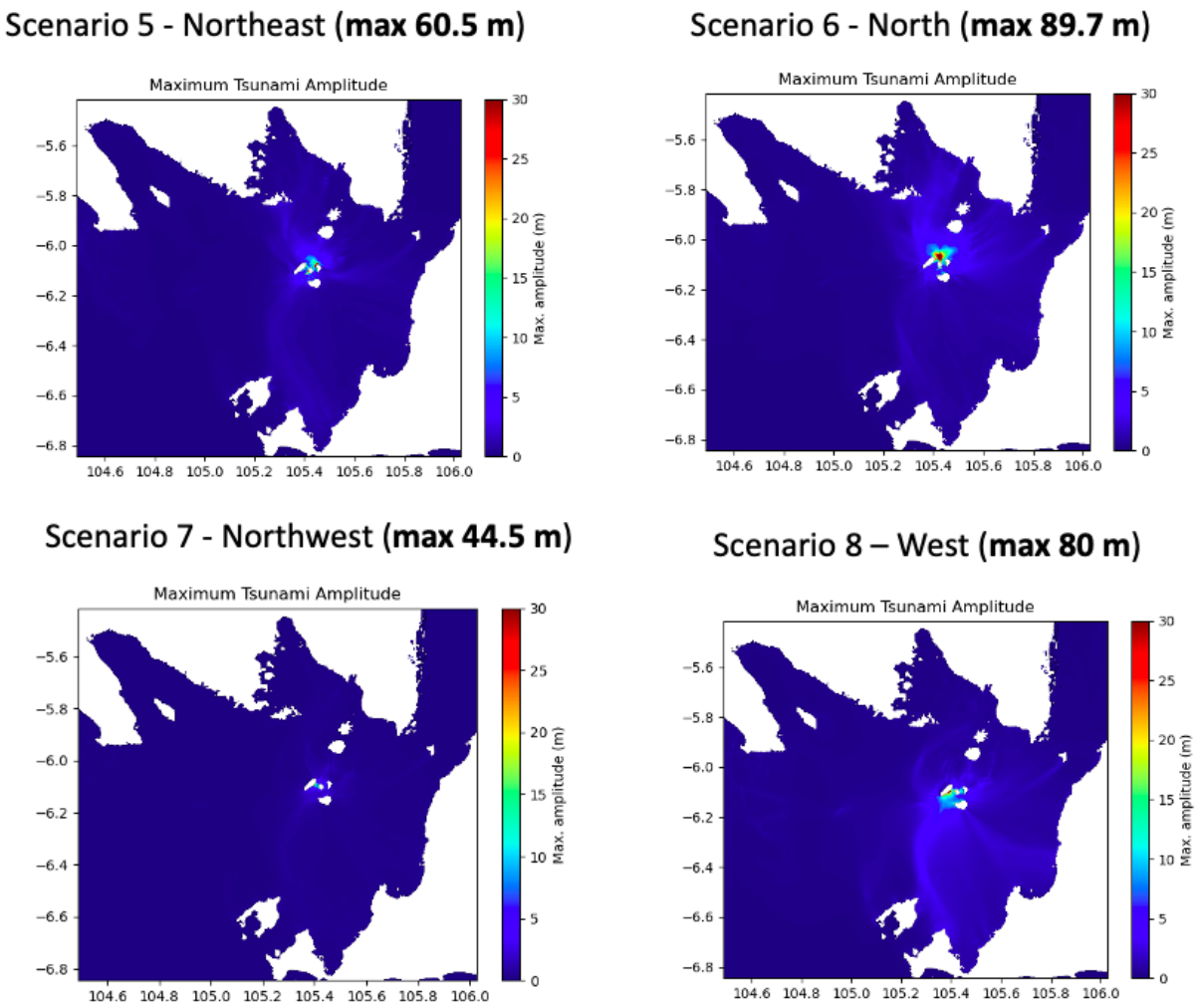


Figure 5.19 Maximum tsunami amplitude in the northeast, north, northwest, and west directions

The simulation results reveal substantial variation in maximum tsunami amplitudes across different landslide directions, emphasizing the strong directional dependence of tsunami generation in volcanic settings. This directional sensitivity is primarily attributed to the interaction between the landslide movement vector, seafloor morphology, and coastline orientation, factors that collectively influence how wave energy is transferred and amplified.

For instance, a landslide directed toward the southwest results in a peak amplitude of 45 meters, while the southward direction leads to a slightly higher wave height of 51.8 meters. Interestingly, a collapse toward the southeast produces a notably lower maximum amplitude of 27.7 meters, possibly due to wave energy being dispersed away from major coastal impact zones. In contrast, a collapse toward the east generates the highest recorded amplitude at 73.4 meters, suggesting a favorable alignment of energy focus toward shallow coastal regions in that direction.

Continuing clockwise, the northeast direction yields a significant amplitude of 60.5 meters, followed by the north, which produces the overall maximum of 89.7 meters. This could reflect constructive wave interactions or channeling effects influenced by bathymetric features in the northern basin. The northwest direction results in a more moderate amplitude of 44.5 meters,

while the west direction once again shows an elevated peak of 80 meters, indicating another potential high-impact zone due to wave refraction and coastal topography.

These results demonstrate that, even when key landslide parameters such as volume and geometry are held constant, the direction of movement alone can significantly influence the resulting tsunami intensity. The analysis further emphasizes the spatial variability and directional dependence of tsunami energy release from volcanic landslides. The detailed amplitude values presented here provide important context for assessing the potential severity of coastal impacts under different collapse scenarios.

Although this directional analysis is not directly integrated into the predictive modeling phase, it offers valuable insights for tsunami hazard assessment and risk communication, particularly in regions surrounding volcanic islands such as Anak Krakatau. Understanding these directional effects is essential for developing comprehensive multi-scenario planning and enhancing situational awareness in tsunami-prone areas.

5.4.4 Waveform time series

This section presents the simulation results of tsunami waveform time series for selected scenarios, as illustrated in Figure 5.20. These time series depict the temporal evolution of tsunami wave heights, measured in meters, over time, recorded at both original and synthetic tide gauge stations along the coasts of Sumatra and Java. By providing a one-dimensional representation of wave propagation, the time series offers detailed insights into the dynamics of tsunami waves, including their arrival times, amplitudes, and durations at various observation points.

Analyzing these waveform time series enhances our understanding of how tsunami waves evolve and interact with coastal regions over time. This temporal perspective is crucial for identifying patterns in wave behavior, such as the timing of peak amplitudes and the sequence of wave arrivals, which are essential for assessing potential impacts on coastal communities.

Importantly, the waveform time series generated in this study serve as a foundational dataset for the predictive modeling efforts discussed in Chapter 6. By capturing the nuanced temporal characteristics of tsunami waves resulting from various landslide scenarios, these time series provide the necessary input for training and validating machine learning models aimed at forecasting tsunami impacts. The integration of this detailed simulation data into predictive frameworks is vital for developing early warning and enhancing tsunami preparedness strategies.

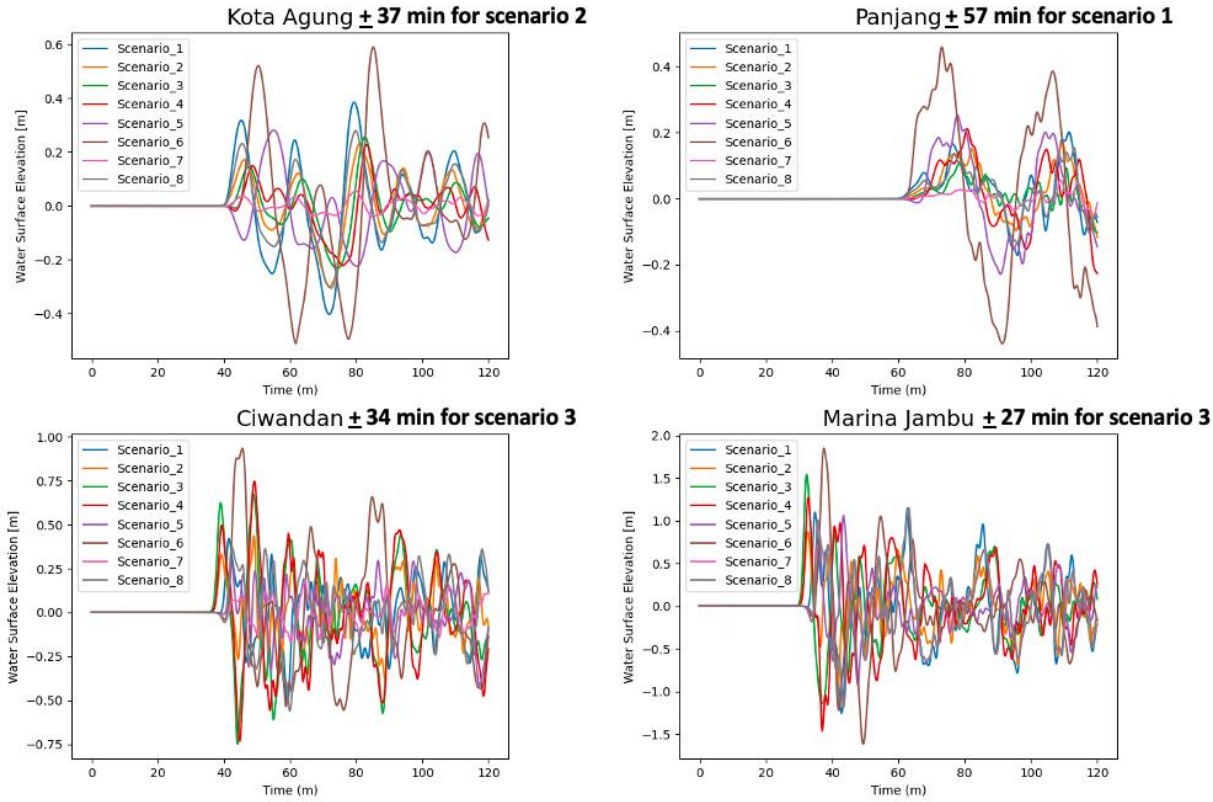


Figure 5.20 The first tsunami wave arrived within minutes recorded on tide gauges

An examination of the waveform time series reveals significant temporal variability in tsunami arrival times across different tide gauge stations and simulated scenarios. These variations reflect the complex interplay between tsunami source dynamics, propagation pathways, and coastal topography, all of which shape the timing and intensity of tsunami impact along the coast.

At Kota Agung station, the earliest arrival was recorded under Scenario 2, with waves reaching the shoreline approximately +37 minutes after the collapse event. In contrast, the Panjang station under Scenario 1 registered a later arrival time of +57 minutes, suggesting differences in wave travel distances and energy dispersion due to directional variation. The Ciwandan station exhibited a relatively fast arrival in Scenario 3, at +34 minutes, while Marina Jambu experienced the quickest overall arrival, also under Scenario 3, at just +27 minutes.

These results emphasize the scenario-specific characteristics of tsunami propagation and arrival timing. Even when input parameters such as landslide volume or geometry remain consistent, differences in collapse direction and source location result in markedly different wave behaviors at the coast. Such temporal nuances are critical for effective tsunami risk assessment and early warning operations.

Importantly, the waveform time series generated from these simulations not only enhances scientific understanding of tsunami dynamics but also serve as crucial input data for the predictive modeling approach presented in the next chapter, as mentioned before. By capturing realistic wave arrival patterns across multiple scenarios and tide gauge locations, these time series directly support the development of data-driven models that aim to improve forecast accuracy and response time in future tsunami events.

5.5 Summary

This chapter presented the development of a synthetic tsunami scenario database based on numerical simulations of 1,000 landslide-induced events around the Anak Krakatau volcanic complex. The selection of 1,000 scenarios aimed to capture a broad spectrum of potential collapse configurations, addressing the uncertainty inherent in volcanic tsunami generation while remaining computationally feasible.

The simulation process was conducted using COMCOT software, which implements the nonlinear shallow water equations. Key collapse parameters, volume, geometry, slope angle, and direction, were derived from empirical observations and integrated into the modeling framework. These parameters, identified in Chapter 4, provided the physical foundation for scenario construction.

A complete simulation workflow was implemented, encompassing initial condition setup, landslide forcing, and wave propagation modeling using a network of both original and synthetic tide gauge stations. The resulting outputs include spatial wavefield patterns, maximum tsunami heights, and full time-series waveforms. Model validation against the 2018 Anak Krakatau tsunami confirmed that the simulations could replicate essential wave characteristics with reasonable accuracy across multiple stations.

The constructed scenario database represents a comprehensive range of tsunami behaviors under varying collapse conditions and source dynamics. It establishes a solid basis for downstream predictive modeling efforts. In particular, the simulated waveform data will be utilized in Chapter 6 for the development and evaluation of predictive models aimed at full tsunami waveform forecasting and warning level classification.

To reinforce the key insights presented in this chapter, Table 5.3 provides a concise summary of the major simulation results and their relevance to the overall research framework.

Table 5.7 Summary of Key Results

Aspect	Summary of Findings
Number of Scenarios	1,000 simulated scenarios of collapse-induced tsunami events
Simulation Tool	COMCOT v1.7 using Nonlinear Shallow Water Equations (NSWE)
Grid Resolution	0.2 arc minutes (~370 meters); balancing detail and computational cost
Collapse Parameter Range	Volume: 0.04–0.33 km^3; Dip: 5°–9°; Direction: 8 azimuths (e.g., N, NE, E, S, SW, etc.)
Tide Gauge Setup	4 real tide gauges (MJ, CIW, PAN, KA) and 6 synthetic gauges near volcano
Validation	Good match with 2018 event at MJ, CIW, and KA stations; minor discrepancy at PAN
Max Tsunami Amplitudes	Range: 27.7 m (SE) to 89.7 m (North); highest from N and W directions
Arrival Time Range	Fastest arrival: 27 mins (MJ); Slowest: 57 mins (PAN)

Key Outputs	Wavefield maps, maximum amplitude maps, and waveform time series for each scenario
Use in Later Chapters	Database serves as input for predictive modelling (Chapter 6)

Chapter 6: Analysis and Implementation

This chapter presents the core analytical and implementation phase of the proposed framework, consisting of two main components: predictive modeling of tsunami maximum amplitude and tsunami waveforms using limited short-term input, and classification of early warning levels based on modeled outputs. Both components utilize the synthetic scenario database developed in the previous chapter and are evaluated through appropriate performance metrics to assess their reliability and practical applicability.

To ensure clarity and coherence, this chapter is structured into three main parts:

- Part I focuses on machine learning approaches for predicting the maximum tsunami amplitudes at multiple coastal stations based on the parameters of landslide scenarios. This section explores various regression algorithms, compares their accuracy, and identifies the most suitable models for early-stage tsunami intensity estimation.
- Part II addresses the prediction of tsunami waveforms using deep learning models, particularly focusing on how short-term observed data (the first 3 minutes of waveforms) can be used to forecast the full evolution of tsunami waves. This part demonstrates the application of sequence-to-sequence learning architectures and evaluates their capability to reconstruct complex near-field tsunami patterns.
- Part III presents a classification model for assigning early warning levels based on the predicted tsunami characteristics. Using a decision tree algorithm, this section proposes a simplified, interpretable framework for categorizing tsunami threats (e.g., “No Warning”, “Minor”, “Tsunami”, “Major”) to support operational decision-making.

Each part includes a consistent structure: Introduction, Methods, Results, Discussion, and Summary, allowing for systematic presentation and comparison of methods and findings. These three parts are integrated to form a comprehensive Chapter 6 that consolidates the analytical and implementation aspects of the proposed volcanic tsunamis forecasting framework.

6.1 Part I. Machine learning approaches for early warning of tsunami Induced by volcano flank collapse and implication for future risk management

6.1.1 Introduction

Tsunamis are most often produced by seabed displacements due to earthquake. However, they may also be produced by many other mechanisms, such as volcanic activity, the second-most frequent tsunamis after earthquakes (Paris et al., 2014). Volcanic tsunamis is a very rare event (low probability), with but create significant hazards (high impacts) in the tsunami domain (Paris et al., 2014). According to Latter (1981), volcanic tsunamis are responsible for more than 20% of deaths over the last 250 years. Furthermore, as highlighted by Day (2015), one of the significant sources of tsunamis from volcanic activity is flank collapse.

Volcanic tsunamis often occur with little or no warning compared to tsunami generated by earthquakes (Maeda et al., 2015), and could be classified as a silent tsunami (Mulia et al., 2020), making it difficult to provide early warning. One of the most recent and significant instances is the tragic event that occurred on December 22, 2018, when the Anak Krakatau volcano, one of the world's most dangerous volcanoes (Dogan et al., 2021; Zorn et al., 2022), unexpectedly produced a tsunami due to volcanic activity, killing hundreds of people due to lack of notice of early warning (Grilli et al., 2019).

The promising approaches for providing tsunami early warnings currently include machine learning and deep learning algorithms. Machine learning can solve complex numerical modelling of tsunamis that require high computational cost and time, becoming rapid at a very fast speed and low cost (Fauzi and Mizutani, 2020; Siek and Rafles, 2022). This is becoming a noteworthy advancement, particularly in terms of how it will affect the development of a tsunami early warning system that will act as an evacuation prompt in a real-time manner.

In the field of tsunami research, several studies have already explored these approaches. Fauzi and Mizutani (2020) applied two machine learning models, a convolutional neural network and a multilayer perceptron, to estimate tsunami inundation in real-time which results better in term of time than the previous method, NearTIF, by proposed Gusman et al. (2014). Kamiya et al. (2022) predict tsunami flow depth distribution in real-time using data from earthquake-induced tsunamis in the Nankai Trough Japan were constructed by numerical simulations by adopting the machine learning algorithm namely regression and k-means for clustering. Mulia et al., (2022) proposed a deep network architecture of deep learning to estimate high-resolution tsunami inundation and achieving a result of approximately a 90% reduction in real-time computational efforts compared to conventional physics-based. Moreover, Dharmawan et al. (2023) developed tsunami tide prediction in shallow water using recurrent neural networks to improve the model implementation of tsunami early warning in Indonesia with the result of time prediction under 0.1 second. All the previous works have shown that machine learning and deep learning algorithms produce better and more rapid results compared to the conventional method. However, the previous works mentioned above only focused on the tsunamis triggered by earthquakes.

Unlike the tsunami triggered by earthquakes which the mechanism well understood, volcanic tsunami event such as volcanic collapse, have complex trigger mechanisms. Thus, the tsunami

generation scenario and also the parameters used for the database generation from the numerical simulation are different from the tsunamis caused by earthquakes. Therefore, the study for development of early warning for volcanic tsunamis using the learning approach is paramount to complement the tsunami early warning ecosystem from various sources. Ratnasari et al. (2023) developed the early warning system using real-time forecasting method based on a pre-computed database for future tsunamis accompanied by collapse of volcano. This method considered spatiotemporal changes in water surface level and flux around volcano. However, this method and database of scenario covered less of the parameters and the future potential of others mechanism of volcano collapse, such as collapse directions which is expected to landslide in all directions.

In this study, we focus on tsunamis generated by volcanic collapse, using the 2018 Anak Krakatau volcanic eruption as a case study to advance tsunami early warning. By generating 320 collapse scenarios with diverse parameters, we aimed to predict the maximum tsunami amplitudes observed at coastal locations. Six synthetic observation stations were strategically positioned around the Anak Krakatau volcano to capture critical data. Machine learning algorithms, including Random Forest, Gradient Boosting, and Neural Networks, were employed to link the highest tsunami amplitudes at these stations with those observed at the coast. Furthermore, we innovatively leveraged two deep learning algorithms—Long Short-Term Memory (LSTM) and Complex LSTM—to forecast full waveform time series at coastal observation points, enabling a more comprehensive understanding of tsunami behavior. The study also highlights the practical implications of these predictions, particularly for enhancing risk management strategies, such as optimizing evacuation decision-making based on maximum amplitude and arrival time forecasts.

6.1.2 Method

6.1.2.1 Machine learning model flow chart

Figure 6.1 illustrates the machine learning model development flowchart for volcanic tsunami early warning. The process begins with the volcanic tsunami dataset, followed by data preprocessing, where simulation results are adjusted, including interpolating two-hour simulation data into minute-level intervals for consistency. Feature engineering is applied to extract critical features for model configuration, such as the number of trees in Random Forest, hyperparameters for Gradient Boosting, or architectural details for models like LSTM and Complex LSTM. Machine learning optimization involves training the model iteratively while evaluating performance against predefined accuracy thresholds. If the accuracy is insufficient, hyperparameter tuning is conducted, and the cycle repeats until the desired accuracy is achieved, resulting in a trained algorithm ready for deployment.

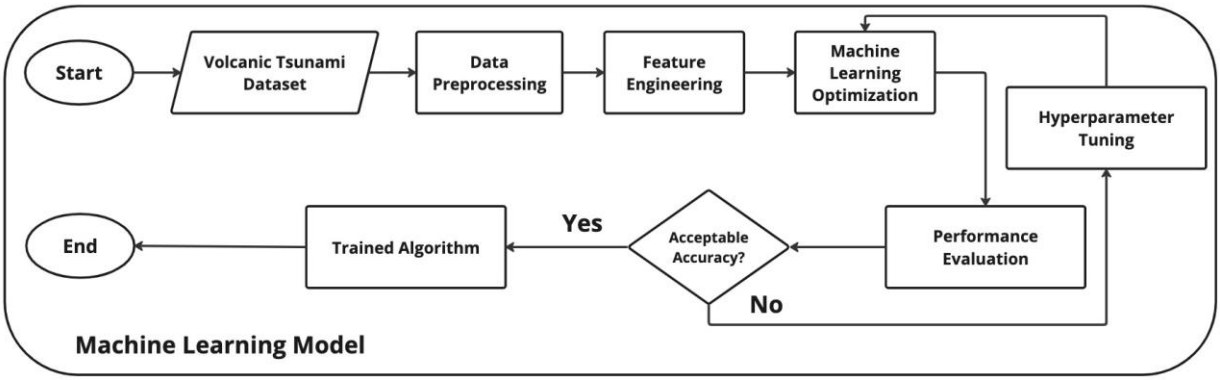


Figure 6.1 Machine learning model flow chart

6.1.2.2 Data Preprocessing

The dataset consists of 320 scenarios, each representing a two-hour tsunami simulation generated from numerical modeling. These scenarios provide detailed time-series data of tsunami wave heights, but the raw output of simulation often contains irregular time intervals or non-uniform data points, requiring standardization before model development. To address this, data preprocessing is applied to ensure consistency and usability for machine learning models. During data preprocessing, the simulation results are adjusted by interpolating the two-hour simulation data into uniform, minute-level intervals. This step transforms the raw data into a structured format with 120 data points per scenario, ensuring temporal consistency across all scenarios. The interpolation process not only improves data uniformity but also facilitates seamless integration with downstream machine-learning workflows. By standardizing the temporal resolution, the dataset becomes suitable for feature extraction, model training, and evaluation.

6.1.2.3 Feature Engineering and Machine Learning Optimization

Feature engineering and machine learning optimization focus on identifying and extracting critical parameters to enhance model performance. To predict maximum tsunami elevations at coastal observation points (tide gauges), algorithms such as Random Forest, Gradient Boosting, and Neural Networks are utilized. In addition, models like LSTM and Complex LSTM are designed to capture the full tsunami waveform, leveraging their ability to process sequential data effectively. These approaches will be detailed further in subsequent sections. Model performance is assessed using Root Mean Squared Error (RMSE) for maximum tsunami elevation predictions and both Mean Squared Error (MSE) and Mean Absolute Error (MAE) for the LSTM algorithm, ensuring a comprehensive evaluation of accuracy and consistency.

6.1.2.3.1 Random Forest

Random Forest, an ensemble learning technique, was utilized to predict maximum tsunami elevations based on multi-scenario time series data. The dataset was carefully prepared, extracting features related to the highest elevations recorded during tsunami events. This data was then split into training and testing sets using an 80:20 ratio to ensure robust model evaluation. The Random Forest Regressor from the sklearn.ensemble module was chosen due to its ability to handle high-dimensional data and capture non-linear relationships effectively. A key parameter, the number of decision trees ($\eta estimatiors$), was set to 100. This choice

strikes a balance between model complexity and computational efficiency, allowing the model to generalize well without overfitting to the training data.

The performance of the Random Forest model was evaluated using the Root Mean Squared Error (RMSE) metric on the testing data, providing insights into its predictive accuracy and generalization ability. The trained Random Forest model was saved for future deployment in tsunami elevation forecasting tasks.

6.1.2.3.2 Gradient Boosting

In addition to Random Forest, Gradient Boosting was employed to further enhance the accuracy of maximum tsunami elevation predictions. The Gradient Boosting Regressor from sklearn.ensemble was chosen for its capability to sequentially improve model predictions by minimizing errors. Several parameters were carefully selected to optimize the Gradient Boosting model's performance. The number of decision trees ($\eta estimatiors$) was also set to 100, ensuring a sufficiently complex model without excessive computational overhead. A learning rate of 0.1 was chosen to control the contribution of each tree to the ensemble, preventing overfitting and promoting model stability. The maximum tree depth ($maxdepth$) was limited to 3 to avoid excessively complex trees that could lead to overfitting.

Similar to Random Forest, the performance of the Gradient Boosting model was evaluated using RMSE on the testing data, providing a comprehensive assessment of its predictive accuracy. The trained model was saved for future use in the forecasting framework.

6.1.2.3.3 Neural Network

To complement the ensemble methods, a Neural Network model was integrated to capture complex non-linear relationships in the data. The input features were standardized using StandardScaler to ensure uniformity and facilitate model convergence during training. The Neural Network architecture consisted of an input layer with 64 neurons, a hidden layer with 64 neurons, both utilizing the Rectified Linear Unit (ReLU) activation function, and an output layer with 1 neuron for regression tasks. The choice of 64 neurons in each layer was based on experimentation and empirical evidence, balancing model complexity and capacity to capture intricate patterns in the data.

The model was compiled with the Adam optimizer and a learning rate of 0.001 to ensure stable and efficient training. Training was conducted for 100 epochs with a batch size of 32, allowing the model to learn complex patterns in the data gradually. The model's performance was evaluated using RMSE on the testing data, providing insights into its predictive accuracy and generalization ability. Both the trained Neural Network model and the scaler used for data standardization were saved for future forecasting tasks within the predictive framework.

6.1.2.3.4 Long-Short Term Memory (LSTM) and Complex Long-Short Term Memory (LSTM)

LSTM, a specialized type of recurrent neural network (RNN), stands out for its ability to capture long-term dependencies in sequential data (Van Houdt, Mosquera and Nápoles, 2020) (Figure 6.2). Unlike traditional neural networks, LSTM is designed to overcome the vanishing gradient problem, allowing it to retain and utilize information over extended sequences, a

critical feature for modeling time-series data. One key aspect of LSTM's architecture is its intricate system of memory cells, input gates, forget gates, and output gates. LSTM is widely used for many application including the very important area such as health (Chae et al., 2022). However, is still limited application to disaster prediction, especially tsunami waveform trigger by volcano collapse.

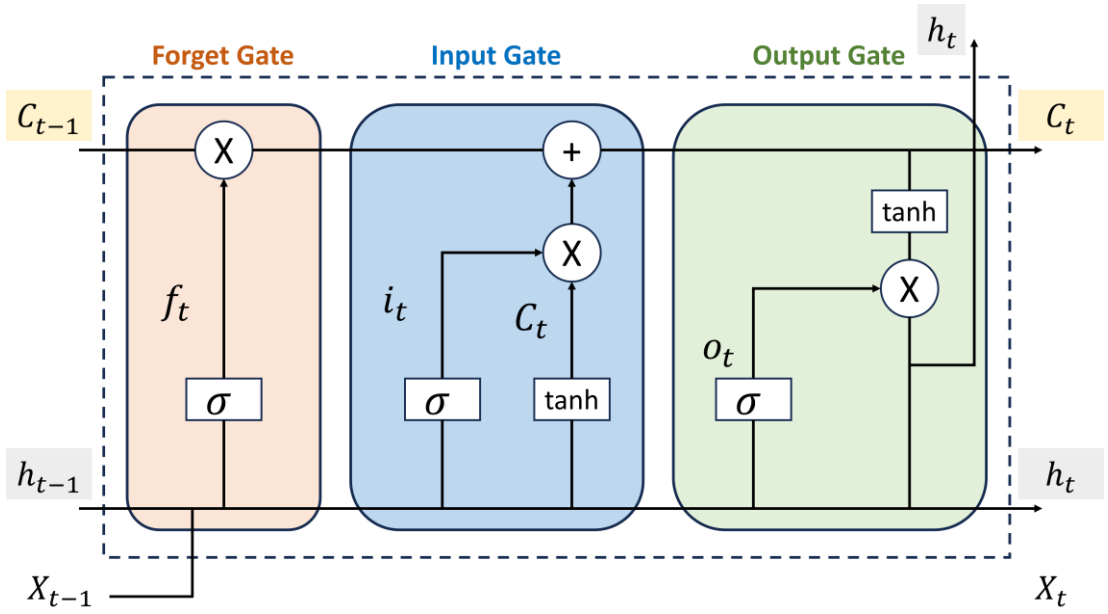


Figure 6.2 LSTM architecture block

Moreover, the detail of proposed early warning model outlined in this research involves the strategic recording of initial tsunami waves triggered by volcanic landslides, utilizing them as input for the subsequent prediction model. The novelty lies in the ability to extrapolate the entire tsunami wave pattern from these early observations. We used the first 9 minutes as best input based on the trial-and-error method. The performance of the LSTM was evaluated using MSE and MAE on the testing data, providing a comprehensive assessment of its predictive accuracy. In addition, we also compared the LSTM with the Complex LSTM model. In complex LSTM, we added Dropout option in every layer of LSTM. More detail of layers between LSTM and Complex LSTM shown in Table 6.1.

Table 6.1 LSTM and Complex LSTM model layer definition

Model	**Layers**
LSTM	LSTM (50) -> Dense (1)
Complex LSTM	LSTM (50) -> Dropout (0.2) -> LSTM (50) -> Dropout (0.2) -> LSTM (50) -> Dropout (0.2) -> Dense (1)

The performance of the LSTM was evaluated using both Mean Squared Error (MSE) and Mean Absolute Error (MAE) metrics. MSE was selected for its sensitivity to larger errors. It is particularly useful for emphasizing significant discrepancies between predicted and actual values. MAE providing a balanced perspective on the model's prediction accuracy by averaging

the absolute differences between predicted and actual values. The complementary use of MSE and MAE ensures a comprehensive evaluation of the LSTM's performance.

6.1.3 Result

6.1.3.1 Snapshots of wavefield

The presented some examples of initial tsunami wavefield snapshot from simulation as illustration (Figure 6.3), specifically focused on the initial 1-5 minutes following potential landslides across diverse directions with tsunami-generating potential. This investigation spans eight distinct landslide directions, namely east, northeast, northwest, west, north, southwest, southeast, and south, providing a comprehensive exploration of the various geological triggers that could lead to a tsunami event.

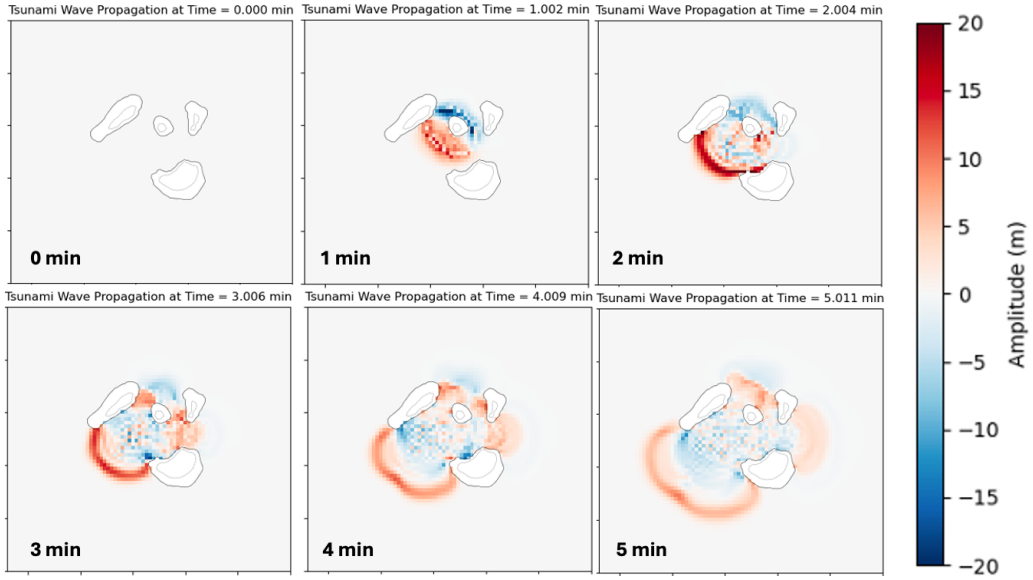


Figure 6.3 Example of tsunami wavefield snapshots for one simulation scenario during the first 1–5 minutes

6.1.3.2 Tsunami maximum amplitude prediction

We evaluated the performance of the tsunami maximum amplitude (elevation) by comparing the predicted max elevation with actual max elevation from database simulation result using three machine learning algorithms: Random Forest, Gradient Boosting, and Neural Network. We plotted the prediction result (Figure 6.4-6.7) on s scatter diagram for each coastal observation station, Marina Jambu, Ciwandan, Panjang, and Kota Agung, in order. Moreover, we also plotted the residuals to visualize the result of frequency distribution.

