REPUBLIC OF TURKEY
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF NANOTECHNOLOGY AND ENGINEERING SCIENCES

PREDICTION OF LANDSLIDE TSUNAMI RUN-UP THROUGH ANN-BASED MODELS

SAVAŞ YAĞUZLUK
MASTER OF SCIENCE

SUPERVISOR
ASST. PROF. DR. BARAN AYDIN

CO-SUPERVISOR
ASST. PROF. DR. MUSTAFA AÇIKKAR

ADANA 2022

PREDICTION OF LANDSLIDE TSUNAMI RUN-UP THROUGH ANN-BASED MODELS

submitted by **Savaş YAĞUZLUK** in partial fulfillment of the requirements for the degree of **Master of Science in Department of Nanotechnology and Engineering Sciences, Graduate School of Natural and Applied Sciences, Adana Alparslan Türkeş Science and Technology University** by,

Assoc. Prof. Dr. Ahmet BEYCİOĞLU
Graduate School Director

Asst. Prof. Dr. Hatice İmge OKTAY BAŞEĞMEZ
Department Chair

Asst. Prof. Dr. Baran AYDIN
Supervisor

Asst. Prof. Dr. Mustafa AÇIKKAR
Co-Supervisor

Thesis Committee

Asst. Prof. Dr. Baran AYDIN
Adana Alparslan Türkeş Science and Technology University

Assoc. Prof. Dr. Utku KÂNOĞLU
Middle East Technical University

Asst. Prof. Dr. Fatih KILIÇ
Adana Alparslan Türkeş Science and Technology University

Thesis Defense Date

15/02/2022

I hereby declare that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all information that is not original to this work.

[Signature]

Savaş YAĞUZLUK

ABSTRACT

PREDICTION OF LANDSLIDE TSUNAMI RUN-UP THROUGH ANN-BASED MODELS

Savaş YAĞUZLUK
Department of Nanotechnology and Engineering Sciences
Supervisor: Asst. Prof. Dr. Baran AYDIN
Co-Supervisor: Asst. Prof. Dr. Mustafa AÇIKKAR
February 2022, 65 pages

New prediction models based on multilayer perceptron are proposed which successfully predict the maximum run-up of landslide-generated tsunami waves and the roles of parameters affecting the maximum run-up are assessed. The input to the models consists of approximately 55,000 rows of data employing six predictors, namely cross section of the landslide profile, the beach slope angle, the initial slide submergence, the vertical thickness and the horizontal length of the sliding mass, and the time of the maximum run-up. The target variable is the maximum tsunami run-up. An optimization study was first performed and a subset of 9,000 randomly sampled rows was determined to represent the whole data. A total number of 63 models consisting all possible combinations of the six parameters were then constructed for the 9,000-row data set, in an attempt to predict the maximum run-up. The prediction capability of the models was assessed by calculating Root Mean Square Error, Mean Absolute Error, Mean Absolute Percentage Error (MAPE), and Multiple Correlation Coefficient (R) as performance metrics. The MLP-based models led predictions with a minimum MAPE of approximately 1.1% and with R=1, revealing that the slide thickness has the largest impact on the maximum tsunami run-up, whereas the slide cross-section and the slope angle have minimal effect. Further, a backward elimination algorithm equipped with ReliefF feature importance method is utilized as a practical tool. Comparison with existing literature showed the reliability and applicability of the offered models. This approach can be therefore used as a fast and accurate methodology to support or when there is lack of analytical or numerical modeling.

Keywords: artificial neural networks, feature selection, landslide tsunami, maximum run-up, multilayer perceptron

ÖZET

HEYELAN KAYNAKLI TSUNAMİ TIRMANMASININ ANN TABANLI MODELLERLE TAHMİNİ

Savaş YAĞUZLUK
Nano Teknoloji ve Mühendislik Bilimleri Anabilim Dalı
Danışman: Dr. Öğr. Üyesi Baran AYDIN
İkinci Danışman: Dr. Öğr. Üyesi Mustafa AÇIKKAR
Şubat 2022, 65 sayfa

Heyelan kaynaklı tsunami dalgalarının maksimum tırmanmalarını başarılı bir şekilde tahmin eden ve maksimum tırmanmayı etkileyen parametrelerin rollerini inceleyen çok katmanlı algılayıcı tabanlı tahmin yöntemleri önerilmiştir. Modellerin girdisi heyelan profilinin kesiti, sahilin eğim açısı, heyelanın başlangıçtaki derinliği, düşeydeki kalınlığı ve yataydaki genişliği ile maksimum tırmanmanın gerçekleştiği zaman olmak üzere 6 kestiricili yaklaşık 55.000 satırlık veri setinden oluşmaktadır. Hedef değişkeni maksimum tsunami tırmanmasıdır. Öncelikle bir optimizasyon çalışması gerçekleştirilmiş ve tüm veri setini temsil etmek üzere rastgele seçilmiş 9.000 satırdan oluşan bir veri seti belirlenmiştir. Daha sonra maksimum tırmanmayı tahmin edebilmek amacıyla, 6 kestiricinin tüm kombinasyonlarını içerecek şekilde 63 model oluşturulmuştur. Modellerin tahmin kabiliyetleri, ortalama karesel hata, ortalama mutlak hata, ortalama mutlak yüzde hata (MAPE) ve çoklu korelasyon katsayısı (R) performans göstergeleri hesaplanarak belirlenmiştir. MLP tabanlı modeller maksimum tırmanmayı yaklaşık %1,1 minimum MAPE ve R=1 ile tahmin etmiş ve heyelanın düşey kalınlığının tırmanma üzerindeki en etkili parametre olduğunu, heyelan profilinin kesiti ile sahilin eğim açısının ise etkisi en az seviyede olan parametreler olduğunu ortaya çıkarmıştır Ayrıca, pratik bir araç olarak ReliefF öznitelik önem yöntemi destekli bir geri eleme algoritması kullanılmıştır. Literatür sonuçları ile yapılan karşılaştırma önerilen modellerin güvenilirliğini ve uygulanabilirliğini göstermiştir. Bu nedenle, önerilen yaklaşım analitik ya da sayısal modelleri desteklemek amacıyla, ya da yoksunlukları durumunda hızlı ve hassas bir yöntem olarak kullanılabilecektir.

Anahtar Kelimeler: çok katmanlı algılayıcı, heyelan kaynaklı tsunami, maksimum tırmanma, özellik seçimi, yapay sinir ağları

To my dearest mother Gülperi and my father Fuat,

and to my lovely sisters...

ACKNOWLEDGEMENTS

First and foremost, I am extremely grateful to my esteemed supervisor, Asst. Prof. Dr. Baran Aydın, for his invaluable guidance, continuous support, and almighty patience throughout my master study. Genuinely, this study would not be possible without his immense knowledge, numerous revisions and guiding resolve the confusion I had in the field of engineering sciences, as well as plentiful experience that have encouraged me all the time of my academic research and daily life. He has been a magnifical mentor whenever I needed help and motivation professionally and personally. It would be a cliche without mentioning his splendid teaching which has provided me insight into ordinary differential equations which I scared of whole my academic life.

I express my great appreciation to Asst. Prof. Dr. Mustafa Açıkkar, who accepted to be my co-supervisor in his tight schedule, for his patience and being a source of new ideas as well as opening a new path in which I was totally an alien. Thank you for offering advice with a perfect blend of insight and humor while treating tasty coffee at the office.

I would like to thank the Thesis Committee Members Assoc. Prof. Dr. Utku Kânoğlu, who is my academic ancestor, and Asst. Prof. Dr. Fatih Kılıç for the valuable comments, suggestions and their kindness.

I would like to offer my gratitude to some of the great mathematicians, Prof. Dr. Ali Akgül, Assoc. Prof. Dr. Mehmet Gülbahar, and Asst. Prof. Dr. Kadir Karataş; to some of the marvelous linguists, Assoc. Prof. Dr. Ömer Özer, Assoc. Prof. Dr. Eser Ördem, and Asst. Prof. Dr. Duygu İşpınar Akçayoğlu, for their incomparable teaching skills with courteous attitude.

Adana Science and Technology University also deserves praise for its student-friendly academic culture which helped me in times of need. I have spent 4 academic years filled with full of joy, accomplishments, including bringing the first sports award of the university, and unforgettable memories with awesome people.

Many people have been supportive during this long-drawn journey. Especially, I am thankful to my cousins Hüseyin, M. Sait, Yunus, Hebun, Serdar, and Özgür; to my friends, Abdullah Baykan, Demet Sert, Dilar and Zümra Dinç, Feramez Aka, İsa Timurtaş, and Zülfü Gül.

My precious parents, Gülperi and Fuat, and my sisters, Gülçin, Burçin, and Büşra: Thank you all for always believing in me and making available of your dedicated support in all aspects of life. Literally, this thesis would not have been possible without your support and unconditional love even in times of my unbearable attitude. This work, as with all of my work, belongs to you…

TABLE OF CONTENTS

TABLE OF CONTENTS ..... **x**

LIST OF FIGURES ..... **xii**

LIST OF TABLES ..... **xiv**

NOMENCLATURE ..... **xv**

1. INTRODUCTION **1**

1.1. Large-Scale Tsunamigenic Landslide Disasters Around the World 2

1.2. Artificial Neural Networks 4

1.3. Scope of the Thesis 6

2. ANN-BASED PREDICTION MODELS **7**

2.1. General Structure of ANNs 8

2.2. Regression 9

2.3. Multilayer Perceptron (MLP) 11

2.4. General Regression Neural Network (GRNN) 12

2.5. Radial Basis Function Neural Network (RBFNN) 13

2.6. Feature Selection 13

2.7. K-fold Cross-validation 14

2.8. Performance Metrics 14

3. IMPLEMENTATION OF PREDICTION MODELS **17**

3.1. Data Set Generation 17

3.2. Network Selection 19

3.3. Implementation of MLP Models 19

4. RESULTS AND DISCUSSIONS **22**

4.1. Data Set Selection 22

4.2. Model Results 25

4.3. Application of the Feature Selection Algorithm 30

4.4. Comparison with the Literature 33

5. CONCLUSIONS......35
6. RECOMMENDATIONS......36
REFERENCES......37
APPENDIX......45
Analytical Modeling of Landslide Tsunamis......45
CURRICULUM VITAE......50

LIST OF FIGURES

Figure 1.1. Biological neuron of human brain (After Vu-Quoc, 2012)................................... 5

Figure 2.1. Machine learning algorithms.. 8

Figure 2.2. Regression estimation (After Smola and Vishwanathan, 2008)........................... 10

Figure 2.3. Structure of an MLP network with one hidden layer...12

Figure 2.4. Structure of a typical GRNN. ……………………….................................… 13

Figure 2.5. Implementation of the 10-fold cross-validation algorithm. For each of the 10 test sets, 10-fold cross-validation is employed on the training set. The accuracy of each prediction is the average of the folds...14

Figure 3.1. Definition sketch for the landslide problem (not to scale). δ and L respectively indicate the maximum vertical thickness and the maximum horizontal length of the sliding mass at its initial location (x_0). The undisturbed water depth is $h_0(x) = x \tan \beta$, where β is the beach angle with the horizontal. SWL stands for the still water level. (Modified from Aydın, 2021)…………………………………………………………………………... 18

Figure 3.2. (a) Flowchart showing implementation steps of the MLP-based prediction models. (b) Flowchart showing implementation steps of the feature selection algorithm.................... 21

Figure 4.1. Variation of average MSE with the data set size.………….................................23

Figure 4.2. Scatter plots of the output variable for *(a)* the whole data set with 54,952 rows, and *(b)* the 9K data set. The predictors from top to bottom are the initial slide submergence, the maximum slide thickness, the maximum slide length, and the time of the maximum run-up. 24

Figure 4.3. Box-whiskers plot for the descriptive statistics of the 9K data set.......................25

Figure 4.4. Average RMSE values of the combination models tabulated in Tables 4.3-4.7. The gray numbers indicate the number of predictors..31

Figure 4.5. Performance metrics of the ReliefF models tabulated in Table 4.9. …………… 33

Figure A.1. Comparison of the initial bottom profiles given in Equations (A.12) and (A.15), respectively, (*a*) on the slope as functions of x, and, (*b*) on a horizontal plane as functions of ξ. The initial slide submergence is chosen as $\xi_0 = 6$. ……………...… 47

Figure A.2. Dependence of run-up on landslide horizontal length for a fixed slide thickness of $\delta = 20$ m (left), and, slide thickness for a fixed slide length of $L = 200$ m (right). Results for three different values of the initial slide submergence, i.e. $x_0 = 0, 3, 6,$ (dimensionless) are presented and the maximum run-ups are also indicated with colored circles. The slope angle is chosen as $\beta = 10°$. Note that the dimensional initial submergence parameter is $\tilde{x}_0 = Lx_0$.... 49

LIST OF TABLES

Table 3.1. Performance metrics and execution times (reported in hh:mm:ss format) for the three widely used networks are compared in an aim to select the optimum network..............19

Table 4.1. Descriptive statistics of the variables in the 9K data set except for the bottom profile code. The means and the standard deviations for the whole data set are also provided.23

Table 4.2. Example model codes and considered and disregarded predictors........................ 26

Table 4.3. Average performance metrics for the models with a single predictor. The results are sorted in increasing RMSE values..26

Table 4.4. Average performance metrics for the models with two predictors. The results are sorted in increasing RMSE values... 27

Table 4.5. Average performance metrics for the models with three predictors. The results are sorted in increasing RMSE values... 28

Table 4.6. Average performance metrics for the models with four predictors. The results are sorted in increasing RMSE values... 29

Table 4.7. Average performance metrics for the models with five and six predictors. The results are sorted in increasing RMSE values.. 30

Table 4.8. Predictors of the 9K data set are ranked according to their ReliefF scores............30

Table 4.9. Average performance metrics of the models constructed for the 9K data set in view of the ReliefF feature importance method..32

NOMENCLATURE

b_p	: Code of the landslide bottom profile
x_0	: Initial submergence of the sliding mass
β	: Beach slope angle
δ	: Maximum (vertical) thickness of the sliding mass
L	: Horizontal length of the sliding mass
t	: Time of the maximum tsunami run-up
$R_{\max}$	: Maximum tsunami run-up
MAE	: Mean Absolute Error
MAPE	: Mean Absolute Percentage Error
MSE	: Mean Squared Error
RMSE	: Root Mean Squared Error
R	: Multiple Correlation Coefficient

1. INTRODUCTION

Tsunamis are ocean surface waves which originate after sudden displacement of the sea bottom. The word tsunami is a borrowing from Japanese language and its literal meaning is harbor wave. A tsunami propagates long distances across oceans in the form of a series of long gravity waves (Murty, 1975). Upon arriving a coast, it pushes huge amount of water onto the shore. This change in the mean sea level at the shoreline is called tsunami run-up. In other words, run-up is the vertical height onshore above the mean sea level reached by a tsunami wave.

Earthquakes are the most significant source of tsunamis, the frequency being followed by landslides and volcanoes (Gusiakov, 2009). Thus and rightly so, tsunami research has concentrated mostly on the hazard posed by seismic sources. Because these natural movements are happening randomly, extremely rapid and severely destructive, it is obviously vital to predict run-up and inundation of tsunamis in order to evacuate people from the potential impact areas and minimize economical damage. Although larger earthquakes and landslides have direct impact, it is also important to know or predict effects of waves away from the source region, in coastal areas in particular (Ward, 2000).

Besides in the seas, oceans and bays, tsunamis can also occur in lakes and reservoirs and, besides earthquakes, they could be generated by landslides, rock falls, snow avalanches, and glaciers. There occurred hazardous tsunamis triggered by near-shore submarine landslides along subduction zones (Ruff, 2003). There are also several well-documented cases with complex generation mechanisms, involving landslides as a secondary hidden mechanism alongside an earthquake, and resulted in a surprise tsunami, i.e. a totally unexpected event or an event for which abnormally high run-up values were recorded at some sites.

