

Article

Tsunami Distribution Functions along the Coast: Extended

Ira Didenkulova ^{1,*} , Andrey Zaitsev ² and Efim Pelinovsky ^{3,4}

¹ Department of Mathematics, University of Oslo, 0851 Oslo, Norway

² Special Research Bureau for Automation of Marine Researches, Far Eastern Branch of Russian Academy of Sciences, Yuzhno-Sakhalinsk 693023, Russia

³ Institute of Applied Physics, Russian Academy of Science, Nizhny Novgorod 603950, Russia

⁴ V.I. Il'ichev Pacific Oceanological Institute, Far Eastern Branch of Russian Academy of Sciences, Vladivostok 690041, Russia

* Correspondence: irinadi@math.uio.no

Abstract: The distribution of tsunami runup heights along the coast is studied both theoretically and experimentally using observation data of historical tsunami from 1992 to 2018. The physical mechanisms leading to the lognormal distribution of tsunami runup heights along the coast are discussed, and its statistical moments are calculated. It is shown that the lognormal distribution describes well the measurements of tsunami characteristics over the past 30 years. Special attention is paid to the multi-source 2018 Palu–Sulawesi tsunami, which was generated by an earthquake with magnitude 7.5 and numerous subsequent landslides. It is shown that even in this special case the lognormal distribution is a rather good approximation.

Keywords: lognormal distribution; tsunami runup heights; distribution functions



Citation: Didenkulova, I.; Zaitsev, A.; Pelinovsky, E. Tsunami Distribution Functions along the Coast: Extended. *J. Mar. Sci. Eng.* **2022**, *10*, 1137. <https://doi.org/10.3390/jmse10081137>

Academic Editors: Han Soo Lee, Kyeong Ok Kim, Seung-Buhm Woo, Jin-Hee Yuk and Youngjin Choi

Received: 18 July 2022

Accepted: 16 August 2022

Published: 18 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A lot of data has been accumulated on the distribution of tsunami heights along the coast, both from historical data and from numerical modelling of historical and hypothetical events. In particular, a lot of data was obtained after the catastrophic Indian Ocean 2004 tsunami due to its global propagation and impact. An analysis of the distribution of wave heights along the coast allows zoning of areas according to the degree of tsunami hazard and to plan measures to prevent and mitigate the consequences of natural disasters. Taking into account the recurrence of earthquakes, it is possible to obtain long-term estimates of tsunami heights for specific coastal points with a given probability; such assessments are now being carried out as part of the PTHA approach (Probabilistic Tsunami Hazard Assessment) [1].

Meanwhile, even for one specific event, the distribution of tsunami heights along the coast is extremely inhomogeneous, and for its analysis one can use methods of probability theory and mathematical statistics. Van Dorn [2] was the first to apply a statistical approach to analyze observed tsunami runup heights. He found that a lognormal distribution was the best fit for tsunamis on the coast of the Hawaiian Islands. This analysis was continued by [3] with a special focus on the Japanese coast. He analyzed six events: Meiji Sanriku of 15 June 1896, Showa Sanriku of 3 March 1933, Nankaido of 21 December 1946, Chile of 23 May 1960, Niigata of 16 June 1964, and Tokashi-oki of 16 May 1968. The theoretical interpretation of the lognormal distribution associated with random seafloor heterogeneity was given by Go Chan Nam (his original paper was published in Russian as a preprint, while English translation is given in [4]). Subsequently, the lognormal distribution was used to describe many real tsunamis that occurred in 1992–2011, including the catastrophic tsunamis of 2004 and 2011 [5–8]. Meanwhile, when analyzing real data, deviations from the lognormal distribution have also been found, see papers cited above. There may be several reasons for this. First, this could be related to the strong nonlinearity of tsunami waves

approaching the coast in the form of a bore. Then, due to the nonlinear dissipation of the wave energy at the front, the bore height also changes. Second, the properties of different sections of the coast differ from each other, violating the main assumption of the central limit theorem on the homogeneity of random variables. This idea is developed by [9–12], who believed that the generalized Pareto distribution would be a better approximation of the tsunami wave heights distribution on the coast. Third, the measurements of historical tsunamis are not statistically homogeneous, with more detailed data (with a small step) often being obtained in the region of the largest runups. Fourth, the seabed bathymetry is not statistically homogeneous and contains extended deterministic sections (continental slope, shelf) on which wave transforms according to the classical (deterministic) methods of the long wave theory.

In this paper, we discuss the lognormal distribution of tsunami runup heights along the coast, which is often used to interpret field survey data (Section 2). This analysis is supplemented by calculating the exact statistical moments of this distribution. In Section 3 we analyze the field survey data for a number of historical tsunamis. The data from the multi-source catastrophic 2018 Palu–Sulawesi tsunami is analyzed in Section 4 using distribution functions. The results are summarized in Discussion.