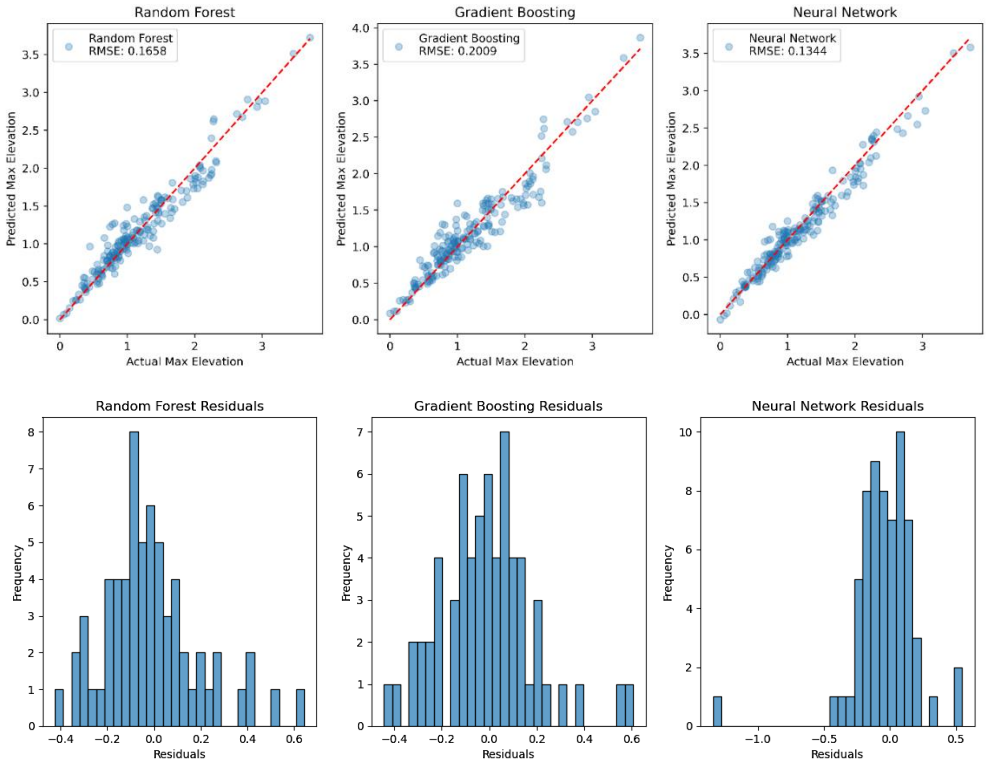


Figure 6.4 Scatter diagram and residuals histogram of tsunami max amplitude prediction on Marina Jambu station

The scatter plots in Fig 8 for all three models show a strong positive correlation between the actual and predicted maximum elevations, indicated by the proximity of data points to the red dashed line representing the ideal 1:1 relationship. Among the models, Gradient Boosting demonstrated the lowest Root Mean Squared Error (RMSE) of 0.1945, suggesting its superior accuracy in predicting maximum elevations. The Random Forest model followed closely with an RMSE of 0.2058, while the Neural Network model exhibited a slightly higher RMSE of 0.2457. These results imply that while all three models are effective, Gradient Boosting provides the most precise predictions in this context. The residual histograms further elucidate the performance of these models by displaying the distribution of prediction errors. For the Random Forest model, the residuals are symmetrically distributed around zero, with a slight skew towards underestimation. The Gradient Boosting model's residuals exhibit a similar symmetrical distribution but with a more pronounced peak near zero, indicating fewer large errors and higher consistency in predictions. The Neural Network model, although generally effective, shows a wider spread of residuals, suggesting more frequent larger errors compared to the other models.

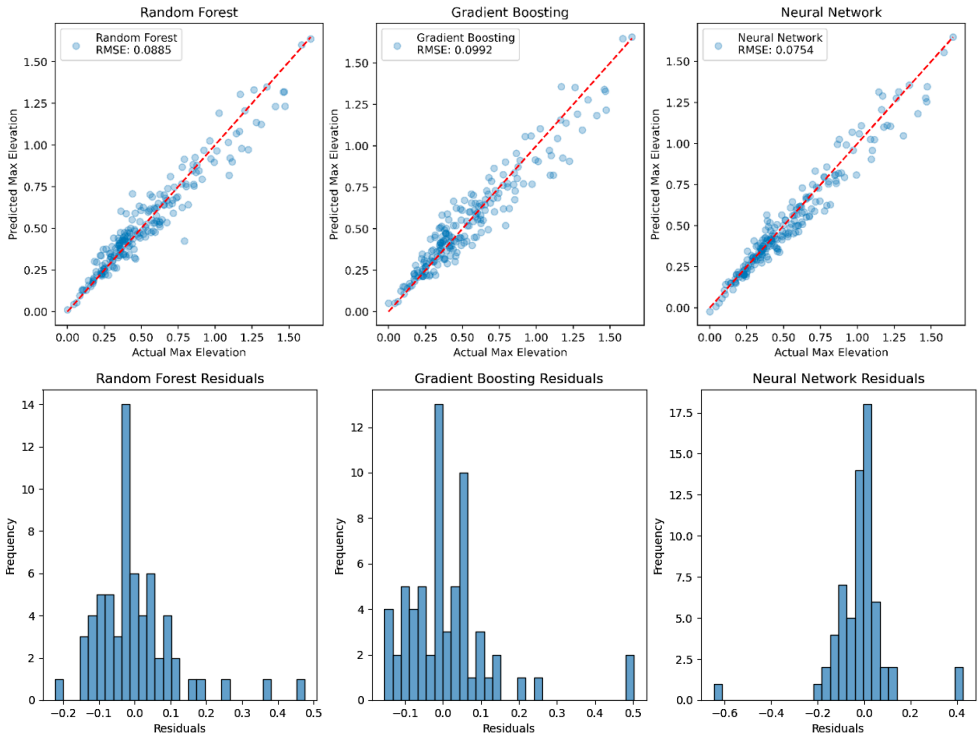


Figure 6.5 Scatter diagram and residuals histogram of tsunami max amplitude prediction on Ciwandan station

For result of Ciwandan station shown in Fig 9, the scatter plots for all three models show a strong positive correlation between the actual and predicted maximum elevations compared to Marina Jambu station. Among the models, Random Forest demonstrated the lowest Root Mean Squared Error (RMSE) of 0.1126, suggesting its superior accuracy in predicting maximum elevations. The Gradient Boosting model followed closely with an RMSE of 0.1200, while the Neural Network model exhibited a slightly higher RMSE of 0.1288. These findings suggest that while all three models work well, Random Forest makes the most accurate predictions in this situation. The residual histograms, which show the distribution of prediction errors, shed additional light on how well these models perform. In comparison to the other models, the Random Forest model exhibits a wider distribution of residuals of small errors, despite being the most effective. However, the errors are still relatively small. The residuals of the Gradient Boosting model show a symmetrical distribution that is similar to this, but with a more noticeable peak near zero, which denotes fewer large errors and better prediction consistency. The residuals in the Neural Network model have a symmetric distribution around zero, with a slightly wider range of errors.

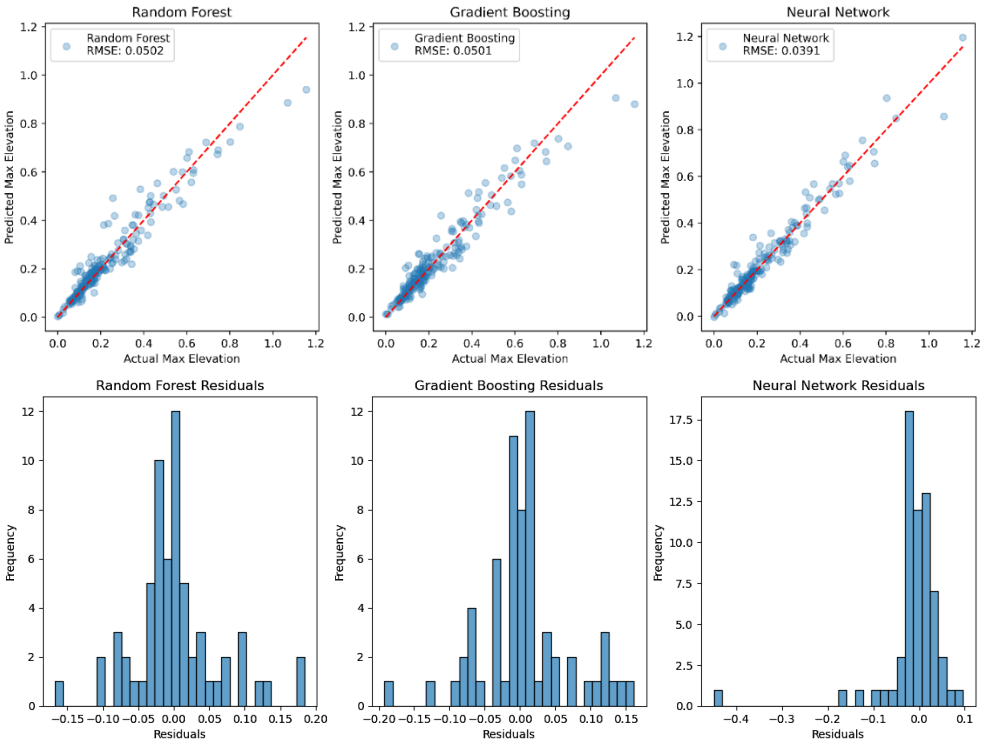


Figure 6.6 Scatter diagram and residuals histogram of tsunami max amplitude prediction on Panjang station

The scatter plots on Panjang shown in Fig 10 demonstrate a strong correlation between the actual and predicted maximum elevations for all three models compared to previous stations. This is evidenced by the concentration of data points along the red dashed line, which represents a perfect 1:1 correlation. Among the models, Random Forest achieved the lowest Root Mean Squared Error (RMSE) of 0.0609, indicating its superior performance in this particular prediction task. Gradient Boosting followed with an RMSE of 0.0631, and the Neural Network model had a slightly higher RMSE of 0.0702. These results suggest that while all three models are highly effective, Random Forest provides the most accurate predictions, closely followed by Gradient Boosting. The residual histograms further clarify the performance by illustrating the distribution of prediction errors. For the Random Forest model, the residuals show a symmetric distribution around zero with a slight skew towards underestimation, which is a positive indicator of model reliability. The Gradient Boosting model exhibits a similar symmetrical distribution of residuals, with a notable peak around zero, signifying fewer large errors and greater consistency. The Neural Network model, although performing well, displays a wider spread of residuals, indicating more frequent larger errors compared to the other two models.

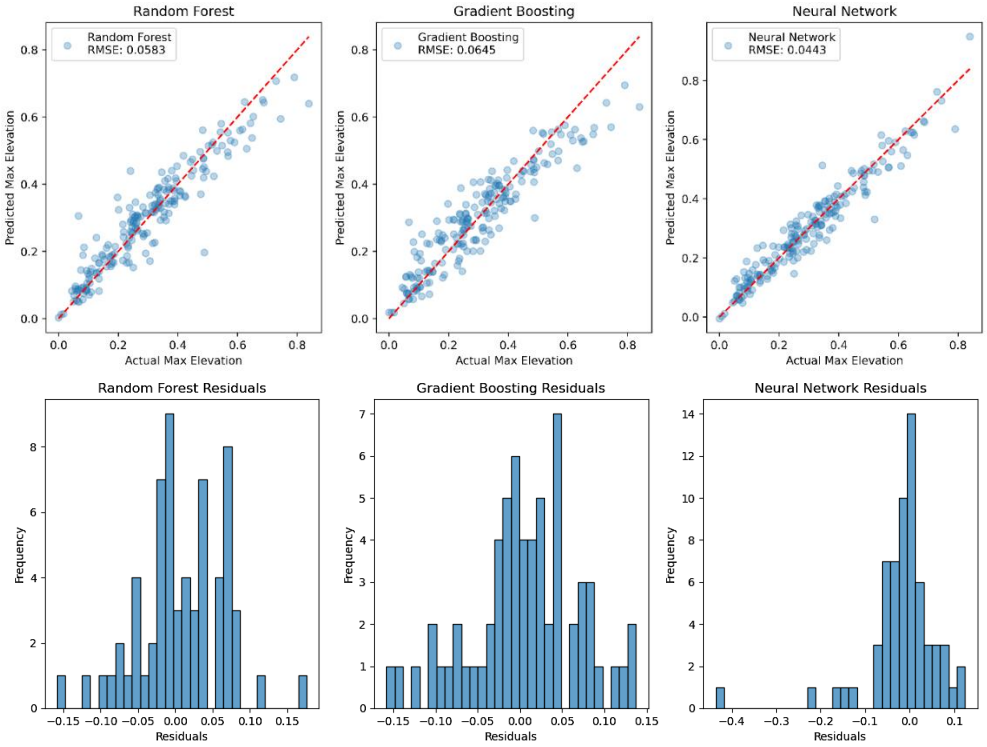


Figure 6.7 Scatter diagram and residuals histogram of tsunami max amplitude prediction on Kota Agung station

For Kota Agung station shown in Fig 11, the scatter plots show a strong correlation between the actual and predicted maximum elevations for all three models. This is evidenced by the concentration of data points along the red dashed line, which represents a perfect 1:1 correlation. Among the models, Random Forest achieved the lowest Root Mean Squared Error (RMSE) of 0.0586, indicating its superior performance in this particular prediction task. Gradient Boosting followed with an RMSE of 0.0644, and the Neural Network model had a slightly higher RMSE of 0.0809. These findings imply that while all three models are very successful, Random Forest and Gradient Boosting offer the most accurate predictions. The residual histograms, which show the distribution of prediction errors, shed additional light on the performance. The residuals for the Random Forest model exhibit a symmetric distribution around zero with a small underestimation skew, which is a sign of the model's reliability. The residual distribution of the Gradient Boosting model is similarly symmetrical and has a prominent peak near zero, which denotes fewer large errors and higher consistency. Compared to the other two models, the Neural Network model exhibits a wider residual spread, which suggests a higher frequency of larger errors, despite its good performance.

Notably, two data points in the Neural Network prediction result appears to deviate significantly from the general trend and is visually detached from the main cluster (see Figure 6.6 and 6.7, top right). This outlier may reasonably be considered an extreme event candidate. As this study focuses on high-impact, low-probability events, identifying such anomalies is crucial. The corresponding point was found at the Panjang and Kota Agung tide gauge stations, indicating that future analysis and model refinement may incorporate the detection of such extreme patterns as part of the forecasting framework.

6.1.3.3 Tsunami full waveform prediction

In this section, we present a comparative analysis of the performance of two Long Short-Term Memory (LSTM) models: a standard LSTM and a Complex LSTM as shown in Figure 6.8, with respect to their training and validation losses, as well as their prediction accuracy on wave height data. From six synthetic observation stations, we only used one station, Rakata 6 (see Fig 3B), as a test of tsunami full time series prediction.

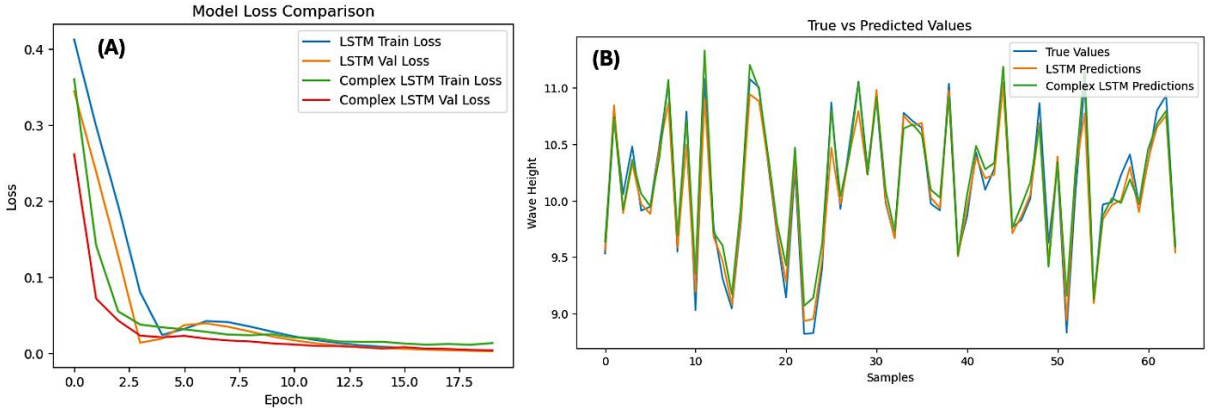


Figure 6.8 Comparative analysis of the performance of two Long Short-Term Memory (LSTM) models. (A) LSTM and Complex LSTM mode loss comparison. (B) true vs predicted values result

The left panel of Figure 6.8 (A) illustrates the loss curves for both the standard LSTM and the Complex LSTM models over 20 epochs. This loss curves useful to understand the model performance of time series prediction. The reason we used 20 epochs is trial and error method. First, we used few numbers of epochs and then use high number, but the model getting better when running the around 20 epochs.

The blue and orange curves represent the training and validation losses for the standard LSTM, while the green and red curves depict the training and validation losses for the Complex LSTM. Initially, both models exhibit a rapid decrease in loss, indicating effective learning from the data. The standard LSTM shows a sharp decline in training loss, dropping below 0.1 within the first 5 epochs, and then stabilizing with minimal fluctuations around 0.05. The validation loss for the standard LSTM follows a similar trend but stabilizes at a slightly higher value compared to the training loss, suggesting slight overfitting. Conversely, the Complex LSTM demonstrates a slightly better performance in terms of both training and validation losses. The training loss for the Complex LSTM drops more sharply than the standard LSTM, reaching values close to zero around the 5th epoch. The validation loss for the Complex LSTM also stabilizes at a lower value compared to the standard LSTM, indicating better generalization and less overfitting.

The right panel of Figure 6.8 (B) compares the true wave heights (blue curve) with the predictions from the standard LSTM (orange curve) and the Complex LSTM (green curve). Both models are able to capture the general trend and variability of the true wave heights, indicating their efficacy in predicting the target variable. Upon closer examination, it is evident that the Complex LSTM predictions align more closely with the true values across the entire sample range. This is particularly noticeable in the regions with higher variability, where the Complex LSTM tracks the peaks and troughs of the true wave heights more accurately than the Complex LSTM. The standard LSTM, while generally accurate, shows slightly larger

deviations from the true values, particularly in regions of rapid change. The comparative analysis clearly demonstrates the superior performance of the Complex LSTM model over the standard LSTM in the case of training and testing model. The Complex LSTM not only achieves lower training and validation losses but also provides more accurate predictions of wave heights.

6.1.3.3.1 Testing with short duration observation input

Given the promising accuracy results from our initial tests, we proceeded to evaluate the model using short-duration observation inputs. This approach aims to determine whether the model can generate reliable predictions from brief observation periods, a crucial capability for tsunami early warning systems. Figure 6.9 (A) presents the outcomes of forecasting wave heights using a short duration observation input of 9 minutes, with predictions generated by both a standard Long Short-Term Memory (LSTM) model and a Complex LSTM model. The 9 minutes achieved from the trial-and-error method while finding the optimal input for full waveform prediction.

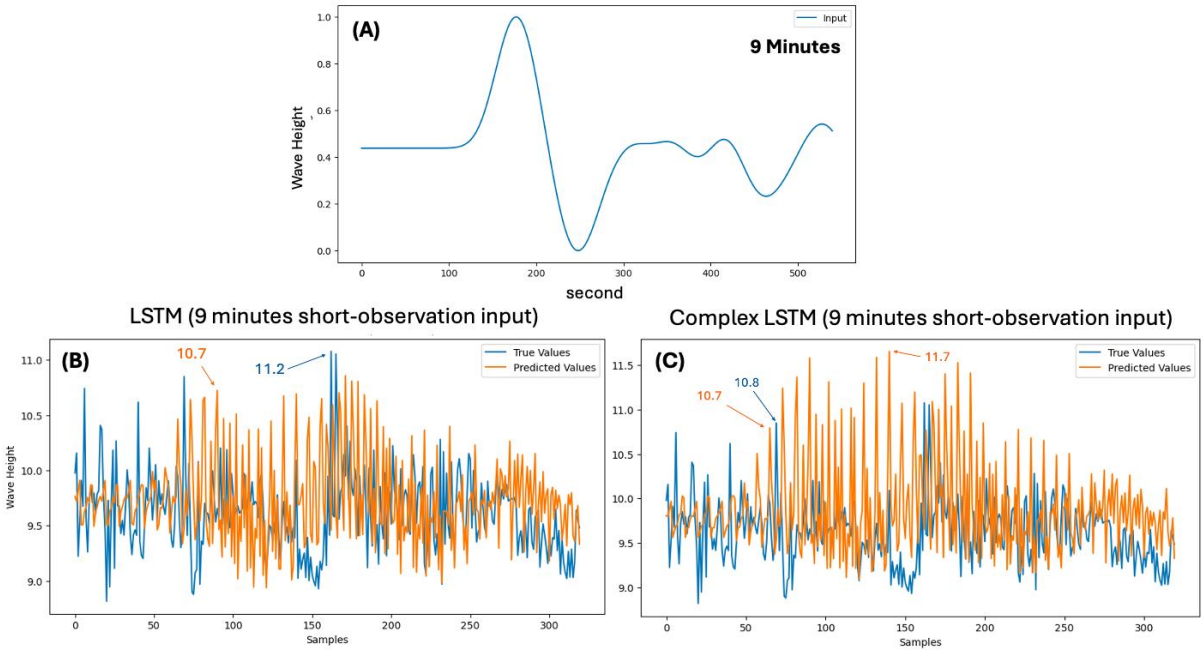


Figure 6. 9 Testing with short duration observation input. (A) Testing the model with short-duration input (9 min) to fit the goal of early warning. (B) LSTM result of full waveform prediction (C) Complex LSTM result of full waveform prediction

Figure 6.9 (B) and (C) showcase the prediction results, with the left panel illustrating the standard LSTM model's predictions and the right panel showing the Complex LSTM model's predictions. In both panels, the blue curves denote the true wave heights, while the orange curves represent the predicted values.

Upon comparative analysis, the standard LSTM model demonstrates a relatively consistent performance, with its predictions closely mirroring the true wave heights. It captures the overall trend and variability, although there are some discrepancies, particularly in areas with rapid changes. Conversely, the Complex LSTM model exhibits a less consistent performance with the short observation input. Its predictions show higher variability and larger deviations from

the true wave heights compared to the standard LSTM. While the Complex LSTM still captures certain trends, its predictions are less stable and exhibit more significant errors.

Further detailed observations reveal that the standard LSTM's predictions are more stable and consistent across the entire sample range. The Complex LSTM predictions, however, exhibit greater fluctuations and are less reliable in regions with high variability. Both models face challenges in accurately predicting the peaks and troughs, but the standard LSTM shows a slight advantage. For instance, around sample 100, the standard LSTM's predictions are closer to the true peak value compared to the Complex LSTM. Additionally, the errors in the Complex LSTM's predictions are more pronounced, indicating that the added complexity of the model does not necessarily translate to better performance with limited input data. In conclusion, when using short duration observation input, the standard LSTM model outperforms the Complex LSTM model in terms of prediction accuracy and stability. The results suggest that the additional complexity in the Complex LSTM may not be beneficial for scenarios with limited input data. Moreover, the practical implications of these results, including potential limitations, will be discussed in detail in the following section.

6.1.3.3.2 Evaluation of short duration observation input testing

Figure 6.10 presents a comparison of Mean Squared Error (MSE) and Mean Absolute Error (MAE) for LSTM and Complex LSTM model. Each model is represented by a unique symbol with specific colors and shapes to differentiate their test results.

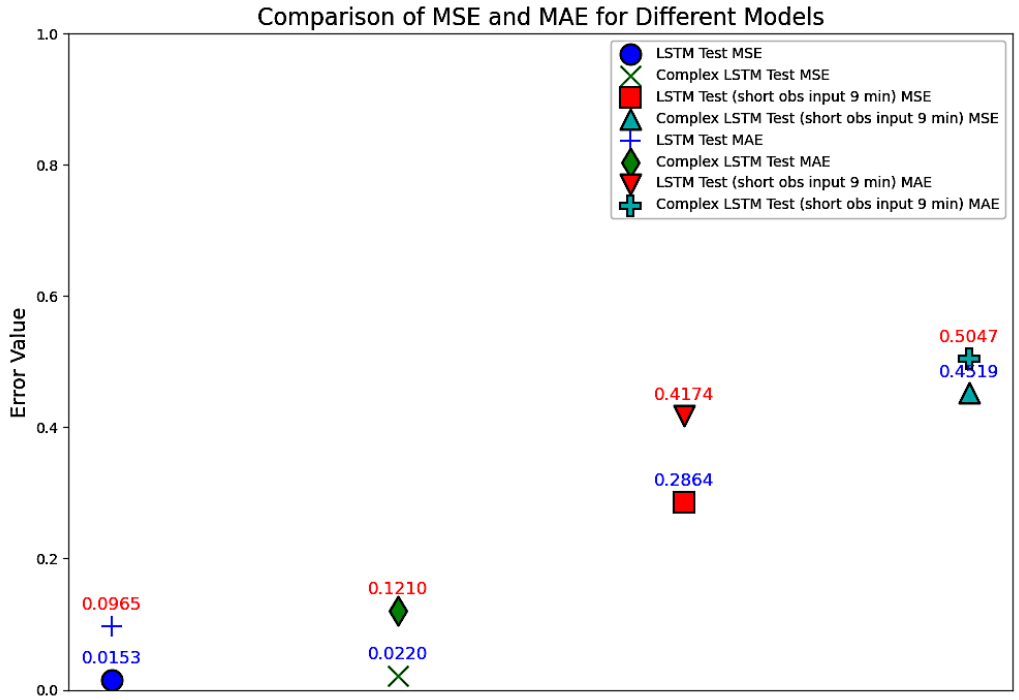


Figure 6.10 Short duration prediction comparison of MSE and MAE for LSTM and Complex LSTM

From the figure, we observe that the LSTM Test model has the lowest error values, with an MSE of 0.0153 (indicated by the blue circle) and an MAE of 0.0965 (indicated by the blue plus sign). This suggests that the LSTM Test model is highly accurate compared to the other models. The Complex LSTM Test model shows slightly higher error values, with an MSE of 0.0220 (indicated by the green cross) and an MAE of 0.1210 (indicated by the green diamond). Despite

the increased complexity, this model performs well but not as accurately as the LSTM Test model.

The LSTM Test model with short duration observation input (9 minutes) exhibits poorer performance compared to the previous two models, with an MSE of 0.2864 (indicated by the red square) and an MAE of 0.4174 (indicated by the red inverted triangle). This could be due to the shorter input observation duration, which provides less data for the model to make accurate predictions. Furthermore, the Complex LSTM Test model with short observation input (9 minutes) has the highest error values, with an MSE of 0.4519 (indicated by the cyan triangle) and an MAE of 0.5047 (indicated by the cyan plus sign). This indicates that increasing the model's complexity does not necessarily lead to better performance, especially when the input data is insufficient. In short, Figure 6.10 indicates that the LSTM Test model with full observation data yields the best predictive accuracy, while reducing the observation duration and increasing model complexity do not necessarily improve performance.

6.1.4 Discussion

6.1.4.1 Maximum amplitude prediction accuracy

The comparative analysis of tsunami maximum amplitude prediction models across four coastal stations—Marina Jambu, Ciwandan, Panjang, and Kota Agung—highlights varying performance of the Gradient Boosting, Random Forest, and Neural Network models.

At Marina Jambu, the Gradient Boosting model demonstrates the highest accuracy with the lowest RMSE of 0.1945, indicating its effectiveness in predicting maximum elevations at this station. The Random Forest model also performs well with an RMSE of 0.2058, while the Neural Network model shows a slightly higher error with an RMSE of 0.2457. The Gradient Boosting model's residuals show a more pronounced peak near zero, indicating fewer large errors and higher consistency compared to the other models. In contrast, the Ciwandan station results favor the Random Forest model, which achieves the lowest RMSE of 0.1126, suggesting it is the most accurate predictor in this context. The Gradient Boosting and Neural Network models follow with RMSEs of 0.1200 and 0.1288, respectively. Despite the Random Forest model showing a wider distribution of small residuals, it remains the most reliable model at this station, with the Gradient Boosting model also providing consistent and accurate predictions.

For Panjang station, the Random Forest model again proves to be the most effective, achieving the lowest RMSE of 0.0609. The Gradient Boosting model closely follows with an RMSE of 0.0631, and the Neural Network model records a slightly higher RMSE of 0.0702. The residual histograms for the Random Forest model demonstrate a symmetric distribution around zero with a slight underestimation skew, indicating high reliability. The Gradient Boosting model shows similar performance, while the Neural Network model has a broader spread of residuals, suggesting more frequent larger errors. At Kota Agung, the Random Forest model maintains its superior performance with the lowest RMSE of 0.0586, followed by the Gradient Boosting model with an RMSE of 0.0644, and the Neural Network model with an RMSE of 0.0809. The Random Forest model's residuals exhibit a symmetric distribution around zero, reinforcing its accuracy and reliability. The Gradient Boosting model also performs well with a notable peak

near zero, while the Neural Network model, although generally effective, shows a wider spread of residuals, indicating a higher frequency of larger errors.

6.1.4.2 Advantages and disadvantages of each model for maximum amplitude prediction

Each machine learning model used in the study demonstrated unique strengths and weaknesses in predicting tsunami maximum amplitudes across four coastal stations. Random Forest excelled at handling high-dimensional data and capturing non-linear relationships (Junior, Murti and Rahmawati, 2023), making it a strong candidate for static feature-based predictions. Its relatively fast training time also ensured efficient model development. However, Random Forest struggled with sequential data such as tsunami waveform data, leading to less smooth predictions that could introduce inaccuracies in forecasting maximum amplitudes. Additionally, its reliance on static features limited its application in dynamic, real-time prediction systems required for evolving tsunami scenarios.

Gradient Boosting, on the other hand, offered superior accuracy through iterative error minimization and effectively modeled complex, non-linear interactions (Iaccarino et al., 2021; Rabbani, Adytia and Husrin, 2023). This made it well-suited for capturing the intricate relationships between input features and maximum tsunami amplitudes. Nevertheless, its computationally intensive nature and sensitivity to hyperparameter tuning posed challenges. The slower training time also hindered its use in scenarios where rapid updates and predictions are essential, such as real-time monitoring at the coastal stations.

For Neural Networks, their ability to model complex non-linear relationships made them particularly effective in capturing the nuances of tsunami amplitude predictions. Their flexible architecture allowed for customization based on the specific requirements of the coastal stations. However, these models required significant computational resources and were prone to overfitting without careful regularization, making them less practical for resource-constrained environments (Novianty et al., 2022; Song and Cho, 2024). Despite their strengths, deploying Neural Networks for real-time tsunami prediction at multiple coastal stations remains challenging due to their high resource demands. These findings highlight the need to carefully balance accuracy, computational efficiency, and adaptability when choosing a machine learning model for practical tsunami amplitude prediction.

6.1.4.3 Limitation and practical challenge of LSTM algorithm for short duration prediction input

The analysis of model performance in tsunami full waveform prediction reveals distinct differences in predictive accuracy among various LSTM-based models. The LSTM Test model, with its full observation data, emerges as the most accurate, achieving the lowest error values. It records a Mean Squared Error (MSE) of 0.0153 and a Mean Absolute Error (MAE) of 0.0965. These low error metrics suggest that this model effectively captures the necessary data patterns to make precise predictions, outperforming other models in this comparison. Moreover, the Complex LSTM Test model, although more sophisticated, shows slightly higher error values with an MSE of 0.0220 and an MAE of 0.1210. While still performing well, this model's increased complexity does not translate to superior accuracy compared to the simpler LSTM

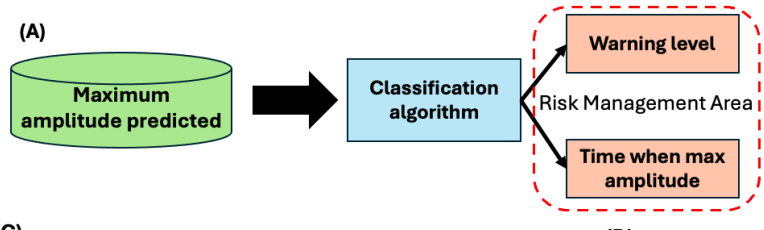
Test model. This implies that the added complexity may not be fully beneficial without corresponding improvements in prediction performance.

Regarding short duration observation inputs (9 minutes), the models demonstrate significantly poorer performance. The LSTM Test model with this shortened input duration shows a substantial increase in error values, with an MSE of 0.2864 and an MAE of 0.4174. This decline in accuracy likely results from the reduced amount of input data, which limits the model's ability to learn and predict accurately. Furthermore, the Complex LSTM Test model with short duration inputs performs the worst, highlighting the limitations of added complexity under data constraints. It records the highest error values with an MSE of 0.4519 and an MAE of 0.5047. These results clearly indicate that more complex models do not necessarily perform better, especially when they are provided with insufficient data.

For a challenge in the practical application, while this approach facilitates rapid initial forecasting, it limits the ability to capture the cases that involve longer-duration events or secondary wave amplifications (Williams, Rowley and Garthwaite, 2019), which are not addressed in this study. Furthermore, this limitation becomes important for coastal regions located farther from the tsunami source, where later-stage waveforms are essential for assessing the overall impact. Consequently, relying solely on short-term inputs may result in less accurate and incomplete forecasts under these conditions, reducing the effectiveness of early warning in practical applications.