Landslides are naturally triggered by volcanic eruptions, coastal erosion, intense rainfall, and earthquakes besides human triggers such as deforestation, mining, stream and ocean current alteration, and are very serious geological hazards that often cause devastating disasters. Landslides frequently occur worldwide but some certain hotspots are notably opportune. Asia underlined by the United Nations' Food and Agriculture Organization as a highly susceptible to landslides since it has vulnerable soils, heavy rainfall, steep terrain, and earthquake activity. Central, South, and Northwestern America can be remarked as other susceptible locations (Sassa et al., 2017)

1.1. Large-Scale Tsunamigenic Landslide Disasters Around the World

In tsunami science researchers want to know the tsunami generation potential of seismic or landslide source to produce adequate physical and mathematical models. In the case of earthquake-generated (i.e. tectonic) tsunamis, only a small fraction (usually less than 1% for typical shallow earthquakes) of the earthquake energy is known to be converted into tsunami wave energy, although even such a small energy is enough to generate strong tsunamis (Ruff, 2003). In the case of landslide-tsunamis, on the other hand, there is a dynamic interaction between landslide motion and tsunami wave generation. Although the amount of energy release is usually small compared to earthquakes, landslides have shown their potential as a ubiquitous and costly natural hazard. The resulting impulse waves, which are generated by landslides, can cause disaster due to run-up along to shoreline.

In the recent history, the 1958 Lituya Bay, Alaska (USA) landslide is remembered as one of the most spectacular examples of a rarely happening subaerial landslide that caused large localized waves. It was set off from the great strike-slip earthquake in south eastern Alaska on 10 July 1958 (Miller, 1960; Fritz, 2001; Fritz and Moser, 2003; Fritz et al., 2003a; Fritz et al., 2003b; Fritz et al., 2004; Fritz, 2009; Weiss et al., 2009; Xenakis et al., 2017). The landslide, which fell into a narrow fjord, caused a slosh wave which reached about 520 m height and then continued out the fjord where it was at about 10 meters' height when it first met the ocean. The wave amplitude decreased quickly as it spread out in the open ocean (Ruff, 2003). Even though the total displaced water volume was too small to cause a destructive trans-oceanic tsunami, the sui generis Lituya Bay event indicated that landslides can generate localized tsunamis that can be extremely hazardous.

Shortly afterwards, another notable event which is one of the most studied landslide cases in the literature occurred in Europe. On 9 October 1963, a block of nearly 270 million cubic meters was detached from the reservoir of the Vajont Dam in Northern Italy, caused an impulse wave which over-topped the dam and caused to wiped out a community of Longarone village with an unconfirmed number of more than 2,000 casualties, making it the deadliest landslide in Europe (Petley, 2008). With its 270 m height, the Vajont event is also the second highest tsunami run-up in recorded history that was generated by a landslide (Muller, 1964, Crosta et al., 2016).

Catastrophic events also occurred more recently. In 2007, a rock fell into Lake Lucerne, Switzerland, created an impulse wave that caused slight damage on the opposite shore of the lake when it flowed up into the village of Weggis (Fuchs and Boes, 2010). In 2010, a tsunami that ran up nearly 40 m on the opposite shore and destroyed trees, roads and campsite facilities at Chehalis Lake, Canada (Brideau et al., 2011; Wang et al., 2015). Many other examples of landslide-tsunamis and impulse waves are summarized in the literature. While Slingerland and Voight (1979) address events occurred in the Americas, Huber (1982) focuses on events that occurred in Switzerland in the last 600 years, and Panizzo et al (2005a) analyze events occurred in Italy. Reviews of these studies expose that landslide tsunamis are more frequent than thought.

The 17 July 1998 Papua New Guinea tsunami should also be mentioned as one of the most interesting events, with the tsunami actually triggered by a submarine landslide that is generated by the earthquake. This was a unique event which became a milestone in attracting researchers' interest and, in direct proportion to this, number of research focusing on different aspects of landslide generated tsunamis has increased in the literature in the last two decades. From an analytical point of view, several studies that are now considered as benchmark in the literature such as Lynett and Liu (2003) and Sammarco and Renzi (2008) are published. In addition, Postacioglu and Ozeren (2008), Renzi and Sammarco (2010, 2012, 2016), Seo and Liu (2013), Ozeren and Postacioglu (2011) and Ramadan et al. (2014) worth mentioning as remarkable analytical studies in the field. Lynett and Liu (2002), Okal and Synolakis (2004), Lynett and Liu (2005), Enet and Grilli (2007) and Beisel et al. (2007) can be counted among numerical studies. The number of publications on physical modeling is not less either. Liu et al. (2005), Lynett and Liu (2005), Enet and Grilli (2007), Di Risio (2009a, 2009b), Mohammed and Fritz (2012, 2013), Heller and Spinneken (2015), Romano et al. (2016) and McFall and Fritz (2017) contributed in understanding effects of different parameters on subsequent waves generated by subaerial or submarine landslides. Heller et al. (2009) and Evers et al. (2019) published detailed manuscripts on the landslide generated impulse waves in reservoirs. Roberts et al. (2014) compiled a catalogue which is composed of landslide-generated waves that were triggered by subaerial landslides including 254 events from the 14th century AD to 2012. Yavari-Ramshe and Ataie-Ashtiani (2016) compiled another list including 34 historical events together with their properties which include the slide volume that varies from small-scaled of about 100 m^3 to large-scaled of about 100 km^3.

1.2. Artificial Neural Networks

Human being has been fascinated with the idea of prediction of the future from their early existence to present. Processing and analyzing big data by applying customary techniques are attainable yet formidable. Besides successful precision, accurate analysis is the other key role since remarkable mistakes in some areas may result large loss of lives, property or money. Thus, many researchers have struggled with successful analysis of temporal data during the last decades. Today in science, however, the neural networks which process big data successfully provides miscellaneous practices with better precision for the researchers.

Artificial neural networks (ANNs) are a system that simulates the learning and finding correlation behaviors of human brain by analyzing some given data and were originally engineered by McCulloch and Pitts (1943) with the idea of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. ANNs can briefly be described as a simplified mathematical model of the human brain, inspired by the biological neuron (Figure 1.1). They have their own unique intelligence practice that is not working by applying a procedure step-by-step, which is frequently the case that happens in most techniques. They rather consist of the combination of simple interconnected units which are neurons. These neurons work in parallel in order to carry out multiple tasks, such as prediction, optimization, pattern recognition or control. Due to its communication ability, the neuron is the key element in neural networks which are inspired by the structure of nervous systems. In order to explain the analogies between the biological synaptic activity and ANN, it can be considered that signals arrive to the synapse as the neuron's inputs and can be whether abated or amplified utilizing an associated weight. These input signals can whether prompt the neuron on condition that a positive weighted synapsis is executed or, on the contrary, they can obstruct it if the weight is negative. Eventually, the neuron is activated providing that the sum of the weighted inputs is equal or greater than a certain threshold. Neurons consequently present, binary results: activation or not activation.

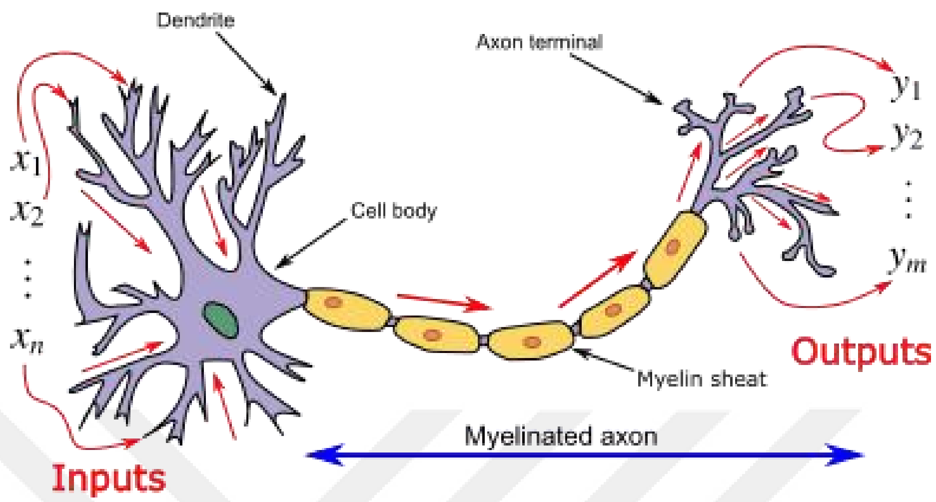

Figure 1.1. Biological neuron of human brain (After Vu-Quoc, 2012).

The architecture of neural networks consists in the organization and position of the neurons with respect to the input or output of the network. In this regard, the number of layers, the number of neurons each layer, the connection grade and the type of connections among neurons are the fundamental parameters of the network. With reference to the number of layers, ANN can be classified into monolayer, or multilayer networks (MLP) as in the case of the present study. The former has only one input layer and one output layer, whereas the multilayer networks are a generalization of the monolayer ones, which add hidden layers between the input and the output (Rosenblatt, 1958; Martínez-Álvarez, 2015).

Artificial intelligence's roots date back to archaic age. It is known from historical documents that Greek myths incorporated the idea of intelligent robots and artificial beings. Philosophers and mathematicians such as Aristotle, Al-Jazari, Descartes, Hobbes, Pascal, Leibniz, Bolzano, and many more, mentioned about the autonomous machines in their works. They realized that rational thoughts could be made systematic as algebra which the basis of the computer science.

1.3. Scope of the Thesis

It is obviously vital to predict period, run-up height and inundation distance of tsunamis in order to facilitate timely evacuation of coastal populations and minimize economical damage. Physical model studies play an important role in revealing dependence of maximum tsunami run-up on certain landslide parameters; however, they are extensive studies which require infrastructure, time and money. This study aims to suggest a computational framework to predict the maximum landslide tsunami run-up and to investigate the relationship between the geometrical parameters of the landslide and the maximum tsunami run-up. Since such an analysis involving many parameters is extremely challenging, if not impossible, using analytical or even numerical modeling, prediction models based on ANNs are suggested.

For this purpose, first, a generalized version of the linear analytical solution developed by Liu et al. (2003) was used to calculate the run-up of submarine landslides over a uniformly sloping beach and a data set was eventually generated. Then, ANNs-based prediction models are proposed in order to predict the maximum run-up of landslide-generated tsunami waves. The parameters that are effective on the maximum run-up are also determined through these models. The predictions were obtained through a multi-layer feedforward network which is described in Chapter 2.

There are a limited number of studies in the literature utilizing neural networks for prediction of tsunami waves. Erdurmaz (2004), which appears to be the first such study, used ANNs for prediction of the average tsunami height induced on the shore by earthquakes, given depth and magnitude of the earthquake. Panizzo and Girolamo (2005) carried out an ANNs analysis and eventually suggested new empirical relationships for their experimental results on principal features of the waves generated by landslide movement in a three-dimensional water body. Even though there are other ANNs-based studies on forecasting mean wave heights, periods, and tidal levels at some coastal areas, to the best of our knowledge, there is no such work using machine learning techniques and also utilizing feature importance methods to predict run-up of landslide- induced tsunamis, which is the main goal of the present study.

2. ANN-BASED PREDICTION MODELS

Machine learning, one of the sub-fields of soft computing, teaches computers naturally such as to humans and animals; that is, teaches computers learn from data instead of the experience as humans do. Machine learning algorithms use computational methods to “learn” information directly from data, which is a training data set, without relying on a predetermined equation as a model. The algorithms adaptively improve their performance as the number of samples available for learning increases. These increases are the need for the developing models and appropriate modeling approaches to represent contexts such as reasoning and decision-making techniques in the case of imprecise, incomplete and defective information. As there is no only one teaching and learning method for human beings, there is no the best method or algorithm that fits all the problems as well. Since there are wide range of machine learning algorithms (Figure 2.1), and each takes a different approach to learning, deciding the right algorithm can be remarkably compelling. Algorithm selection does not only depend on the size and type of data but the desired outcomes to get from the data, besides how those outcomes will be used.

The process of finding the right algorithm is relatively trial and error. Thus, even an experienced data scientist cannot decide the best model for a specific problem without trying it out. For the problem in this study, neural networks are opted for as the best model after this trial period. Neural networks that are used are one of the regression methods under the supervised learning which differs from unsupervised learning in terms of data. ANNs are complex mathematical models that are able to reproduce the mechanism of the human mind’s thinking. The fundamental idea that lies behind it is to split the given information over complex and multiple connected networks.

After learning process, the second process starts to apply outputs of the learning process. This is called prediction process trying to predict data that is test data which has not seen by the machine in the training process. These two processes are basically the idea of learning and applying the knowledge which is very similar to human learning experience. Below, phases of the learning experience of the machine are briefly described.

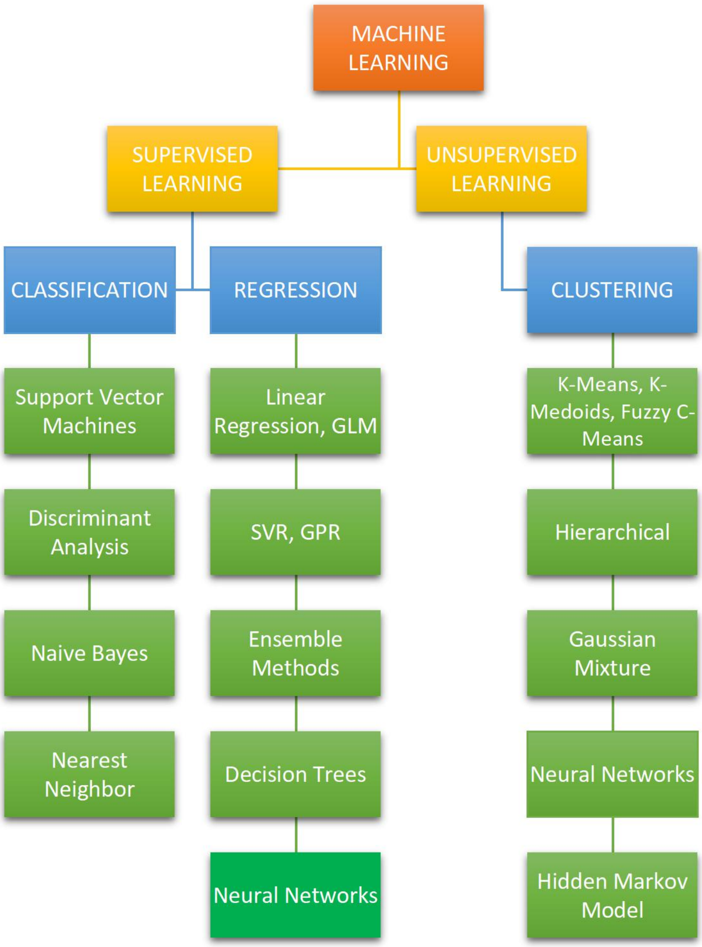

Figure 2.1. Machine learning algorithms.

2.1. General Structure of ANNs

Today's technology leads us to find and use better computation and modeling tools. The term soft computing is a recently coined term describing the symbiotic use of many emerging computing disciplines. Zadeh (1994) described it to be tolerant of imprecision, and uncertainty, in contrast to the traditional hard computing. Soft computing represents the combination of relatively new methods such as fuzzy logic, probabilistic reasoning, genetic algorithms, and ANNs in an incredibly wide variety of application areas, by providing us complementary methods to solve complex real-world problems with cost and time economy.

ANNs are computational structures that can be trained to learn patterns from examples. They were firstly explored by Rosenblatt (1959) and Widrow and Hoff (1960) at the end of 1950s. Artificial neurons, which are designed to imitate biological neurons, are the core components of ANNs, so that the learning process of human brain can be simulated artificially using mathematical methods. With complex architecture capable of finding correlations between observed data without any assumption on the physics of the considered model or phenomenon, they recently started replacing classical mathematical models in an attempt to simulate complex multiparameter engineering problems (Panizzo and Girolamo, 2005). Furthermore, with the massive and ever increasing amount of data becoming available, ANNs are gradually becoming more important and widely implemented in every discipline.

Within machine learning, the two basic learning approaches are unsupervised learning and supervised learning. While unsupervised learning uses machine learning algorithms to analyze and cluster hidden patterns unlabeled data sets without the need for human intervention (hence unsupervised), supervised learning uses labeled data sets designed to train (supervise) algorithms into classifying data or predicting outcomes accurately (Delua, 2021). Unsupervised learning is mainly used for clustering, association and dimensionality reduction. Supervised learning is mainly used for classification and regression. Classification models, as its name signifies, classify data into categories. The problem of predicting landslide-generated tsunami run-up is a regression problem.

2.2. Regression

Regression is a predictive statistical approach which aims to estimate a dependent variable (target variable) with a given set of independent variables (predictors). In mathematical terms, regression is expected to find some function f mapping the observations x_i of a dependent variable, as indicated by black dots in Figure 2.2, to an estimation $\hat{y} = f(x_i)$ which approximates the observed value y, i.e. $y \approx \hat{y}$.

The task of regression consists of modeling the relationships between responses and the covariates of a system. Regression is a fundamental approach used in machine learning, and it appears in a great variety of other research areas and applications, including time-series analysis, solving partial differential equations, control and robotics, optimization, and deep learning applications such as image processing and recognition, autonomous cars, natural language processing, speech-to-text translation, computer games, demographic and election predictions, and so on.

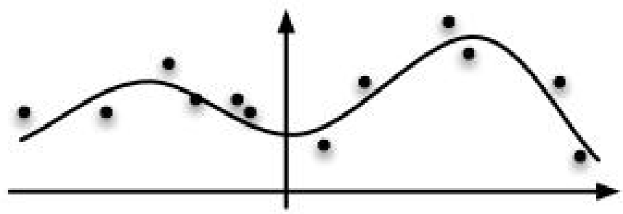

Figure 2.2. Regression estimation (After Smola and Vishwanathan, 2008).

In a simple linear regression (curve fitting) problem, with a single variable x and a single dependent (output, response) variable y, the model is basically

$$y = a_0 + a_1 x + \varepsilon = [a_0 \quad a_1] \cdot \begin{bmatrix} 1 \\ x \end{bmatrix} + [\varepsilon] = \mathrm{A} \cdot \mathrm{X} + \mathcal{E}, \tag{2.1}$$

in which a_0 represents the intercept, a_1 is the slope, and ε is for the error that is also called noise. For n independent variables (x_1, x_2, …,x_n), Equation (2.1) turns out into the general form,

$$y = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + \varepsilon = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \cdot [1 \quad x_1 \quad \cdots \quad x_n] + [\varepsilon] = \mathrm{A} \cdot \mathrm{X} + \mathcal{E}, \tag{2.2}$$

which is called the multiple linear regression. The least squares method is then used to estimate the coefficients a_i $(i = 0,1,2, \dots, n)$ in Equation (2.2).

On the other hand, in a nonlinear regression problem, the model function will be a nonlinear function, for which a polynomial of degree two or more is usually preferred. In this case the model equation through an m^{th} degree polynomial of x will be

$$y_i = f(x_i, m) = \sum_{k=0}^{m} a_k \,.\, x_i^k + \varepsilon_i \qquad (i = 1,2, \dots, n). \tag{2.3}$$

Equation (2.3) can also be written in matrix form as

$$\overbrace{[y_1 \quad y_2 \quad \cdots \quad y_n]}^{\mathrm{Y}} = \overbrace{[a_0 \quad a_1 \quad \cdots \quad a_m]}^{\mathrm{A}} \cdot \overbrace{\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & & x_n \\ \vdots & & \ddots & \vdots \\ x_1^m & x_2^m & \cdots & x_n^m \end{bmatrix}}^{\mathrm{X}} + \overbrace{[\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]}^{\mathcal{E}}, \tag{2.4}$$

or in a compact form as

$$\mathrm{Y} = \mathrm{A} \cdot \mathrm{X} + \mathcal{E}, \tag{2.5}$$

which resembles the forms of Equations (2.1) and (2.2). Methods of linear algebra can then be utilized to solve this matrix equation.

2.3. Multilayer Perceptron (MLP)

There are several types of ANNs introduced in the literature. Among these, the multilayer perceptron (MLP) is a well-known and probably the most used network in which every node in a layer connects to each node in the following layer and exits on the output nodes. An MLP consists of at least three layers: an input layer accepting the inputs (also called features), one or more hidden layers processing the inputs, and an output layer producing the result. The basic structure of an MLP network is illustrated in Figure 2.3.

MLP employs activation functions, also known as transfer functions, in order to find the mathematical relationship between the input and the output patterns. Nonlinear activation functions such as logistic (sigmoid) and hyperbolic tangent (tanh) are usually utilized for data sets that cannot be separated linearly.

Usually, networks with one hidden layer consisting of sufficient number of neurons can adequately fit most finite input-output mapping problems. Therefore, in this study, MLP-based prediction models with one hidden layer are constructed, employing the logistic and the linear activation functions in the hidden and the output layers, respectively. The Levenberg-Marquardt back-propagation algorithm is used for the network training (learning) algorithm and the maximum number of passes of the training data set through the network, also called epochs, is limited to 100 in the model. Number of neurons in the hidden layer is decided by a trial-error approach, as detailed in §3.3.

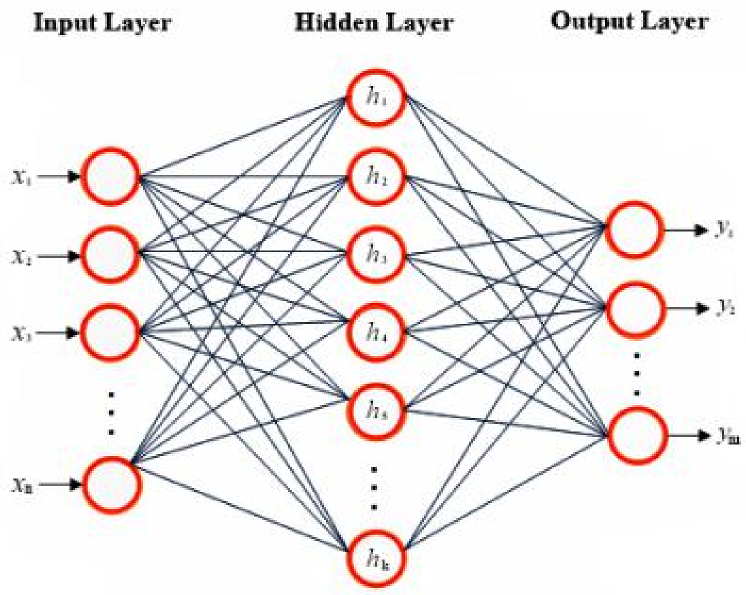

Figure 2.3. Structure of an MLP network with one hidden layer.

2.4. General Regression Neural Network (GRNN)

GRNN is a memory-based network that provides predictions of continuous variables, as in standard regression techniques, that are established based on a single-pass learning algorithm with a highly parallel structure. Since it is a type of supervised feed-forward neural network, the training of GRNN is usually fast because the data only propagates forward once. As seen in Figure 2.4, input, pattern, summation, and output layers form the GRNN which does not require backpropagation procedure to learn error from the training data, and can be used for regression, prediction, and classification (Specht, 1998; Acikkar, 2018; Mahasneh, 2018).

GRNN architecture has the Gaussian kernel function and a smoothing factor σ as the only free parameter in the network to be optimized and has the effect of smoothing the training instances, applied after the network's training is completed. Indeed, the aim of the training process is to find the optimum value of σ, which must be greater than zero, and is typically between 0.01 and 10.

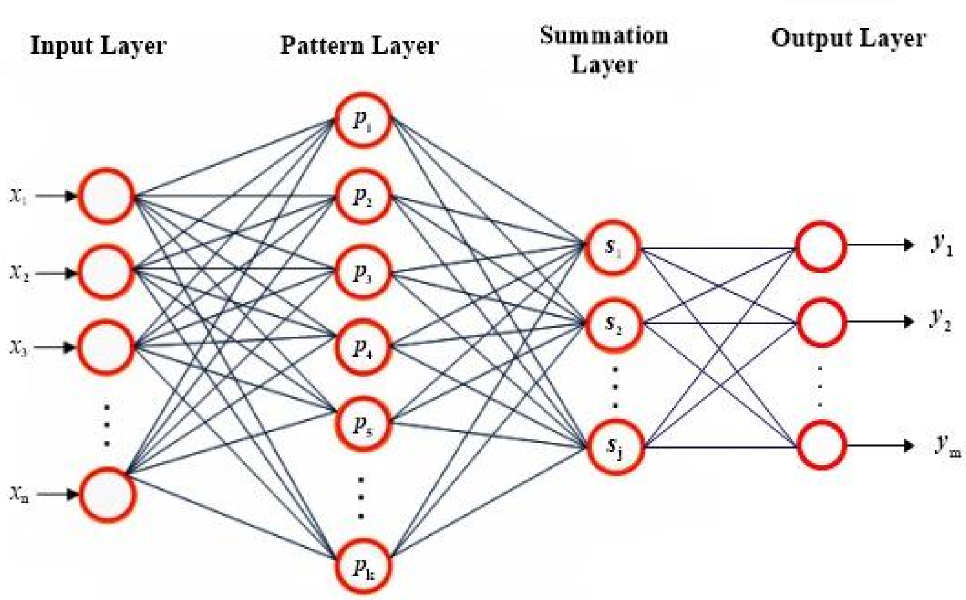

Figure 2.4. Structure of a typical GRNN.

2.5. Radial Basis Function Neural Network (RBFNN)

RBFNN was first introduced by Broomhead and Lowe (1988) as a three-layered alternative to feed-forward back-propagation neural network for both classification and regression problems. The kernel of RBFNN is a positive nonlinear symmetric radial basis function (RBF), the Gaussian function being the most common RBF.

RBFNN monotonically decreases with the distance from the center, mapping high values onto low ones, and mid-range values onto high ones, thanks to its bell-shaped curve which is also known as the normal distribution in statistics. Further, the parameters of RBFNN models are not homogeneous, making its learning process more complex.

2.6. Feature Selection

Features of the data that are used to train the models have great influence on the model's performance. Irrelevant or less relevant features might negatively impact the model (Shaikh, 2018). While feature selection (also known as feature importance) is accepted as an essential algorithm for reducing number of features in a data set in order to increase the speed of the learning process, it also helps understand and manage the data set more easily. Reduction of computational cost and over-fitting possibility and improvement of accuracy are among the other advantages of feature selection (Ding and Peng, 2005). Many systematic methods can be found to handle the selection process. In this study, the ReliefF method is used in order to rank the predictors according to their importance, which is based on their scores (Urbanowicz et al., 2018).

2.7. K-fold Cross-validation

Cross-validation is used to assure the generalization ability of the prediction model. In k-fold cross-validation the data is first segmented into k equally-sized partitions (folds) randomly, out of which k-1 folds are used for learning and the remaining fold is reserved for testing. Subsequently, k iterations of training and testing are performed and the accuracy of the prediction model is assumed to be the average accuracy of these iterations (Refaeilzadeh, 2009). Figure 2.5 illustrates the algorithm for k=10.

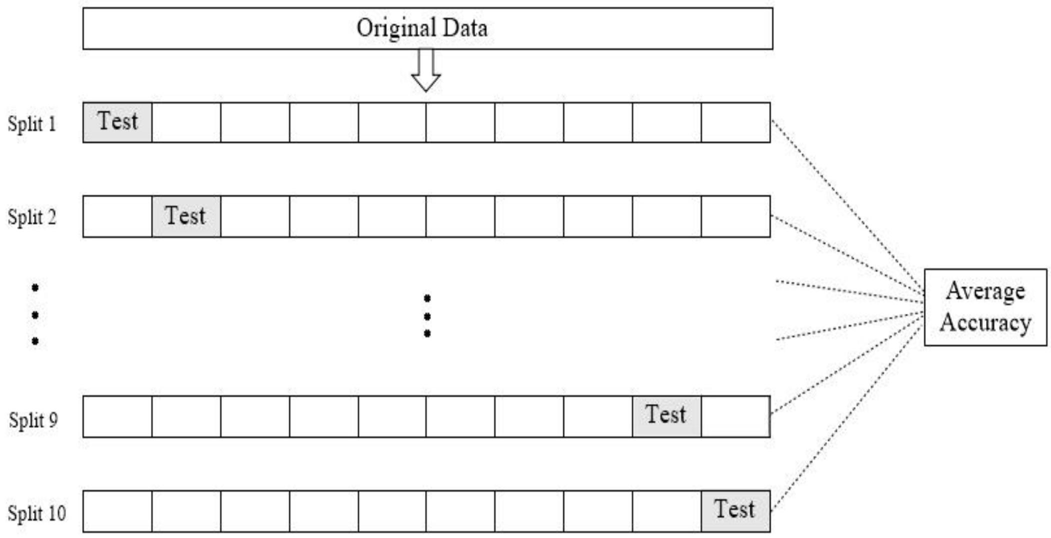

Figure 2.5. Implementation of the 10-fold cross-validation algorithm. For each of the 10 test sets, cross-validation is employed on the training set. The accuracy of each prediction is the average of the folds.

2.8. Performance Metrics

While constructing and testing an appropriate ANNs-based prediction models are real issues, evaluating the outputs is another matter. The model performance cannot be measured precisely unless the results are evaluated correctly. Performance metrics are the tools that evaluate the statistical performance of forecasting ability of the prediction model.

Makridakis et al. (1997) suggested the use of various metrics such as Mean Error, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error, Mean Absolute Percentage Error (MAPE), and Multiple Correlation Coefficient (R). Although there are many studies covering dozens of metrics, MSE, RMSE, MAE and MAPE found as the most popular ones in various scientific surveys (De Gooijer and Hyndman, 2006; Shcherbakov et al., 2013; Botchkarev, 2019). Among those listed above MSE, RMSE, MAE, MAPE, and R will be considered as performance measures in this study. While MSE will be considered in the optimization phase, the others will be used to test the performance of the models.

MSE, also known as square loss, is the most extensively used metric for regression models. It is the average of squares of the differences between predicted and actual values, and interprets how close a regression line is to a set of points. So, for a test data of n rows, if y_i is the actual (measured) value and $\hat{y}_i$ is the predicted value, then MSE is calculated as

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \, . \tag{2.6}$$

RMSE is the square root of MSE:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}. \tag{2.7}$$

MSE/RMSE measures the average magnitude of the error; hence, as expected, smaller MSE/RMSE is preferred.

MAE represents the average absolute difference between the actual and predicted values, i.e.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} | \, y_i - \hat{y}_i | \, , \tag{2.8}$$

while MAPE represents the percentage of average absolute relative error between the actual and predicted values, i.e.

$$\mathrm{MAPE} = 100 \times \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \tag{2.9}$$

As in the case for MSE/RMSE, the smaller MAE/MAPE is, more successful the prediction is.

The multiple correlation coefficient R indicates the strength of the relationship between actual and predicted values. It is given by

$$\mathrm{R} = \sqrt{1 - \left[\sum_{i=1}^{n} (y_i - \hat{y}_i) \right] \Big/ \left[\sum_{i=1}^{n} (y_i - \bar{y}) \right]}, \tag{2.10}$$

where $\bar{y}$ is the average of the actual values. Note that it is desired to have $\mathrm{R} \approx 1$.