6.1.4.4 Warning level and future risk management

Considering our result, we noticed the result predicted of tsunami maximum amplitude from using machine learning algorithms explained in Chapter 4, might be developed and integrated into the risk management area. The max tsunami amplitude is predicted on each coastal stations for all scenarios. By storing into the database and creating a specific code program, the result could be executed very fast to resulting the warning level. The warning level follows the policy from the local agency to defined threshold. This warning level information then might be main risk information to support the decision-making, such as evacuation (Figure 6.11).



(B) Warning level threshold

Warning Level	Wave Height (WH) Range [m]
No Warning	0,0 = WH < 0,1
Minor Tsunami	0,1 = WH < 0,5
Tsunami	0,5 = WH <3,0
Major Tsunami	WH ≥ 3,0

Defined Warning levels in Indonesia-Tsunami Early Warning System (Meteorology, Climatology, and Geophysical Agency, 2008)

(C) Example of warning level and time on each coastal station

Station	Scenario	Max Amplitude (m)	Time (min)	Warning Level
Marina Jambu	east_1	0.27	49.7	Minor Tsunami
Ciwandan	north_21	0	0	No Warning
Panjang	northeast_43	0.39	76.2	Minor Tsunami
Kota Agung	south_82	0.06	111.6	No Warning
Marina Jambu	west_317	2.0	35.0	Tsunami
Ciwandan	southwest_298	0.4	41.6	Minor Tsunami
Panjang	southwest_124	0.16	111.8	Minor Tsunami
Kota Agung	west_154	0.26	75.1	Minor Tsunami

(D) Warning counts and arrival average time

Station	Tsunami (avg time)	Minor Tsunami (avg time)	Major Tsunami (avg time)	No Warning (avg time)
Marina Jambu	271 (42.5)	41 (53.9)	6 (31.1)	2 (23.1)
Ciwandan	147 (45.6)	169 (57.9)	-	4 (45.6)
Panjang	26 (72.4)	223 (89.0)	-	71 (10.2)
Kota Agung	26 (72.4)	223 (89.0)	-	71 (102.8)

Figure 6.11 Risk management concept based for support decision-making on prediction result. (A) Flow of risk management (B) Warning level threshold-based agency threshold (C) Example output of warning level and time (D) Summary of warning counts and arrival average

Risk management concept on Figure 6.11 provides a comprehensive overview of the tsunami warning system, breaking down the process into several key components. Part (A) illustrates the initial step, where the maximum amplitude of the predicted tsunami wave is calculated. This predicted amplitude is then fed into a classification algorithm, which categorizes the risk management area based on two main criteria: the warning level and the time when the maximum amplitude is expected to occur. Part (B) details the thresholds for different warning levels based on wave height (WH). These thresholds, defined by the Indonesia Tsunami Early Warning System (2008), are as follows: "No Warning" for wave heights less than 0.1 meters, "Minor Tsunami" for wave heights between 0.1 and 0.5 meters, "Tsunami" for wave heights between 0.5 and 3.0 meters, and "Major Tsunami" for wave heights equal to or greater than 3.0 meters.

Part (C) presents examples of warning levels and times at different coastal stations under various scenarios. For instance, Marina Jambu under the east_1 scenario predicts a maximum amplitude of 0.27 meters in 49.7 minutes, resulting in a "Minor Tsunami" warning. Other scenarios, like Ciwarandan north_21, predict no warning with a maximum amplitude of 0 meters. The data demonstrates variability in warning levels and times depending on the station and scenario, highlighting the importance of location-specific predictions in the warning system. Part (D) summarizes the warning counts and average arrival times for different levels of tsunami warnings across various stations. For example, Marina Jambu has experienced 271 tsunamis with an average arrival time of 42.5 minutes, 41 minor tsunamis with an average time of 53.9 minutes, 6 major tsunamis with an average time of 31.1 minutes, and 2 instances of no warning with an average time of 23.1 minutes. This data provides insight into the frequency and timeliness of tsunami events at different stations, which is critical for assessing the effectiveness of the warning system and improving preparedness measures.

6.1.5 Summary of Part I

This sub chapter presents a novel method to enhance early warning systems for tsunamis triggered by volcanic landslides, addressing the complexities of volcanic collapse mechanisms. Unlike well-understood earthquake-induced tsunamis, volcanic tsunamis occur with little to no warning, making early detection crucial. Building on Ratnasari et al. (2023), who developed a real-time forecasting method based on a pre-computed database, our study addresses their limitations by incorporating more parameters and potential future scenarios that close to realistic. Using the 2018 Anak Krakatau volcanic tsunami as a case study, we generated 320 collapse scenarios to predict the highest tsunami amplitude at coastal observations. Six synthetic observation stations around Anak Krakatau were used, with inputs paired with the highest coastal amplitudes, allowing for the prediction of complete tsunami waves from short-duration initial waveform inputs.

This method aims to fill the detection gap and provide crucial lead time for coastal communities, enabling timely evacuation and reducing the impact of unexpected tsunamis. We employed three machine learning algorithms—Random Forest, Gradient Boosting, and Neural Network—to predict the maximum tsunami amplitude at coastal stations. Additionally, we explored two deep learning algorithms, Long-Short Term Memory (LSTM) and Complex LSTM, to forecast full waveform series at coastal observations.

The model capability for predicting tsunami maximum amplitude on coastal observation stations shows the Random Forest model consistently provides the most accurate predictions across three out of the four stations (Ciwandan, Panjang, and Kota Agung), making it the most reliable model overall. The Gradient Boosting model performs exceptionally well at Marina Jambu and also delivers consistent predictions at other stations, proving to be a strong contender. The Neural Network model, while effective, tends to have higher RMSE and broader residual spreads, indicating more frequent larger errors and slightly less reliability compared to the other two models. This analysis highlights that model performance can vary significantly depending on the observation station, emphasizing the importance of selecting the appropriate model based on specific station characteristics and prediction requirements in the future.

Regarding tsunami full waveform, the result underscores that the LSTM Test model with full observation data provides the best predictive accuracy among all the models tested. On the other hand, both the reduction of observation duration and the increase in model complexity do not inherently enhance performance and, in fact, can significantly degrade it when come to short duration input or when the data is limited. This analysis still highlights the importance of sufficient input data and optimal model complexity in achieving accurate predictions, even when using deep learning model such LSTM compared using the traditional machine learning. Additionally, recent developments in deep transformer models have demonstrated state-of-the-art performance in time series forecasting (Wu et al., 2020). These models utilize the self-attention mechanism, effectively addressing the "short-term memory" issue in long sequences (Zhou et al., 2020). Future research should incorporate this approach to enhance tide prediction models, thereby improving both efficiency and accuracy.

Furthermore, the prediction results from this study can be extended and integrated into future risk management to support decision-making. The proposed risk management concept on the previous chapter emphasizes a structured approach to tsunami risk management, highlighting the importance of accurate predictions, classification, and timely dissemination of warnings to reduce the impact of tsunamis on coastal communities (Juanara and Lam, 2024). We developed a code using Python and Jupyter Notebook to execute the maximum tsunami amplitude predictions at coastal stations. The warning levels for each coastal station are stored in supporting information, facilitating rapid and effective response during tsunami events.

While this research focuses on the tsunami caused by the Anak Krakatau volcano, its implications extend beyond this specific case. We believe the insights gained from this study can be applied to other volcanoes with similar risks of collapse-induced tsunamis. Anak Krakatau is just one of several significant volcanoes in Southeast Asia identified as sources of subaerial volcanic tsunami hazards, as noted by Zorn et al. (2022). Other significant volcanoes include Ritter Island in Papua New Guinea, Lliwerung in Indonesia, and Didicas in the Philippines. The broader applicability of our study's methodologies and findings enhances our understanding of volcanic tsunamis, helping coastal areas vulnerable to such geological events become better informed and prepared across various volcanic landscapes.

6.2 Part II. Deep learning-based waveform forecasting and physical simulations for volcano collapse tsunami early warning

6.2.1 Introduction

Volcanic activity has played a significant role in shaping the Earth's geological landscape, forming islands and altering coastal environments. Unlike mountains, which grow through gradual uplift and erosion, volcanoes form rapidly due to successive eruptions (Kent, 2016). While volcanic activity is essential for land formation, it also poses serious hazards, including pyroclastic flows, lava flows, and sector collapses. Among these hazards, volcanic tsunamis represent a particularly devastating and often underestimated threat. Historical records indicate that over 100,000 people have perished due to volcanic island eruptions (Gaudru, 2005), with tsunamis being a major indirect cause of fatalities. According to the National Geophysical Data Center's tsunami database, volcanic activity has triggered at least 139 tsunamis worldwide since 2100 BCE (Mutaqin *et al.*, 2019).

A striking example of a volcanic tsunami occurred on December 22, 2018, when the Anak Krakatau volcano in the Sunda Strait, Indonesia, suffered a catastrophic sector collapse, generating a deadly tsunami. This event resulted in hundreds of fatalities, with runup heights reaching 13.5 meters and inundation distances extending up to 330 meters (Grilli *et al.*, 2019; Heidarzadeh, Putra, *et al.*, 2020). Unlike earthquake-induced tsunamis, which are typically detected by seismic networks, the Anak Krakatau tsunami lacked any seismic precursor, revealing a critical weakness in existing tsunami early warning systems (TEWS) (Grilli *et al.*, 2021). Although volcanic collapse-induced tsunamis are less frequent than earthquake-generated tsunamis, their unpredictability and absence of early warning mechanisms make them particularly hazardous.

Efforts to improve volcanic tsunami forecasting of Anak Krakatau volcano in Indonesia have primarily focused on numerical modeling and physical experiments (Giachetti *et al.*, 2012; Borrero *et al.*, 2020; Mulia *et al.*, 2020; Omira and Ramalho, 2020; Paris *et al.*, 2020; Perttu *et al.*, 2020; Ren *et al.*, 2020; Dogan *et al.*, 2021; Grilli *et al.*, 2021). Following the 2018 collapse, Anak Krakatau has been rapidly regrowing (Mulia *et al.*, 2020), underscores the importance of developing more advanced forecasting models.

However, most existing TEWS (Tsunami Early Warning Systems) still focus exclusively on earthquake-generated tsunamis, relying on seismic networks, deep-ocean buoys, and tide gauges to detect tsunami waves (Maeda et al., 2015). In Indonesia, for example, the Indonesian Tsunami Early Warning System (InaTEWS) has been developed to monitor seismic activity and issue alerts (Harig *et al.*, 2020). Despite these advancements, volcanic tsunamis remain largely unaccounted for in current warning protocols, highlighting an urgent need for alternative detection and forecasting methods.

Although volcanic collapse-induced tsunamis share similarities with landslide-generated tsunamis as categorized as non-seismic tsunami (Lahcene *et al.*, 2021), they pose unique challenges for early warning. Subaerial and submarine landslide tsunamis have been extensively studied through experimental and numerical modeling (Bullard *et al.*, 2019; Heller and Ruffini, 2023). While these studies have significantly advanced our understanding of

tsunami generation mechanisms, they do not directly address real-time detection and forecasting, especially for volcanic tsunamis.

Numerical modelling has been instrumental in understanding landslide-generated tsunami dynamics, particularly for subaerial and submarine landslides. Panizzo, De Girolamo and Petaccia (2005) focused on forecasting impulse waves generated by subaerial landslides, providing key insights into wave amplitude, velocity, and energy transfer dynamics. Later, Yavari-Ramshe and Ataie-Ashtiani (2016) conducted a comprehensive review of recent advances in landslide-generated tsunami modeling, emphasizing the complexity of wave generation, propagation, and run-up processes. While these models have greatly improved our ability to simulate tsunami characteristics, they remain largely focused on landslide-generated tsunamis and do not fully incorporate the challenges associated with volcanic collapse scenarios, particularly in real-time forecasting applications.

In recent years, machine learning (ML) and artificial intelligence (AI) have emerged as powerful tools for tsunami early warning and forecasting. Various AI-based models have been explored to enhance accuracy and speed in tsunami prediction. For instance, Convolutional Neural Networks (CNNs) have been used to predict tsunami inundation patterns in Nankai, Japan (Fauzi and Mizutani, 2020), while Recurrent Neural Networks (RNNs) have been applied to model tsunami waveform dependencies (Dharmawan *et al.*, 2023). Additionally, decision tree algorithms have been utilized to classify tsunami early warning levels (Juanara and Lam, 2024).

Beyond seismic tsunamis, AI has also been applied to landslide-generated tsunami forecasting. Meng, Hu and Ancey (2020) introduced a data-driven approach for predicting wave generation from mass flows, while Wu *et al.* (2022) employed physics-informed deep learning to model landslide surge propagation. More recently, Jenkins, Heller and Giannakidis (2025) proposed a machine learning framework for landslide-tsunami prediction, integrating slide geometry and fluid interactions into an AI-based predictive model. These studies demonstrate the potential of AI-driven tsunami forecasting for landslide-induced tsunamis. However, while these models have successfully predicted tsunami wave patterns from landslides, they do not account for the unique variability of collapse directions in volcanic settings, where mass failures can occur in multiple directions rather than a single downslope motion.

Unlike typical landslides, which generally follow predictable gravitational paths, volcanic sector collapses can occur asymmetrically, with material falling in different directions depending on the structure of the volcano, previous eruptions, and magma intrusion patterns. This directional variability significantly affects tsunami wave propagation, as different collapse angles can generate different wave patterns and impact different coastal regions. Traditional early warning approaches struggle with this variability, as they typically rely on precomputed tsunami databases that assume a single or dominant landslide direction.

Furthermore, most AI-based tsunami forecasting studies rely on single observation points, which fail to capture the spatial complexity of tsunami wave propagation (Liu *et al.*, 2021). Different from seismic tsunamis, which propagate radially from a point source, volcanic tsunamis involve multiple wave interactions and varying collapse dynamics, necessitating a multi-station observation approach for improved prediction accuracy.

While these prior studies provide a foundation for AI-driven tsunami forecasting, they do not directly address the challenges posed by unpredictable collapse directions in volcanic characteristics. This gap in research underscores the need for an AI-driven framework tailored specifically for volcanic tsunami forecasting, integrating spatially distributed synthetic observation stations for more effective prediction since the tsunami early warning need a multidisciplinary approach (Jin and Lin, 2011).

This part II aims to develop a deep learning technique for volcanic collapse tsunami waveform forecasting for early warning applications using Anak Krakatau volcano collapse as case study. Previous studies have successfully applied machine learning (ML) models to predict tsunami wave characteristics, but they are often trained on idealized laboratory conditions, simplified slide geometries, and limited datasets, which reduce their generalization capability for real-world applications. Furthermore, existing ML-based tsunami models rely on single observation points (e.g., Liu *et al.*, 2021) and limited collapse direction (Ratnasari *et al.*, 2023), which do not fully capture the spatial complexity of tsunami propagation from volcanic sector collapses.

To address these limitations, this part II introduces spatially distributed synthetic observation stations and utilizes short-term tsunami waveform data as input for prediction. By leveraging Deep Learning architectures, Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM), this part seeks to improve forecasting accuracy and provide a data-driven alternative for non-seismic tsunami warning systems. This part involves the following objectives:

- Developing tsunami forecasting framework using multiple synthetic observation stations, overcoming the limitations of single-point observations.
- Utilizing short-term tsunami waveform observations to enhance quick early warning, predicting wave evolution from the initial few minutes of recorded data.
- Evaluating LSTM and BiLSTM performance to ensure accurate and interpretable tsunami forecasts for integration into early warning systems.

The remainder of this part is structured as follows: Section 2 describes the dataset generation and method, including the synthetic observation stations and collapse tsunami simulation parameters, as well as the deep learning architectures and evaluation protocols. Section 3 presents and quantifies the forecasting performance of the LSTM and BiLSTM models. Section 4 provides an in-depth discussion of the results, including model interpretability, the impact of short-term observations, and implications for tsunami early warning. Section 5 summarizes the key findings, outlines future research directions, and discusses the limitations of this part, including the challenges of real-world generalization and potential biases in synthetic data.

6.2.2 Method

6.2.2.1 Deep learning flow

The deep learning flowchart presented in this part (Figure 6.12) illustrates the integrated framework combining physical simulation and data-driven modelling to support real-time tsunami waveform forecasting. This hybrid approach consists of two primary stages: (1) the

generation of synthetic tsunami data through numerical simulations, and (2) the application of deep learning techniques to predict future tsunami waveforms based on early observations.

In the physical simulation phase, synthetic observation stations are established and simulation parameters, such as landslide geometry, direction, and source characteristics, are configured. Landslide-generated tsunami scenarios are then simulated using COMCOT. The resulting synthetic waveforms are validated against real-world events, such as the 2018 Anak Krakatau tsunami, to ensure physical realism and reliability. Only datasets that meet acceptable validation thresholds proceed to the next modelling phase.

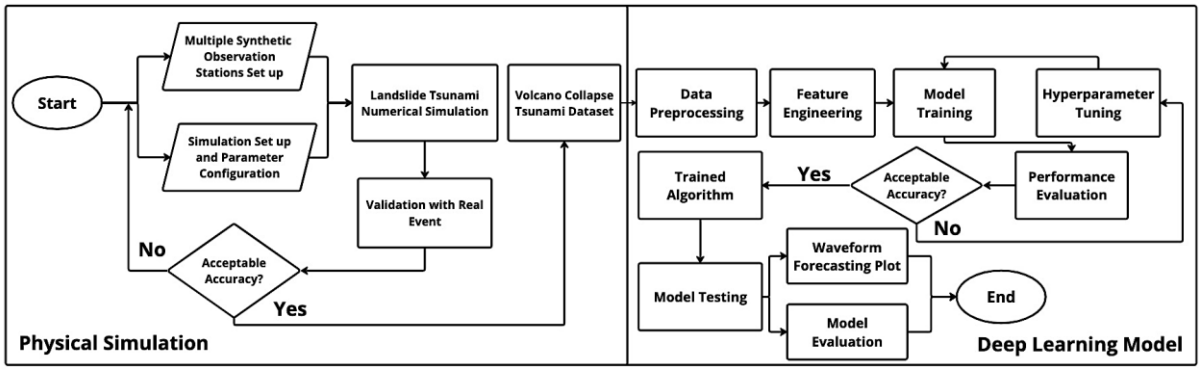


Figure 6.12 Research flow chart for integrating the physical simulation and deep learning model for waveform forecasting

In the deep learning phase flow, the model is systematically processed through several key stages:

1. Data Preprocessing: Raw simulation outputs are cleaned, normalized, and reshaped to ensure compatibility with model input structures. This step ensures temporal alignment and scales amplitude values into appropriate ranges for neural network training.
2. Feature Engineering: From the pre-processed waveforms, features such as the first 10 minutes of observed wave heights are extracted to serve as input sequences. These features are designed to mimic early tsunami detection and enable short-term forecasting.
3. Model Training: Using these inputs, the deep learning model (e.g., LSTM or BiLSTM) is trained to learn the underlying temporal patterns in the tsunami data. The training process is monitored using a validation set to detect overfitting and optimize generalization.
4. Hyperparameter Tuning: As described in Section 6.2.2.3, an exhaustive grid search using KerasTuner is conducted to find the optimal model configuration, including the number of hidden units, activation functions, and learning rate.
5. Performance Evaluation: Once trained, the model is evaluated using multiple metrics (e.g., Mean Squared Error, R^2) on a held-out test set. This evaluation quantifies the model's ability to generalize beyond the training data.
6. Model Testing and Forecast Visualization: The best-performing model is then applied to generate waveform forecasts, which are plotted and compared against simulated "ground truth" to assess predictive accuracy and temporal alignment.

7. Model Evaluation: Final assessments are made to determine whether the model can reliably capture the complex wave dynamics of tsunami propagation. This step includes qualitative waveform comparisons and quantitative statistical evaluations.

This flow, from physical simulation to deep learning prediction, ensures a scalable pipeline that can be adapted for real-time tsunami early warning systems. The integration of physics-based validation and data-driven learning improves both the reliability and responsiveness of forecast outputs. Each step of this flow will be explained in further detail in the subsequent sections.

6.2.2.2 Long Short-Term Memory (LSTM) and Bi LSTM Neural Networks

The primary goal of this study is to predict tsunami waveforms over time using a sequential modeling approach. Such a problem involves analyzing time-series data, where preserving the order of observations is critical. Recurrent Neural Networks (RNNs) are well-suited for this task due to their ability to process sequential data. However, traditional RNNs face limitations, such as the vanishing gradient problem, which reduces their effectiveness in learning long-term dependencies. To address these limitations, the Long Short-Term Memory (LSTM) network is utilized. LSTMs introduce memory cells that can store information for extended periods, allowing the model to retain and leverage long-term dependencies in the data (Hochreiter and Schmidhuber, 1997). This makes LSTM particularly effective in modeling the temporal dependencies inherent in tsunami waveforms.

The use of LSTM in this study was a deliberate choice based on the sequential nature of tsunami data and the objective of the prediction task. Tsunami waveforms are inherently time-dependent, with each time step influenced by preceding wave characteristics. LSTM is specifically designed to handle such time series data through its memory cells, which enable the model to retain information across long input sequences. This capability supports the learning of complex temporal dependencies, such as the rising and falling wave patterns, that are essential for accurate forecasting based on early-stage observations.

In this research, the model was trained on a short-term segment of tsunami waveform input (e.g., the first few minutes) to predict the subsequent evolution of the waveform. This formulation aligns with a sequence-to-sequence learning problem, for which LSTM is particularly well-suited.

The architecture of Vanilla LSTM cell consists of three gates; Forget Gate (f_t), Input Gate (i_t), and Output Gate (o_t). Forget gate determines which information from the previous cell state (C_{t-1}) should be discarded. Input gate decides what new information should be added to the cell state (C_t). Output gate controls which information from the cell state is passed to the hidden state (h_t). These gates regulate the flow of information and enable the network to selectively retain relevant patterns while discarding noise. The core equations governing the LSTM operations are shown in Equation 6.1.

Equation 6. 1 LSTM Architecture

$$f_t = \sigma(W_f \,.[h_{t-1}, x_t] + \, b_f)$$

$$i_t = \sigma(W_i \,.\, [h_{t-1}, x_t] + \, b_i)$$

$$\tilde{C}_t = tanh(W_C \,.\, [h_{t-1}, x_t] + \, b_C)$$

$$C_t = f_t * \, C_{t-1} + i_t * \, \tilde{C}_t$$

$$o_t = \sigma(W_o \,.\, [h_{t-1}, x_t] + \, b_o)$$

$$h_t = \, o_t * \tanh\,(C_t) \tag{6.1}$$

Figure 6.13 (a) illustrates the internal structures of an LSTM cell, highlighting how these gates interact to manage information flow. Figure 6.13 (b) demonstrates how multiple LSTM cells are connected sequentially, with the output of one cell feeding into the next, enabling effective processing of sequential data.

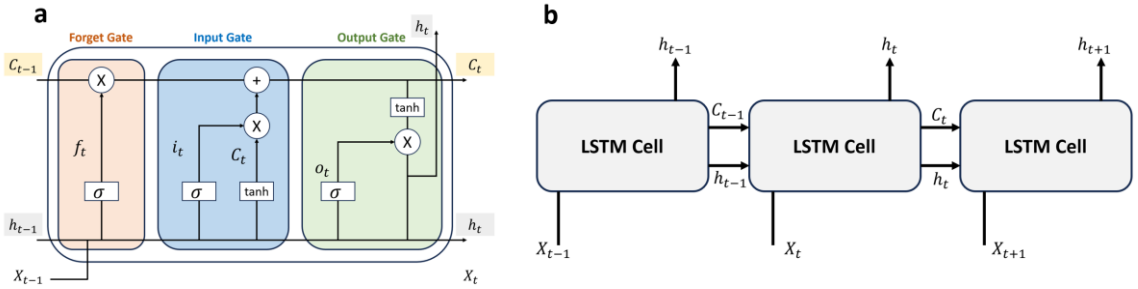


Figure 6.13 Structure of Vanilla LSTM cell and connection of the cell. (a) internal structures of an LSTM cell consist of forget, input, and output gate. (b) LSTM cell sequential connection

While Vanilla LSTM processes sequence in a single direction (from past to future), Bidirectional LSTM (BiLSTM) enhances the networks capability by processing the sequence in both directions, forward and backward (Graves, Fernández and Schmidhuber, 2005). This dual-directional processing provides richer contextual information by capturing dependencies from both the past and the future. The forward LSTM processes the sequence as $x_1, x_2, \ldots, x_t$ while backward LSTM processes the sequence in reverse $x_t, x_{t-1}, \ldots, x_1$. Recent studies have demonstrated that BiLSTM outperforms conventional LSTM models in time-series forecasting tasks due to its ability to process sequential data from both past and future contexts simultaneously. For instance, Kang *et al.* (2020) applied BiLSTM for wastewater flow prediction and demonstrated that optimizing hyperparameters significantly enhanced the model's predictive accuracy. Similarly, Said *et al.* (2021) utilized BiLSTM for multivariate COVID-19 case forecasting and reported superior performance compared to traditional approaches, further highlighting the importance of selecting optimal hyperparameters.

The outputs from both directions ($h_t^{forward}$ *and* $h_t^{backward}$) are concatenated to form a final output representation for each time step. Figure 6.14a illustrates the structure of a Vanilla LSTM layer, while Figure 6.14b extends this to BiLSTM, showing how the two directional LSTMs are combined to produce a comprehensive representation.

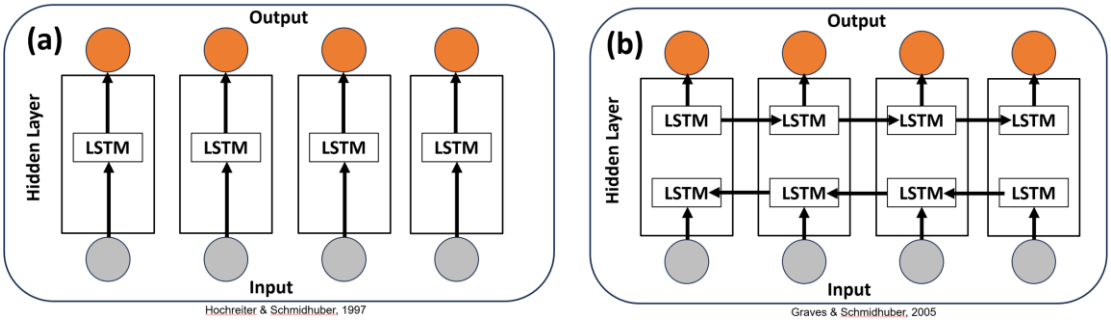


Figure 6.14 Architecture of Vanilla layer (a) and Bidirectional LSTM layer (b)

6.2.2.3 Model architecture

The Vanilla LSTM model consists of a single LSTM layer with 32 units, followed by a fully connected Dense layer for waveform prediction. Meanwhile, the BiLSTM model incorporates a bidirectional LSTM layer with 64 units, enabling the model to capture dependencies in both forward and backward directions. The architectural differences between these models are summarized in Table 6.2.

Table 6.2 LSTM and BiLSTM architectures

Layer	**LSTM Model**	**BiLSTM Model**	**Number of Units**	**Activation Function**
1	LSTM	BiLSTM	32/64	ReLU
2	Dense	Dense	Output sequence length	Linear

6.2.2.4 Model hyperparameter tuning

To enhance the predictive capability of the model, a comprehensive hyperparameter tuning procedure was conducted using Grid Search with the KerasTuner framework. This process systematically explored a range of hyperparameter values to identify the optimal configuration for both Vanilla LSTM and Bidirectional LSTM (BiLSTM) models. Key hyperparameters investigated included the number of hidden units, activation functions, and learning rates.

Hyperparameter tuning was carried out by exhaustively evaluating all possible combinations within a predefined search space, as outlined in Table 6.3. This allowed for the identification of the configuration that yielded the lowest validation loss on the training set. By minimizing the validation loss, the selected configuration not only improved generalization performance but also reduced the risk of overfitting, critical for robust tsunami waveform prediction in unseen scenarios.

The final model training utilized the Adam optimizer and the Mean Squared Error (MSE) as the loss function. Each model was trained for 500 epochs with a batch size of 32. During the tuning phase, the validation loss was continuously monitored, and early stopping was employed to prevent overfitting where applicable. The best-performing hyperparameters were selected based on the lowest validation loss achieved during tuning.

Table 6.3 Hyperparameter search space and best configuration

	Hyperparameter	**Search Space**	**Selected Value**
LSTM	Units	{32, 64, 128}	32

	Activation Function	{ReLU, Tanh}	ReLU
	Learning Rate	{0.01, 0.001, 0.0001}	0.01
BiLSTM	Units	{32, 64, 128}	64
	Activation Function	{ReLU, Tanh}	ReLU
	Learning Rate	{0.01, 0.001, 0.0001}	0.01

The chosen configuration, consisting of a moderate LSTM unit size (32), a larger BiLSTM unit size (64), ReLU activation function, and a learning rate of 0.01, demonstrated a balance between learning capacity and training stability. This setup enabled the model to capture complex spatiotemporal patterns in tsunami waveform data while maintaining computational efficiency.

In summary, the tuning process significantly improved model performance by providing an optimal trade-off between accuracy and generalization. This ensured that the model could effectively learn from the training data while maintaining robust predictive capability on new, unseen tsunami scenarios. Subsequent model performance evaluation using quantitative metrics (e.g., MSE, RMSE, MAE) is discussed in the following section to assess prediction quality and model reliability.

6.2.2.5 Model training strategy

After selecting the optimal hyperparameters, the model was trained using a structured strategy to ensure efficient convergence and prevent overfitting. The dataset was split into 80% training and 20% validation to evaluate the generalization capability of the model during training. To enhance training efficiency, an EarlyStopping mechanism was implemented, monitoring validation loss with a patience of 10 epochs. If no improvement was observed, training was halted, and the model weights from the best epoch were restored. Additionally, ReduceLROnPlateau was applied to dynamically adjust the learning rate, reducing it by a factor of 0.5 if no improvement occurred for 5 consecutive epochs, with a minimum learning rate threshold of 1e-6.

The training process was monitored using the learning curve, and the performance of the trained LSTM model was assessed through waveform predictions. The results of the training process and model performance evaluation are presented in Section 3.

6.2.2.6 Waveform forecasting model using multiple synthetic station

The proposed waveform prediction and early warning model introduces an approach to tsunami waveform forecasting by utilizing short-term observation data from multiple synthetic stations (Station 1–5) as input to predict the complete waveform at a specific station. This methodology leverages minimal initial data to efficiently generate accurate tsunami waveform predictions, demonstrating the potential for early warning systems with reduced computational requirements.

As illustrated in Figure 6.14, the top panel shows the initial waveform data recorded during the first 3 minutes from six synthetic observation stations, where each colored line represents the wave height recorded at a specific station, serving as the input to the model. The decision to use the first 3 minutes of observation data as input is based on the critical timing of tsunami

wave propagation. By the 3-minute mark, the waves have typically reached the islands where the synthetic stations are located, ensuring that sufficient waveform information is recorded. This duration provides an adequate representation of the wave dynamics near the source, capturing essential features such as amplitude, direction, and energy distribution, which are crucial for accurate forecasting. Thus, the 3-minute observation window balances efficiency with the need for reliable data to initialize the waveform prediction model. To further support this decision, we also conducted a sensitivity analysis comparing different input durations (1, 3 and 5 minutes). A summary of this comparison is provided in Section 2.7.

The bottom panel displays the forecasted waveform at one of the synthetic stations (Station 1) for the remaining simulation duration of 2 hours. This output, generated by the trained LSTM and BiLSTM models, predicts the evolution of the tsunami waveform based on the initial short-term observations.