3. IMPLEMENTATION OF PREDICTION MODELS

3.1. Data Set Generation

The data set to be processed by ANN models is generated through the analytical solution (A.7) described in the Appendix section below. This data set employs six input variables, often called predictors or features, and one response variable, namely the maximum run-up, $R_{\max}$.

Two different sliding mass profiles are imposed over the sea bottom. These profiles are indicated by the code b_p in the data set: $b_p = 1$ designates the exponential (i.e. Gaussian) profile defined as $h_e(\xi, t) = \exp\left(-(\xi - \xi_0 - t)^2\right)$, while $b_p = 2$ designates the hyperbolic profile defined as $h_s(\xi, t) = \operatorname{sech}^2(\xi - \xi_0 - t)$.

The other predictors considered in the prediction models are the beach slope angle β, the initial slide submergence x_0, the landslide vertical thickness δ, the landslide horizontal length L, and the instant (time) t when the maximum run-up is observed (Figure 3.1). The beach slope angle is varied from $2°$ to $20°$ to scatter mild to moderate slopes. The choice of initial slide submergence determines whether the landslide is subaerial or submarine. Since this study focused on submarine landslides, x_0 is varied between 1.5 and 6 (dimensionless). While dimensionless parameters are used in the analytical model for convenience (see the Appendix for details), the data has been converted into dimensional form in order to be processed by prediction models. Accordingly, the spatial variable $\xi = 2\sqrt{\mu x / \tan\beta}$ introduced in Equation (A.5) for the sake of a convenient analytical solution is converted back into x. As a consequence, the submergence parameter $\xi_0 = 2\sqrt{\mu x_0 / \tan\beta}$ used in the definition of the bottom profiles is replaced in our data set by the dimensional parameter in the x-coordinate, defined by

$$x_0 = \frac{\tan\beta}{\delta}\left(\frac{1}{2}L\,\xi_0\right)^2. \tag{3.1}$$

Figure A.2 in the Appendix demonstrates the dependence of the target variable, i.e. the maximum tsunami run-up ($R_{\max}$) on the predictors defined above, namely β, x_0, δ, L, and t, for the exponential bottom profile.

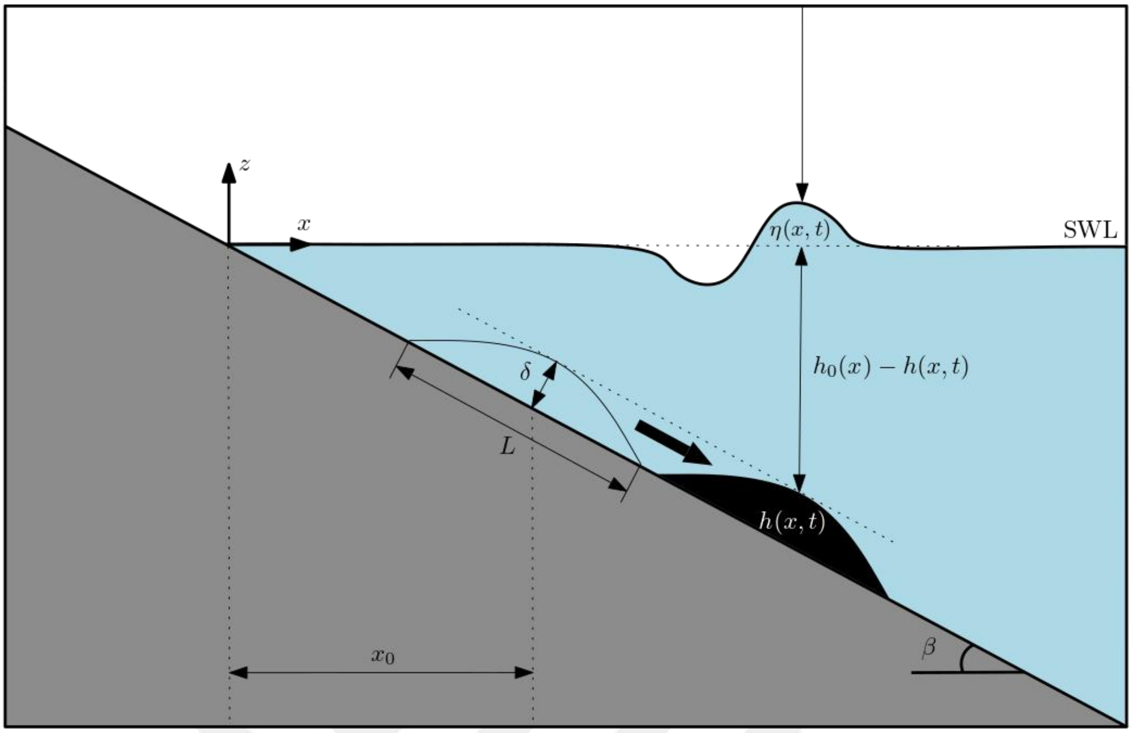

Figure 3.2. Definition sketch for the landslide problem (not to scale). δ and L respectively indicate the maximum vertical thickness and the maximum horizontal length of the sliding mass at its initial location (x_0). The undisturbed water depth is $h_0(x) = x \tan \beta$, where β is the beach angle with the horizontal. SWL stands for the still water level. (Modified from Aydın, 2021).

In our methodology the vertical landslide thickness δ is calculated by multiplying a randomly selected horizontal slide length $10 \text{ m} < L < 500 \text{ m}$ with the slide aspect ratio chosen in the interval $0.01 \leq \mu \leq 0.2$, which is determined based on the thin-slide assumption $\mu = \delta/L \ll 1$.

This methodology generated a data set consisting of 57,926 rows and 7 columns. The maximum run-up, $R_{\max}$, is further narrowed down to vary between 0.5 m and 25 m so that our data set reflects a realistic range observed in the measurements from historical tsunamis (Harbitz, 1992; Yeh et al., 1993; Assier-Rzadkiewicz et al., 2000; Fine et al., 2005; Bondevik et al., 2005; Tinti et al., 2005; Ten Brink et al., 2006; Tappin et al., 2008; Geist et al., 2009), although smaller and larger values were obtained for the predictor ranges defined above. Under this limitation the data set size eventually reduced to 54,952 rows.

3.2. Network Selection

For a given machine learning problem, there is the possibility of choosing among a vast number of methods. Even within ANNs, a variety of networks is available for use. In this section, results of a preliminary network selection study carried out to determine the optimum method in terms of error magnitude and computational cost is presented. Computations are performed with the three networks introduced in Chapter 2, namely MLP, GRNN and RBFNN, over a 10,000-row subset of the whole data set which is randomly selected. Each network is executed for 10-folds and 4-subfolds (see the next section for implementation details) and the results are tabulated in Table 3.1. Although the performance metric results are comparable to a certain degree, there is significant difference in execution times. Considering that each model will be run multiple times, details of which are presented in Chapter 4, proceeding with GRNN or RBFNN would require much longer computer time and would not bring any advantages in terms of model performances. Therefore, based on the results in Table 3.1, it is decided to proceed with MLP, which is superior in terms of both error magnitude and also computer time.

Table 3.1. Performance metrics and execution times (reported in hh:mm:ss format) for the three widely used networks are compared in an aim to select the optimum network.

Network	RMSE (m)	MAE (m)	MAPE (%)	Execution Time
MLP	0.0630	0.0435	1.4029	00:03:27
GRNN	0.4049	0.2427	3.8617	03:25:33
RBFNN	0.1167	0.0789	2.0820	00:43:50

3.3. Implementation of MLP Models

The flowchart of MLP-based prediction models and the feature selection algorithm is presented in Figure 3.2. As stated before, MLP is preferred in the present study due to its ability to adequately fit finite input-output mapping problems (see §2.1). The number of predictors (input variables) is six, and the maximum run-up is the only output of the data set (see §3.1).

Implementation steps of the MLP-based prediction models are illustrated in Figure 3.2(a). The cross-validation algorithm with k = 10 is first used to randomly split the data set into ten subsets, nine of which are used as training data and the remaining tenth is used as test data as illustrated in Figure 2.5. Then, for each fold, the training data is used to determine the optimum hidden layer size, i.e. the optimum number of neurons used in the hidden layer, which is a crucial step in the MLP. Afterwards, the prediction model constructed with the optimized number of neurons in the hidden layer is tested against the test data and RMSE, MAE, MAPE and R are calculated for the fold. Finally, the performance of the prediction model is determined by calculating each metric's average over the folds.

In a similar fashion, while optimizing the hidden layer size, each training data set is further split into four subsets (i.e. 4-fold cross-validation), three of which are used for training and the remaining quarter is reserved for test. The number of neurons is then changed between 2 and 20, the average MSE is recorded in each case and the neuron number with minimum average MSE is assigned to be the optimum number of neurons in the hidden layer. It is important here to note that the fold's test data is not used in this optimization phase.

Implementation of the feature selection algorithm is described in Figure 3.2(b). Six models are constructed for the six predictors, by disregarding the feature with the worst score in each following model (see §4.3 for the details).

All computations are performed in the MATLAB™ environment (Release 2019a, Update 4) on a Microsoft Windows 10 (64-bit) workstation.

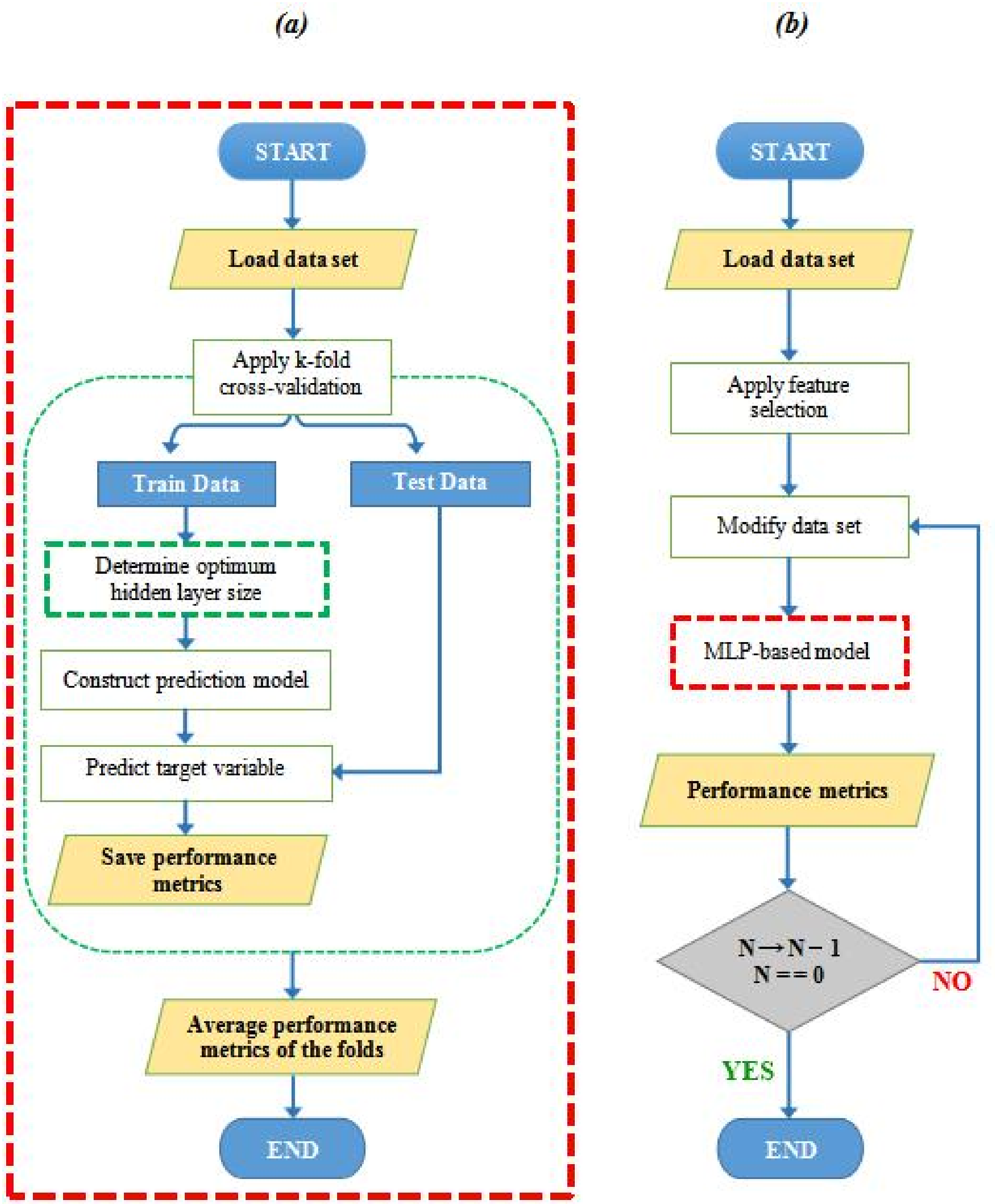

Figure 3.2. (a) Flowchart showing implementation steps of the MLP-based prediction models. (b) Flowchart showing implementation steps of the feature selection algorithm.

4. RESULTS AND DISCUSSIONS

4.1. Data Set Selection

In order to reduce the computational cost, subsets of various size are sampled out of the 54,952-row whole data set and a preliminary data set selection study is performed. For this purpose, eight data sets having 3,000 to 24,000 randomly chosen rows with 3,000 increments are generated, corresponding approximately to 5% to 40% of the whole data set, respectively, and a smaller representative data set is selected by means of the method described in §3.3, based on the average MSE value of the subsets, as the use of MSE as performance metric in the optimization phase is a standard approach.

The average MSE values from five repeated runs are plotted against the size of the data set in Figure 4.1. The results show that MSE is not improving for subsets larger than 9,000 rows. The scatter plots (i.e. maximum run-up distributions with respect to each feature, except the bottom profile and the beach slope angle) are also compared in Figure 4.2 for the whole 54,952-row data set and the 9,000-row subset, which will be referred to as 9K hereinafter. The figure shows that the 9K subset mimics the whole data set for all features and hence it can be used to represent the whole data set. The descriptive statistics of the variables in the 9K subset is tabulated in Table 4.1 and a box-whiskers plot is provided in Figure 4.3. The mean and the standard deviation of the variables in the 9K subset are in great agreement with those of the whole data set, as shown in the relevant columns of Table 4.1.

Based on the statistical experiments above, it is consequently decided to perform calculations with the 9K data set for the purpose of proposing prediction models with low computational cost. This approach brought great time economy since the models are run multiple times to assure stable results.