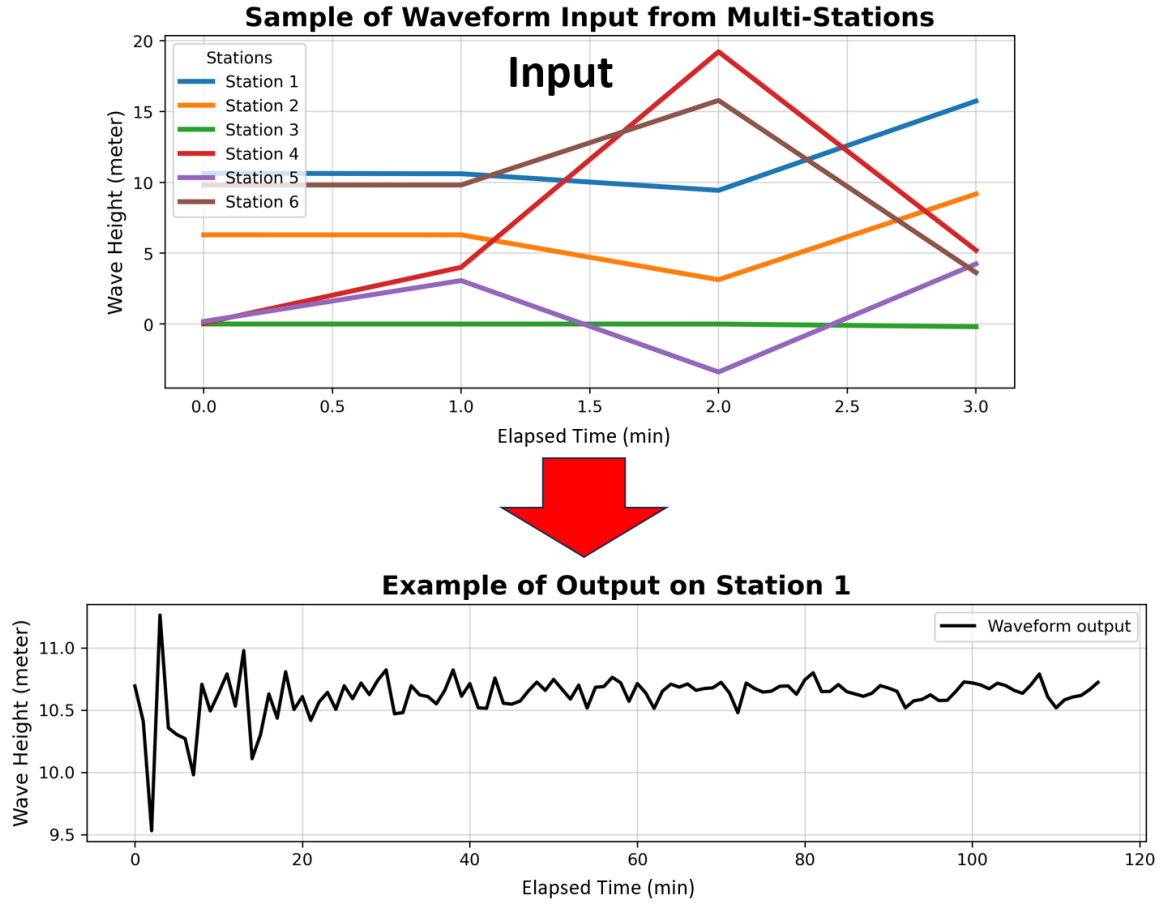


Figure 6.15 Illustration of waveform forecasting model using multiple input

For this research, 1000 simulation scenarios of volcanic tsunami waveforms were generated by varying collapse parameters, such as volume, dip angle, and direction, providing time-series data recorded at multiple tide gauges (synthetic stations). The Vanilla LSTM model is employed to learn the temporal dependencies of individual waveforms, focusing on capturing long-term trends caused by volcanic flank collapses. In contrast, the BiLSTM model enhances this approach by leveraging bidirectional dependencies, allowing the model to better capture complex interactions between past and future wave behaviours. This is particularly advantageous for near-source tide gauges, where wave dynamics exhibit intricate bidirectional

patterns. By integrating these architectures, the proposed research aims to develop a tsunami early warning that combines the predictive strengths of LSTM and BiLSTM for precise and efficient waveform forecasting.

6.2.2.7 Performance evaluation

To objectively assess the predictive performance of the deep learning models, three standard error metrics were employed: Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) (see equation 6.5 - 6.7). These metrics quantify the discrepancy between the predicted tsunami waveforms and the ground-truth waveforms obtained from numerical simulations. The selection of these metrics is based on their widespread application and robustness in time-series forecasting tasks, particularly in geophysical modelling where the accuracy of both amplitude and timing is critical.

- Mean Absolute Error (MAE) measures the average magnitude of absolute differences between predicted and actual values, without considering their direction. It is a linear score, meaning all individual differences are weighted equally. This metric provides an intuitive interpretation of the model's average forecasting error in meters and is less sensitive to outliers than squared-error metrics (see Equation 6.5).

Equation 6. 2 MAE

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (6.2)$$

- Mean Squared Error (MSE) calculates the average of the squares of the differences between the predicted and actual values. By squaring the errors, MSE penalizes larger errors more heavily than MAE. This metric is useful for identifying models that may produce occasional large deviations, which are particularly relevant in tsunami modelling where rare high peaks can have serious consequences (see Equation 6.6).

Equation 6.3 MSE

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (6.3)$$

- Root Mean Squared Error (RMSE) is the square root of the MSE, which brings the error metric back to the same unit as the original data (meters in this study). RMSE provides a more interpretable magnitude of the average error and retains the MSE's sensitivity to larger deviations (see Equation 6.7).

Equation 6.4 RMSE

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (6.4)$$

These metrics were calculated for each trained model using a held-out test dataset that was not involved in the training process. Lower values of MAE, MSE, and RMSE indicate higher predictive accuracy and better generalization capability. In the context of tsunami early warning systems, reducing RMSE is particularly important, as even small underestimations or

overestimations in wave height can significantly affect risk assessment and emergency response decisions.

Through the combined use of these three metrics, the evaluation framework provides a comprehensive perspective on model performance, balancing robustness to interpretability and sensitivity to large errors.

6.2.2.8 Sensitivity analysis of input duration

To evaluate the effect of input duration on model performance, a sensitivity analysis was conducted by varying the length of the initial waveform segment used as input to the forecasting model. The tested input windows were 1, 3, and 5 minutes, with the aim of identifying the optimal balance between forecast accuracy and early warning timeliness.

As presented in Table 4, shorter input durations (e.g., 1 minute) resulted in the highest errors (MSE 0.425013, RMSE 0.651930), indicating that limited waveform data may not adequately capture critical wave dynamics such as directionality and energy build-up. On the other hand, the 5-minute input yielded the lowest errors (MSE 0.027561, RMSE 0.166014), benefiting from more complete wave characteristics but potentially delaying timely warnings.

The 3-minute input window demonstrates a compelling trade-off, with substantially lower errors than the 1-minute input (MSE 0.092578, RMSE 0.304266). This represents a 53% reduction in RMSE compared to the 1-minute input, while still enabling early warning issuance within a practical time frame. Therefore, based on the findings summarized in Table 4, the 3-minute window is adopted as the optimal configuration for this study, offering a reliable balance between prediction accuracy and operational responsiveness.

Table 6.4 Sensitivity analysis of different input durations on Vanilla LSTM performance (Station 4)

Input Duration (min)	MSE	RMSE
1 min	0.425013	0.65193
3 min	**0.092578**	**0.304266**
5 min	0.027561	0.166014

The use of a 3-minute input is also justified by wave arrival timing: within this duration, tsunami waves typically reach the synthetic stations, ensuring that sufficient initial waveform features are available for accurate prediction.

6.2.3 Result

This section presents the learning curves during training and validation, along with the results of the trained model for one sample scenario. It also provides an illustration of tsunami wave propagation during the first 1–3 minutes through snapshots, which serve as input for the deep learning model. Additionally, the section plot and evaluates the performance of the waveform forecasting model using short-term input observations.

6.2.3.1 Deep learning training and validation

The training process of the deep learning model is evaluated using the learning curve, as depicted in Figure 6.16 (a). The learning curve illustrates the evolution of training and validation loss over 500 epochs, demonstrating a rapid decrease in loss during the initial epochs

before stabilizing. The optimal model was identified at epoch 141, where validation loss ceased to improve, leading to the application of the EarlyStopping mechanism to halt training and restore the best model weights. Additionally, the ReduceLROnPlateau strategy dynamically adjusted the learning rate, reducing it by a factor of 0.5 if no improvement was observed for 5 consecutive epochs, ensuring efficient convergence while mitigating overfitting.

The progressive reduction in loss during training reflects how the model learns from data through a process called backpropagation and gradient descent. After making predictions in each training step, the model calculates the difference (or error) between the predicted and actual values. Backpropagation is then used to determine how much each part of the model contributed to the error. Based on this information, the model updates its internal parameters using gradient descent, which gradually adjusts the weights in a direction that reduces the error. This cycle continues over many training epochs, allowing the model to improve its accuracy step by step.

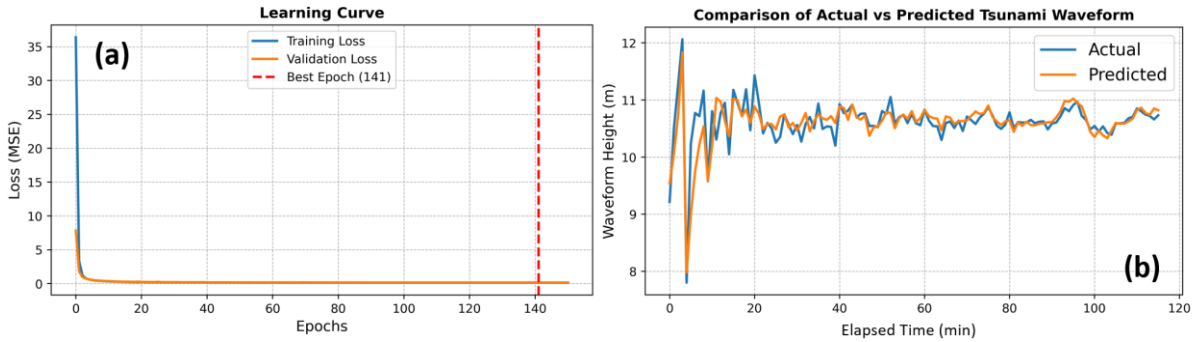


Figure 6.16 Model training validation. (a) Learning curve training and validation loss during training (b) Comparison the trained model with the one sample of database scenario

To validate the performance of the trained deep learning model, tsunami waveform predictions were assessed by comparing the predicted and actual waveforms, as shown in Figure 6.16 (b). The model successfully captured the waveform dynamics, particularly in the later time steps, where predicted values closely followed the actual observations. While some discrepancies were observed in the early phase, the overall performance demonstrates the model capability to generalize tsunami waveform evolution. This result highlights the effectiveness of integrating deep learning-based forecasting with numerical simulations for tsunami early warning.

6.2.3.2 Snapshots of waveform propagation for short-term input

Figure 6.17 illustrates tsunami wave propagation during the first 1–3 minutes, captured at three distinct time steps: 1 minute, 2 minutes, and 3 minutes. These snapshots provide a clear visualization of the early-stage dynamics of the tsunami triggered by the Anak Krakatau volcanic collapse. In the 1-minute snapshot, the wave remains concentrated near the source, displaying the initial dispersion of energy as the tsunami begins to propagate outward. By the 2-minute snapshot, the wavefront expands, with increasing amplitude observed in areas directly adjacent to the collapse zone, reflecting the outward transfer of energy. By the 3-minute snapshot, the wave has propagated further, exhibiting more complex interactions as it reaches the islands where the synthetic observation stations are located.

These snapshots highlight the importance of early-stage wave observations for forecasting, as they capture critical information on wave direction, energy distribution, and interactions with the surrounding environment. Based on these observations, 3 minutes was selected as the optimal input duration because it represents the point at which the wave reaches the islands where the synthetic stations are positioned. This dataset serves as the input for the waveform forecasting model to predict the long-term evolution of the tsunami at synthetic observation points.

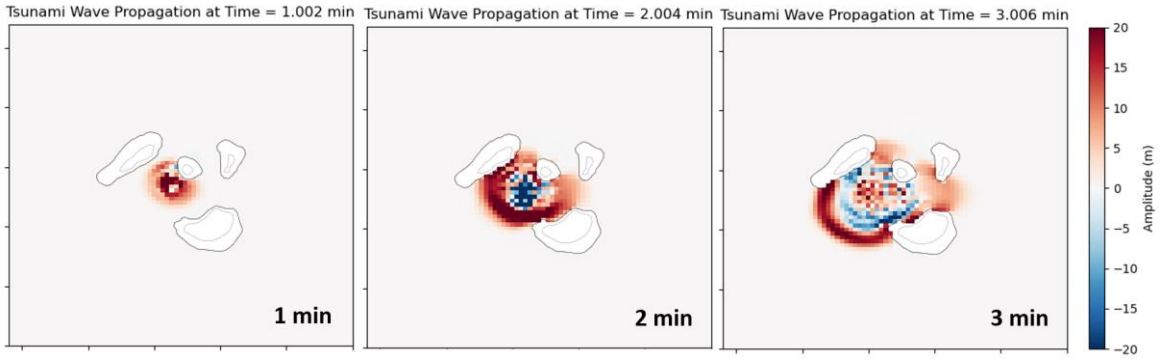


Figure 6.17 Example snapshots of tsunami waveform propagation during the first 1–3 minutes following a southwest-direction collapse

6.2.3.3 Full waveform forecasting using 3 minutes short-term input

Figure 6.18 illustrates the waveform forecasting results using the Vanilla LSTM model across six synthetic stations (Station 1 to Station 6), displaying selected scenario samples from testing database. Figure 6.18 presents similar results using the Bidirectional LSTM (BiLSTM) model. Each figure compares the predicted waveform (red line) with the observed waveform (black line) across different test scenarios. The dashed vertical line at the 3-minute mark separates the input observation window from the forecasted waveform. Figures 6.18 and 6.19 provide a detailed evaluation of the performance and capabilities of both models, while the remaining waveform forecasting plots will be included in the supporting information.

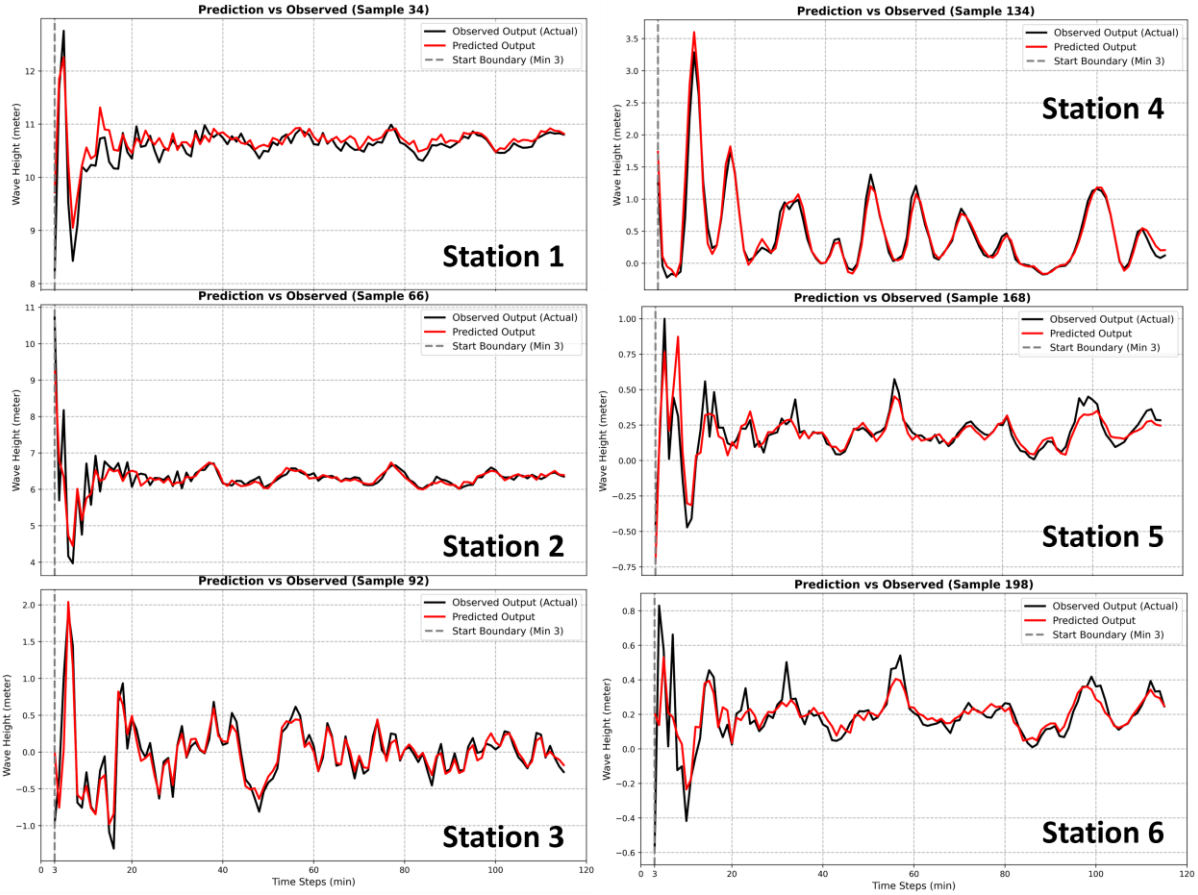


Figure 6.18 Comparison of predicted and observed tsunami waveforms at six stations using the Vanilla LSTM model

The Vanilla LSTM model demonstrates a strong capability in capturing the overall trends and amplitudes of the observed tsunami waveforms using only three minutes of input waveform propagation. This is particularly evident at stations closer to the tsunami source, where the model effectively learns temporal dependencies in wave propagation with minimal influence from external factors such as bathymetric variations, coastal interactions, and nonlinear dispersive effects.

At Station 1 and Station 2, which are situated near the tsunami generation zone, the predicted waveforms closely align with the observed data. The model accurately reproduces peak wave amplitudes and subsequent oscillatory patterns, which are crucial in understanding wave energy dissipation and secondary wave interactions. The high level of agreement suggests that the Vanilla LSTM effectively captures primary tsunami dynamics, making it a reliable tool for short-range forecasting.

At Station 3 and Station 4, positioned at an intermediate distance from the source, the model maintains reasonable accuracy in predicting waveform structures. The gradual decay of wave amplitudes aligns well with the expected energy attenuation as the tsunami propagates further. However, minor deviations in oscillatory behavior appear in later timesteps, indicating challenges in capturing nonlinear dispersion effects and coastal reflections. These discrepancies highlight the model's capability to generalize to mid-range locations, albeit with some limitations.

At Station 5 and Station 6, which are the farthest from the tsunami source, discrepancies between predicted and observed waveforms become more pronounced. The underestimation of wave amplitudes and phase shift errors suggest that as tsunami waves travel over longer distances, additional environmental factors—such as coastal topography interactions and energy dispersion—become increasingly significant. The inability of the Vanilla LSTM to fully account for these long-range propagation complexities indicates that while the model performs well for near-source and mid-range regions, further refinements are required to improve accuracy for distant tide gauge stations.

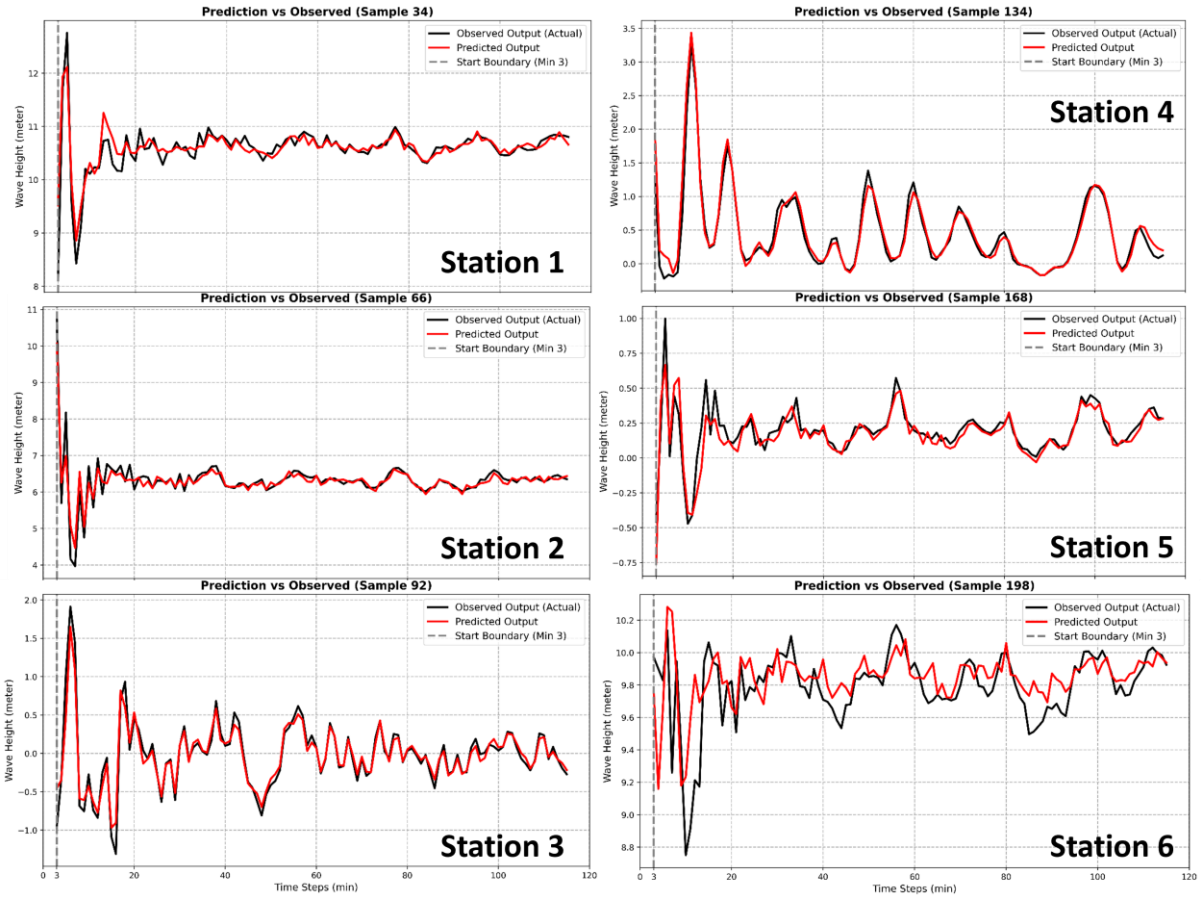


Figure 6.19 Comparison of predicted and observed tsunami waveforms at six stations using the Bi LSTM model

In addition to Vanilla LSTM, the BiLSTM model (Figure 6.19) demonstrates improved performance in most cases compared to the Vanilla LSTM (Figure 6.18) in predicting tsunami waveforms across different stations. The BiLSTM model demonstrates an enhanced capability in capturing intricate waveform structures, reducing prediction deviations, and improving phase accuracy. This improvement can be attributed to the ability of BiLSTM to process both past and future temporal dependencies, enabling more precise waveform reconstructions compared to the unidirectional Vanilla LSTM.

At Station 1 and Station 2, which are located closer to the tsunami source, the predicted waveforms exhibit a high degree of alignment with the observed data. The model effectively captures peak amplitudes and subtle oscillatory patterns following the initial peak. Compared to the Vanilla LSTM, the BiLSTM better preserves oscillation frequencies and reduces phase

shifts, indicating a more refined understanding of tsunami wave propagation in near-source regions where wavefront dynamics dominate and external influences remain minimal.

At Station 3 and Station 4, positioned at an intermediate distance from the source, the BiLSTM model maintains superior performance over the Vanilla LSTM. The predicted waveform structure follows the observed data more closely, and the model effectively mitigates wave amplitude decay errors. Notably, oscillatory trends and secondary wave interactions are better preserved, which are critical for understanding wave reflections and energy dissipation. Compared to the Vanilla LSTM, the BiLSTM improves phase accuracy and reduces deviations in later time steps, making it more reliable for mid-range tsunami forecasting.

For Station 5 and Station 6, which are farther from the tsunami source, the BiLSTM continues to demonstrate notable improvements over the Vanilla LSTM. These stations are characterized by complex wave interactions due to coastal reflections, seabed variations, and non-linear dispersion effects. Despite these challenges, the BiLSTM successfully mitigates amplitude underestimation and reduces phase misalignment, particularly in Station 6, where Vanilla LSTM previously exhibited significant prediction errors. The improvement in waveform fidelity across these distant stations highlights the model's superior generalization capability across different spatial domains.

A key factor contributing to the BiLSTM's superior performance is its ability to leverage information from both past and future time steps, unlike the Vanilla LSTM, which relies solely on past information as also demonstrated in previous studies such as (Adytia *et al.*, 2024). This bidirectional processing mechanism enables the model to better reconstruct waveform sequences, particularly in later-stage oscillations and energy dissipation trends, which are crucial for long-term tsunami forecasting applications. Therefore, the BiLSTM model significantly enhances the representation of wave amplitudes, phase alignment, and oscillatory trends, making it a more stable choice for tsunami early warning systems.

Table 6.5 Evaluation results of waveform forecasting models (Vanilla LSTM and BiLSTM) across six synthetic stations using MSE, MAE, and RMSE metrics

	Vanilla LSTM			**Bidirectional LSTM**		
	Avg MSE	**Avg MAE**	**Avg RMSE**	**Avg MSE**	**Avg MAE**	**Avg RMSE**
Station 1	0.1148	0.1614	0.3090	0.1225	0.1582	0.3111
Station 2	0.1906	0.2008	0.4044	0.1868	0.1902	0.3935
Station 3	0.1814	0.1961	0.3807	0.1790	0.1854	0.3627
Station 4	0.0925	0.1275	0.2604	0.0878	0.1269	0.2545
Station 5	0.0410	0.0903	0.1827	0.0390	0.0905	0.1784
Station 6	0.0834	0.1330	0.2522	0.0633	0.1157	0.2159

To interpret model performance under varying wave dynamics, Table 6.5 summarizes the error metrics and highlights distinct spatial patterns. It presents the Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) for both Vanilla LSTM and Bidirectional LSTM (BiLSTM) models across six synthetic observation stations. The

results indicate that the BiLSTM model generally outperforms the Vanilla LSTM, particularly in capturing waveform characteristics at stations affected by complex wave interactions.

The Vanilla LSTM model shows relatively consistent performance, with RMSE values ranging from 0.1827 (Station 5) to 0.4044 (Station 2). The lowest error values at Station 5 (RMSE: 0.1827, MAE: 0.0903, MSE: 0.0410) suggest that the model performs best in locations where wave propagation dynamics are relatively simple. However, higher RMSE values observed at Station 2 (RMSE: 0.4044, MAE: 0.2008, MSE: 0.1906) and Station 3 (RMSE: 0.3807, MAE: 0.1961, MSE: 0.1814) highlight the model's difficulty in accurately predicting waveforms at greater distances from the tsunami source, where wave reflections, dispersion effects, and coastal interactions become more prominent.

In contrast, the BiLSTM model consistently outperforms the Vanilla LSTM model across most stations, particularly in cases involving complex wave reflections and interactions. For instance, at Station 1, the BiLSTM model achieves an RMSE of 0.3111 and MAE of 0.1582, compared to the Vanilla LSTM's RMSE of 0.3090 and MAE of 0.1614. Similarly, at Station 4, the BiLSTM exhibits lower RMSE (0.2545) and MAE (0.1269) values, reinforcing its advantage in accurately modeling waveform dynamics in regions influenced by bathymetric variations and multi-directional wave interactions. However, at Station 5, where wave dynamics are relatively simple, both models perform similarly (RMSE: 0.1784 for BiLSTM vs. 0.1827 for Vanilla LSTM), suggesting that BiLSTM provides the most benefit in regions with highly dynamic wave behaviors rather than in locations with straightforward wave propagation.

6.2.3.4 Forecast model test with the real observational 2018 event

To validate the AI-based forecasting models under real-world conditions, we conducted an experiment using actual tide gauge data recorded during the 2018 Anak Krakatau tsunami event. This real-event validation was designed to assess the model's generalization capability beyond synthetic simulations and to evaluate its practical application in operational early warning scenarios.

We tested both the trained Vanilla LSTM and Bidirectional LSTM (BiLSTM) models on four tide gauge locations that recorded tsunami waveforms during the 2018 event: Kota Agung, Ciwandan, Panjang, and Marina Jambu. As shown in Figure 12, the red lines represent the model predictions starting from minute 3, while the blue lines show the real observed waveforms. This setup simulates the short-term early warning context, where only limited initial data is available for forecasting.

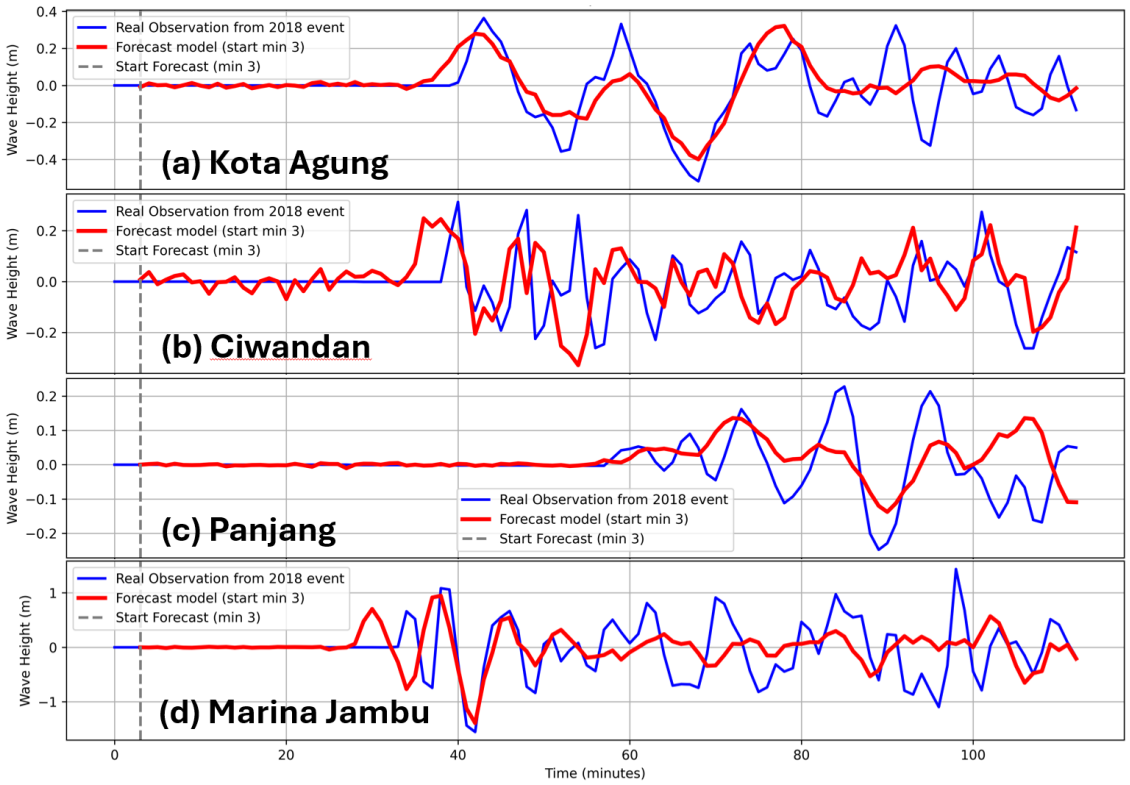


Figure 6.20 Forecast model results test using real observational data from the 2018 Anak Krakatau tsunami event at four tide gauge locations: (a) Kota Agung, (b) Ciwandan, (c) Panjang, and (d) Marina Jambu. The red line indicates the model prediction (starting from

The prediction results at Kota Agung showed good alignment with observed waveforms using the Vanilla LSTM model, achieving a Mean Absolute Error (MAE) of 0.1202 and a Root Mean Square Error (RMSE) of 0.1767. Similarly, the Ciwandan station yielded MAE = 0.1658 and RMSE = 0.2569. For Panjang and Marina Jambu, the BiLSTM model was used. While Panjang achieved low errors (MAE = 0.0788, RMSE = 0.1316), the performance at Marina Jambu was notably poorer (MAE = 0.4553, RMSE = 0.6300), possibly due to stronger noise and waveform variability at that location.

A summary of these results is presented in Table 6.6, which reports the MAE and RMSE values across all four stations. The average values, MAE of 0.2050 and RMSE of 0.2985, demonstrate that the model was able to generalize reasonably well when applied to real tsunami data, even though it was trained entirely on synthetic scenarios.

Table 6.6 Performance evaluation of the forecasting model test

Tide gauge station	Model used	Mean MAE	Mean RMSE
Kota Agung	Vanilla LSTM	0.1202	0.1767
Ciwandan	Vanilla LSTM	0.1658	0.2569
Panjang	BiLSTM	0.0788	0.1316
Marina Jambu	BiLSTM	0.4553	0.6300
Average		**0.2050**	**0.2988**

These results provide encouraging evidence of the model's potential for real-world deployment, particularly for early warning purposes where rapid and reliable forecasting is

critical. This validation with real observational data supports the model's applicability and lays the groundwork for future integration into operational systems.

6.2.3.5 Model scalability and computational efficiency

To assess the model's scalability and suitability for real-time early warning systems, we evaluated the computational efficiency of both Vanilla LSTM and BiLSTM architectures during training and inference. The experiments were conducted on a standard desktop PC (Intel i7-7700 CPU @ 3.60GHz, 16 GB RAM, without GPU acceleration).

Table 6.7 Comparison of training and inference time between Vanilla LSTM and BiLSTM models

Tide gauge station	Model used	Training Time (s)	Average Inference Time per Scenario (s)
Station 1	Vanilla LSTM	14.43	0.0012
Station 2	Vanilla LSTM	6.47	0.0014
Station 3	Vanilla LSTM	7.17	0.0022
Station 4	Vanilla LSTM	10.25	0.0027
Station 5	Vanilla LSTM	6.28	0.0025
Station 6	Vanilla LSTM	13.57	0.0012
Average		**9.7**	**0.0019**
Station 1	BiLSTM	13.17	0.0027
Station 2	BiLSTM	9.59	0.0031
Station 3	BiLSTM	16.12	0.0104
Station 4	BiLSTM	9.65	0.003
Station 5	BiLSTM	11.29	0.0069
Station 6	BiLSTM	11.63	0.0022
Average		**11.91**	**0.0047**

As shown in Table 6.7, the average training time per station for Vanilla LSTM was approximately 9.70 seconds, compared to 11.74 seconds for BiLSTM. Although BiLSTM required moderately longer training durations due to its bidirectional processing, the inference time remained within acceptable real-time bounds. The average inference time per scenario was 0.0019 seconds for Vanilla LSTM and 0.0047 seconds for BiLSTM.

These results indicate that both models are computationally lightweight enough to be deployed in near-real-time settings, even without high-end hardware. To further enhance efficiency, future work may explore the application of model optimization strategies such as pruning, quantization, and knowledge distillation to reduce model size and latency without compromising accuracy.

6.2.3.6 Sensitivity analysis to the number of input gauges

To evaluate the model robustness under limited observational coverage, a sensitivity analysis was performed by varying the number of input gauges used during forecasting. This test reflects real-world scenarios in which not all sensors may be available. Table 6.8 and Figure 6.21 present the model's forecasting performance using one, two, three, and all six virtual gauges as input, targeting Station 3 as a representative location.

Table 6.8 Comparison of number of gauges configuration with forecasting error

Number of Gauges	Configuration (Example)	RMSE (m)
1	Gauge 1 only	0.468
2	Gauges 1 & 2	0.409
3	Gauges 1, 3 & 5	0.396
6	All gauges	0.3807

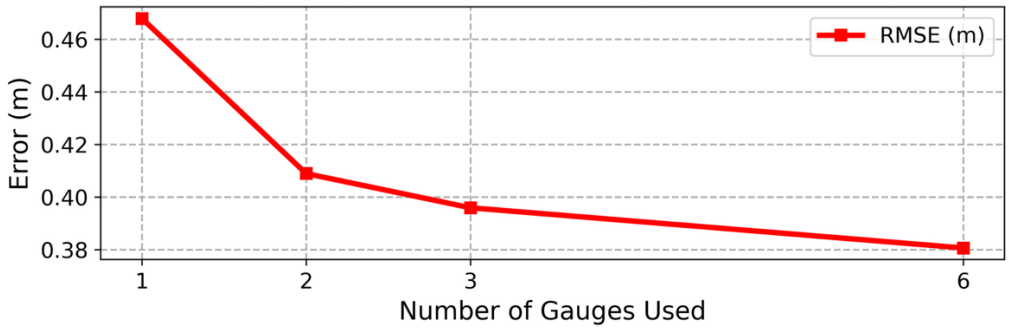


Figure 6.21 Number of Gauges vs Forecasting Error

The results show that although prediction accuracy declines as the number of gauges decreases, the model still produces meaningful forecasts using only three strategically selected gauges. Using one or two gauges led to higher RMSE values (0.468 and 0.409 m, respectively), while three gauges reduced the RMSE to 0.396 m, only slightly higher than the full-input configuration (0.3807 m). These findings demonstrate that the model remains effective even under reduced input conditions, supporting its applicability in early warning systems with sparse instrumentation.

This evaluation highlights a critical trade-off between gauge availability and forecast accuracy. It also reinforces the model's practical value in resource-constrained regions, where the number of operational tide gauges may be limited due to geographical and logistical challenges, limited resources, or damage from tsunami waveforms.

6.2.3.7 Evaluation metrics for early warning relevance

To evaluate the operational utility of the forecasting models, two additional metrics were introduced: Peak Wave Amplitude Error (PWAE) and Time-to-Peak Error (TPE). PWAE measures the absolute error between predicted and observed peak wave heights, while TPE captures the time discrepancy in peak wave arrival. These metrics are particularly relevant for tsunami early warning, where both the amplitude and timing of the peak wave are critical.

Table 6.9 presents a statistical summary across 30 representative example test scenarios. The PWAE ranged from 0.01 m to 0.13 m, with a mean value of 0.057 m and standard deviation of 0.036 m. Meanwhile, TPE values ranged from 0 to 68 minutes, with a mean of 7.32 minutes and standard deviation of 14.25 minutes. Notably, the time-to-peak error was below 5 minutes in most scenarios, though some long-period waveforms resulted in larger deviations.

Table 6.9 Peak Wave Amplitude Error (PWAE) and Time-to-Peak Error (TPE) across 30 example scenarios

Metric	Minimum	Maximum	Mean	Standard Deviation
PWAE (m)	0.01	0.13	0.057	0.036
TPE (minutes)	0	68	7.32	14.25

These results highlight the forecasting model general reliability in predicting both the magnitude and timing of peak tsunami waves. The relatively low average PWAE across scenarios indicates that the model can capture critical peak wave heights with acceptable precision, essential for estimating potential inundation and damage. Likewise, although the average TPE includes some outliers, most scenarios achieved timing errors within the acceptable threshold for real-time early warning. The variability observed, particularly in long-period or complex waveforms, emphasizes the importance of integrating task-specific metrics.

Such metrics offer a more nuanced assessment of the model operational readiness, complementing general-purpose error measures like MAE or RMSE, and ensuring alignment with practical decision-making needs during emergency response.

6.2.4 Discussion

The findings of this study highlight the effectiveness of utilizing LSTM-based architectures for tsunami waveform forecasting using short-term observational data from multiple synthetic stations. The proposed models, Vanilla LSTM and Bidirectional LSTM (BiLSTM), demonstrated their ability to predict long-term tsunami waveforms accurately, underscoring their potential for real-time tsunami early warning applications. This section explores the implications of these results, assesses the performance of the models, and identifies limitations and directions for future research.

6.2.4.1 Performance Analysis

The Bidirectional LSTM consistently outperformed the Vanilla LSTM across most performance metrics and stations, particularly in regions with complex wave dynamics. The ability of the BiLSTM to leverage bidirectional temporal dependencies allowed it to capture both past and future interactions, resulting in more accurate predictions. This finding aligns with prior studies that emphasize the advantages of bidirectional architectures in sequence modeling tasks. However, in regions where wave dynamics are simpler, such as Station 5, the performance of the two models was comparable, suggesting that simpler scenarios may not significantly benefit from the added complexity of bidirectional architectures.

The use of short-term observational data (first 3 minutes) as input proved sufficient for capturing critical features of tsunami waveforms, including amplitude. By using only minimal input data, the proposed approach minimizes computational overhead, making it suitable for early warning applications. The seamless transition from observed input to predicted waveforms underscores the robustness of the models in handling diverse scenarios. Despite the promising results, slight deviations in predictions for distant stations, such as Station 3 and Station 6, highlight the challenges of accurately modeling wave interactions.

Considering our result, we also noticed that other proposed methods for tsunami early warning systems for Anak Krakatau Volcano. Several studies have investigated alternative approaches to this challenge. Mulia *et al.* (2020) demonstrated the effectiveness of a ship-based height positioning method for detecting tsunami amplitudes. Wang *et al.* (2023) utilized high-frequency radar data assimilation for tsunami prediction. Ratnasari *et al.* (2023) employed precomputed results within identified scenarios as input for rapid tsunami propagation simulations. Despite the complexity of early warning systems for non-seismic tsunamis, continuous exploration of various methods is deemed beneficial. Therefore, we posit that the proposed method in this study represents a novel approach, utilizing minimal input from synthetic or virtual observations placed on islands surrounding Anak Krakatau. This approach is envisioned as a step towards establishing an efficient early warning system. With advancements in simulation and data analytics for prediction, we believe this study marks the initial strides towards generating further research ideas aimed at addressing the challenges of early warning systems for non-seismic tsunamis.

However, it was also observed that when the model was tested using the real observational data from the 2018 event, the results at Marina Jambu were notably poor. After rerunning the model multiple times and thoroughly verifying each step of the analysis, the predictions at this station remained consistently suboptimal. Based on these evaluations, the following possible reasons have been identified to explain the model poor performance at Marina Jambu.

The models relatively poor performance at Marina Jambu, as reflected by its higher MAE (average 0.4553) and RMSE (average 0.6300) compared to other stations, warrants specific attention. Upon further examination, two possible contributing factors can be identified.

First, the geographical location of Marina Jambu might play a critical role. It is likely situated in close proximity to the source of the tsunami-generating flank collapse. As such, the station is more directly exposed to the initial and highest-energy wavefronts, which may arrive earlier and exhibit stronger, more abrupt waveform characteristics. Because the model was trained solely on synthetic scenarios without incorporating source-station distances, it could not adapt its prediction to this site-specific dynamic. The generalized training data, while diverse in terms of scenario parameters, did not provide the model with the necessary sensitivity to differentiate how location relative to the source affects waveform behaviour.

Second, the waveform recorded at Marina Jambu exhibits high amplitudes (more than 1 meter) and rapid fluctuations (8 min earlier than Ciwandan, 10 min earlier than Kota Agung, and 30 min earlier than Panjang) compared to other stations, which may indicate the presence of local effects such as early wave reflections caused by complex coastal geometry or seabed topography (see Figure below). Since the LSTM model was trained on relatively regular and idealized synthetic waveforms, it likely over-smoothed the predictions and failed to accurately reproduce the sharper and earlier-arriving features observed in the real data.

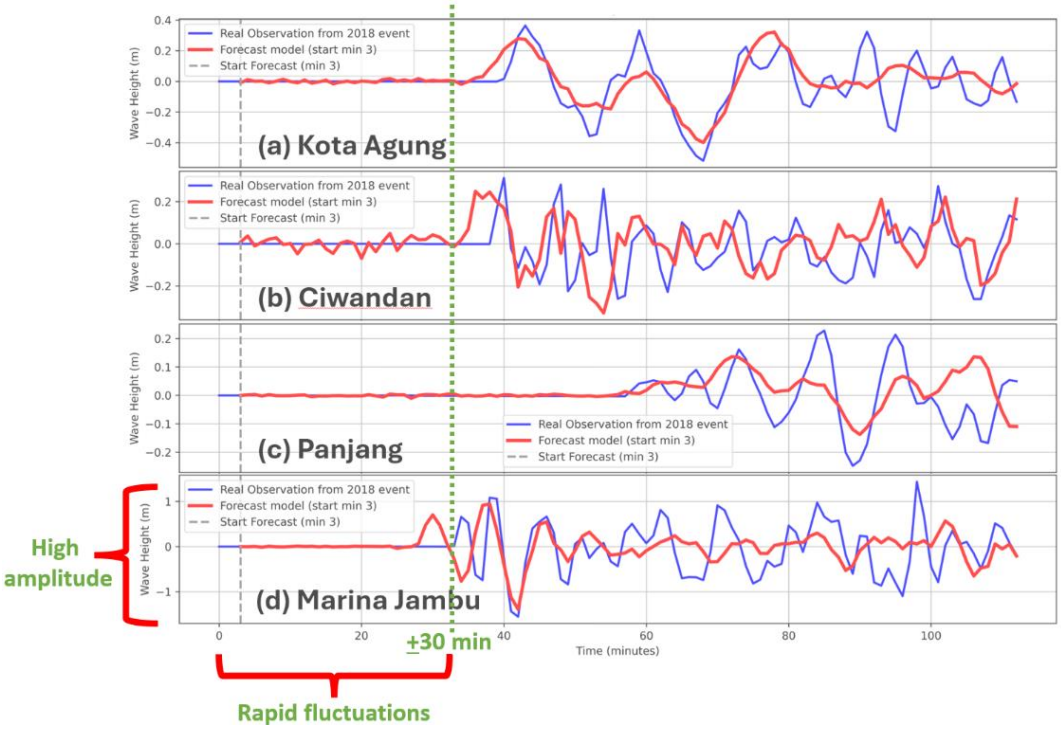


Figure 6. 22 Explanation and Analysis of Model Deficiency at Marina Jambu Station

6.2.4.2 Practical Implications

The proposed approach has significant implications for early warning systems, demonstrating the feasibility of leveraging deep learning algorithms to predict tsunami waveforms efficiently and accurately. The ability to generate reliable forecasts based on limited initial data is crucial for decision-making in disaster response and mitigation. In particular, the application of BiLSTM highlights its potential to enhance tsunami prediction accuracy, thereby reducing false alarms and improving response times.

Furthermore, the findings suggest that integrating LSTM-based models with existing numerical simulations could significantly improve the accuracy and speed of tsunami forecasting. By complementing traditional physical models and generating databases using data-driven approaches, early warning systems could become more robust and scalable for real-time applications. These hybrid systems would be especially beneficial for regions with limited computational resources or real-time monitoring capabilities.

Additionally, while this study focuses on tsunamis triggered by the Anak Krakatau volcano, its findings extend beyond this specific case. The forecasting model, developed through the integration of deep learning and physical simulations, can be adapted for other volcanoes with similar collapse-induced tsunami risks. Anak Krakatau is just one case of several significant volcanoes in Southeast Asia recognized as potential sources of subaerial volcanic tsunamis, as highlighted by Zorn *et al.* (2022). Other notable examples include Ritter Island in Papua New Guinea and Didicas in the Philippines.

6.2.4.3 Limitations and Future Work

Despite the success of the proposed models, several limitations must be acknowledged. First, while both Vanilla LSTM and BiLSTM models demonstrated feasibility for real-time

forecasting with low inference time, further optimization is recommended for large-scale deployment. Future research may explore techniques such as model pruning, quantization, or the design of lightweight architectures to reduce computational load and improve response time. These strategies are particularly important for deployment in edge-computing environments or resource-constrained early warning systems.

Second, the models were tested on a limited number of tide gauge stations, which may not fully capture the variability of coastal and bathymetric features. This limits the ability to generalize findings to other regions with different geophysical characteristics. In addition, the synthetic training dataset may introduce biases stemming from assumptions in the COMCOT simulation inputs—such as collapse volume, dip angle, and source direction. These parameters, along with simplifications in bathymetry and coastal topography, can affect waveform generation and model performance. Future work should explore more diverse and high-resolution bathymetric data, vary simulation parameters more broadly, and conduct sensitivity analyses to assess their impact on prediction accuracy.

Third, to enhance robustness under operational conditions, future studies could apply noise injection and data augmentation techniques during training. These methods simulate waveform variability and imperfections typically present in real observational data. Moreover, combining synthetic and limited real data via hybrid training, or leveraging transfer learning approaches, may improve model adaptability, particularly for data-scarce events such as volcanic tsunamis.

Fourth, future research could explore advanced deep learning architectures, such as Transformers or BERT, to improve temporal pattern recognition. Expanding the model inputs to incorporate complementary data sources, such as remote sensing, infrasound measurements, or ocean-bottom pressure sensors, could further enhance forecasting reliability and broaden the application of tsunami early warning systems.

Lastly, in this study, we focused on tsunami generation triggered by volcanic flank collapse, using Anak Krakatau as a case study. While this mechanism is highly relevant and physically well-characterized, it is important to acknowledge that volcanic tsunamis can result from multiple processes, including explosive eruption dynamics, air-sea shock coupling, and atmospheric pressure waves, as seen in the 2022 Hunga Tonga-Hunga Ha'apai event (Lynett *et al.*, 2022). These processes may interact and produce compound signals in the near field.

The current model isolates collapse-induced tsunamis to evaluate the feasibility of waveform prediction using early short-term synthetic observation. Future work should expand this framework to incorporate multi-mechanism simulations, which could enhance the model's applicability to a broader range of volcanic tsunami scenarios. This is crucial for capturing the full complexity of explosive eruptions where atmospheric coupling and pressure-driven waves dominate.

6.2.5 Summary of Part II

This sub chapter proposed and evaluated the use of Long Short-Term Memory (LSTM) based architectures, including Vanilla LSTM and Bidirectional LSTM (BiLSTM), for tsunami waveform forecasting using short-term observational data from multiple synthetic stations. The results demonstrated that both models effectively predict long-term tsunami waveforms, with

the BiLSTM consistently outperforming the Vanilla LSTM in terms of accuracy, particularly in regions with complex wave dynamics. The use of only the first 3 minutes of observational data as input highlights the potential for efficient and computationally feasible early warning systems. The findings emphasize the capability of LSTM-based models to capture critical temporal dependencies in tsunami waveforms, offering an approach to complement traditional numerical simulations. The integration of these machine learning models into operational early warning systems could enhance real-time prediction accuracy, improve disaster response and prevention, and reduce false alarms, thereby minimizing the potential impact of tsunami events on coastal region.

While promising, this study also identified areas for further research, including the need to validate the models with more real-world observational data, optimize computational efficiency, and expand their applicability to regions with diverse bathymetric and topographic features. Future exploration of advanced deep learning architectures and additional data sources, such as remote sensing, volcanic thermal, and infrasound, could further improve the scalability of the proposed approach. These advancements would represent a significant step toward developing reliable and efficient tsunami early warning systems that contribute to disaster risk reduction and coastal resilience.

Although this research focuses on the volcano tsunami caused by Anak Krakatau volcano, the consequences go beyond this particular case study. We believe and expect that the knowledge gained from this study can be generalized to other volcanoes that have a comparable risk of collapse-induced tsunamis. Anak Krakatau is just one of several notable volcanoes in Southeast Asia that rank among the subaerial volcanic tsunami hazard sources, as highlighted by Zorn *et al* (2022) research. Other notable volcanoes include Ritter Island in Papua New Guinea, Lliwerung in Indonesia, and Didicas in the Philippines. The wider applicability of the study's methodologies and findings paves the way for a more thorough comprehension of the dynamics of volcanic tsunamis, which in turn helps coastal areas that are susceptible to these geological phenomena in a variety of volcanic landscapes become more aware of and prepared for them.

6.3 Part III. Classification of non-seismic tsunami early warning level using decision tree algorithm

6.3.1 Introduction

High-impact, low-probability (HILP) events present significant challenges. 'High consequence' refers to events with severe impacts or outcomes resulting in major damages. Conversely, 'low probability' indicates events that are unlikely to occur frequently. The consequences of HILP events extend rapidly across various sectors, often causing second or third-order impacts. Recent trends show that HILP tragedies are increasing in severity and frequency (Lee et al., 2012b), yet historical data are scarce, making prediction difficult (Lakshita & Nair, 2021). One example of a HILP is tsunamis (Madeira et al., 2020). Tsunamis are commonly generated by seabed displacements caused by earthquakes. However, they can also result from other mechanisms, such as volcanic activity, which is categorized as non-seismic tsunamis (R. Paris et al., 2014). Although non-seismic tsunamis are rare events (low probability), there have been approximately 100 notable such as volcanic tsunamis recorded in the past three centuries, yet they pose significant hazards due to their high consequences within the tsunami domain (R. Paris et al., 2014). Furthermore, as highlighted by (Day, 2015), one significant source of tsunamis from volcanic activity is flank or sector collapse.

Anak Krakatau, part of the renowned Krakatau volcanic complex located in the Sunda Strait between Java and Sumatra in Indonesia, has a storied history of volcanic activity. In December 2018, Anak Krakatau (AKV) experienced a major lateral collapse during a period of eruptive activity (Grilli et al., 2019). While flank instability and sector collapse are common occurrences on volcanic islands, the collapse of the AKV posed a particularly severe threat, triggering a deadly tsunami (Walter et al., 2019). This catastrophic tsunami struck the Sunda Strait region, resulting in hundreds of fatalities, and thousands of injuries (Grilli et al., 2019; Heidarzadeh, Ishibe, et al., 2020; Heidarzadeh, Putra, et al., 2020; Muhari et al., 2019). The volcano is mainly located together with the caldera wall (Deplus et al., 1995; R. Paris et al., 2014) which will be contributing to the flank collapse of AKV and the triggering tsunamis. As happened in the event of December 2018, the thickness, volume, and velocity of collapse become important points in tsunami generation from volcano collapse (R. Paris et al., 2014).

In order to save people's lives from volcanic tsunamis, especially for coastal society, an early warning system is significant for taking decisions and action (United Nations Office for Disaster Risk Reduction, 2023). However, currently, there is no standardized approach for early warning systems particularly for non-seismic tsunamis (Selva et al., 2021). Most tsunami warning systems are developed for seismogenic tsunamis (e.g., earthquake-generated) and not for non-seismic tsunamis (Titov et al., 2005). Therefore, the traditional tsunami early warning method based on the seismic sensor may not be appropriate for detecting volcanic tsunamis, as they do not produce seismic waves. Hence, detecting these kinds of non-seismic tsunamis and issuing early warnings become very challenging tasks in the tsunami domain. Previous studies have only focused on numerical modelling to understand the mechanisms of volcanic tsunamis (Cutler et al., 2022; Grilli et al., 2021; Omira & Ramalho, 2020; Perttu et al., 2020; Sabeti & Heidarzadeh, 2024; Terry et al., 2022), yet still less emphasis on developing effective early warning systems.

Moreover, other past research has studied various Machine Learning (ML) algorithms for disaster prediction. For instance, Support Vector Machines (SVM) are widely used in classification tasks since this algorithm handles high dimensional data and is able to make an optimal decision (Linardos et al., 2022; Sun et al., 2024). Even though the accuracy of SVM is high, in most cases, SVM is considered as black-box model. The reason because the model cannot give the appropriate insight about how the decision occurs. Similar to SVM, k-Nearest Neighbors (kNN) is also usually for classification tasks in many research and application fields, including disaster prediction (Abraham et al., 2024; Gulghane et al., 2024). kNN has a strength point such as being easy to interpretable and simple to implement. However, the computational cost is very high if the study uses a large dataset since the kNN will be storing all training data dan have to calculate the distance for each new prediction. This is incompatible with the purpose of tsunami early warning systems that require a real-time application.

Alternatively, the Decision Tree method is one of the ML algorithms that has a simple and interpretable model for dealing with the decision process (Kappi & Mallikarjuna, 2024; Pota et al., 2022). The interpretability makes the Decision Tree suitable for the aim of immediate response such as tsunami early warning. By splitting the dataset using some features and determined thresholds, this model forms a structure like a tree structure where each branch stands for a decision rule, and each leaf illustrates a classification score or output. This model gives two main advantages, namely transparency and efficiency. Transparency is very important in the process of decision-making. Decision Tree can visualize the simple process and make it easier to understand if-then rules. Regarding efficiency, Decision Trees can deal with large datasets and are capable of reducing the computational cost. This is very important in the case of early warning issuing. Moreover, Decision Trees are normally appropriate for handling imbalanced datasets and non-linear relationships between given features, which usually appear in the processing of tsunami datasets.

In this research, a Decision Tree model and a tsunami numerical model were combined to optimize the classification of early warning levels for non-seismic tsunamis. By integrating these two approaches, the numerical model plays an important role in simulating tsunami events and creating a dataset while the Decision Tree algorithm classifies the outcome data for improving the decision-making capabilities of tsunami early warning systems. This combination allows for both accurate simulation of complex, non-seismic tsunami events, such as the Anak Krakatau collapse, and efficient classification for real-time decision-making. By focusing on interpretability and operational efficiency, this integrated approach ensures that early warning provides actionable insights promptly.

6.3.2 Method

The methodology of this study employs a comprehensive warning decision model that integrates tsunami numerical simulations, input sequence consistency, and model construction for supervising the database within a decision-making framework (Figure 6.22). Tsunami numerical simulation utilizes COMCOT (Cornell Multi-grid Coupled Tsunami Model) software (Woo & Cho, 1998), which simulates tsunami scenarios based on various volcanic collapse parameters.

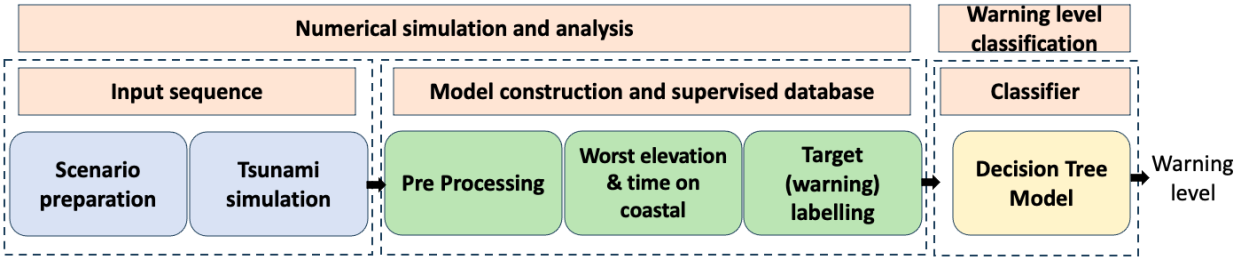


Figure 6.23 Flowchart of warning decision model

6.3.2.1 Supervised database and classifier model

After confirming that the simulation results reasonably match with the observed waveforms, we stored them in the database as the simulated results. These simulation results were used to construct a supervised database, forming the foundation for the decision support model. In this study, a supervised database refers to a dataset specifically structured for use in supervised learning. It comprises input features derived from simulation results, such as maximum wave elevation and collapse thickness, paired with labeled outputs representing tsunami warning levels. These labels enable a classifier model to learn the relationship between the input attributes and the correct warning levels. The classifier was trained using this supervised database, utilizing the learned relationship to predict warning levels from new simulation outputs by determining the appropriate warning level based on the key attributes. The definition of warning levels was based on the guidelines from the Indonesian Meteorology, Climatology, and Geophysical Agency (Baumert et al., 2018) as shown in Table 6.11: No Warning if wave height < 0.1 meter, Minor Tsunami issued if wave height range between 0.1 and < 0.5 meter, Tsunami if wave height between 0.5 and < 3.0 meter, and Major Tsunami if wave height range $\geq$ 3.0 meter.

These wave height thresholds were selected based on the official classification system used by the Indonesian Meteorology, Climatology, and Geophysical Agency (BMKG), which underpins the operational framework of the Indonesian Tsunami Early Warning System (InaTEWS). The defined ranges (as shown in Table 6.11) reflect BMKG's official guidance for tsunami warning dissemination and are used nationally in both real-time alerts and disaster drills. They are designed to balance scientific thresholds with practical considerations for evacuation and public safety in coastal regions of Indonesia. Thus, the classification system employed in this study aligns with local standards and supports real-world applicability.

Table 6.10 Supervised database with warning level for classifier model

Scenario	Collapse direction	Angle degree	Thickness (m)	Length (m)	Width (m)	Volume (km^3)	Worst Elevation (WE)	Time at WE (min)	Warning Level
1	East	9°	176	2906	78.2	0.04	0.27	49.72	Minor Tsunami
22	north	9°	62	1250	903.2	0.07	0.18	41.00	Minor Tsunami
43	northeast	9°	127	1395	846.7	0.15	0.76	41.13	Tsunami
66	northwest	9°	130	3288	444.5	0.19	1.30	32.94	Tsunami
91	south	9°	127	3490	541.5	0.24	1.14	41.32	Tsunami
114	southeast	9°	63	1499	2859.0	0.27	1.24	32.85	Tsunami
137	southwest	9°	43	1439	4848.3	0.30	0.76	41.62	Tsunami
160	west	9°	106	3727	835.3	0.33	0.83	42.72	Tsunami

Table 6.11 Define warning levels refer to Indonesia Meteorology, Climatology, and Geophysical Agency

Warning Level	Wave Height (WH) Range [m]
No Warning	0.0 = WH < 0.1
Minor Tsunami	0.1 = WH < 0.5
Tsunami	0.5 = WH <3.0
Major Tsunami	WH ≥ 3.0

6.3.2.2 Decision tree algorithm

The Decision Tree is a supervised learning method applicable to both classification and regression tasks, primarily used for classification problems. This method employs a tree-structured model where internal nodes signify dataset features, branches indicate decision rule, and leaf nodes signify outcomes. Within a Decision Tree, there are two types of nodes: Decision Nodes facilitate decision-making and have multiple branches, while Leaf Nodes represent final outcomes without further branching. Decisions within the tree are based on dataset features, presenting a graphical representation of all potential solutions or decisions given specific conditions. The tree starts with a root node and branches out, forming a hierarchical structure, hence the name 'Decision Tree'.

The tree is constructed using the Classification and Regression Tree (CART) algorithm, which iteratively splits the dataset based on specific criteria. The Decision Tree algorithm operates by asking a series of yes/no questions and splitting the tree into sub-trees based on the answers (Figure 6.23). Decision Trees are preferred for several reasons. Firstly, they mimic human decision-making processes, making them easy to understand (Quinlan, 1986). Secondly, their logic is transparent and straightforward to interpret due to their tree-like structure (Letham et al., 2015). To predict the class of a given dataset, the Decision Tree algorithm starts at the root node and compares the root attribute values with the actual dataset attributes. Based on this comparison, the algorithm follows the corresponding branch to the next node. This process continues recursively until it reaches a leaf node.

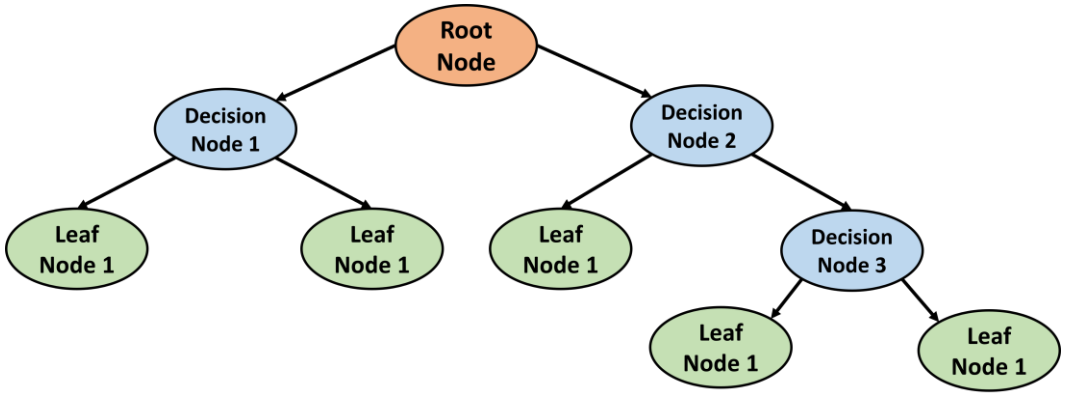


Figure 6.24 Decision tree model

Attribute Selection Measures (ASMs) are crucial for determining the best attribute for the root and sub-nodes. Two popular ASMs are Information Gain and the Gini Index. Information Gain helps estimate the change in entropy following dividing the data set based on the attributes

present. It examines the sensitivity of information given to a feature present about a class. The splitting in the Decision Tree algorithm starts with the highest gain first in order to reach the maximum Information Gain. The calculation of how Information Gain and entropy occurred in the Decision Tree are shown in equation 6.5 – 6.7 below.

Equation 6.5 – 6.7 Equation Information gain and entropy

$$\text{Information Gain} = \text{Entropy (S)} - (\text{Weighted Avg} * \text{Entropy of each feature}) \quad (6.5)$$

$$\text{Entropy (S)} = -\text{P(yes)} \log_2 \text{P (yes)} - \text{P(no)} \log_2 \text{P(no)} \quad (6.6)$$

On the other hand, the Gini Index is used to measure how good and optimal a node is in selecting features and determining the separation point based on its target class. Attributes with a small Gini Index indicate something better (purer) and, it is calculated as shown in equation 5 (3) below.

$$\text{Gini Index} = 1 - \sum \text{jP}_\text{j}^2 \quad (6.7)$$

6.3.2.3 Evaluation method

In using machine learning algorithms, evaluation methods are very important to measure the success of a model, in this case the ability to classify. In the case of Decision Tree, confusion matrices, F1 score, macro, and weighted average accuracies are used as evaluation methods. Confusion metrics will evaluate the comparison of actual versus predicted tsunami warning levels at four coastal stations. While F1 is used to assess the accuracy of the model for each class, especially when imbalanced datasets occur. To treat all classes equally, macro accuracy is used, and to find out the frequency level in each class, weighted average accuracy is applied. The metrics mentioned above are expected to be able and well proper to evaluate the performance of the classification model for volcanic tsunami early warning level.

6.3.3 Result

This result session will display tsunami waveform timeseries at each coastal tide gauge station along with information related to the specific time the tsunami will arrive at the coast and its worst elevation information. In addition, the calculation results of the decision tree algorithm and evaluation of the model classification performance will be visualized.

6.3.3.1 Tsunami waveform on each gauge or station

Tsunami waves resulting from 160 scenario simulation datasets were recorded at four coastal tide gauges, namely Marina Jambu, Ciwandan, Kota Agung, and Panjang. These four gauges are located on the islands of Java and Sumatra in the Sunda Strait. Marina Jambu and Ciwandan are located on the coast of Java (Banten) and Kota Agung and Panjang are located on the southern coast of Sumatra. The recorded tsunami wave recordings are important as input information called water surface elevation (WSE). The waveform is then visualized into a plot that depicts the waveform over time indicates tsunami behavior and its potential impact on coastal regions (Figure 6.24).