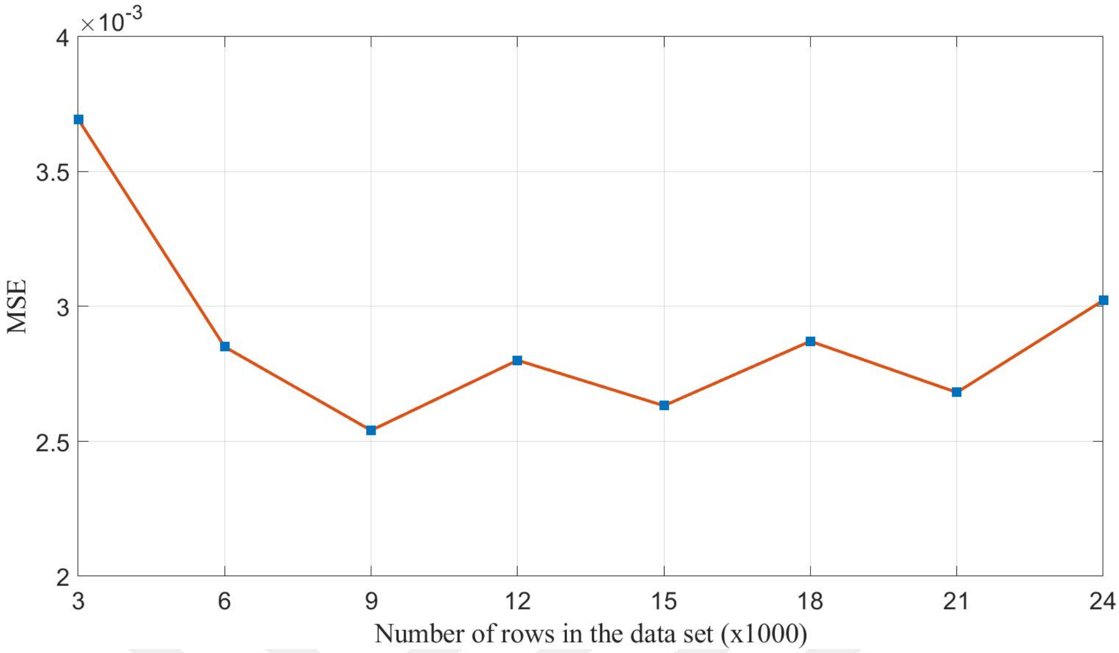

Figure 4.1. Variation of average MSE with the data set size.

Table 4.1. Descriptive statistics of the variables in the 9K data set except for the bottom profile code. The means and the standard deviations for the whole data set are also provided.

Variable	Minimum	Maximum	Mean		Standard Deviation	
	9K	9K	9K	Whole Data	9K	Whole Data
x_0 (m)	15.29	2999.84	836.48	841.70	680.30	684.05
β (degrees)	2	20	9.86	9.84	5.87	5.88
δ (m)	0.76	65.57	18.33	18.42	12.58	12.60
L (m)	10.14	499.99	218.94	220.17	131.23	131.29
t (sec)	7.42	292.51	72.63	73.03	39.25	39.43
$R_{\max}$ (m)	0.50	24.99	9.64	9.69	6.84	6.85

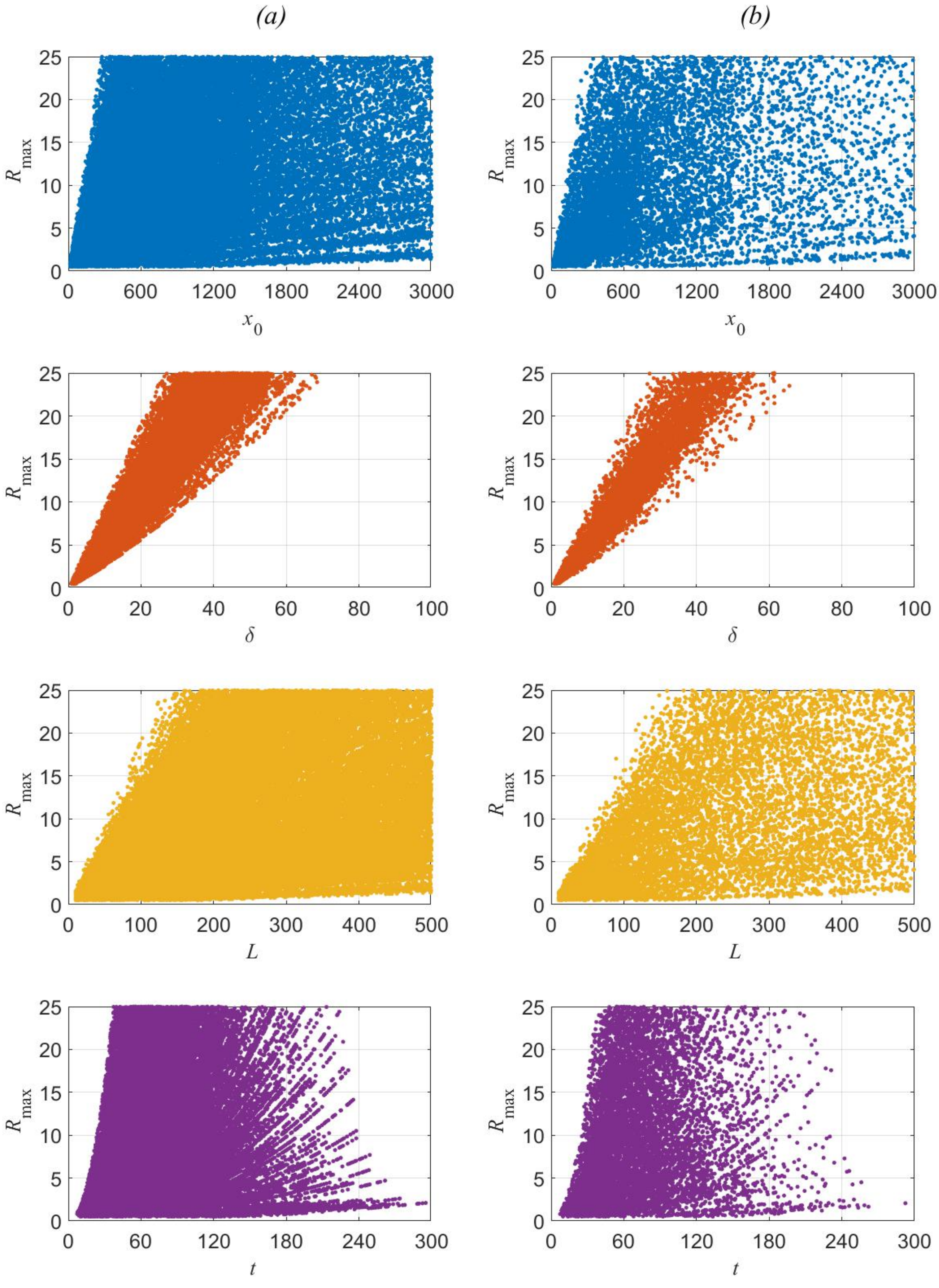

Figure 4.2. Scatter plots of the output variable for *(a)* the whole data set with 54,952 rows, and *(b)* the 9K data set. The predictors from top to bottom are the initial slide submergence, the maximum slide thickness, the maximum slide length, and the time of the maximum run-up.

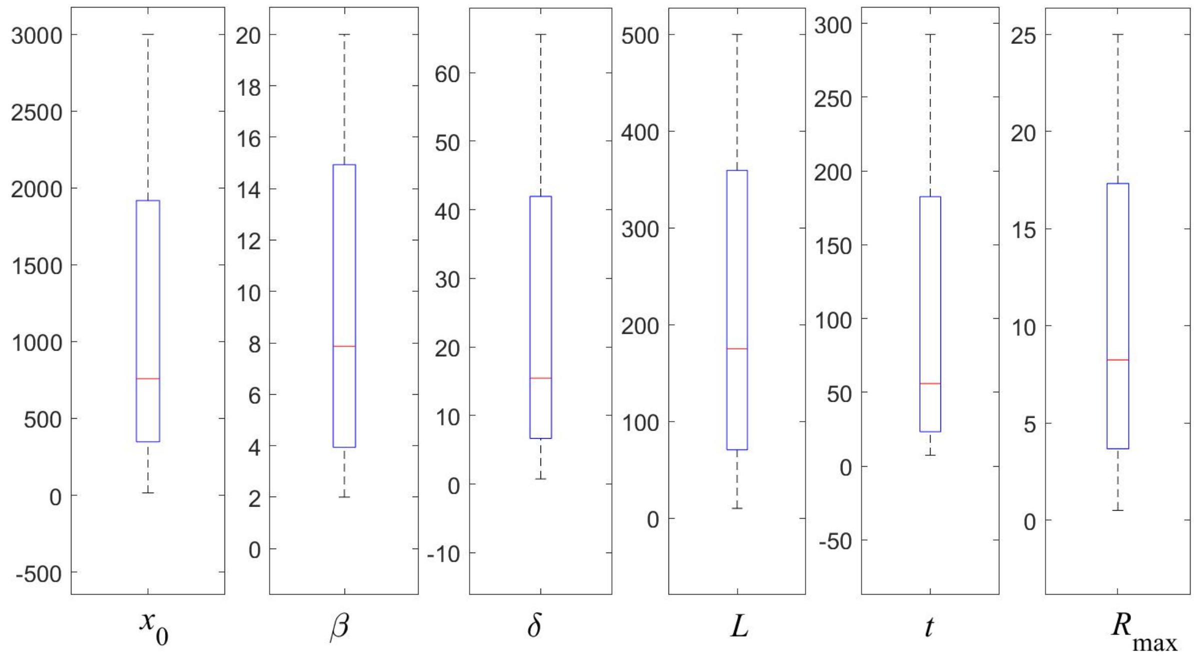

Figure 4.3. Box-whiskers plot for the descriptive statistics of the 9K data set.

4.2. Model Results

A total number of 63 models are constructed for all possible combinations of the six predictors. The models are named in such a way that, in addition to the model number, they also indicate the number of predictors considered in the model (the digit before the dash, see Table 4.2). Each model is run 30 times, with 10-folds and 4-subfolds and average RMSE, MAE, MAPE and R values are recorded as performance metrics. The results reported for each prediction model are the average results of 30 repeated executions of the code. This approach minimizes any possible abnormal increase or decrease in the performance metrics due to randomness and stables the results. The results are tabulated in Tables 4.3-4.7.

The results of single predictor models presented in Table 4.3 clearly indicate that the vertical landslide thickness (δ) has the best performance metrics, and hence is the most influential parameter on the maximum tsunami run-up ($R_{\max}$), compared to the other predictors. The initial slide submergence (x_0), the horizontal length (L), and the instant of maximum run-up (t) seem to be comparable, as they have similar performance metric values. On the other hand, the bottom profile (b_p) and the beach slope angle (β) have almost no correlation with $R_{\max}$, as can be seen from their multiple correlation coefficient (R) values.

Table 4.2. Example model codes and considered and disregarded predictors.

Model Code	Considered Predictors	Disregarded Predictors	Model Code	Considered Predictors	Disregarded Predictors
M1-1	b_p	x_0, β, δ, L, t	M2-7	b_p, x_0	β, δ, L, t
M1-2	x_0	b_p, β, δ, L, t	M3-22	b_p, x_0, β	δ, L, t
M1-3	β	b_p, x_0, δ, L, t	M4-42	b_p, x_0, β, δ	L, t
⋮	⋮	⋮	M5-57	$b_p, x_0, \beta, \delta, L$	t
M1-6	t	$b_p, x_0, \beta, \delta, L$	M6-63	$b_p, x_0, \beta, \delta, L, t$	-

Table 4.3. Average performance metrics for the models with a single predictor. The results are sorted in increasing RMSE values.

Model	Predictors	RMSE	MAE	MAPE	R
M1-4	δ	1.8507	1.2917	15.7997	0.9626
M1-5	L	5.7978	4.6905	106.7620	0.5295
M1-2	x_0	5.9818	4.8467	113.9974	0.4837
M1-6	t	6.3529	5.2529	134.5559	0.3628
M1-1	b_p	6.8413	5.8229	164.0551	0.0122
M1-3	β	6.8463	5.8268	164.1301	0.0032

The results for models with two predictors in Table 4.4 also confirm the influence of the slide thickness, as absence of δ causes significant decrease in performance metrics. x_0, L and t again have comparable results. However, when combined with the same predictor, slightly better metric values are observed in the models including L, compared to those including x_0 and t. Finally, b_p and β appear as the predictors with worst metric results, as above.

Similar trends are further observed in three- and four-predictor models (Tables 4.5 and 4.6, respectively). Whilst models with δ have better performance metrics, a sudden increase in RMSE, MAE and MAPE and decrease in R is observed in the models excluding δ. For instance, models employing δ have R > 0.95, while models disregarding it have R < 0.75.

Among the models without δ, the second group is now formed with t, while it is difficult to distinguish the priority of L or x_0. The models with worst metrics again include b_p and β.

As an exceptional case, the model M4-54 (first row in Table 4.6), which includes the predictors other than x_0 and β, yields almost the same results as M6-63 with all six predictors (Table 4.7). Both models having an R value of 1.00 indicates that the model M4-54 explains the maximum run-up behavior equally well as the model M6-63.

Table 4.4. Average performance metrics for the models with two predictors. The results are sorted in increasing RMSE values.

Model	Predictors	RMSE (m)	MAE (m)	MAPE (%)	R
M2-12	β, δ	1.5973	1.1461	14.2413	0.9723
M2-16	δ, L	1.7049	1.1541	13.4427	0.9684
M2-11	x_0, δ	1.7542	1.2220	15.3875	0.9665
M2-20	δ, t	1.7731	1.2491	15.3920	0.9657
M2-10	b_p, δ	1.7742	1.2458	15.1986	0.9657
M2-21	L, t	5.5293	4.4420	95.3700	0.5878
M2-18	x_0, t	5.6701	4.5437	101.7133	0.5581
M2-14	x_0, L	5.6988	4.5776	103.0864	0.5521
M2-15	β, L	5.7905	4.6880	106.8441	0.5312
M2-13	b_p, L	5.7993	4.6963	107.0326	0.5276
M2-7	b_p, x_0	5.9789	4.8459	113.9975	0.4841
M2-9	x_0, β	5.9820	4.8473	114.3393	0.4835
M2-19	β, t	6.0381	4.9027	119.9511	0.4685
M2-17	b_p, t	6.2985	5.1964	131.6817	0.3881
M2-8	b_p, β	6.8457	5.8241	163.9666	0.0079

Table 4.5. Average performance metrics for the models with three predictors. The results are sorted in increasing RMSE values.

Model	Predictors	RMSE	MAE	MAPE	R
M3-41	δ, L, t	1.1973	0.8420	10.0777	0.9845
M3-24	b_p, β, δ	1.3233	0.9702	12.2200	0.9811
M3-31	β, δ, L	1.4082	1.0012	11.8356	0.9785
M3-30	x_0, δ, L	1.4239	0.9423	10.8694	0.9780
M3-25	x_0, β, δ	1.4905	1.0774	14.0514	0.9759
M3-37	β, δ, t	1.5428	1.0982	13.4889	0.9742
M3-35	b_p, δ, t	1.5796	1.1217	14.2766	0.9729
M3-29	b_p, δ, L	1.6203	1.1102	12.9140	0.9715
M3-23	b_p, x_0, δ	1.6657	1.1740	14.8054	0.9698
M3-36	x_0, δ, t	1.6685	1.1718	15.4083	0.9697
M3-34	x_0, β, t	4.6243	3.5337	73.2624	0.7362
M3-40	β, L, t	4.9510	3.8338	74.4543	0.6890
M3-39	x_0, L, t	5.0365	3.9457	81.7642	0.6760
M3-38	b_p, L, t	5.5181	4.4305	95.3526	0.5899
M3-32	b_p, x_0, t	5.6365	4.5075	101.1317	0.5656
M3-28	x_0, β, L	5.6662	4.5464	102.4423	0.5592
M3-26	b_p, x_0, L	5.6984	4.5816	103.0223	0.5521
M3-27	b_p, β, L	5.7937	4.6941	107.1909	0.5296
M3-22	b_p, x_0, β	5.9675	4.8262	113.8302	0.4874
M3-33	b_p, β, t	6.0161	4.8845	119.7479	0.4746

Table 4.6. Average performance metrics for the models with four predictors. The results are sorted in increasing RMSE values.