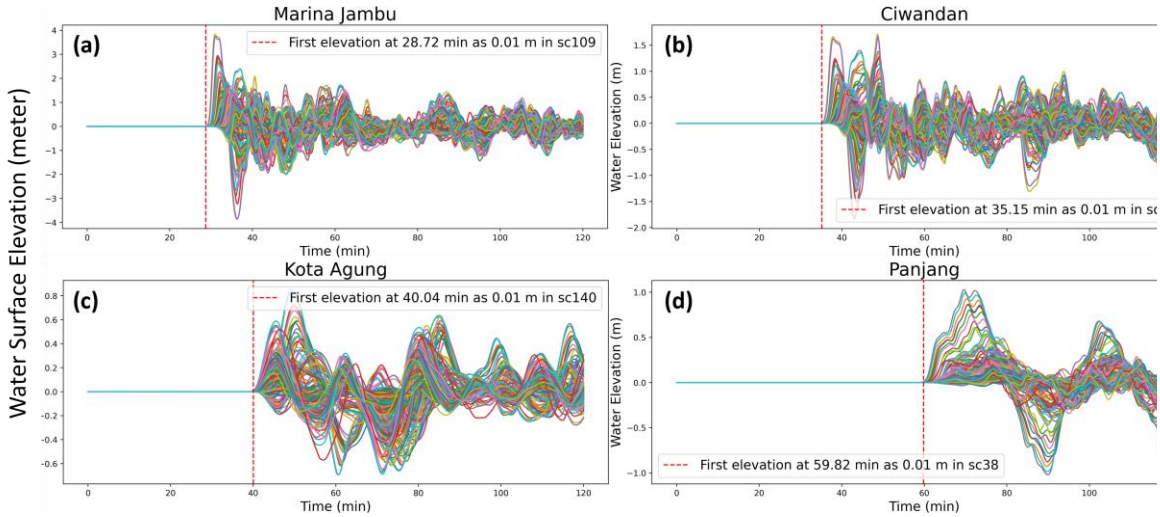


Figure 6.25 Waveforms of water elevation at four tide gauges during multiple scenarios. (a) Marina Jambu gauge (b) Ciwandan gauge (c) Kota Agung gauge (d) Panjang gauge

At Marina Jambu tide gauge, the tsunami waveform is the most important to be noticed. By searching in the database, scenario 109 becomes the very first significant elevation that surpasses the 0.01-meter threshold occurring at around 28.72 minutes. This timing shows at what minute in the worst-case scenario from the database, the tsunami starts to reach the coastal area and causes coastal inundation (indicated by the red dashed vertical line). This early detection is critical, given that Marina Jambu is one of the closest points on the coast of Java to the tsunami source. The relatively short time frame of approximately 27-28 minutes before the WSE exceeds the threshold underscores the urgent need for rapid evacuation and early warning systems in this area. The Ciwandan gauge shows a similar pattern, with the first elevation surpassing the 0.01-meter threshold at 35.15 minutes in scenario 109. The data indicates that while Ciwandan has a slightly longer lead time compared to Marina Jambu, the difference is not substantial, highlighting the need for swift action in both locations.

In Kota Agung, the first significant elevation above the 0.01-meter threshold occurs later, at 40.04 minutes in scenario 140, as marked by the red dashed vertical line in the plot. This delay compared to the Java coastal gauges suggests that communities in Sumatra, such as Kota Agung, have a slightly longer window to respond to tsunami warnings. However, this extended time frame should not lead to complacency, as the waveform data still indicates a pressing need for efficient and effective evacuation procedures. For Panjang gauge, the first elevation surpassing the 0.01-meter threshold occurs at 59.82 minutes in scenario 38. The waveform data here shows that Panjang, while experiencing the longest lead time among the four gauges, still faces significant tsunami risks that require prompt and coordinated response efforts.

Each subplot in the visualization not only highlights the initial exceedance of the 0.01-meter threshold but also depicts the overall dynamics of water surface elevation changes across different coastal locations and scenarios. These detailed waveforms offer a clear and comprehensive picture of how tsunamis evolve and impact various regions, providing essential data for improving tsunami early warning systems and emergency response strategies.

6.3.3.2 Specific time for worst elevation of tsunami waveform on each gauge

The visualization of tsunami waveform data from 160 scenarios at the four coastal gauges—Marina Jambu, Ciwandan, Kota Agung, and Panjang—provides a detailed look at the specific time intervals and maximum elevations recorded at these locations. Each subplot in the image represents one gauge and plots the Water Surface Elevation (WSE) in meters against time in minutes (Figure 6.25). This detailed representation is crucial for decision-making when determining warning levels, as it helps pinpoint the periods of highest tsunami elevation.

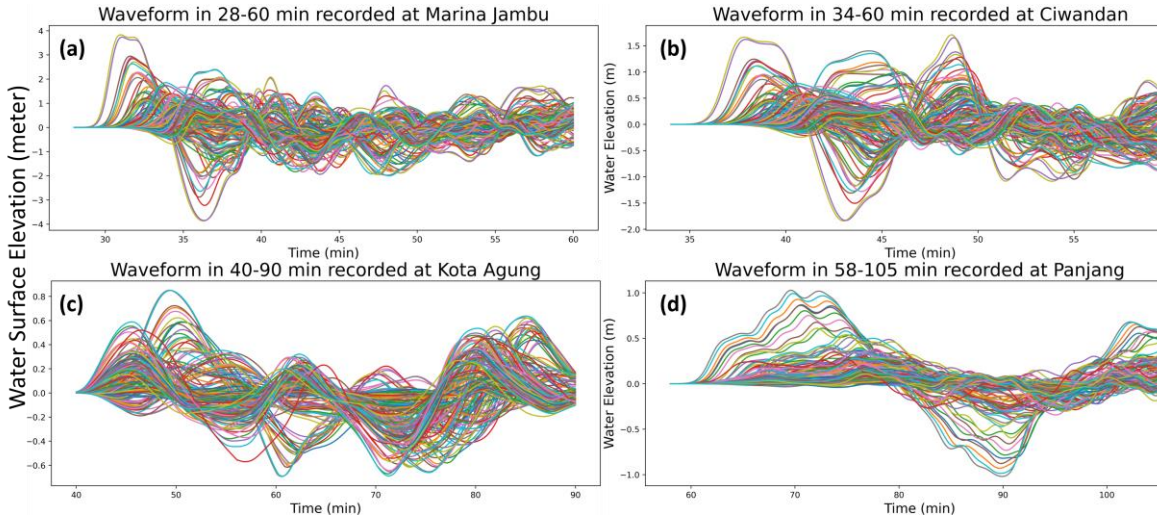


Figure 6.26 Specific time and worst elevation when reached coastal gauges (a) Marina Jambu record in 26-60 min (b) Ciwandan record in 34-60 min (c) Kota Agung record in 50-90 min (d) Panjang record in 58-105 min

For Marina Jambu, the worst elevation occurs between 26 and 60 minutes. This is the fastest time span for a tsunami to reach the coast compared to the other three coastal gauges because it arrived in less than 30 minutes. Similarly, the Ciwandan gauge records the worst elevation between 34 and 60 minutes. The waveform data within this timeframe emphasizes significant peaks in water surface elevation, underscoring the importance of this interval as tsunami worst elevation input into the database. This targeted approach to warning and evacuation can significantly enhance the effectiveness of tsunami response efforts. For Kota Agung, the worst elevation occurs between 40 and 90 minutes. The wave shape recorded in Kota Agung tends to be similar to Marina Jambu with dynamic up and down fluctuations even though the height is below one meter. For Panjang gauge, records the worst elevation between 58 and 105 minutes with the average waveform in a flat form.

The plot visualization of waveform data on each coastal gauge above revealed the specific time intervals for deciding the tsunami worst elevations. By identifying the periods when the WSE is highest, the early warning model can be developed as one of the significant parameters of warning level classification. This targeted approach ensures that warning issue strategies are timely and effective, reducing the potential impact of tsunamis on coastal communities.

6.3.3.3 Decision tree model visualization

Figure 6.26 and 6.27 illustrate the results of decision tree classification applied to data from four coastal tide gauges: Kota Agung, Panjang, Ciwandan, and Marina Jambu. Each decision

tree represents a series of splits based on different features related to tidal measurements, ultimately categorizing the potential tsunami warning levels at each gauge.

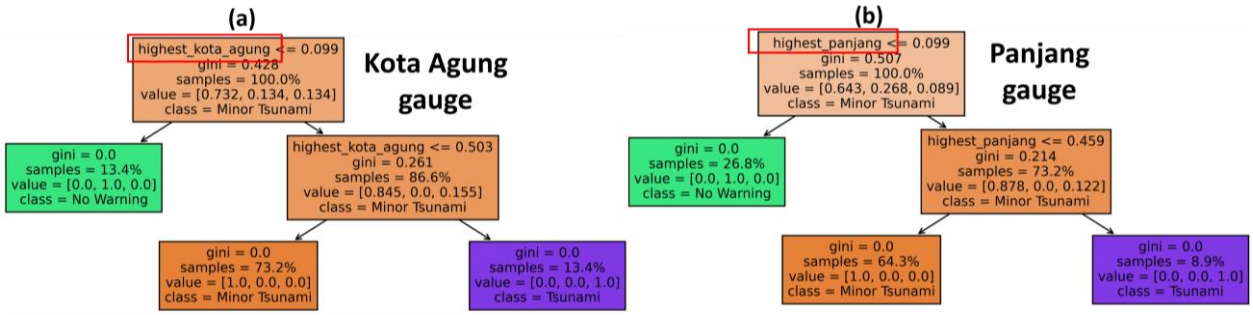


Figure 6.27 Decision tree classification results. (a) Kota Agung gauge (b) Panjang gauge

As shown in Figure 6.26, for the Kota Agung gauge (a), the decision tree classifier begins by splitting based on the feature highest_kota_agung. If this value is less than or equal to 0.099, the event is classified as a "Minor Tsunami" with a Gini impurity of 0.428, indicating some uncertainty in the classification. This branch further subdivides into pure nodes (Gini = 0) that classify as "No Warning" if the specified criterion is met. On the other hand, if highest_kota_agung is greater than 0.099, the model results in further subdivisions, eventually classifying the observations as either "Minor Tsunami" or "Tsunami." For the Panjang gauge (b), the initial split is also based on the feature highest_panjang. If then rule also applied to the value is less than or equal to 0.099, then the event is classified as a "Minor Tsunami". The next splitting on highest_panjang with different thresholds and samples leads to three classifications namely "No Warning," "Minor Tsunami," and "Tsunami". Naturally, the nodes will become purer as well as the Gini values decrease, demonstrating to become more confidence in classifications at the leaves of the tree such as "Minor Tsunami" and "Tsunami".

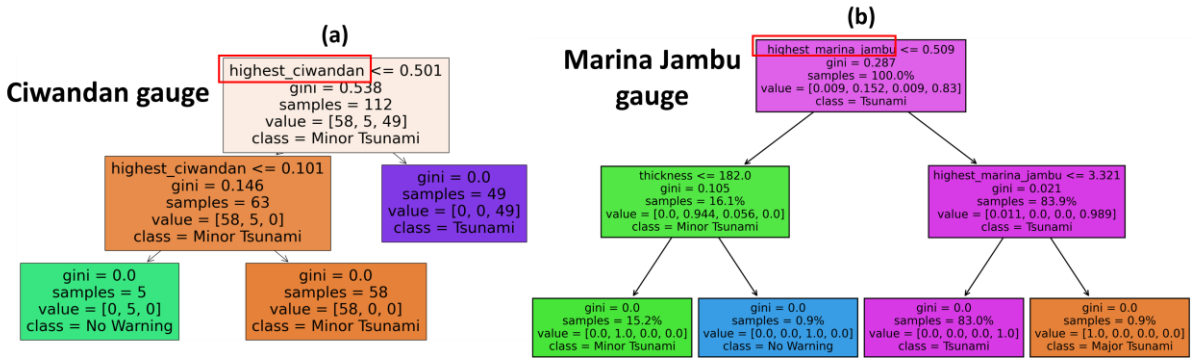


Figure 6.28 Decision tree classification results. (a) Ciwandan gauge (b) Marina Jambu gauge

Based on Figure 6.27, at Ciwandan gauge (a), the decision tree classifier starts the process with highest_ciwandan as a feature. Then splitting at 0.501 is categorized as "Minor Tsunami". More splitting based on the different thresholds classify into either "No Warning" or "Minor Tsunami". For Marina Jambu gauge (b), similar to Ciwandan, the decision tree begins with the feature highest_marina_jambu. An initial split at 0.509 "Tsunami" is the most important feature. The next splitting leads to a "Minor Tsunami" or "Major Tsunami" based on the threshold.

The decision tree classifiers for each gauge demonstrate a logical approach to categorizing tsunami early warning levels. In addition, from all four gauges visualized above, the decision

trees revealed that the highest of waveform becomes a primary feature over the other features such as thickness. In order to improve the classification, trees regularly split on important features at different thresholds to find the lowest gini impurity. As a result, a consistent reduction of impurity confirms that the model predictions become increasingly accurate and reliable. The classifications range from "No Warning" to "Minor Tsunami," "Tsunami," and "Major Tsunami," demonstrating how various feature values determine these results. This structured approach allows the model to handle the complexity of varying tsunami threats, enhancing its reliability and effectiveness for early warning.

6.3.3.4 Model evaluation: confusion matrix, f1, macro, and weighted average accuracies score

Figure 6.28 below shows the evaluation results of decision tree classification using confusion matrices at four coastal stations. As widely used for assessing classification performance, confusion matrices provide detailed insights into the number of true and false predictions produced by the model, categorized by the actual and predicted classes.

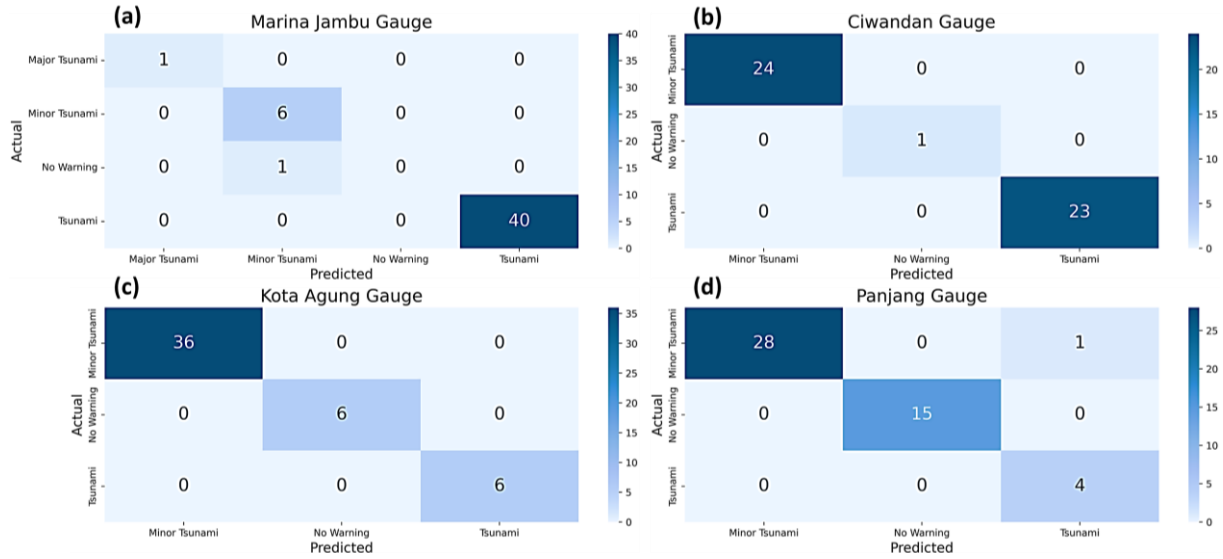


Figure 6.29 Model evaluation on each gauge. (a) confusion matrix on Marina Jambu (b) confusion matrix on Ciwandan (c) confusion matrix on Kota Agung (d) confusion matrix on Panjang

Based on Figure 6.28, the confusion matrix on Marina Jambu gauge (a) shows the predictions result by the model. It correctly predicted 1 "Major Tsunami", 6 "Minor Tsunami", and 40 "Tsunami" cases. Only 1 misclassification in this gauge that predicted "Minor Tsunami" that should be "No Warning". Similarly, the Ciwandan gauge (b) revealed 24 correct predictions for "Minor Tsunami", 23 correct predictions for "Tsunami" and 1 correct for "No Warning". There are no misclassifications in this gauge. Overall, confusion metrics for Marina Jambu and Ciwandan demonstrated the model's proficiency in classifying between "Minor Tsunami" and "Tsunami" with high accuracy, despite a single misclassification suggesting some room for improvement.

The confusion matrix for the Kota Agung gauge (c) shows high accuracy similar to the Ciwandan gauge previously, with 36 correct predictions for "Minor Tsunami" 6 predictions accurate for "Tsunami", 6 correct predictions for "No Warning". Importantly, there are no misclassifications in this gauge. For the Panjang gauge (d), the matrix shows 28 correct

predictions for "Minor Tsunami," 15 correct for "No Warning," and 4 for "Tsunami". However, there is 1 misclassification where a "Minor Tsunami" is predicted as "Tsunami".

Furthermore, F1, micro, and weighted average accuracy scores were used for evaluation (Table 6.12). These metrics provided an understanding of model performance by combining precision and recall into a single score for each class.

Table 6.12 F1-score, macro and weighted average accuracies of model evaluation metrics

F1-score	Marina Jambu	Ciwandan	Kota Agung	Panjang
No warning	0.00	1.00	0.00	1.00
Minor tsunami	0.92	1.00	0.92	0.98
Tsunami	1.00	1.00	1.00	0.89
Major tsunami	1.00	-	1.00	-

Accuracy	Marina Jambu	Ciwandan	Kota Agung	Panjang
Macro average	0.73	1.00	0.73	0.96
Weighted average	0.97	1.00	0.97	0.96

For the Marina Jambu gauge, the F1 scores are as follows: 0.00 for "No Warning," 0.92 for "Minor Tsunami," 1.00 for "Tsunami," and 1.00 for "Major Tsunami." The macro average of 0.73 and the weighted average of 0.97. These high scores indicate the model's strong performance in predicting "Minor Tsunami," "Tsunami," and "Major Tsunami" events. However, the score of 0.00 for "No Warning" reflects significant challenges in predicting this class, consistent with the previously noted misclassifications. For the Ciwandan gauge, the F1 scores are 1.00 for both "No Warning" and "Minor Tsunami," and 1.00 for "Tsunami." Since the gauge does not have data for "Major Tsunami", no score is provided for that class. Both macro and weighted averages are also at 1.00. This perfect score demonstrates the model's exceptional capability in predicting all relevant classes for this station, with no errors recorded in the confusion matrix.

For the Kota Agung gauge, the F1 scores are 0.00 for "No Warning," 0.92 for "Minor Tsunami," 1.00 for "Tsunami," and 1.00 for "Major Tsunami." Similar to Marina Jambu, the macro average of 0.73 and a weighted average of 0.97. While the model effectively predicts "Minor Tsunami," "Tsunami," and "Major Tsunami" events, it struggles significantly with the "No Warning" class, as indicated by the score of 0.00. Moving to the Panjang gauge, the F1 scores are 1.00 for "No Warning", 0.98 for "Minor Tsunami", and 0.89 for "Tsunami". Since there is no data for "Major Tsunami", no score is provided for that class. The macro average of 0.96 and the weighted average of 0.96. These scores indicate a generally strong performance, with the model performing slightly less effectively for the "Tsunami" class compared to the other gauges.

6.3.4 Discussion

This study simulating multi-scenario dataset encompassing all possible collapse directions and varying volumes has proven to be a suitable approach in anticipating future events similar to the 2018 Anak Krakatau Volcanic tsunami. Based on this database, the selected decision tree

algorithm is able to classify the warning level of volcanic tsunamis, enhancing its predictive capabilities and solving challenges for early warning systems as will be discussed below.

6.3.4.1 Analysis of significant features and influences from geographical context

The results of the study show that the integration of numerical simulations and Decision Tree classification is effective in predicting tsunami warning levels. The input features in the dataset included collapse thickness, maximum amplitude on the coast, collapse volume, direction, dip, width, length, and time to maximum wave elevation. The study finding is the identification of maximum wave amplitude (highest waveform) as the most significant feature followed by collapse thickness by referring to information gain and gini index in the Decision Tree model. Information gain measures the reduction in uncertainty (entropy) when the dataset is split while gini index measures the purity of the splits generated by an attribute. Information gain and Gini index procedurally evaluated each feature contribution to improving the predictive accuracy. Information gain was calculated by assessing the entropy before and after the split, with attributes like maximum amplitude and collapse thickness. This result indicates that these two features were the most critical in distinguishing between tsunami events and non-tsunami or minor tsunamis or even no warning.

Regarding Gini index, in this model, maximum amplitude achieved the lowest gini values at most of the tide gauges showing they produced the purest splits. While other attributes, such as collapse volume, are also relevant, they contributed less significantly to only on Marina Jambu gauge. The reason because that Marina Jambu has many input features since this gauge is the closest to the volcano. By focusing on these significant attributes, the Decision Tree model was able to effectively predict whether a tsunami would occur and classify the warning level. For example, the high predictive accuracy at Marina Jambu and Ciwandan reflects the shorter tsunami arrival times in these two locations, due to their closeness to the volcanic collapse. This also discovered the new insight into how the geographical context plays a role in the model predictions, providing information for future evacuation planning.

6.3.4.2 Model performance across different locations

The model performance differed among various coastal gauges, mainly due to the geographical context of the collapse site as previously mentioned. Marina Jambu and Ciwandan on Java Island revealed faster tsunami arrival times, with Marina Jambu having a critical time of less than 30 minutes. The reason because these gauges are closer to the Anak Krakatau volcanic collapse site. On the other hand, gauges on Sumatra Island, such as Kota Agung, had longer lead times of about 40 minutes due to their greater distance from the collapsed volcano. It emphasizes the importance of geographic context when issuing tsunami warnings such as distance from the collapse site. Moreover, the simulation not only identifies the arrival time but also the worst-case wave elevations. For example, Ciwandan had a time for the worst wave heights between 34 and 60 minutes, while Panjang in Sumatra saw these between 60 and 75 minutes. This information can help local decision makers with evacuation planning based on when the highest risk and time of arrival is expected.

6.3.4.3 Class imbalance during the model learning

Another challenge of the model is the imbalance class which had limited representation between the different categories in the dataset, especially "No Warning" category. The class imbalance problem appears when tsunamis events dominate the dataset while no warning occusr less frequently. This imbalance can lead to bias toward predicting the more common class such as "Tsunami", "Major Tsunami", and "Minor Tsunami" while struggling to predict the less frequent "No Warning" class. Even though the dataset used in this study included a variety of scenarios, there was less observance of the "No Warning" category. This made it difficult for the model to learn from these examples and resulted in weaker performance in predicting when no tsunami would occur. Oversampling or undersampling techniques can be used to address this issue. Alternatively, cost-sensitive learning can be implemented, where the model is penalized more heavily for misclassifying the minority class. This forces the model to pay closer attention to the underrepresented class, improving its ability to correctly predict less frequent events like "No Warning."

6.3.4.4 Comparative analysis with other studies

To point out the strength of this model, comparison to other studies using machine learning models such as Regression Trees and integrated networks in tsunami early warning systems are conducted. For example, (Cesario et al., 2024) demonstrated the use of regression trees for tsunami wave forecasting, emphasizing the interpretability and computational efficiency of decision-tree-based methods. This aligns with our findings, where the Decision Tree algorithm enhances interpretability and provides real-time classifications, offering advantages over other approaches that struggle with high training times and limited transparency, such as neural networks. Additionally, (B. Zhou et al., 2024) proposed a real-time tsunami monitoring system using submerged mooring systems, which we have referenced to highlight the value of combining real-time data collection with machine learning models like ours to improve tsunami early warning systems.

6.3.4.5 Model improvement strategies

To further improve the performance of the Decision Tree model in handling class imbalance, several strategies can be implemented. Ensemble methods such as boosting and bagging are highly effective techniques for improving model robustness. Boosting trains several weak classifiers, each one correcting the mistakes of the previous, then improves the model ability to handle hard-to-predict cases like the "No Warning" category. Furthermore, in the case of a tsunami dataset, bagging can be used by training multiple models on random portions of the data, such as wave height or time intervals to reduce errors and avoid overfitting to specific tsunami scenarios. By integrating the predictions of multiple classifiers, ensemble methods can handle the class imbalance of challenge. Besides to ensemble methods, adding another significant feature dataset in the tsunami domain could enhance its predictive capabilities. For example, seismic activity, atmospheric pressure, or collapse precursor signs may provide the model with a wider perspective for predicting non-tsunami events.

6.3.5 Summary of Part III

The decision tree classifier model used in this sub chapter presents a solution to dealing with early warning decisions, especially for non-seismic tsunamis such as volcanic tsunamis. Volcanic tsunamis, caused by events like submarine and subaerial collapse, pose a distinct challenge for early warning systems. Unlike seismic tsunamis that can be identified through seismic activity monitoring, volcanic tsunamis demand a different early warning strategy because of their unique detection causes. Given the current lack of early warning for such events, this study tried to fill a critical gap in preliminary non-seismic early warning system research by providing a classification warning model for anticipating and responding to decision-making. The model addresses this challenge by simulating different collapse scenarios and storing the resulting data in a comprehensive database for use in a model classifier. As a result, the model accuracy performance shows the reliability and effectiveness of the classification.

Moreover, to enhance the robustness and reliability of the decision model, it is important to explore additional collapse parameters, such as seismic activity and collapse precursor signs. Including these parameters can lead to a more detailed and accurate model that can handle a wider range of cases and give better predictions. A stronger grasp of these factors will help build a full and effective early warning system. For example, the seismic activity affects the collapse timing, and knowing this link can sharpen the model ability to predict. Furthermore, it is also important to continue developing models and frameworks that integrate technical early warning systems with robust decision-making processes. The current model is capable of classifying the warning level with high accuracy. This ability could be integrated into the decision-making process to provide a more effective and timely early warning system for volcanic tsunamis.

Additionally, improving the performance in predicting the "No Warning" class is also an important area for model development. To cope with this situation, it is essential to balance the dataset by adding more scenarios that simulate very small volume collapses, which may trigger the "No Warning" classification in gauge recordings. Training the model with a more balanced dataset, can strengthen the ability to accurately "No Warning" class. Balancing the dataset prevents bias toward frequent tsunami events, enabling the model to effectively identify all potential scenarios. Moreover, training the model capabilities to handle outlier volumes, an exceptionally, large or small collapse, is very important. Outliers can significantly impact the model predictive accuracy. Ensuring the model can effectively account for these extremes in all collapse possibilities.

In summary, this sub chapter highlights the contribution of machine learning algorithm application, specifically the Decision Tree, in classifying tsunami early warning levels based on non-seismic tsunami or volcanic tsunamis. Furthermore, it also advances by integrating numerical simulations with machine learning to improve real-time prediction and classification of tsunami warning levels. Although this research primarily focuses on the case study of a tsunami triggered by the Anak Krakatau volcano, its implications extend beyond this specific event. Anak Krakatau is only one of several key volcanoes in Southeast Asia identified as potential sources of subaerial volcanic tsunami hazards, as highlighted by (Zorn et al., 2022).

Other significant volcanoes include Ritter Island in Papua New Guinea, Lliwerung in Indonesia, and Didicas in the Philippines. The methodology can be adapted to other volcanic settings, providing a foundation for enhancing tsunami early warning systems in regions with similar risks of collapse-induced tsunamis.

6.4 Summary

This chapter explored the application of machine learning and deep learning models for tsunami early warning, specifically in the context of volcanic tsunamis triggered by the collapse of Anak Krakatau. Building upon the synthetic scenario database developed in the previous chapter, various models were implemented to predict maximum tsunami amplitudes and waveform patterns using limited early observations from synthetic tide gauge stations.

In the first part, three machine learning algorithms—Random Forest, Gradient Boosting, and Neural Network—were applied to predict the highest tsunami amplitudes at coastal locations. Among these, Random Forest demonstrated the most consistent and accurate performance across multiple observation points, highlighting its reliability for non-seismic tsunami warning applications.

In the second part, deep learning architectures including Vanilla LSTM and Bidirectional LSTM were employed to forecast complete tsunami waveforms using only the first few minutes of observational data. BiLSTM outperformed the standard LSTM, particularly in regions with complex wave behavior, confirming the potential of sequence-based models for time series prediction in early warning systems.

The final section introduced a Decision Tree classification model designed to assign early warning levels based on simulated inputs from volcanic collapse scenarios. The model demonstrated high classification accuracy and provided a practical mechanism for non-seismic tsunami warning based on precomputed database inputs.

Overall, this chapter demonstrated how machine learning models, when integrated with physical simulations, can contribute to the development of rapid and reliable tsunami early warning systems for volcanic events. These efforts complement the scenario generation process in Chapter 5 and are grounded in the conceptual understanding introduced in Chapter 4.

To reinforce the key findings presented in this chapter, Table 6.13 summarizes the main results and insights across the three modelling approaches.

Table 6.13 Summary of Key Results

Aspect	Summary of Findings
Prediction Target	Maximum tsunami amplitude at coastal stations and full waveform at synthetic stations
Scenario Input	320 scenarios for amplitude prediction; 1,000 scenarios for waveform forecasting
Best Models (Max Amplitude)	Random Forest at 3 of 4 stations (Ciwandan, Panjang, Kota Agung); Gradient Boosting at Marina Jambu
Best Models (Waveform Forecasting)	Bidirectional LSTM consistently outperformed Vanilla LSTM across most synthetic stations
Input Duration	9 minutes (for max amplitude); 3 minutes (for waveform forecasting)
Performance Metrics	RMSE, MAE, and MSE used for model evaluation across stations

Model Limitations	Complex LSTM showed poorer performance on short input; errors increase for distant stations
Practical Use	Results integrated into real-time early warning system with classified warning levels
Risk Framework	Warning thresholds (No Warning, Minor, Tsunami, Major) set per station and used for rapid decision-making
Transferability	Methodology applicable to other volcanoes (e.g., Ritter Island, Didicas, Lliwerung)

Chapter 7: Discussion and Conclusion

This chapter synthesizes the key findings of the study and evaluates the implications of the proposed forecasting framework for High-Impact, Low-Probability (HILP) disasters, with a particular emphasis on volcanic tsunamis. The research presents a comprehensive integration of simulation-based data generation, machine learning prediction, and warning classification to overcome the limitations of traditional seismic-based early warning systems. Using the 2018 Anak Krakatau tsunami as a case study, this study explored a new path to predict tsunami waveform using synthetic observations and data-driven models. The discussion below unifies critical insights, addresses methodological strengths and limitations, and outlines future directions.

7.1 Main contributions and novelty

- The main contribution of this research lies in addressing a critical gap in tsunami forecasting by introducing a forecasting pipeline specifically designed for non-seismic HILP events. Unlike conventional tsunami databases that are predominantly generated from earthquake-induced scenarios, this study focuses on events triggered by mechanisms such as volcanic flank collapse (Table 7.1). These events require a different approach due to their localized impact zones, highly variable source characteristics, and lack of standardized collapse direction-based parameters.
- Although the proposed structure may resemble Japan advanced tsunami warning systems, the focus, modelling assumptions, and source mechanisms are fundamentally different. Japan system is designed primarily for tectonic earthquakes, whereas this study extends early warning capabilities to volcanic tsunami events that occur without strong seismic precursors (see “Reference early warning system” in the Aspect column in the Table 7.1 below).
- Moreover, a key novelty of this study is the construction of a scenario-based synthetic database, consisting of 1,000 systematically generated non-seismic tsunami simulations. By varying critical parameters such as collapse volume, geometry, slope angle, and directionality, the database captures the complexity and uncertainty inherent in HILP events, especially volcanic tsunamis. This forms the foundation for subsequent predictive modelling and analysis.

Table 7. 1 Summary of Contributions and Novelty of This Study

Aspect	Conventional approach (earthquake-induced)	This study novel contribution (volcanic collapse-induced)
Hazard source	Earthquake-induced tsunamis	Focus on volcanic tsunamis (volcanic flank collapse and landslide)
Scenario parameterization	Based on moment magnitude, fault length, rupture depth, slip	Based on collapse volume, geometry, dip angle, direction, location-specific settings

Database generation	Precomputed scenarios from tectonic fault rupture using uniform grid setups	Synthetic database of 1,000 scenarios generated using varied landslide/collapse parameters
Simulation model	Often use seismic tsunami models with predefined rupture models	Uses COMCOT with nonlinear shallow water equations tailored for landslide-type sources
Forecasting input	Earthquake origin time, magnitude, and location	Short-term waveform input (first 3 minutes) for real-time prediction using machine learning
Machine learning application	Limited, mostly on seismic datasets	Novel machine learning models trained on synthetic non-seismic waveform data for early warning
Warning system implication	Optimized for common tsunamis in seismic-prone zones	Expands early warning capability to rare but high-impact volcanic tsunamis
Target area	Regions with seismic history and dense instrumentation	Include areas with limited seismic activity but high-impact volcanic tsunamis hazard potential
Reference early warning system	Based on Japan earthquake monitoring and fault rupture model, optimized for tectonic sources	New framework targets non-seismic tsunami events (volcanic flank collapse) that lack seismic precursors and cannot be captured by conventional fault-based like Japan, by using dynamic slope collapse parameters such as volume, geometry, and direction

Additional technical and practical strengths of the proposed framework, including short-term forecasting, model scalability, and real-time warning implementation, are elaborated in Section 7.2.

Regarding contribution to Knowledge Science, this study at least give three contributions, namely:

- Identification of Specific Feature of HILP as Natural Phenomena

Through a comprehensive review and analysis of historical HILP disasters, the study reveals specific feature and significant parameters that inform the contextual and modelling basis for forecasting. This supports knowledge creation in underexplored high-risk domains.

- Integrated Novel Forecasting Framework for HILP Events

Develops a novel framework that integrates physical simulations and machine learning models. It demonstrates knowledge integration across geophysics and disaster science to address the complexity and uncertainty of non-seismic tsunami events.

- Data-Driven Utilization for Early Warning

By leveraging short-term tsunami waveform observations, the model enables rapid and localized early warning outputs. This supports near-field tsunami forecasting and enhances decision-making, contributing to disaster resilience through the effective utilization of actionable knowledge.