Model	Predictors	RMSE	MAE	MAPE	R
M4-54	b_p, δ, L, t	0.0514	0.0361	1.1097	1.0000
M4-50	x_0, β, δ, t	0.7508	0.5245	7.7532	0.9939
M4-46	x_0, β, δ, L	1.0432	0.7333	8.9397	0.9883
M4-45	b_p, β, δ, L	1.0539	0.7679	9.1652	0.9880
M4-56	β, δ, L, t	1.1330	0.7963	10.0969	0.9862
M4-42	b_p, x_0, β, δ	1.1627	0.8434	11.6082	0.9854
M4-55	x_0, δ, L, t	1.1662	0.8115	9.7060	0.9853
M4-49	b_p, β, δ, t	1.2484	0.9038	11.5121	0.9832
M4-44	b_p, x_0, δ, L	1.2797	0.8884	10.3090	0.9823
M4-53	x_0, β, L, t	1.4078	1.0606	22.8188	0.9784
M4-48	b_p, x_0, δ, t	1.4251	0.9860	14.0556	0.9780
M4-47	b_p, x_0, β, t	4.6062	3.5330	73.7903	0.7386
M4-52	b_p, β, L, t	4.9884	3.8891	75.5164	0.6832
M4-51	b_p, x_0, L, t	4.9894	3.8975	80.6489	0.6832
M4-43	b_p, x_0, β, L	5.6742	4.5586	102.9691	0.5573

Table 4.7 further shows that replacing L by t results in noticeably worse metrics (see models M5-57 and M5-58, respectively). It is not surprising to see that the model with the worst metrics is the one disregarding δ, i.e. M5-59, which, one again, resembles the pattern above.

The average RMSE values reported in the third columns of Tables 4.3-4.7 are also plotted in Figure 4.4. Decrease is observed in the error as the number of predictors in the models increases, as expected. It is also clearly seen from the figure that, for each group with a given number of predictors, the models with minimum RMSE value employ δ (indicated by a filled marker).

Table 4.7. Average performance metrics for the models with five and six predictors. The results are sorted in increasing RMSE values.

Model	Predictors	RMSE	MAE	MAPE	R
M6-63	$b_p, x_0, \beta, \delta, L, t$	0.0513	0.0369	1.1084	1.0000
M5-60	b_p, x_0, δ, L, t	0.0556	0.0395	1.2029	1.0000
M5-61	b_p, β, δ, L, t	0.0576	0.0419	1.2735	1.0000
M5-57	$b_p, x_0, \beta, \delta, L$	0.0767	0.0557	1.5613	0.9999
M5-58	$b_p, x_0, \beta, \delta, t$	0.2513	0.1805	4.1333	0.9993
M5-62	x_0, β, δ, L, t	0.3288	0.2332	5.5907	0.9988
M5-59	b_p, x_0, β, L, t	1.3223	0.9756	25.2243	0.9806

4.3. Application of the Feature Selection Algorithm

In an effort to see the influence of each feature on the maximum run-up, the features are ranked by using the ReliefF feature importance method. In this method the predictors ranked according to their influence on the output by assigning a score to each predictor in such a way that predictor with a smaller score has less effect on the output compared to another predictor with a higher score. Table 4.8 tabulates the ReliefF scores the six predictors considered in this study.

A backward elimination procedure is then utilized and six models are constructed based on the predictors' ReliefF scores. While all features are employed in the first model (this is actually the model M6-63 above), the feature with the worst score is disregarded in the subsequent model (Table 4.9). The performance metrics of each model are plotted in Figure 4.5.

Table 4.8. Predictors of the 9K data set are ranked according to their ReliefF scores.

Predictor	δ	t	b_p	x_0	L	β
ReliefF Score	0.0268	0.0011	0	-0.0010	-0.0016	-0.0042

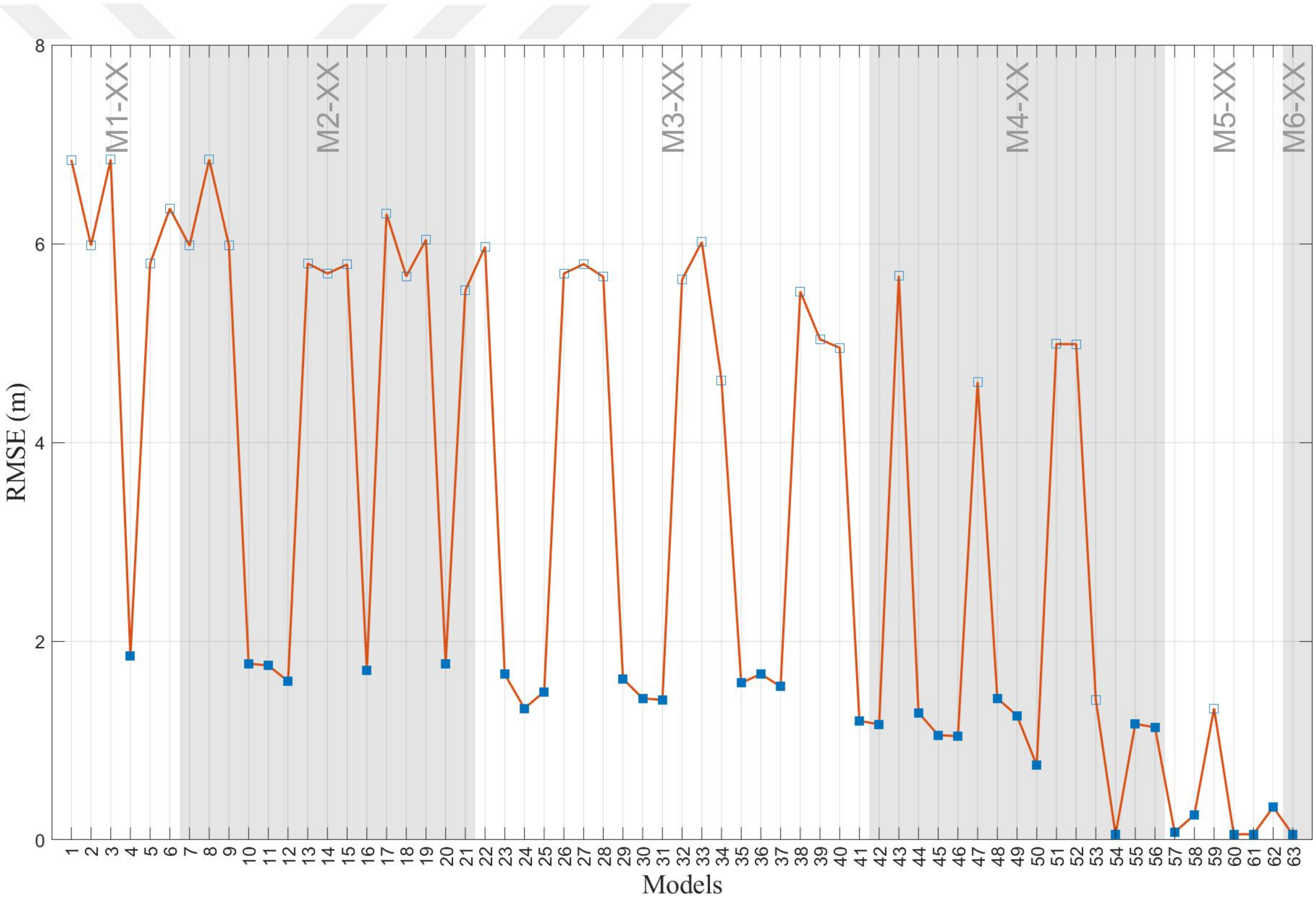

Figure 4.4. Average RMSE values of the combination models tabulated in Tables 4.3-4.7.

The ReliefF model results tabulated in Table 4.9 partly differs from those reported in the previous section. The model M6-63, which includes all features, and the model M5-60, which disregards the predictor with the worst score (the beach slope angle β) as listed in Table 4.8, are at the same order, confirming that β is the least effective parameter on the maximum run-up of landslide-generated tsunamis, $R_{\max}$. This result is in accordance with the previous results reporting β as one of the two predictors that have minimal effect on $R_{\max}$, along with b_p.

When the landslide horizontal length (L) is further excluded (M4-48), however, there are momentous changes in the performance metrics. This shows the significance of L among the considered landslide parameters. Nevertheless, the ReliefF algorithm ranks L as the second worst predictor and b_p as a moderate predictor. This appears to be the major difference between utilizing and not utilizing feature selection.

As expected, performance of the models gets worse when the initial slide submergence (x_0), the bottom profile code (b_p), and the time of maximum run-up (t) are further ignored in M3-35, M2-20, and M1-4, respectively. Based on the results presented in Tables 4.8 and 4.9, it is apparent that the ReliefF algorithm predicts the parameter having the greatest impact on $R_{\max}$ to be the landslide vertical thickness (δ), among the predictors considered in this study, as in the previous section.

Table 4.9. Average performance metrics of the models constructed for the 9K data set in view of the ReliefF feature importance method.

Model	Predictors	RMSE (m)	MAE (m)	MAPE (%)	R
M6-63	$b_p, x_0, \beta, \delta, L, t$	0.0513	0.0369	1.1084	1.0000
M5-60	b_p, x_0, δ, L, t	0.0556	0.0395	1.2029	1.0000
M4-48	b_p, x_0, δ, t	1.4251	0.9860	14.0556	0.9780
M3-35	b_p, δ, t	1.5796	1.1217	14.2766	0.9729
M2-20	δ, t	1.7731	1.2491	15.3920	0.9657
M1-4	δ	1.8507	1.2917	15.7997	0.9626

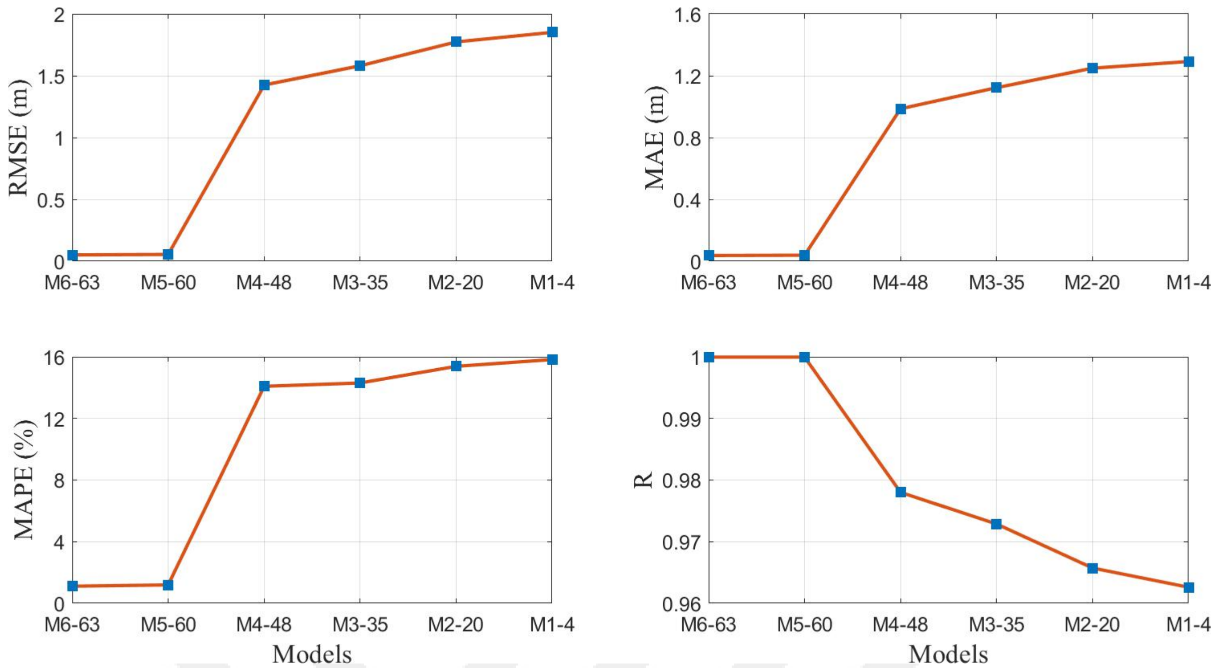

Figure 4.5. Performance metrics of the ReliefF models tabulated in Table 4.9.

4.4. Comparison with the Literature

In spite of the simple character of the analytical model employed to generate the data source, the results presented above have shown good agreement with the literature. Comprehensive experimental studies for solid-block (Lynett and Liu, 2005) and deformable granular landslides (Mohammed and Fritz, 2011) sliding over an inclined plane revealed the significance of slide thickness as a primary factor governing the maximum run-up, while it is influenced secondarily by slide volume, implying the role of slide length (and width in the case of three-dimensional analysis). A similar conclusion was also drawn for circular geometry (i.e., conical island) by Romano et al. (2016), who found out that landslide thickness significantly affects the wave amplitudes in the near field.

Very recently, Sabeti and Heidarzadeh (2020) carried out a sensitivity analysis to determine the impact of landslide parameters such as initial submergence, length, width, and thickness of the slide as well as slope angle on the initial wave amplitude. Their results indicated that slide volume and initial submergence are the most dominant parameters and that slide thickness is also very effective while slope angle and slide length have limited impact on the initial wave amplitude. Their findings confirm those of Jiang and Le Blond (1992), who investigated the relationship of a submarine slide and surface water waves that the slide generates and found that the two major parameters determining the interaction between the slide and the water waves are the density of sliding material and the initial submergence of the slide. Although these studies focused on to predict the maximum initial wave amplitude given the volume and the initial submergence of the slide, their findings support our results as well.

One final remark to make is on the beach slope angle (β), which has shown to have minimal effect on the maximum run-up ($R_{\max}$). Analytical (Liu et al., 2003; Aydın, 2021) and experimental studies (Lynett and Liu, 2005) have shown that, unlike the case of subaerial slides, the maximum run-up depends on the beach slope in the case of submarine slides. This does not contradict with the results presented above. Our findings can rather be interpreted as an indication of a limited impact of β on $R_{\max}$, compared to the other parameters considered in the analysis.

5. CONCLUSIONS

In this thesis study, prediction models based on multilayer perceptron (MLP) are proposed in an attempt to predict the maximum run-up of landslide-generated tsunami waves and assess the role of parameters affecting the maximum run-up. For this purpose, a data set of approximately 55,000 rows is first generated from an analytical solution. This data set employs six landslide parameters which are the cross section of landslide profile, the beach slope angle, the initial slide submergence, the vertical thickness and the horizontal length of the sliding mass, and the time of the maximum run-up, along with the maximum tsunami run-up as the target variable. The whole data set was randomly sampled into multiple subsets which then underwent a preliminary optimization analysis and finally reduced into a 9,000-row data set. This data set successfully represents the statistical properties, yet is smaller than one sixth of the original data set, and hence is economical in terms of computational load.

A total number of 63 models consisting all possible combinations of the six parameters were then constructed for the 9,000-row data set, in order to predict the maximum run-up. The prediction capability of the models was assessed by calculating RMSE, MAE, MAPE, and R as performance metrics. The MLP-based models led predictions with a minimum MAPE of approximately 1.1% and with R=1. A backward elimination algorithm furnished with ReliefF feature importance is also utilized as a practical approach and similarities and differences with the results of the combination models are presented.

Consequently, based on the methodology introduced, the slide thickness (δ) was shown to be the most effective parameter on the maximum run-up of landslide tsunamis, while the cross-section of the slide profile (b_p) and the beach slope angle (β) appeared as the least effective parameters. The initial slide submergence (x_0), the horizontal slide length (L) and the time of maximum run-up (t) had moderate effects.

Comparison with existing literature showed the reliability and applicability of the offered models. This approach can be therefore suggested as a fast and accurate methodology to support, or when there is lack of, analytical or numerical modeling tools.

6. RECOMMENDATIONS

For future work, application of the proposed methodology to data collected in the field or generated in the laboratory may be suggested. Moreover, inclusion of further parameters accounting for steepness of the slide cross section and slide material would enhance the prediction capability of the models. If a three-dimensional analysis is attempted, slide volume and cross-section in the transverse direction can be considered as additional landslide features.

REFERENCES

Assier-Rzadkiewicz, S., Heinrich, P., Sabatier, P. C., Savoye, B., & Bourillet, J. F. (2000). Numerical modelling of a landslide-generated tsunami: The 1979 Nice event. *Pure and Applied Geophys*, *157*, 1717–1727.

Aydın, B. (2021). An analytical study on tsunami run-up due to submarine landslides from different bottom profiles. *Afyon Kocatepe University Journal of Science and Engineering 21(2)* 426–433. doi: 10.35414/ akufemubid.819348.

Beisel, S., Chubarov, L., Fedotova, Z., & Khakimzyanov, G. (2007). On the approaches to a numerical modeling of landslide mechanism of tsunami wave generation. *Communications on Pure and Applied Analysis, 11*(1) 121–135.

Bondevik, S., Løvholt, F., Harbitz, C., Mangerud, J., Dawson, A., & Svendsen, J. (2005). The Storegga slide tsunami—comparing field observations with numerical simulations. *Marine Petroleum Geolology*, *22*, 195-208. doi: 10.1016/j. marpetgeo.2004.10.003.

Botchkarev, A. (2019). A New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms. *Interdisciplinary Journal of Information, Knowledge, and Management,* 14. 45-79. doi: 10.28945/4184.

Brideau, M.-A., Sturzenegger, M., Stead, D., et al. (2011). Stability analysis of the 2007 Chehalis Lake landslide based on long-range terrestrial photogrammetry and airborne LiDAR data. *Landslides, 9*(1) 75–91. doi: 10.1007/s10346-011-0286-4.

Broomhead, D.S. & Lowe, D. (1988). Multi-Variable Functional Interpolation and Adaptive Networks. *Complex Systems, 2*, 327-355

Crosta, G. B., Imposimato, S., & Roddeman, D. (2016). Landslide spreading, impulse water waves and modelling of the Vajont rockslide. *Rock Mechanics and Engineering Geology, 49*(6), 2413–2436. doi: 10.1007/s00603-015-0769-z.

De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting, 22*(3), 443–473. doi: 10.1016/j.ijforecast.2006.01.001.

Delua, J. (2021). Supervised vs. Unsupervised Learning: What's the Difference? https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning. Accessed on February 06, 2022.

Di Risio, M., Bellotti, G., Panizzo, A., & Girolamo, P. D. (2009a). Three-dimensional experiments on landslide generated waves at a sloping coast. *Coastal Engineering, 56*(5), 659–671. doi: 10.1016/j.coastaleng.2009.01.009.

Di Risio, M., Girolamo, P. D., Bellotti, G., Panizzo, A., Aristodemo, F., Molfetta, M., Petrillo, A. (2009). Landslide-generated tsunamis runup at the coast of a conical island: New physical model experiments. *Journal of Geophysical Research, 114*(C01009). doi: 10.1029/2008JC004858.

Ding, C. & Peng, H. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology, 3*(2) 185–205.

Enet, F. & Grilli, S. (2007). Experimental study of tsunami generation by three dimensional rigid underwater landslides. *Journal of Waterway, Port, Coastal, and Ocean Engineering, 133*(6), 442–454. doi: 10.1061/(ASCE)0733-950X(2007)133:6(442).

Erdurmaz, M. S. (2004). Neural network prediction of tsunami parameters in the Aegean and Marmara Seas. Master Thesis, The Graduate School of Natural and Applied Sciences, Middle East Technical University.

Evers, F., Heller, V., Fuchs, H., Hager, W., & Boes, R. (2019). *Landslide generated impulse waves in reservoirs - Basics and Computation* (2nd ed.). Laboratory of Hydraulics, Hydrology and Glaciology VAW, ETH Zurich.

Fine, I. V., Rabinovich, A. B., Bornhold, B. D., Thomson, R. E., & Kulikov, E. A. (2005). The Grand Banks landslide-generated tsunami of November 18, 1929: Preliminary analysis and numerical modelling. *Marine Geology*, *215*, 45–57. doi: 10.1016/j.margeo.2004.11.007.

Fritz, H. M., Hager, W. H., Minor, H.-E. (2001). Lituya Bay case rockslide impact and wave run-up. *Science of Tsunami Hazards*, *19*(1), 3–22.

Fritz, H. M. & Moser, P. (2003). Pneumatic Landslide Generator. *International Journal of Fluid Power, 4*(1), 49–57. doi: 10.1080/14399776.2003.10781155.

Fritz, H. M., Hager, W. H., & Minor, H.-E. (2003a). Landslide generated impulse waves. *Experiments in Fluids, 35*(6), 505–519. doi: 10.1007/s00348-003-0659-0.

Fritz, H. M., Hager, W. H., & Minor, H.-E. (2003b). Landslide generated impulse waves, Part 2: Hydrodynamic impact craters. *Experiments in Fluids*, 35, 520–532.

Fritz, H. M., Hager, W. H., & Minor, H.-E. (2004). Near Field Characteristics of Landslide Generated Impulse Waves. *Journal of Waterway, Port, Coastal, and Ocean Engineering, 130*(6), 287–302. doi: 10.1061/(ASCE)0733-950X(2004)130:6(287).

Fritz, H. M., Mohammed, F., Yoo, J. (2009). Lituya Bay landslide impact generated mega-tsunami 50th anniversary. *Pure and Applied Geophysics, 166*(1), 153–175. doi: 10.1007/s00024-008-0435-4.

Fuchs, H. & Boes, R. (2010). Computation of rockfall-induced impulse waves in the Lake of Lucerne. *Wasser Energie Luft*, *102*(3), 215–221 (in German).

Geist, E. L., Lynett, P. J., & Chaytor, J. D. (2009). Hydrodynamic modeling of tsunamis from the Currituck landslide. *Marine Geology*, *264*, 41–52.

Gusiakov, V. K. (2009). Chapter 2, Tsunami history: Recorded. In E. Bernard, A. Robinson (Ed.), *The Sea, Volume 15: Tsunamis, pp. 23–53*. Harvard University Press.

Harbitz, C. (1992). Model simulations of tsunamis generated by the Storegga slides. *Marine Geology, 105*(1-4), 1–21. doi: 10.1016/0025-58227(92) 90178-K.

Heller, V. & Spinneken, J. (2015). On the effect of the water body geometry on landslide tsunamis: Physical insight from laboratory tests and 2D to 3D wave parameter transformation. *Coastal Engineering, 104*(118), 113–134. doi: 10.1016/j.coastaleng.2015.06.006.

Heller, V., Hager, W., & Minor, H. E. (2009). *Landslide generated impulse waves in reservoirs: Basics and Computation*, *VAWMitteilung 211* (*Editor: R. Boes*). Laboratory of Hydraulics, Hydrology and Glaciology VAW, ETH Zurich, 8093 Zurich.

Huber, A. (1982). Quantifying impulse wave effects in reservoirs. In *Commission Internationale des Grands Barrages, Rio de Janeiro, Brazil.*

Jiang, L. & LeBlond, P. H. (1992). The coupling of a submarine slide and the surface waves which it generates. *Journal of Geophysical Research, 97*(C8), 12731–12744. doi: 10.1029/92JC00912.

Liu, P. L.-F., Lynett, P., & Synolakis, C. E. (2003). Analytical solutions for forced long waves on a sloping beach. *Journal of Fluid Mechanics, 478*, 101–109. doi: 10.1017/S0022112002003385.

Liu, P. L.-F., Wu, T.-R., Raichlen, F., Synolakis, C. E., & Borrero, J. C. (2005). Runup and rundown generated by three-dimensional sliding masses. *Journal of Fluid Mechanics, 536*, 107–144. doi: 10.1017/S0022112005004799.

Lynett, P. & Liu, P. (2002). A numerical study of submarine landslide generated waves and runup. *Proceedings of the Royal Society of London, A 458*, 2885–2910.

Lynett, P. & Liu, P. L.-F. (2005). A numerical study of the run-up generated by three-dimensional landslides. *Journal of Geophysical Research: Oceans, 110*(C03006). doi: 10.1029/2004JC002443.

Makridakis, S., Wheelwright, S., Hyndman, R. (1997). *Forecasting: Methods and Application (3rd ed.)*. John Wiley and Sons, New York.

Martínez-Álvarez, F., Troncoso, A., Cortés, G., & Riquelme, J. (2015). A Survey on Data Mining Techniques Applied to Electricity-Related Time Series Forecasting. *Energies, 8,* 13162-13193. doi: 10.3390/en81112361.

McCulloch, W.S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics, 5*, 115–133.

McFall, B. C., & Fritz, H. M. (2017). Runup of granular landslide-generated tsunamis on planar coasts and conical islands. *Journal of Geophysical Research: Oceans, 122*, 6901–6922. doi: 10.1002/2017JC012832.

Miller, D. J. (1960). The Alaska earthquake of July 10, 1958: Giant wave in Lituya Bay. *Bulletin of the Seismological Society of America,* (50) 252–266.

Mohammed, F. & Fritz, H. M. (2012). Physical modeling of tsunamis generated by three-dimensional deformable granular landslides. *Journal of Geophysical Research: Oceans, 117*(C11015). doi: 10.1029/2011JC00785.

Mohammed, F. & Fritz, H. M. (2013). Correction to "Physical modeling of tsunamis generated by three-dimensional deformable granular landslides". *Journal of Geophysical Research: Oceans, 118*, 3221–3221. doi: 10.1002/jgrc.20218.

Muller, L. (1964). The rock slide in the Vajont Valley. *Rock Mechanics and Engineering Geology, 2*(3-4), 148–212.

Murty, T. S. (1975). The dynamics of tsunamis. *OCEANS '75 conference, San Diego, CA, USA* (pp. 515-522). doi: 10.1109/OCEANS.1975.1154156.

Okal, E. A. & Synolakis, C. E. (2004). Source discriminants for near-field tsunamis. *Geophysical Journal International, 158*(3), 899–912. doi: 10.1111/j.1365-246X.2004.02347.x.

Ozeren, S. & Postacioglu, N. (2011). Nonlinear landslide tsunami run-up. *Journal of Fluid Mechanics, 691*, 440–460. doi: 10.1017/jfm.2011.482.

Panizzo, A. & Girolamo, P. (2005). Forecasting impulse waves generated by subaerial landslides. *Journal of Geophysical Research, 100*(C12025). doi: 10.1029/ 2004JC002778.

Panizzo, A., De Girolamo, P., Di Risio, M., Maistri, A., & Petaccia, A. (2005a). Great landslide events in Italian artificial reservoirs. *Natural Hazards and Earth System Sciences, 5*(5), 733–740. doi: 10.5194/nhess-5-733-2005.

Petley, D. (2008). *The Vaiont (Vajont) landslide of 1963*. https://blogs.agu.org/landslideblog/2008/12/11/the-vaiont-vajont-landslide-of-1963/. Accessed on February 06, 2022.

Postacioglu, N. & Ozeren, M. S. (2008). A semi-spectral modelling of landslide tsunamis. *Geophysical Journal International, 175*, 1–16.

Ramadan, K. T., Omar, M. A., & Allam, A. A. (2014). Modeling of tsunami generation and propagation under the effect of stochastic submarine landslides and slumps spreading in two orthogonal directions. *Ocean Engineering, 75*, 90–111. doi: 10.1016/j.oceaneng.2013.11.013.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-Validation. In Liu, L. & Özsu, M. T. (Ed.), *Encyclopedia of Database Systems* (pp. 532–538). US, Boston, MA: Springer. doi: 10.1007/978-0-387-39940-9_565.

Renzi, E. & Sammarco, P. (2010). Landslide tsunamis propagating around a conical island. *Journal of Fluid Mechanics, 650*, 251–285. doi: 10.1017/ S0022112009993582.

Renzi, E. & Sammarco, P. (2012). The influence of landslide shape and continental shelf on landslide generated tsunamis along a plane beach. *Natural Hazards and Earth System Sciences, 12*(5), 1503–1520. doi: 10.5194/nhess-12-1503-2012.

Renzi, E. & Sammarco, P. (2016). The hydrodynamics of landslide tsunamis: Current analytical models and future research directions. *Landslides, 13*(6) (2016) 1369–1377. doi: 10.1007/s10346-016-0680-z.

Roberts, N. J., McKillop, R., Hermanns, R.L., Clague, J. J., & Oppikofer, T. (2014). Preliminary global catalogue of displacement waves from subaerial landslides. In *World Landslide Forum 3*, *2-6 June 2014, Beijing, China*. doi: 10.1007/978-3-309-04996-0_104.

Romano, A., Di Risio, M., Bellotti, G., Molfetta, G., Damiani, L., & De Girolamo, P. (2016). Tsunamis generated by landslides at the coast of conical islands: Experimental benchmark dataset for mathematical model validation. *Landslides, 13*(6), 1379–1393. doi: 10.1007/s10346-016-0696-4.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review, 65*, 386–408.

Rosenblatt, F. (1959). *Two theorems of statistical separability in the perceptron. In: Mechanization of thought processes*. HM Stationary Office, London, England.

Ruff, L. J. (2003). Some aspects of energy balance and tsunami generation by earthquakes and landslides, *Pure and Applied Geophysics,* (160) 2155–2176.

Sabeti, R. & Heidarzadeh, M. (2020). Semi-empirical predictive equations for the initial amplitude of submarine landslide-generated waves: Applications to 1994 Skagway and 1998 Papua New Guinea tsunamis. *Natural Hazards, 103*, 1591–1611. doi: 10.1007/s11069-020-04050-4.

Sammarco, P. & Renzi, E. (2008). Landslide tsunamis propagating along a plane beach. *Journal of Fluid Mechanics, 598*, 107–120. doi: 10.1017/S0022112007009731.

Sassa, K., Mikos, M. & Yin, Y. (2017). *Advancing culture of living with landslides*. Volume 1, Springer International Publishing, ISDR-ICL Sendai Partnerships.

Seo, S.-N. & Liu, P. L.-F. (2013). Edge waves generated by the landslide on a sloping beach. *Coastal Engineering,* 73, 133–150. doi: 10.1016/j.coastaleng.2012.10.008.

Shaikh, R. (2018). *Feature selection techniques in machine learning with python*. https://towardsdatascience.com/feature-selection-techniques-in-machine-learning-with-python-f24e7da3f36e. Accessed on February 06, 2022.

Shcherbakov, M., Brebels, A., Shcherbakova, N.L., Tyukov, A., Janovsky, T.A. & Kamaev, V.A.. (2013). A survey of forecast error measures. *World Applied Sciences Journal,* 24, 171-176. doi: 10.5829/idosi.wasj.2013.24.itmies.80032.

Slingerland, R. & Voight, B. (1979). Occurrences, properties and predictive models of landslide-generated impulse waves. *Rockslides and Avalanches, 2*, 317–397. doi: 10.1016/B978-0-444-41508-0.50017-X.

Smola, A. & Vishwanathan, S. V. N. (2008). *Introduction to Machine Learning.* Cambridge University Press, ISBN: 0521825830.

Tappin, D. R., Watts, P., & Grilli, S. (2008). The Papua New Guinea tsunami of 17 July 1998: Anatomy of a catastrophic event. *Natural Hazards and Earth System Sciences*, *8*, 1–24.

Ten Brink, U., Geist, E., & Andrews, B. (2006). Size distribution of submarine landslides and its implication to tsunami hazard in Puerto Rico. *Geophysical Research Letters, 33*(L11307). doi: 10.1029/2006GL026125.

Tinti, S., Manucci, A., Pagnoni, G., Armigliato, A., & Zaniboni, F. (2005). The 30 December 2002 landslide-induced tsunamis in Stromboli: Sequence of the events reconstructed from the eyewitness accounts. *Natural Hazards and Earth System Sciences, 5*, 763–775.