7.2 Framework strengths

Several unique strengths of the proposed framework in this study include:

- **End-to-End Pipeline Integration**: The framework stands out for its comprehensive, end-to-end design, encompassing conceptualization, physical simulation, data generation, machine learning-based forecasting, and classification. This seamless integration contrasts with fragmented approaches in the literature and provides a replicable and modular structure applicable to other disaster contexts. The ability to handle the full modelling cycle enhances usability in both academic and operational domains.
- **Scenario-Based Synthetic Database**: A key strength is the creation of a synthetic database consisting of 1000 diverse landslide-induced tsunami scenarios. These scenarios were systematically generated by varying critical parameters such as collapse volume, geometry, dip angle, and direction. This extensive dataset allows the model to learn from a wide range of plausible events, effectively capturing the uncertainty and complexity of HILP (High Impact, Low Probability) volcanic tsunami events.
- **Short Observation Input Forecasting**: The model ability to forecast the complete tsunami waveform using only 3 to 9 minutes of initial observation data represents a major advancement in early warning capability. This feature enables rapid and timely forecasts that are critical for emergency response, particularly in near-field tsunami situations where lead times are extremely limited. It also supports practical deployment in real-time early warning systems.
- **Model Scalability and Adaptability**: Although the framework was developed based on the Anak Krakatau case study, it is highly adaptable and can be applied to other geophysical settings. Its modular design supports implementation in various collapse-prone offshore volcanoes and enclosed basins, including non-seismic tsunami scenarios in lake environments. Locations such as Ritter Island (Papua New Guinea), Liliwerung (Indonesia), and Didicas (Philippines) serve as potential targets for future deployment, supporting regional hazard assessment and preparedness.
- **Multi-Functional Use Case**: Beyond waveform prediction, the framework supports additional applications such as scenario classification, sensitivity analysis, and data-driven decision support. This flexibility opens possibilities for coupling with risk maps, vulnerability indices, and evacuation planning tools, expanding its use in interdisciplinary disaster resilience initiatives.
- **Reproducibility and Open Science Potential**: The systematic design and documentation of simulation parameters, modelling procedures, and learning architectures contribute to reproducibility. This foundation supports future scientific collaboration, benchmarking, and continuous development under open science principles.

7.3 Scientific and practical implications

This section emphasizes the broader implications of the research, with a focus on how the proposed framework can influence scientific advancement and practical disaster management.

From a scientific perspective, the integration of physical numerical modelling and machine learning within this study marks a significant advancement in tsunami forecasting methodologies. By combining high-fidelity tsunami simulation (e.g., landslide-induced wave generation) with data-driven prediction models, this research provides a hybrid framework that bridges the gap between physics-based understanding and AI applications. This is particularly relevant for near-field volcanic tsunamis triggered by sector collapse, a relatively underexplored domain in tsunami engineering.

The study characterizes and analyses thousands of synthetic waveforms generated near the volcano source. These waveforms capture a wide range of behaviours, enabling a better understanding of near-field tsunami dynamics, including arrival patterns and multi-peak structures correlations. Such insights can serve as a foundation for future work in defining early warning parameters specific to volcanic collapse events, especially in geologically similar environments.

On the practical side, this study offers actionable knowledge for disaster preparedness, particularly regarding the strategic deployment of observation instruments. By placing virtual or synthetic tide gauges near the volcano in the simulation setup, the study provides an operational reference for evaluating the effectiveness and feasibility of installing real tide gauges in such hazardous, offshore locations. This approach allows authorities to assess the cost-benefit and safety trade-offs in monitoring near-source waveforms, which are crucial for early warning in high-risk volcanic islands.

Its short-term input prediction capability also provides critical lead time that can be harnessed for issuing public alerts and evacuation decisions. The compatibility with standard disaster warning levels (e.g., BMKG, BNPB) facilitates smoother integration with national systems and supports rapid, spatially targeted responses.

7.4 Discussion of limitations

Despite its promising results, this research acknowledges several limitations that warrant further investigation. First, the predictive model developed in this study relies exclusively on time-series waveform data, without integrating 2D or 3D hydrodynamic inundation models. While the model can estimate offshore wave amplitudes and arrival times with considerable accuracy, it lacks the capability to simulate flood extent, inland water penetration, and infrastructure impact onshore. This limitation arises due to the software ability of the tsunami simulation tool used, which does not support full inundation dynamics. Addressing this shortcoming through the integration of hydrodynamic solvers in future work would enable more comprehensive impact assessments and support decision-makers in evacuation planning and coastal zoning.

Second, one inherent limitation of the current collapse or landslide simulation software is the absence of explicit consideration for the physical properties of landslide materials, such as density, viscosity, or rheology. The COMCOT model, while capable of simulating tsunami generation due to submarine mass movement, represents landslides primarily as time-dependent bathymetric deformations without incorporating the mechanical characteristics of the sliding material. This simplification allows for computational efficiency and flexible

scenario generation but does not capture the complex dynamics associated with different types of mass failure, such as cohesive rockslides, granular debris flows, or muddy slumps. As a result, the current simulation model assumes that the tsunami-generating potential of a landslide is sufficiently described by the geometric and kinematic parameters of the seafloor deformation, namely, volume, slope angle, direction, and duration, while treating the material as a homogenous moving mass. This limitation may affect the accuracy of modelled wave generation in cases where the landslide material properties significantly influence the energy transfer between the sliding mass and the overlying water column.

Third, while the framework demonstrates strong potential for transferability, its current implementation remains limited to volcanic tsunami scenarios. The use of synthetic data generation, feature engineering, and supervised learning techniques has yet to be validated across different hazard types. Future studies could explore the application of this framework to other high-impact low-probability (HILP) events such as glacial lake outburst floods (GLOFs), storm surges, and compound or cascading events, including NaTech (natural-technological) disasters. Each of these scenarios may require domain-specific input features, temporal resolution adjustments, and hazard-specific calibration.

Forth, there exists a significant challenge in the class imbalance problem encountered during classification tasks. The “No Warning” or “Non-Tsunami” class in the training data was underrepresented, leading to reduced classification accuracy, particularly an increase in false negatives. This imbalance potentially undermines the reliability of early warning outputs, especially when minor flank collapses that do not generate tsunamis are incorrectly flagged. Future iterations of this model should incorporate advanced data balancing strategies, such as SMOTE (Synthetic Minority Over-sampling Technique) or cost-sensitive learning. Nevertheless, caution must be exercised when implementing such models in real-world contexts, for example at Anak Krakatau in Indonesia, where underestimating a potential tsunami could lead to severe consequences. Ensuring conservative and safety-oriented application is therefore essential.

Fifth, one limitation of the proposed framework, as previously discussed in Chapter 6, is the noticeable decrease in model performance when using only the first 9 minutes of waveform data as input. While this setup was designed to simulate the constraints of real-time early warning, particularly in near-coast tsunami scenarios, it revealed a significant trade-off between forecasting urgency and predictive accuracy. This limitation highlights the challenge of producing reliable predictions within very short observation windows (e.g., less than 3 minutes), which are often encountered in high-impact, fast-onset events. Future work should address this issue by incorporating more advanced forecasting algorithms, such as encoder-decoder architectures, attention-based models like Transformer, or hybrid CNN-LSTM frameworks that have shown improved performance in time series prediction tasks. Additionally, expanding the training dataset could involve synthesizing more diverse scenario variations (e.g., different collapse angles, volumes, and directions), and integrating historical real-event data or noise-augmented simulations to improve model generalizability and robustness.

Lastly, the framework assumes idealized boundary conditions and does not currently account for real-time environmental factors such as tide levels, bathymetric noise, or observational errors in tide gauges. These practical uncertainties could influence waveform predictions in real deployment scenarios. Introducing probabilistic modelling or uncertainty quantification techniques would allow for better risk-informed decision-making under real-world constraints.

7.5 Recommendations for future work

To expand and strengthen the proposed framework, future research should consider the following directions:

1. Incorporating Landslide Material Properties
   Future work may benefit from coupling COMCOT with more advanced rheological models or adopting multiphase or solid-fluid interaction frameworks, which allow for more realistic representation of diverse landslide materials. Integrating such models could improve the physical realism of tsunami generation simulations, especially for case studies involving highly variable or poorly constrained material types.

2. Integration with Inundation Modelling
   The current study focuses on one-dimensional (1D) offshore tsunami waveform prediction at predefined tide gauge locations. This 1D approach emphasizes temporal prediction at specific points offshore, without resolving spatial variations along the coastline. While it serves as a valuable early warning component, it does not yet simulate inundation or interaction with land-based topography. As part of future work, the integration of this 1D prediction framework with two-dimensional (2D) or three-dimensional (3D) inundation models is highly recommended. This coupling would enable a more comprehensive assessment of coastal impact by modelling how predicted waveforms propagate over land and interact with complex coastal features, including terrain, infrastructure, and population exposure. Such integration is essential for estimating spatial inundation extents, water depths, and arrival times, thereby enhancing the applicability of the model for evacuation planning and emergency response.

3. Incorporate Real-Time Data Streams
   Enhance the operational capability of the early warning framework by incorporating real-time data sources, including tide gauge sensors, satellites, GPS-based displacement data, and bottom pressure recorders. Integrating these live data streams would enable real-time model updating through techniques such as data assimilation or streaming input preprocessing. This could facilitate near-instantaneous decision-making, improving forecast reliability and reducing latency between tsunami generation and public alert issuance.

4. Advance Deep Learning Techniques and Metrics Evaluation
   Future work should explore the application of more advanced neural architectures beyond LSTM and BiLSTM. Transformer-based models or hybrid architectures that

combine convolutional and recurrent layers may improve the model's ability to capture long-range temporal dependencies from limited observation inputs. Moreover, incorporating attention mechanisms could enable the model to focus on critical waveform segments, thereby further enhancing prediction accuracy. Such developments may support reliable forecasting with shorter input windows, which is crucial for real-time applications. In addition, the selection of evaluation metrics should be optimized; reporting both RMSE and MSE is redundant, and only the most appropriate metric should be used.

5. Multi-Hazard Extension
   The modular structure of the proposed framework makes it adaptable for forecasting other types of High-Impact, Low-Probability (HILP) events beyond volcanic tsunamis. Modular structure means that the framework is designed in separate per stage, flexible and customizable components. Future studies could extend its applicability to events such as earthquake-triggered tsunamis, cascading infrastructure failures, or cyber-physical disruptions. Embedding such flexibility will promote broader disaster resilience applications and facilitate integrated risk assessment tools that can operate under various hazard scenarios in multi-risk environments.

6. Broader Implementation in Diverse Volcanic Settings
   To extend the real-world impact of the proposed framework, future implementation should focus not only on volcanic systems located in offshore islands or lake environments with known flank collapse potential, such as Ritter Island in Papua New Guinea, Lliwerung in Indonesia, or Didicas in the Philippines, but also explore its adaptability to other high-impact, low-probability (HILP) disaster scenarios. These include submarine landslides, glacial calving-induced tsunamis, or meteorological-driven coastal flood surges.
   The framework can be adapted to these different disaster types by modifying the input parameters used in scenario generation (e.g., replacing volcanic collapse volume and slope angle with landslide mass and run-out velocity, or meteorological forcing variables), and by incorporating alternative datasets such as seafloor sediment profiles, historical flood extents, or ice-sheet deformation models.
   The machine learning component of the model remains flexible and can be retrained using synthetic datasets relevant to the target hazard, ensuring predictive accuracy across domains. Scenario-specific physical modelling can also be tailored using adjusted source mechanisms and propagation models appropriate to each hazard type.
   To further enhance its applicability in data-sparse environments, the proposed framework can continue to operate using generalized or low-resolution topographic and bathymetric inputs, especially when high-resolution local datasets are unavailable. Pre-computed synthetic scenario databases, developed based on plausible geophysical assumptions and regional analogs, can serve as a practical alternative to real-time data streams. By combining these generalized models with early waveform observations from even a single tide gauge or low-cost sensor, the classifier can still identify likely scenarios and issue timely warnings. This adaptability enables the deployment of the

framework in under-resourced regions, making it a viable solution for strengthening early warning capabilities across diverse hazard settings.

7. Incorporating Extreme Value Evaluation Metrics
   While the current study employs Mean Squared Error (MSE) as a loss function due to its widespread use and optimization convenience, it is acknowledged that MSE may not be fully appropriate for the primary objective of HILP, which correlate with prediction of extreme values, as peak amplitudes. Future research should investigate the use of alternative or complementary loss functions and evaluation metrics that emphasize peak errors or high-impact wave characteristics.

References

Abdurrachman, M., Widiyantoro, S., Priadi, B., & Ismail, T. (2018). Geochemistry and structure of krakatoa volcano in the Sunda Strait, Indonesia. *Geosciences (Switzerland)*, *8*(4). https://doi.org/10.3390/geosciences8040111

Abraham, K., Abdelwahab, M., & Abo-Zahhad, M. (2024). Classification and detection of natural disasters using machine learning and deep learning techniques: A review. In *Earth Science Informatics* (Vol. 17, Issue 2, pp. 869–891). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s12145-023-01205-2

Ahmed, A. (2021). Urgent call for a global enforcement of the public sharing of health emergencies data: Lesson learned from serious arboviral disease epidemics in Sudan. *International Health*, *12*(4), 238–240. https://doi.org/10.1093/INTHEALTH/IHZ122

Ambraseys, N., & Synolakis, C. (2010). Tsunami catalogs for the eastern Mediterranean. *Journal of Earthquake Engineering*, *14*(3), 309–330. https://doi.org/10.1080/13632460903277593

Analytica, O. (2024). High Impact, Low Probability events: Why the improbable can't be ignored. In *Oxford Analytica Insight*. https://www.linkedin.com/pulse/high-impact-low-probability-events-why-improbable-cant-ikone/

Angove, M., Arcas, D., Bailey, R., Carrasco, P., Coetzee, D., Fry, B., Gledhill, K., Harada, S., von Hillebrandt-Andrade, C., Kong, L., McCreery, C., McCurrach, S. J., Miao, Y., Sakya, A. E., & Schindelé, F. (2019). Ocean observations required to minimize uncertainty in global tsunami forecasts, warnings, and emergency response. In *Frontiers in Marine Science* (Vol. 6, Issue JUN). Frontiers Media S.A. https://doi.org/10.3389/fmars.2019.00350

Annunziato, A. (2022). Tsunami Detection Model for Sea Level Measurement Devices. *Geosciences (Switzerland)*, *12*(10). https://doi.org/10.3390/geosciences12100386

Annunziato, A., Prasetya, G., & Husrin, S. (2019). Science of Tsunami Hazards Anak Krakatau Volcano Emergency Tsunami Early Warning System. *Journal of Tsunami Society International*. http://tides.big.go.id:8888/dash/prov/Banten.html

Apotsos, A., Jaffe, B. and Gelfenbaum, G. (2011) 'Wave characteristic and morphologic effects on the onshore hydrodynamic response of tsunamis', *Coastal Engineering*, 58(11), pp. 1034–1048. Available at: https://doi.org/10.1016/j.coastaleng.2011.06.002.

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., & Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. *Mechanical Systems and Signal Processing*, *147*, 107077. https://doi.org/10.1016/j.ymssp.2020.107077

Aven, T. (2003). Foundations of Risk Analysis Foundations of Risk Analysis. In *A Knowledge*

and Decision- oriented Prespective (Vol. 1).

Aven, T. (2013). On the meaning of a black swan in a risk context. *Safety Science*, *57*, 44–51. https://doi.org/10.1016/j.ssci.2013.01.016

Baumert, N., Mück, M., & Muhari, A. (2018). *Guideline for Tsunami Risk Assessment in Indonesia Scientific Indonesian-German Working Group on Tsunami Risk Assessment*. https://doi.org/10.13140/RG.2.2.20247.11685

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. In *Frontiers in Earth Science* (Vol. 9). Frontiers Media S.A. https://doi.org/10.3389/feart.2021.628772

Ben, C. (2024, August 20). Tornado-like waterspout: 'black swan' event that sank the Bayesian. *The Times*. https://www.thetimes.com/uk/science/article/how-climate-change-is-causing-lethal-waterspouts-in-the-mediterranean-0k2vt7r6j?utm_source=chatgpt.com&region=global

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

Blaser, L., Ohrnberger, M., Riggelsen, C., Babeyko, A., & Scherbaum, F. (2011). Bayesian networks for tsunami early warning. *Geophysical Journal International*, *185*(3), 1431–1443. https://doi.org/10.1111/j.1365-246X.2011.05020.x

Borrero, J. C., Solihuddin, T., Fritz, H. M., Lynett, P. J., Prasetya, G. S., Skanavis, V., Husrin, S., Kushendratno, Kongko, W., Istiyanto, D. C., Daulat, A., Purbani, D., Salim, H. L., Hidayat, R., Asvaliantina, V., Usman, M., Kodijat, A., Son, S., & Synolakis, C. E. (2020). Field Survey and Numerical Modelling of the December 22, 2018 Anak Krakatau Tsunami. *Pure and Applied Geophysics*, *177*(6), 2457–2475. https://doi.org/10.1007/s00024-020-02515-y

Brown, S. K., Jenkins, S. F., Sparks, R. S. J., Odbert, H., & Auker, M. R. (2017). Volcanic fatalities database: analysis of volcanic threat with distance and victim classification. *Journal of Applied Volcanology*, *6*(1). https://doi.org/10.1186/s13617-017-0067-4

Calvari, S., Di Traglia, F., Ganci, G., Bruno, V., Ciancitto, F., Di Lieto, B., Gambino, S., Garcia, A., Giudicepietro, F., Inguaggiato, S., Vita, F., Cangemi, M., Inguaggiato, C., Macedonio, G., Mattia, M., Miraglia, L., Nolesini, T., Pompilio, M., Romano, P., … Belviso, P. (2022). Multi-parametric study of an eruptive phase comprising unrest, major explosions, crater failure, pyroclastic density currents and lava flows: Stromboli volcano, 1 December 2020–30 June 2021. *Frontiers in Earth Science*, *10*. https://doi.org/10.3389/feart.2022.899635

Carrara, A., Guzzetti, F., Cardinali, M., & Reichenbach, P. (1999). Use of GIS technology in

the prediction and monitoring of landslide hazard. *Natural Hazards*, *20*(2–3), 117–135. https://doi.org/10.1023/a:1008097111310

Cesario, E., Giampá, S., Baglione, E., Cordrie, L., Selva, J., & Talia, D. (2024). Machine Learning for Tsunami Waves Forecasting Using Regression Trees. *Big Data Research*, *36*. https://doi.org/10.1016/j.bdr.2024.100452

Chae, M., Gil, H. W., Cho, N. J., & Lee, H. (2022). Machine Learning-Based Cardiac Arrest Prediction for Early Warning System. *Mathematics*, *10*(12). https://doi.org/10.3390/math10122049

Chamola, V., Hassija, V., Gupta, S., Goyal, A., Guizani, M., & Sikdar, B. (2021). Disaster and Pandemic Management Using Machine Learning: A Survey. *IEEE Internet of Things Journal*, *8*(21), 16047–16071. https://doi.org/10.1109/JIOT.2020.3044966

Cheemarla, N. R., Hanron, A., Fauver, J. R., Bishai, J., Watkins, T. A., Brito, A. F., Zhao, D., Alpert, T., Vogels, C. B. F., Ko, A. I., Schulz, W. L., Landry, M. L., Grubaugh, N. D., van Dijk, D., & Foxman, E. F. (2023). Nasal host response-based screening for undiagnosed respiratory viruses: a pathogen surveillance and detection study. *The Lancet Microbe*, *4*(1), e38–e46. https://doi.org/10.1016/S2666-5247(22)00296-8

Cheung, K.F. *et al.* (2011) 'Modeling of 500-year tsunamis for probabilistic design of coastal infrastructure in the Pacific Northwest', *Coastal Engineering*, 58(10), pp. 970–985. Available at: https://doi.org/10.1016/j.coastaleng.2011.05.003.

Comes, T., Mayag, B., & Negre, E. (2014). Decision Support for Disaster Risk Management: Integrating Vulnerabilities into Early-Warning Systems. *Information Systems for Crisis Response and Management in Mediterranean Countries*.

Cornell, C. A., & Newmark, N. M. (1978). Seismic reliability of nuclear power plants. *Probabilistic Analysis of Nuclear Reactor Safety*, XIV.1.1-XIV.1.14. https://inis.iaea.org/records/dmacn-y9106

Corradino, C., Ramsey, M. S., Pailot-Bonnetat, S., Harris, A. J. L., & Negro, C. Del. (2023). Detection of Subtle Thermal Anomalies: Deep Learning Applied to the ASTER Global Volcano Dataset. *IEEE Transactions on Geoscience and Remote Sensing*, *61*. https://doi.org/10.1109/TGRS.2023.3241085

Creppage, K. E., Gallaway, M. S., Russell, D. A., Early, J. M., Smith, H. J., Mooney, A. C., Hydrick, A. M., & Kasper, M. R. (2024). Global Emerging Infections Surveillance Program Contributions to Pandemic Preparedness and Response. *Emerging Infectious Diseases*, *30*(14), 9–12. https://doi.org/10.3201/eid3014.240305

Cutler, K. S., Watt, S. F. L., Cassidy, M., Madden-Nadeau, A. L., Engwell, S. L., Abdurrachman, M., Nurshal, M. E. M., Tappin, D. R., Carey, S. N., Novellino, A., Hayer, C., Hunt, J. E., Day, S. J., Grilli, S. T., Kurniawan, I. A., & Kartadinata, N. (2022). Downward-propagating eruption following vent unloading implies no direct magmatic trigger for the 2018 lateral collapse of Anak Krakatau. *Earth and Planetary Science Letters*, *578*, 117332. https://doi.org/10.1016/J.EPSL.2021.117332

Dang, H. Van *et al.* (2023) 'Development of the global tsunami forecasting system considering the dynamic interaction of tide-tsunami around the Korean Peninsula', *Frontiers in Marine Science*, 10(November), pp. 1–24. Available at: https://doi.org/10.3389/fmars.2023.1258552.

Day, S. J. (2015). Volcanic Tsunamis. In *The Encyclopedia of Volcanoes* (Second Edition, pp. 993–1009). Elsevier.

De Angeli, S., Malamud, B. D., Rossi, L., Taylor, F. E., Trasforini, E., & Rudari, R. (2022). A multi-hazard framework for spatial-temporal impact analysis. *International Journal of Disaster Risk Reduction*, *73*(February), 102829. https://doi.org/10.1016/j.ijdrr.2022.102829

Deplus, C., Bonvalotb, S., Dahrin, D., Diament, M., Harjonoc, H., & Dubois, J. (1995). Inner structure of the Krakatau volcanic complex (Indonesia) from gravity and bathymetry data. In *Journal of Volcanology and Geothermal Research* (Vol. 64).

Dharmawan, W., Diana, M., Tuntari, B., Astawa, I. M., Rahardjo, S., & Nambo, H. (2023). Tsunami tide prediction in shallow water using recurrent neural networks: model implementation in the Indonesia Tsunami Early Warning System. *Journal of Reliable Intelligent Environments*. https://doi.org/10.1007/s40860-023-00214-8

Di Traglia, F., Calvari, S., Borselli, L., Cassanego, L., Giudicepietro, F., Macedonio, G., Nolesini, T., & Casagli, N. (2023). Assessing flank instability of Stromboli volcano (Italy) by reappraising the 30 December 2002 tsunamigenic landslides. *Landslides*, *20*(7), 1363–1380. https://doi.org/10.1007/s10346-023-02043-5

Djalante, R., Lassa, J., Setiamarga, D., Sudjatma, A., Indrawan, M., Haryanto, B., Mahfud, C., Sinapoy, M. S., Djalante, S., Rafliana, I., Gunawan, L. A., Surtiari, G. A. K., & Warsilah, H. (2020). Review and analysis of current responses to COVID-19 in Indonesia: Period of January to March 2020. *Progress in Disaster Science*, *6*. https://doi.org/10.1016/j.pdisas.2020.100091

Dogan, G. G., Annunziato, A., Hidayat, R., Husrin, S., Prasetya, G., Kongko, W., Zaytsev, A., Pelinovsky, E., Imamura, F., & Yalciner, A. C. (2021). Numerical Simulations of December 22, 2018 Anak Krakatau Tsunami and Examination of Possible Submarine Landslide Scenarios. *Pure and Applied Geophysics*, *178*(1), 1–20. https://doi.org/10.1007/s00024-020-02641-7

Du, Y., Ding, Y., Li, Z., & Cao, G. (2015). The role of hazard vulnerability assessments in disaster preparedness and prevention in China. *Military Medical Research*, *2*(1), 1–7. https://doi.org/10.1186/s40779-015-0059-9

Ellingwood, B., M., J. G., MacGregor, F., T., V. Galambos, F., & C.Allin Cornell, M. (1982). Probability Based Load Criteria: Load Factors and Load Combinations. *ASCE Journal of the Structural Division*, *108*(5). https://doi.org/doi/10.1061/JSDEAG.0005959

Fauzi, A., & Mizutani, N. (2020). Machine Learning Algorithms for Real-time Tsunami Inundation Forecasting: A Case Study in Nankai Region. *Pure and Applied Geophysics*,

177(3), 1437–1450. https://doi.org/10.1007/s00024-019-02364-4

Filip, R., Gheorghita Puscaselu, R., Anchidin-Norocel, L., Dimian, M., & Savage, W. K. (2022). Global Challenges to Public Health Care Systems during the COVID-19 Pandemic: A Review of Pandemic Measures and Problems. *Journal of Personalized Medicine*, *12*(8). https://doi.org/10.3390/jpm12081295

Francis, R., & Colli, A. (2014). Information-based reliability weighting for failure mode prioritization in photovoltaic (PV) module design. *PSAM 2014 - Probabilistic Safety Assessment and Management*, *November*.

Frankel, J., & Saravelos, G. (2012). Can leading indicators assess country vulnerability? Evidence from the 2008-09 global financial crisis. *Journal of International Economics*, *87*(2), 216–231. https://doi.org/10.1016/j.jinteco.2011.12.009

Fukurai, M., & Ludwig, L. G. (2025). Risk perception to preparedness: 2011 pre- and post-great Tōhoku Japan earthquake. *International Journal of Disaster Risk Reduction*, *116*(November 2024), 105074. https://doi.org/10.1016/j.ijdrr.2024.105074

Funabashi, Y., & Kitazawa, K. (2012). Fukushima in review: A complex disaster, a disastrous response. *Bulletin of the Atomic Scientists*, *68*(2), 9–21. https://doi.org/10.1177/0096340212440359

Gao, C., Lan, X., Li, N., Yuan, Y., Ding, J., Zhou, Z., Xu, F., & Li, Y. (2024). Large language models empowered agent-based modeling and simulation: a survey and perspectives. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-03611-3

Giachetti, T., Paris, R., Kelfoun, K., & Ontowirjo, B. (2012). Tsunami hazard related to a flank collapse of Anak Krakatau Volcano, Sunda Strait, Indonesia. *Geological Society Special Publication*, *361*(1), 79–90. https://doi.org/10.1144/SP361.7

Giordano, G., & De Astis, G. (2021). The summer 2019 basaltic Vulcanian eruptions (paroxysms) of Stromboli. *Bulletin of Volcanology*, *83*(1). https://doi.org/10.1007/s00445-020-01423-2

Gusman, A.R. *et al.* (2019) 'Source model for the tsunami inside palu bay following the 2018 palu earthquake, Indonesia', *Geophysical Research Letters*, 46(15), pp. 8721–8730. Available at: https://doi.org/10.1029/2019GL082717.

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. In *Reviews of Geophysics* (Vol. 55, Issue 4, pp. 1158–1198). Blackwell Publishing Ltd. https://doi.org/10.1002/2017RG000579

Grilli, S. T., Grosdidier, S., & Guérin, C. A. (2016). Tsunami Detection by High-Frequency Radar Beyond the Continental Shelf: I: Algorithms and Validation on Idealized Case

Studies. *Pure and Applied Geophysics*, *173*(12), 3895–3934. https://doi.org/10.1007/s00024-015-1193-8

Grilli, S. T., Guérin, C. A., Shelby, M., Grilli, A. R., Moran, P., Grosdidier, S., & Insua, T. L. (2017). Tsunami Detection by High Frequency Radar Beyond the Continental Shelf: II. Extension of Time Correlation Algorithm and Validation on Realistic Case Studies. *Pure and Applied Geophysics*, *174*(8), 3003–3028. https://doi.org/10.1007/s00024-017-1619-6

Grilli, S. T., Tappin, D. R., Carey, S., Watt, S. F. L., Ward, S. N., Grilli, A. R., Engwell, S. L., Zhang, C., Kirby, J. T., Schambach, L., & Muin, M. (2019). Modelling of the tsunami from the December 22, 2018 lateral collapse of Anak Krakatau volcano in the Sunda Straits, Indonesia. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-48327-6

Grilli, S. T., Zhang, C., Kirby, J. T., Grilli, A. R., Tappin, D. R., Watt, S. F. L., Hunt, J. E., Novellino, A., Engwell, S., Nurshal, M. E. M., Abdurrachman, M., Cassidy, M., Madden-Nadeau, A. L., & Day, S. (2021). Modeling of the Dec. 22nd 2018 Anak Krakatau volcano lateral collapse and tsunami based on recent field surveys: Comparison with observed tsunami impact. *Marine Geology*, *440*, 106566. https://doi.org/10.1016/J.MARGEO.2021.106566

Guérin, C. A., Grilli, S. T., Moran, P., Grilli, A. R., & Insua, T. L. (2018). Tsunami detection by high-frequency radar in British Columbia: performance assessment of the time-correlation algorithm for synthetic and real events. *Ocean Dynamics*, *68*(4–5), 423–438. https://doi.org/10.1007/s10236-018-1139-7

Gulghane, A., Sharma, R. L., & Borkar, P. (2024). A formal evaluation of KNN and decision tree algorithms for waste generation prediction in residential projects: a comparative approach. *Asian Journal of Civil Engineering*, *25*(1), 265–280. https://doi.org/10.1007/s42107-023-00772-5

Gusman, A. R., Tanioka, Y., Macinnes, B. T., & Tsushima, H. (2014). A methodology for near-field tsunami inundation forecasting: Application to the 2011 Tohoku tsunami. *Journal of Geophysical Research: Solid Earth*, *119*(11), 8186–8206. https://doi.org/10.1002/2014JB010958

Harig, S., Immerz, A., Weniza, Griffin, J., Weber, B., Babeyko, A., Rakowsky, N., Hartanto, D., Nurokhim, A., Handayani, T., & Weber, R. (2020). The Tsunami Scenario Database of the Indonesia Tsunami Early Warning System (InaTEWS): Evolution of the Coverage and the Involved Modeling Approaches. *Pure and Applied Geophysics*, *177*(3), 1379–1401. https://doi.org/10.1007/s00024-019-02305-1

Heidarzadeh, M., Ishibe, T., Sandanbata, O., Muhari, A., & Wijanarto, A. B. (2020). Numerical modeling of the subaerial landslide source of the 22 December 2018 Anak Krakatoa volcanic tsunami, Indonesia. *Ocean Engineering*, *195*, 106733. https://doi.org/10.1016/J.OCEANENG.2019.106733

Heidarzadeh, M., Putra, P. S., Nugroho, S. H., & Rashid, D. B. Z. (2020). Field Survey of Tsunami Heights and Runups Following the 22 December 2018 Anak Krakatau Volcano

Tsunami, Indonesia. *Pure and Applied Geophysics*, *177*(10), 4577–4595. https://doi.org/10.1007/s00024-020-02587-w

Helbing, D. (2013). Globally networked risks and how to respond. *Nature*, *497*(7447), 51–59. https://doi.org/10.1038/nature12047

Heller, V. and Ruffini, G. (2023) 'A critical review about generic subaerial landslide-tsunami experiments and options for a needed step change', *Earth-Science Reviews*, 242(May), p. 104459. Available at: https://doi.org/10.1016/j.earscirev.2023.104459.

Heo, K.Y. *et al.* (2019) 'Numerical modeling of meteotsunami-tide interaction in the eastern yellow sea', *Atmosphere*, 10(7). Available at: https://doi.org/10.3390/atmos10070369.

Hidayat, A., Marfai, M. A., & Hadmoko, D. S. (2020). Eruption on Indonesia's volcanic islands: A review of potential hazards, fatalities, and management. *IOP Conference Series: Earth and Environmental Science*, *485*(1). https://doi.org/10.1088/1755-1315/485/1/012061

Hochreiter, S. and Schmidhuber, J. (1997) 'Long Short Term Memory', *Neural Computation*, 9 (8).

International Atomic Energy Agency. (2005). Chernobyl's Legacy: Health, Environmental and Socio-Economic Impacts. *The Chernobyl Forum: 2003-2005*, 1–57.

Izumi, Y. (1981). A geophysical interpretation of the 1883 Krakatau eruption. *Journal of Volcanology and Geothermal Research*, *9*(4), 359–378.

Jain, H., Dhupper, R., Shrivastava, A., Kumar, D., & Kumari, M. (2023). Leveraging machine learning algorithms for improved disaster preparedness and response through accurate weather pattern and natural disaster prediction. *Frontiers in Environmental Science*, *11*(November), 1–14. https://doi.org/10.3389/fenvs.2023.1194918

Jenkins, D.G., Heller, V. and Giannakidis, A. (2025) 'Comprehensive prediction of subaerial landslide-tsunamis via slide model- and water body geometry-invariant machine learning techniques', *Ocean Engineering*, 320(December 2024), p. 120197. Available at: https://doi.org/10.1016/j.oceaneng.2024.120197.