Urbanowicz, R. J., Meeker, M., Cava, W. L., Olson, R. S., & Moore, J. H. (2018). Relief based feature selection: Introduction and review. *Journal of Biomedical Informatics, 85,* 189–203. doi: 10.1016/j.jbi.2018.07.014.

Vu-Quoc, Loc (2012). EGM 4313 Intermediate Engineering Analysis, Spring 2012, University of Florida. https://commons.wikimedia.org/w/index.php?title=File:Neuron3.png&oldid=479262868. Accessed on February 06, 2022.

Wang, J., Ward, S., & Xiao, L. (2015). Numerical simulation of the December 4, 2007 landslide-generated tsunami in Chehalis Lake, Canada. *Geophysical Journal International, 201*, 372–376. doi: 10.1093/gji/ggv026.

Ward, S. N. (2000). *Tsunamis. In Encyclopedia of Physical Science and Technology.* Academic Press, New York.

Weiss, R., Fritz, H. M., & Wunnemann, K. (2009). Hybrid modeling of the megatsunami runup in Lituya Bay after half a century. *Geophysical Research Letters, 36* (L09602). doi: 10.1029/2009GL037814.

Widrow, B. & Hoff, M. E. (1960). Adaptive switching circuits. In *Ire Western Electric Show and Convention Record*, Part 4, 96–104.

Xenakis, A. M., Lind, S. J., Stansby, P. K., & Rogers, B. D. (2017). Landslides and tsunamis predicted by incompressible smoothed particle hydrodynamics (SPH) with application to the 1958 Lituya Bay event and idealized experiment. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science, 473*(2199), 20160674. doi: 10.1098/rspa.2016.0674.

Yeh, H., Imamura, F., Synolakis, C., Tsuji, Y., Liu, P., & Shi, S. (1993). The Flores Island tsunami. *Eos, Transactions American Geophysical Union, 74*(33), 369.

Zadeh, L. A. (1994). Fuzzy logic, neural networks, and soft computing. *Communications of the ACM, 37*(3), 77–84. doi: 10.1145/175247.175255.

APPENDIX

Analytical Modeling of Landslide Tsunamis

The shallow-water wave theory is a widely used theory to simulate evolution of free surface waves. Below an analytical solution based on this theory is presented, which is benefited in order to generate a data set for run-up of tsunamis triggered by a mass of vertical thickness δ and horizontal length L sliding over a plane beach of slope angle β.

The nonlinear shallow-water wave equations in one space dimension consist of the continuity and the momentum equations given as (Tuck and Hwang, 1972)

$$\left(\tilde{\eta}+\tilde{b}\right)_{\tilde{t}}+\left[\left(\tilde{\eta}+\tilde{b}\right)\tilde{u}\right]_{\tilde{x}}=0, \tag{A.1a}$$

$$\tilde{u}_{\tilde{t}}+\tilde{u}\tilde{u}_{\tilde{x}}+g\tilde{\eta}_{\tilde{x}}=0. \tag{A.1b}$$

Here, $\tilde{x}$ is the horizontal distance from the initial shoreline, $\tilde{t}$ is time, $\tilde{z}=\tilde{\eta}(\tilde{x},\tilde{t})$ is the free surface, $\tilde{z}=-\tilde{b}(\tilde{x},\tilde{t})$ is the sea bottom, $\tilde{u}(\tilde{x},\tilde{t})$ is the fluid velocity, g is the gravitational acceleration. Variables in (A.1) are in dimensional form and subscripts indicate partial derivatives. In the presence of a time-dependent perturbation $\tilde{h}(\tilde{x},\tilde{t})$, the bottom is defined as

$$\tilde{b}(\tilde{x},\tilde{t})=\tilde{h}_0(\tilde{x})-\tilde{h}(\tilde{x},\tilde{t}), \tag{A.2}$$

where $\tilde{h}_0(\tilde{x})=\tilde{x}\tan\beta$ is the undisturbed water depth (Figure 3.1).

Introducing δ and L as characteristic scales, the following dimensionless quantities can be defined:

$$x=\tilde{x}/L,\ \ \{b,\eta\}=\{\tilde{b},\tilde{\eta}\}/\delta,\quad u=\tilde{u}/\sqrt{g\delta},\ \ t=\tilde{t}/\sqrt{L^2/(g\delta)}.$$

The dimensionless equations then read

$$(\eta-h)_t+\left[\left(\frac{\tan\beta}{\mu}x-h+\eta\right)u\right]_x=0, \tag{A.3a}$$

$$u_t+uu_x+\eta_x=0, \tag{A.3b}$$

where $\mu=\delta/L$ is the slide aspect ratio.

Further assuming $\{h, \eta\} \ll h_0$ and $u \ll \sqrt{gh_0}$ as per Tuck and Hwang (1972), (A.3) combine into the linear equation (Liu et al., 2003)

$$\eta_{tt} - \frac{\tan\beta}{\mu}(x\eta_x)_x = h_{tt}. \tag{A.4}$$

This equation is nonhomogeneous since the bottom perturbation (h) is time-dependent.

Liu et al. (2003) derived an analytical solution for (A.4) which is valid for thin slides, $\delta \ll L$, or equivalently $\mu \ll 1$, so that μ and $\tan\beta$ are of the same order: $(\tan\beta)/\mu \sim O(1)$. Introducing the change of variables

$$\xi = 2\sqrt{\mu x/\tan\beta}, \tag{A.5}$$

they eliminated the term $(\tan\beta)/\mu$ and rewrote (A.4) as

$$\eta_{tt} - \frac{1}{\xi}\left(\xi\eta_\xi\right)_\xi = h_{tt}. \tag{A.6}$$

Analytical solution of (A.6) can uniquely be determined given two initial conditions defining an initially quiet sea: zero initial wave height, $\eta(\xi, 0) = 0$, and zero initial wave velocity, $u(\xi, 0) = 0$, or $\eta_t(\xi, 0) = h_t(\xi, 0)$ in view of the continuity equation. As expected, the complete solution of this nonhomogeneous initial value problem is the sum of homogeneous and particular solutions, i.e.

$$\eta(\xi, t) = \eta_h(\xi, t) + \eta_p(\xi, t). \tag{A.7}$$

Liu et al. (2003) utilized the Hankel integral transform and obtained the homogeneous solution as

$$\eta_h(\xi, t) = \int_0^\infty w[A(w)\cos(wt) + B(w)\sin(wt)]\, J_0(w\xi)dw, \tag{A.8}$$

in which $A(w)$ and $B(w)$ are to be found from the initial conditions. Moreover, they managed to formulate the particular solution in terms of the bottom perturbation h as

$$\eta_p = \frac{1}{3}\left[h(\xi, t) - \xi h_\xi(\xi, t)\right]. \tag{A.9}$$

Liu et al. (2003) then tested their solution for a Gaussian-type subaerial seafloor deformation,

$$h(\xi, t) = e^{-(\xi - t)^2}. \tag{A.10}$$

Recently, Aydın (2021) generalized Liu et al. (2003)'s solution to further employ submerged landslides by translating the bottom disturbance by an amount

$$\xi_0 = 2\sqrt{\mu x_0/\tan\beta}\,, \tag{A.11}$$

in other words, imposing

$$h_e(\xi,t) = e^{-(\xi-\xi_0-t)^2}. \tag{A.12}$$

It is straightforward to show that (A.12) makes (A.9) a particular solution of (A.6). The spatial distribution of this bottom profile is plotted in Figure A.1. With this modification, slides that are partly or completely submerged can also be considered, by locating the mass on the sloping beach with appropriate choice of the parameter ξ_0 (or x_0, see Figure 3.1).

In addition to the Gaussian disturbance, Aydın (2021) proposed an alternative bottom profile

$$h_s(\xi,t) = \mathrm{sech}^2\,(\xi-\xi_0-t)\,. \tag{A.15}$$

It can easily be verified that (A.15) makes (A.9) another particular solution of (A.6). This soliton-type function has a very similar shape and the same maximum amplitude as the Gaussian profile (A.12), but it is broader, occupying approximately 10% greater area compared to the Gaussian (Figure A.1).

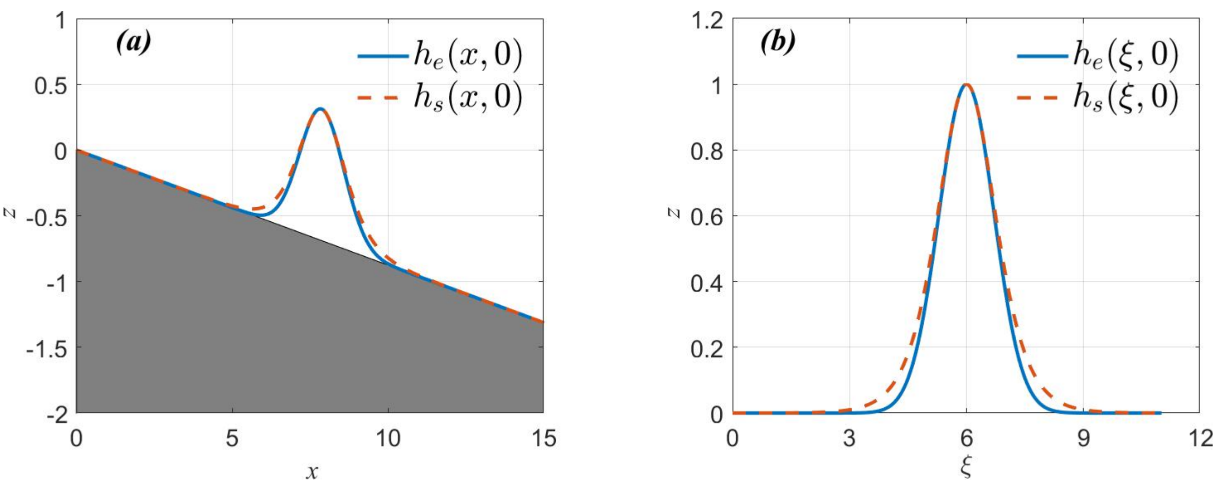

Figure A.1. Comparison of the initial bottom profiles given in Equations (A.12) and (A.15), respectively, (a) on the slope as functions of x, and, (b) on a horizontal plane as functions of ξ. The initial slide submergence is chosen as $\xi_0 = 6$.

Tsunami run-up is defined as the free surface height at the point where the total depth is zero, i.e. $\eta + b = 0$. By definition, $b = x - h$ (recall that the dimensionless undisturbed depth is $h_0 = x$). Therefore $\eta + b = 0$ corresponds to the initial shoreline $x = 0$ under the assumptions of the linear theory ($\{h, \eta\} \ll h_0$) and hence tsunami run-up can be defined as

$$R(t) = \eta(x = 0, t). \tag{A.13}$$

Note that run-up is a time series. The maximum value of run-up, called the maximum run-up,

$$R_{\max} = \max\{R(t), t \geq 0\}, \tag{A.14}$$

is arguably the most important quantity of interest in coastal engineering practice.

In Figure A.2, the landslide tsunami run-up resulting from Equation (A.12) are plotted for three different values of the initial slide submergence. The horizontal slide length and the vertical slide thickness are varied in the left panels, and the right panels, respectively, and the maximum run-up in each case is also indicated. The run-up curves suggest that the maximum run-up variation with δ is much significant than with L.

The analytical model introduced in this section is used to generate the data set needed for the MLP-based prediction models, through which the role of landslide parameters on the maximum run-up is investigated.

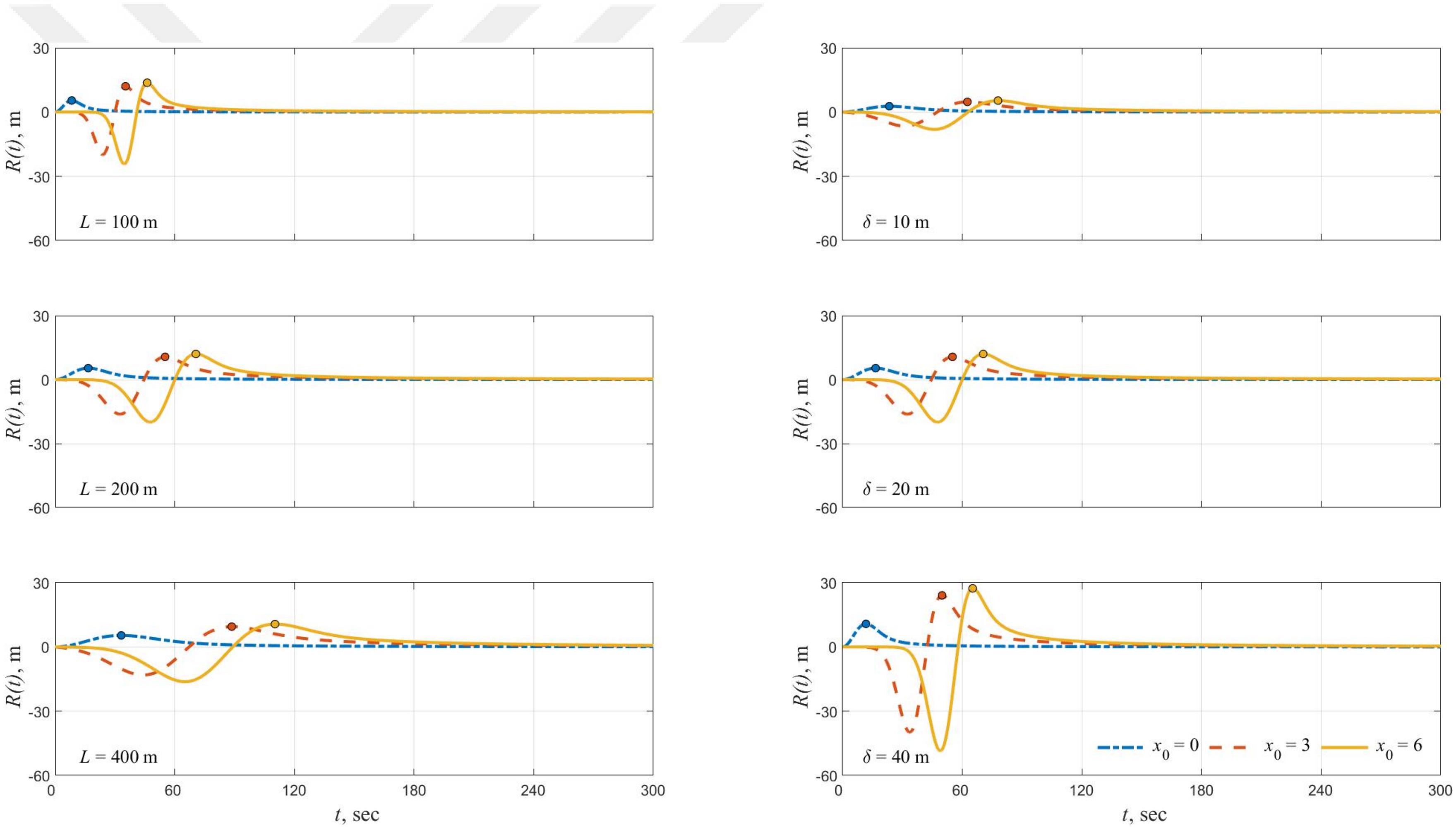

Figure A.2. Dependence of run-up on landslide horizontal length for a fixed slide thickness of $\delta = 20$ m (left), and, slide thickness for a fixed slide length of $L = 200$ m (right). Results for three different values of the initial slide submergence, i.e. $x_0 = 0, 3, 6,$ (dimensionless) are presented and the maximum run-ups are also indicated with colored circles. The slope angle is chosen as $\beta = 10°$. Note that the dimensional initial submergence parameter is $\tilde{x}_0 = Lx_0$.