Jiang, P., & Maria, A. (2019). *Analysis of Earthquake-Induced Natech Events : A Monte Carlo-Markov Chain Methodology for Storage Tank Accident Sequence Modeling and Industrial Downtime Estimation (地震誘発ナテック現象の分析 ： 貯蔵タンク事故シーケンスモデリング及び操業停止 時間推定のためのモンテカルロ ・ マルコフ連鎖手法)*.

Jin, D. and Lin, J. (2011) 'Managing tsunamis through early warning systems: A multidisciplinary approach', *Ocean and Coastal Management*, 54(2), pp. 189–199. Available at: https://doi.org/10.1016/j.ocecoaman.2010.10.025.

Juanara, E. and Lam, C.Y. (2024) 'Classification of Non-Seismic Tsunami Early Warning Level Using Decision Tree Algorithm', *Journal of Information Systems Engineering and*

Business Intelligence, 10(3), pp. 378–391. Available at: https://doi.org/10.20473/jisebi.10.3.378-391.

Juanara, E. and Lam, C.Y. (2025) 'Machine Learning Approaches for Early Warning of Tsunami Induced by Volcano Flank Collapse and Implication for Future Risk Management: Case of Anak Krakatau', *Ocean Modelling*, 194. Available at: https://doi.org/10.1016/j.ocemod.2025.102497.

Joye, B. S. B. (2015). *Deepwater Horizon, 5 years on*. *349*(6248), 592–593.

Kamalrathne, T., Amaratunga, D., & Haigh, R. (2020). Need for effective detection and early warnings for epidemic and pandemic preparedness planning in the context of multi-hazards: Lessons from the COVID-19 pandemic. *International Journal of Disaster Risk Reduction*, *January*. https://doi.org/https://doi.org/10.1016/j.ijdrr.2023.103724

Kamiya, M., Igarashi, Y., Okada, M., & Baba, T. (2022). Numerical experiments on tsunami flow depth prediction for clustered areas using regression and machine learning models. *Earth, Planets and Space*, *74*(1). https://doi.org/10.1186/s40623-022-01680-9

Kappi, M., & Mallikarjuna, B. (2024). Artificial intelligence and machine learning for disaster prediction: a scientometric analysis of highly cited papers. *Natural Hazards*. https://doi.org/10.1007/s11069-024-06616-y

Kato, T., Terada, Y., Nishimura, H., Nagai, T., & Koshimura, S. (2011). Tsunami records due to the 2010 Chile Earthquake observed by GPS buoys established along the Pacific coast of Japan. *Earth, Planets and Space*, *63*(6), e5–e8. https://doi.org/10.5047/eps.2011.05.001

Kent Condie (2016) *Earth as an Evolving Planetary System*. version 1. Socorro, New Mexico: Elsevier Academic Press, Elsevier Inc. Available at: https://www.sciencedirect.com/book/9780128036891/earth-as-an-evolving-planetary-system.

Knight, F. H. (1921). Risk, Uncertainty, and Profit. In *The Rivserside Cambridge*. https://doi.org/10.34156/9783791046006-108

Krausmann, E., & Necci, A. (2021). Thinking the unthinkable: A perspective on Natech risks and Black Swans. *Safety Science*, *139*, 105255. https://doi.org/10.1016/J.SSCI.2021.105255

Lakshita, L., & Nair, N. K. C. (2021). High Impact Low Probability Weatherization Impact Analysis for Electricity Infrastructure. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, *2021-December*, 958–963. https://doi.org/10.1109/TENCON54134.2021.9707369

Lam, C. Y., & Cruz, A. M. (2019). Risk analysis for consumer-level utility gas and liquefied petroleum gas incidents using probabilistic network modeling: A case study of gas incidents in Japan. *Reliability Engineering and System Safety*, *185*, 198–212. https://doi.org/10.1016/j.ress.2018.12.008

Lam, C. Y., & Shimizu, T. (2021). A network analytical framework to analyze infrastructure

damage based on earthquake cascades: A study of earthquake cases in Japan. *International Journal of Disaster Risk Reduction*, *54*. https://doi.org/10.1016/j.ijdrr.2020.102025

Lara, P., Bletery, Q., Ampuero, J. P., Inza, A., & Tavera, H. (2023). Earthquake Early Warning Starting From 3 s of Records on a Single Station With Machine Learning. *Journal of Geophysical Research: Solid Earth*, *128*(11). https://doi.org/10.1029/2023JB026575

Larrayoz, X., Lebeña, N., Casillas, A., & Pérez, A. (2023). Representation Exploration and Deep Learning Applied to the Early Detection of Pathological Gambling Risks. *CEUR Workshop Proceedings*, *3497*, 693–705.

Larson, K. M., Lay, T., Yamazaki, Y., Cheung, K. F., Ye, L., Williams, S. D. P., & Davis, J. L. (2021). Dynamic Sea Level Variation From GNSS: 2020 Shumagin Earthquake Tsunami Resonance and Hurricane Laura. In *Geophysical Research Letters* (Vol. 48, Issue 4). Blackwell Publishing Ltd. https://doi.org/10.1029/2020GL091378

Latter, J. H. (1981). *Tsunamis of Volcanic Origin: Summary of Causes, with Parti-cular Reference to Krakatoa, 1883*.

Layana, S., Aguilera, F., Rojo, G., Vergara, Á., Salazar, P., Quispe, J., Urra, P., & Urrutia, D. (2020). Volcanic Anomalies Monitoring System (VOLCANOMS), a low-cost volcanic monitoring system based on Landsat images. *Remote Sensing*, *12*(10). https://doi.org/10.3390/rs12101589

Ledwoch, A., Brintrup, A., Mehnen, J., & Tiwari, A. (2018). Systemic Risk Assessment in Complex Supply Networks. *IEEE Systems Journal*, *12*(2), 1826–1837. https://doi.org/10.1109/JSYST.2016.2596999

Lee, B., Preston, F., Green, G., & Royal Institute of International Affairs. (2012a). *Preparing for high-impact, low-probability events : lessons from Eyjafjallajökull*. Chatham House.

Lee, B., Preston, F., Green, G., & Royal Institute of International Affairs. (2012b). *Preparing for high-impact, low-probability events*. Chatham House.

Letham, B., Rudin, C., McCormick, T. H., & Madigan, D. (2015). Interpretable classifiers using rules and bayesian analysis: Building a better stroke prediction model. *Annals of Applied Statistics*, *9*(3), 1350–1371. https://doi.org/10.1214/15-AOAS848

Linardos, V., Drakaki, M., Tzionas, P., & Karnavas, Y. L. (2022). Machine Learning in Disaster Management: Recent Developments in Methods and Applications. In *Machine Learning and Knowledge Extraction* (Vol. 4, Issue 2, pp. 446–473). MDPI. https://doi.org/10.3390/make4020020

Liu, C. M., Rim, D., Baraldi, R., & LeVeque, R. J. (2021). Comparison of Machine Learning Approaches for Tsunami Forecasting from Sparse Observations. *Pure and Applied Geophysics*, *178*(12), 5129–5153. https://doi.org/10.1007/s00024-021-02841-9

Liu, Z., Zhang, Z., Sang, J., Yin, Y., & Ding, X. (2024). Development and application of early warning system for heavy metal pollution accident in drinking water source area. *Process Safety and Environmental Protection*, *183*(December 2023), 293–306.

https://doi.org/10.1016/j.psep.2024.01.014

Lorenzoni, N., Stühlinger, V., Stummer, H., & Raich, M. (2020). Long-term impact of disasters on the public health system: A multi-case analysis. *International Journal of Environmental Research and Public Health*, *17*(17), 1–17. https://doi.org/10.3390/ijerph17176251

Lynett, P. *et al.* (2022) 'Diverse tsunamigenesis triggered by the Hunga Tonga-Hunga Ha'apai eruption', *Nature*, 609(7928), pp. 728–733. Available at: https://doi.org/10.1038/s41586-022-05170-6.

Lynett, P.J. *et al.* (2012) 'Observations and modeling of tsunami-induced currents in ports and harbors', *Earth and Planetary Science Letters*, 327–328, pp. 68–74. Available at: https://doi.org/10.1016/j.epsl.2012.02.002.

Løvholt, F., Glimsdal, S., & Harbitz, C. B. (2020). On the landslide tsunami uncertainty and hazard. *Landslides*, *17*(10), 2301–2315. https://doi.org/10.1007/s10346-020-01429-z

M. Elisabeth, P.-C. (1986). Warning Systems in Risk Management. *Risk Analysis*. https://doi.org/https://doi.org/10.1111/j.1539-6924.1986.tb00210.x

Madeira, J., Ramalho, R. S., Hoffmann, D. L., Mata, J., & Moreira, M. (2020). A geological record of multiple Pleistocene tsunami inundations in an oceanic island: The case of Maio, Cape Verde. *Sedimentology*, *67*(3), 1529–1552. https://doi.org/10.1111/sed.12612

Maeda, T., Obara, K., Shinohara, M., Kanazawa, T., & Uehira, K. (2015). Successive estimation of a tsunami wavefield without earthquake source data: A data assimilation approach toward real-time tsunami forecasting. *Geophysical Research Letters*, *42*(19), 7923–7932. https://doi.org/10.1002/2015GL065588

Magalhães, F., Cunha, A., & Caetano, E. (2012). Vibration based structural health monitoring of an arch bridge: From automated OMA to damage detection. *Mechanical Systems and Signal Processing*, *28*, 212–228. https://doi.org/10.1016/j.ymssp.2011.06.011

Mantovani, C., Corgnati, L., Horstmann, J., Rubio, A., Reyes, E., Quentin, C., Cosoli, S., Asensio, J. L., Mader, J., & Griffa, A. (2020). Best Practices on High Frequency Radar Deployment and Operation for Ocean Current Measurement. In *Frontiers in Marine Science* (Vol. 7). Frontiers Media S.A. https://doi.org/10.3389/fmars.2020.00210

Marbukh, V. (2016). Towards systemic risk aware engineering of large-scale networks: Complex systems perspective. *Proceedings - 2016 IEEE 24th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, MASCOTS 2016*, 391–393. https://doi.org/10.1109/MASCOTS.2016.20

Mas, E., Adriano, B., Pulido, N., Jimenez, C., & Koshimura, S. (2014). Simulation of Tsunami inundation in central Peru from future megathrust earthquake scenarios. *Journal of Disaster Research*, *9*(6), 961–967. https://doi.org/10.20965/jdr.2014.p0961

Massimetti, F., Coppola, D., Laiolo, M., Valade, S., Cigolini, C., & Ripepe, M. (2020). Volcanic Hot-Spot detection using SENTINEL-2: A comparison with MODIS-MIROVA

thermal data series. *Remote Sensing*, *12*(5). https://doi.org/10.3390/rs12050820

Masys, A. J., Ray-Bennett, N., Shiroshita, H., & Jackson, P. (2014). High Impact/Low Frequency Extreme Events: Enabling Reflection and Resilience in a Hyper-connected World. *Procedia Economics and Finance*, *18*, 772–779. https://doi.org/10.1016/S2212-5671(14)01001-6

Melgar, D., LeVeque, R. J., Dreger, D. S., & Allen, R. M. (2016). Kinematic rupture scenarios and synthetic displacement data: An example application to the Cascadia subduction zone. *Journal of Geophysical Research: Solid Earth*, *121*(9), 6658–6674. https://doi.org/10.1002/2016JB013314

Meng, Z., Hu, Y. and Ancey, C. (2020) 'Using a data driven approach to predict waves generated by gravity driven mass flows', *Water (Switzerland)*, 12(2), pp. 17–19. Available at: https://doi.org/10.3390/w12020600.

Merz, B., Elmer, F., & Thieken, A. H. (2009). Significance of "high probability/low damage" versus "low probability/high damage" flood events. *Natural Hazards and Earth System Science*, *9*(3), 1033–1046. https://doi.org/10.5194/nhess-9-1033-2009

Mimura, N., Yasuhara, K., Kawagoe, S., Yokoki, H., & Kazama, S. (2011). Damage from the Great East Japan Earthquake and Tsunami - A quick report. *Mitigation and Adaptation Strategies for Global Change*, *16*(7), 803–818. https://doi.org/10.1007/s11027-011-9297-7

Møller, A. P., & Mousseau, T. A. (2006). Biological consequences of Chernobyl: 20 years on. *Trends in Ecology and Evolution*, *21*(4), 200–207. https://doi.org/10.1016/j.tree.2006.01.008

Motta, M., de Castro Neto, M., & Sarmento, P. (2021). A mixed approach for urban flood prediction using Machine Learning and GIS. *International Journal of Disaster Risk Reduction*, *56*(July 2020), 102154. https://doi.org/10.1016/j.ijdrr.2021.102154

Muhari, A., Heidarzadeh, M., Susmoro, H., Nugroho, H. D., Kriswati, E., Supartoyo, Wijanarto, A. B., Imamura, F., & Arikawa, T. (2019). The December 2018 Anak Krakatau Volcano Tsunami as Inferred from Post-Tsunami Field Surveys and Spectral Analysis. *Pure and Applied Geophysics*, *176*(12), 5219–5233. https://doi.org/10.1007/s00024-019-02358-2

Mulia, I. E., Gusman, A. R., & Satake, K. (2020). Applying a Deep Learning Algorithm to Tsunami Inundation Database of Megathrust Earthquakes. *Journal of Geophysical Research: Solid Earth*, *125*(9). https://doi.org/10.1029/2020JB019690

Mulia, I. E., Ueda, N., Miyoshi, T., Gusman, A. R., & Satake, K. (2022). Machine learning-based tsunami inundation prediction derived from offshore observations. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-33253-5

Mulia, I. E., Watada, S., Ho, T. C., Satake, K., Wang, Y., & Aditiya, A. (2020). Simulation of the 2018 Tsunami Due to the Flank Failure of Anak Krakatau Volcano and Implication for Future Observing Systems. *Geophysical Research Letters*, *47*(14).

https://doi.org/10.1029/2020GL087334

Mutaqin, B. W., Lavigne, F., Hadmoko, D. S., & Ngalawani, M. N. (2019). Volcanic Eruption-Induced Tsunami in Indonesia: A Review. *IOP Conference Series: Earth and Environmental Science*, *256*(1). https://doi.org/10.1088/1755-1315/256/1/012023

National Commission on Terrorist Attacks Upon the United States. (2004). *The 9/11 Commission Report*.

National Geophysical Data Center / World Data Service. (2024, October 27). *Global Historical Tsunami Database*. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V5PN93H7

Nations, U. (2005). *International Strategy for Disaster Reduction Evaluation and Strengthening of Early Warning Systems in Countries Affected by the 26 December 2004 Tsunami An initiative supported through the UN Flash Appeal for*. *December 2004*.

Nazli, N., & Kenji, S. (1995). Generation mechanism of tsunamis from the 1883 Krakatau eruption. *Geophysical Research Letters*.

Necmioglu, O., Heidarzadeh, M., Vougioukalakis, G. E., & Selva, J. (2023). Landslide Induced Tsunami Hazard at Volcanoes: the Case of Santorini. *Pure and Applied Geophysics*. https://doi.org/10.1007/s00024-023-03252-8

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., & Agha, R. (2020). The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery*, *78*(March), 185–193. https://doi.org/10.1016/j.ijsu.2020.04.018

Novianto, D., Husrin, S., Nugroho, D., Bramawanto, R., Setiawan, A., Permana, S. M., Sufyan, A., Sianturi, D. S. A., Daniel, D., Ridlo Suhelmi, I., & Fauzah, S. (2021). IDSL (Inexpensive Device for Sea Level) performance analysis for tews (Tsunami Early Warning System) in Sadeng fisheries port. *IOP Conference Series: Earth and Environmental Science*, *860*(1). https://doi.org/10.1088/1755-1315/860/1/012101

Nunes, A. R. (2024). Future-proofing tomorrow: achieving a unified vision through population health and multi-hazard early warning systems. *Perspectives in Public Health*, *144*(6), 336–339. https://doi.org/10.1177/17579139241263708

Omira, R., & Ramalho, I. (2020). Evidence-Calibrated Numerical Model of December 22, 2018, Anak Krakatau Flank Collapse and Tsunami. *Pure and Applied Geophysics*, *177*(7), 3059–3071. https://doi.org/10.1007/s00024-020-02532-x

Paris, A., Heinrich, P., Paris, R., & Abadie, S. (2020). The December 22, 2018 Anak Krakatau, Indonesia, Landslide and Tsunami: Preliminary Modeling Results. *Pure and Applied Geophysics*, *177*(2), 571–590. https://doi.org/10.1007/s00024-019-02394-y

Paris, R. (2015). Source mechanisms of volcanic tsunamis. In *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (Vol. 373, Issue 2053). Royal Society of London. https://doi.org/10.1098/rsta.2014.0380

Paris, R., Switzer, A. D., Belousova, M., Belousov, A., Ontowirjo, B., Whelley, P. L., & Ulvrova, M. (2014). Volcanic tsunami: A review of source mechanisms, past events and hazards in Southeast Asia (Indonesia, Philippines, Papua New Guinea). In *Natural Hazards* (Vol. 70, Issue 1, pp. 447–470). Kluwer Academic Publishers. https://doi.org/10.1007/s11069-013-0822-8

Paris, R., Ulvrova, M., Selva, J., Brizuela, B., Costa, A., Grezio, A., Lorito, S., & Tonini, R. (2019). Probabilistic hazard analysis for tsunamis generated by subaqueous volcanic explosions in the Campi Flegrei caldera, Italy. *Journal of Volcanology and Geothermal Research*, *379*, 106–116. https://doi.org/10.1016/j.jvolgeores.2019.05.010

Paté-Cornell, E. (2012). On "Black Swans" and "Perfect Storms": Risk Analysis and Management When Statistics Are Not Enough. *Risk Analysis*, *32*(11), 1823–1833. https://doi.org/10.1111/j.1539-6924.2011.01787.x

Perdinan, Adi, R. F., Rosya, A., Lestari, L., Harianto, T., Aprilia, S., Basit, R. A., Prihansyah, R., Putra, A. R., Aminingrum, & Jannah, N. (2024). Initiative Collaboration Tool of Early Warning Systems for Early Action to Mitigate Flood Disaster Impacts in Indonesia. *IOP Conference Series: Earth and Environmental Science*, *1359*(1), 2–13. https://doi.org/10.1088/1755-1315/1359/1/012035

Perttu, A., Caudron, C., Assink, J. D., Metz, D., Tailpied, D., Perttu, B., Hibert, C., Nurfiani, D., Pilger, C., Muzli, M., Fee, D., Andersen, O. L., & Taisne, B. (2020). Reconstruction of the 2018 tsunamigenic flank collapse and eruptive activity at Anak Krakatau based on eyewitness reports, seismo-acoustic and satellite observations. *Earth and Planetary Science Letters*, *541*, 116268. https://doi.org/10.1016/J.EPSL.2020.116268

Pota, M., Pecoraro, G., Rianna, G., Reder, A., Calvello, M., & Esposito, M. (2022). Machine learning for the definition of landslide alert models: a case study in Campania region, Italy. *Discover Artificial Intelligence*, *2*(1). https://doi.org/10.1007/s44163-022-00033-5

Quinlan, J. R. (1986). Induction of Decision Trees. In *Machine Learning* (Vol. 1).

Ranghieri, F., & Ishiwatari, M. (2014). *Learning from Megadisasters Great East Japan Earthquake*. https://openknowledge.worldbank.org/entities/publication/db0df170-6101-526e-8fc8-d0e448196fc4

Ratnasari, R. N., Tanioka, Y., Yamanaka, Y., & Mulia, I. E. (2023). Development of early warning system for tsunamis accompanied by collapse of Anak Krakatau volcano, Indonesia. *Frontiers in Earth Science*, *11*. https://doi.org/10.3389/feart.2023.1213493

Ren, Z., Wang, Y., Wang, P., Hou, J., Gao, Y., & Zhao, L. (2020). Numerical study of the triggering mechanism of the 2018 Anak Krakatau tsunami: eruption or collapsed landslide? *Natural Hazards*, *102*(1), 1–13. https://doi.org/10.1007/s11069-020-03907-y

Rice, C. (2002). *National Security Adviser Condoleezza Rice talks with Margaret Warner about Iraq*. Online NewsHour.

Rose, K. M., & Matoza, R. S. (2021). Remote hydroacoustic-infrasonic detection and

characterization of Anak Krakatau eruptive activity leading to, during, and following the December 2018 flank collapse and tsunami. *Bulletin of Volcanology*, *83*(8). https://doi.org/10.1007/s00445-021-01468-x

Sabeti, R., & Heidarzadeh, M. (2024). Estimating maximum initial wave amplitude of subaerial landslide tsunamis: A three-dimensional modelling approach. *Ocean Modelling*, *189*. https://doi.org/10.1016/j.ocemod.2024.102360

Saito, T., Satake, K., & Furumura, T. (2010). Tsunami waveform inversion including dispersive waves: The 2004 earthquake off Kii Peninsula, Japan. *Journal of Geophysical Research: Solid Earth*, *115*(6). https://doi.org/10.1029/2009JB006884

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

Satoshi Fujii. (2014). Oceangraphic Radar Tsunami Measurement System. *New Breeze, Summer* , 8–9.

Sattari, A., Foroumandi, E., Gavahi, K., & Moradkhani, H. (2025). A probabilistic machine learning framework for daily extreme events forecasting. *Expert Systems with Applications*, *265*(November 2024), 126004. https://doi.org/10.1016/j.eswa.2024.126004

Selva, J., Amato, A., Armigliato, A., Basili, R., Bernardi, F., Brizuela, B., Cerminara, M., de' Micheli Vitturi, M., Di Bucci, D., Di Manna, P., Esposti Ongaro, T., Lacanna, G., Lorito, S., Løvholt, F., Mangione, D., Panunzi, E., Piatanesi, A., Ricciardi, A., Ripepe, M., … Zaniboni, F. (2021). Tsunami risk management for crustal earthquakes and non-seismic sources in Italy. In *Rivista del Nuovo Cimento* (Vol. 44, Issue 2, pp. 69–144). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s40766-021-00016-9

Sharifi, A., & Khavarian-Garmsir, A. R. (2020). The COVID-19 pandemic: Impacts on cities and major lessons for urban planning, design, and management. *Science of the Total Environment*, *749*, 1–3. https://doi.org/10.1016/j.scitotenv.2020.142391

Shuto, N. (2002). Tsunami Assessment Method for Nuclear Power Plants in Japan. *Tsunami Evaluation Subcommittee*, *February 2002*, 72.

Siek, M., & Rafles, A. (2022). Data-Driven Modelling For Tsunami Forecasting Using Computational Intelligence. *Proceedings - 2022 IEEE International Conference on Cybernetics and Computational Intelligence, CyberneticsCom 2022*, 185–190. https://doi.org/10.1109/CyberneticsCom55287.2022.9865565

Singh, S., Kumar, R., Panchal, R., & Tiwari, M. K. (2021). Impact of COVID-19 on logistics systems and disruptions in food supply chain. *International Journal of Production Research*, *59*(7), 1993–2008. https://doi.org/10.1080/00207543.2020.1792000

Situ, Z., Zhong, Q., Zhang, J., Teng, S., Ge, X., Zhou, Q., & Zhao, Z. (2025). Attention-based deep learning framework for urban flood damage and risk assessment with improved flood

prediction and land use segmentation. *International Journal of Disaster Risk Reduction*, *116*(100), 105165. https://doi.org/10.1016/j.ijdrr.2024.105165

Sonja van Leeuwen, Paul Tett, David Mills, and J.V.D.M. (2015) 'Impacts of tides on tsunami propagation due to potential Nankai Trough earthquakes in the Seto Inland Sea, Japan', *Journal of Geophysical Research: Oceans*, 120, pp. 1–17. Available at: https://doi.org/10.1002/2014JC010485.Received.

Spiekermann, R., Kienberger, S., Norton, J., Briones, F., & Weichselgartner, J. (2015). The Disaster-Knowledge Matrix - Reframing and evaluating the knowledge challenges in disaster risk reduction. *International Journal of Disaster Risk Reduction*, *13*, 96–108. https://doi.org/10.1016/j.ijdrr.2015.05.002

Stanley, K., & Garrick, B. J. (1980). On the Quantitative Definition of Risk. *Risk Analysis*. https://doi.org/https://doi.org/10.1111/j.1539-6924.1981.tb01350.x

Struelens, M. J., Ludden, C., Werner, G., Sintchenko, V., Jokelainen, P., & Ip, M. (2024). Real-time genomic surveillance for enhanced control of infectious diseases and antimicrobial resistance. *Frontiers in Science*, *2*(April), 1–23. https://doi.org/10.3389/fsci.2024.1298248

Sun, K., Li, Z., Wang, S., & Hu, R. (2024). A support vector machine model of landslide susceptibility mapping based on hyperparameter optimization using the Bayesian algorithm: a case study of the highways in the southern Qinghai–Tibet Plateau. *Natural Hazards*. https://doi.org/10.1007/s11069-024-06665-3

Suppasri, A., Shuto, N., Imamura, F., Koshimura, S., Mas, E., & Yalciner, A. C. (2013). Lessons Learned from the 2011 Great East Japan Tsunami: Performance of Tsunami Countermeasures, Coastal Buildings, and Tsunami Evacuation in Japan. *Pure and Applied Geophysics*, *170*(6–8), 993–1018. https://doi.org/10.1007/s00024-012-0511-7

Syamsidik, Benazir, Luthfi, M., Suppasri, A., & Comfort, L. K. (2020). The 22 December 2018 Mount Anak Krakatau volcanogenic tsunami on Sunda Strait coasts, Indonesia: Tsunami and damage characteristics. *Natural Hazards and Earth System Sciences*, *20*(2), 549–565. https://doi.org/10.5194/nhess-20-549-2020

Taleb, N. N. (2007). The Black Swan: The impact of the highly improbable. In *Random House*. Random House.

Tanabe, F. (2012). Analyses of core melt and re-melt in the Fukushima Daiichi nuclear reactors. *Journal of Nuclear Science and Technology*, *49*(1), 18–36. https://doi.org/10.1080/18811248.2011.636537

Terry, J. P., Goff, J., Winspear, N., Bongolan, V. P., & Fisher, S. (2022). Tonga volcanic eruption and tsunami, January 2022: globally the most significant opportunity to observe an explosive and tsunamigenic submarine eruption since AD 1883 Krakatau. *Geoscience Letters*, *9*(1). https://doi.org/10.1186/s40562-022-00232-z

Tinti, S., Manucci, A., Pagnoni, G., Armigliato, A., & Zaniboni, F. (2005). Natural Hazards

and Earth System Sciences landslide-induced tsunamis in Stromboli: sequence of the events reconstructed from the eyewitness accounts. In *Natural Hazards and Earth System Sciences* (Vol. 5).

Titov, V. V, Gonza´lezgonza´gonza´lez, F. I., Bernard, E. N., Eble, M. C., Mofjeld, H. O., Newman, J. C., & Venturato, A. J. (2005). *Real-Time Tsunami Forecasting: Challenges and Solutions*.

Tokui, J., Kawasaki, K., & Miyagawa, T. (2017). The economic impact of supply chain disruptions from the Great East-Japan earthquake. *Japan and the World Economy*, *41*, 59–70. https://doi.org/10.1016/j.japwor.2016.12.005

Torelló-Sentelles, H., Marra, F., Koukoula, M., Villarini, G., & Peleg, N. (2024). Intensification and Changing Spatial Extent of Heavy Rainfall in Urban Areas. *Earth's Future*, *12*(9). https://doi.org/10.1029/2024EF004505

Tsuboi, M., Sawano, T., Nonaka, S., Hori, A., Ozaki, A., Nishikawa, Y., Zhao, T., Murakami, M., & Tsubokura, M. (2022). Disaster-related deaths after the Fukushima Daiichi nuclear power plant accident - Definition of the term and lessons learned. *Environmental Advances*, *8*(May), 100248. https://doi.org/10.1016/j.envadv.2022.100248

United Nations Office for Disaster Risk Reduction. (2023). *WORDS INTO ACTION: A GUIDE TO MULTI-HAZARD EARLY WARNING SYSTEMS*.

Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review*, *53*(8), 5929–5955. https://doi.org/10.1007/s10462-020-09838-1

Velasco, A., Kosheleva, O., & Kreinovich, V. (2023). High-Impact Low-Probability Events are Even More Important Than it is Usually Assumed. *Fuzzy Information Processing 2023*, *751*. https://doi.org/doi.org/10.1007/978-3-031-46778-3_35

Vesecky, J. F., Laws, K. E., & Paduan, J. D. (2010). A System Trade Model for The Monitoring of Coastal Vessels Using HF Surface Wave Radar and Ship Automatic Identifications Systems (AIS). *International Geoscience and Remote Sensing Symposium (IGARSS)*.

Wachinger, G., Renn, O., Begg, C., & Kuhlicke, C. (2013). The risk perception paradox-implications for governance and communication of natural hazards. *Risk Analysis*, *33*(6), 1049–1065. https://doi.org/10.1111/j.1539-6924.2012.01942.x

Walter, T. R., Haghshenas Haghighi, M., Schneider, F. M., Coppola, D., Motagh, M., Saul, J., Babeyko, A., Dahm, T., Troll, V. R., Tilmann, F., Heimann, S., Valade, S., Triyono, R., Khomarudin, R., Kartadinata, N., Laiolo, M., Massimetti, F., & Gaebler, P. (2019). Complex hazard cascade culminating in the Anak Krakatau sector collapse. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-12284-5

Wang, X. (2009) 'User Manual for COMCOT Version 1.7'. New Zealand: Institute of Geological & Nuclear Science, p. Institute of Geological & Nuclear Science. Available at: https://pdfcoffee.com/comcot-user-manual-v17-pdf-free.html.

Wang, Y., Heidarzadeh, M., Satake, K., Mulia, I. E., & Yamada, M. (2020). A Tsunami Warning System Based on Offshore Bottom Pressure Gauges and Data Assimilation for Crete Island in the Eastern Mediterranean Basin. *Journal of Geophysical Research: Solid Earth*, *125*(10). https://doi.org/10.1029/2020JB020293

Wang, Y., Imai, K., Kusumoto, S., Takahashi, N., Brunel, M. H., London, U., Baba, T., & Mulia, I. E. (2023). Tsunami Early Warning of the Hunga Volcanic Eruption Using an Ocean Floor Observation Network off the Japanese Islands. *Seismological Research Letters*.

Wang, Y., Imai, K., Mulia, I. E., Ariyoshi, K., Takahashi, N., Sasaki, K., Kaneko, H., Abe, H., & Sato, Y. (2023). Data Assimilation Using High-Frequency Radar for Tsunami Early Warning: A Case Study of the 2022 Tonga Volcanic Tsunami. *Journal of Geophysical Research: Solid Earth*, *128*(2). https://doi.org/10.1029/2022JB025153

Watters, R. J., Zimbelman, D. R., Bowman, S. D., & Crowley, J. K. (2000). Rock Mass Strength Assessment and Significance to Edifice Stability, Mount Rainier and Mount Hood, Cascade Range Volcanoes. In *Pure appl. geophys* (Vol. 157).

WHO. (2025). *COVID-19 Epidemiological Update* (Issue January).

Winchester, S. (2003). *Krakatoa: The Day the World Exploded*. HarperCollins.

Woo, .-F, & Cho, S.-B. (1998). *Computer Programs for Tsunami Progpagation and Inundation. National Science Foundation.* https://tsunamiportal.nacse.org/documentation/COMCOT_tech.pdf

Wu, N., Green, B., Ben, X., & O'Banion, S. (2020). *Deep Transformer Models for Time Series Forecasting: The Influenza Prevalence Case*. http://arxiv.org/abs/2001.08317

Wu, Y. *et al.* (2022) 'Numerical modeling of the propagation process of landslide surge using physics-informed deep learning', *Advanced Modeling and Simulation in Engineering Sciences*, 9(1). Available at: https://doi.org/10.1186/s40323-022-00228-6.

Yavari-Ramshe, S. and Ataie-Ashtiani, B. (2016) 'Numerical modeling of subaerial and submarine landslide-generated tsunami waves—recent advances and future challenges', *Landslides*. Springer Verlag, pp. 1325–1368. Available at: https://doi.org/10.1007/s10346-016-0734-2.

Yu, J., Cruz, A. M., Piatyszek, E., Lesbats, M., Tardy, A., Hokugo, A., & Tatano, H. (2017). A survey of impact on industrial parks caused by the 2011 Great East Japan earthquake and tsunami. *Journal of Loss Prevention in the Process Industries*, *50*, 317–324. https://doi.org/10.1016/j.jlp.2017.01.020

Yun, S. J., Kwon, J. W., & Kim, W. T. (2022). A Novel Digital Twin Architecture with Similarity-Based Hybrid Modeling for Supporting Dependable Disaster Management Systems. *Sensors*, *22*(13). https://doi.org/10.3390/s22134774

Zhou, B., Zhang, X., Wan, X., Liu, T., Liu, Y., Huang, H., & Chen, J. (2024). Development and Application of a Novel Tsunami Monitoring System Based on Submerged Mooring.

Sensors, *24*(18), 6048. https://doi.org/10.3390/s24186048

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2020). *Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting*. http://arxiv.org/abs/2012.07436

Zorn, E. U., Orynbaikyzy, A., Plank, S., Babeyko, A., Darmawan, H., Robbany, I. F., & Walter, T. R. (2022). Identification and ranking of subaerial volcanic tsunami hazard sources in Southeast Asia. *Natural Hazards and Earth System Sciences*, *22*(9), 3083–3104. https://doi.org/10.5194/nhess-22-3083-2022