2. Lognormal distribution

The use of lognormal distribution for tsunami runup heights along the coast has been suggested in [2–4], and then revised in [5,13,14]. It is based on the linear theory of wave propagation in the basin of random topography. Even within the linear shallow water wave theory, the wave of maximal height approaching the coast is a result of a complicated process of wave reflection, refraction, diffraction, and resonant effects. Very often, the wave of maximal amplitude is not a leading wave. However, due to the linearity of shallow water equations, the runup height, H , is always proportional to the wave height in the tsunami source, H_0 :

$$H = KH_0, \tag{1}$$

where the coefficient of wave transformation, K , can be computed within a 2D numerical model. In general, coefficient K has no evident physical sense due to the processes mentioned above. If the tsunami wavelength is short compared with the characteristic scale of seabed variations, the ray theory can be used. In this case, 2D equations can be reduced to the set of 1D equations along the propagation path. The ray pattern can be easily computed within the ordinary differential equations of the second order [15]:

$$\frac{d\theta}{dt} = \frac{\cos \zeta}{nR}, \tag{2}$$

$$\frac{d\varphi}{dt} = \frac{\sin \zeta}{R \sin \theta}, \tag{3}$$

$$\frac{d\zeta}{dt} = \frac{\sin \zeta}{n^2 R} \frac{\partial n}{\partial \theta} + \frac{\cos \zeta}{n^2 R \sin \theta} \frac{\partial n}{\partial \varphi} - \frac{\sin \zeta \cos \theta}{nR}, \tag{4}$$

where θ and φ are latitude and longitude of the ray, $n = (gh)^{-1/2}$ is the slowness, g is the gravity acceleration, $h(\theta, \varphi)$ is the water depth, R is the radius of the Earth, and ζ is the ray direction measured counter-clockwise from the South.

An example of ray theory computations from isotropic hypothetical tsunami source in the East Sea is shown in Figure 1, where one can clearly see the complicated character of ray paths due to seafloor variations.

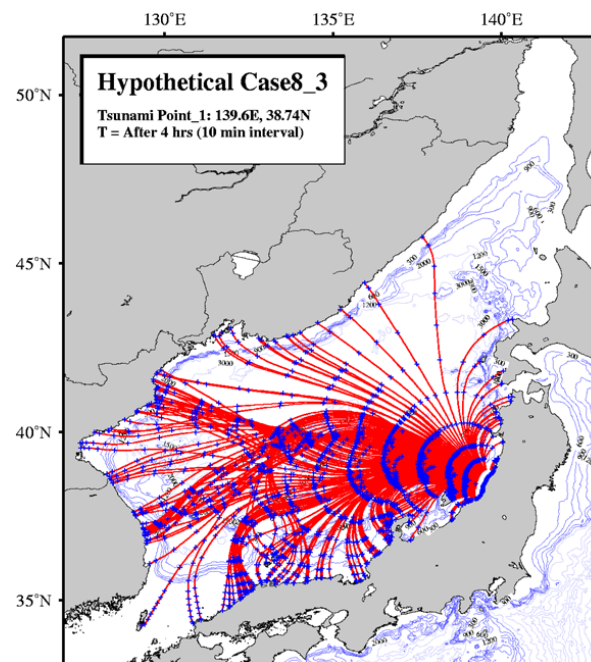


Figure 1. Ray theory computations from isotropic hypothetical tsunami source in the East Sea.

In ray theory, the wave height is described by the famous Green’s law, which is often applied to estimate wave characteristics in the coastal zone.

$$H(l) = H_0 \left[\frac{h_0}{h(l)} \right]^{1/4} \left[\frac{B_0}{B(l)} \right]^{1/2}, \tag{5}$$

where l is the distance along the wave path, $h(l)$ is a local water depth, and $B(l)$ is a differential width of the ray tube (distance between the neighbouring rays). Coefficient K in Equation (1) depends on water depth change along a propagation path and is determined by random bathymetry. After dividing the propagation path into a series of more or less statistically independent segments, the total transformation coefficient becomes a product of the local coefficients of tsunami wave transformation in each segment. In this case, Equation (1) can be rewritten in the logarithmic form:

$$\ln H = \ln H_0 + \sum_i \ln K_i \tag{6}$$

where i characterizes the number of random statistically independent segments along the propagation path and $\ln K_i$ can be considered as independent random variables. The central limit theorem states that the sum of many random independent variables tends to the Gaussian distribution, and therefore $\ln H$ is described by the normal curve. It means that a probability density function (pdf) of the wave height is described by the lognormal distribution:

$$f(H) = \frac{1}{H\sigma_{\ln}\sqrt{2\pi}} \exp \left[-\frac{(\ln H - a)^2}{2\sigma_{\ln}^2} \right]. \tag{7}$$

This distribution has two parameters with evident physical meaning: $a = \langle \ln H \rangle$ is the average value, and σ_{\ln} is the standard deviation of the logarithm of the wave height. For definition, wave height is measured in meters. This function has been widely used to study tsunami characteristics [2,3,5–7].

The lognormal distribution Equation (7) can be reduced to a dimensionless form by introducing a change of variables

$$y = \frac{H^{1/\sigma_{ln}}}{\exp(a/\sigma_{ln})} \quad (8)$$

$$f(y) = \frac{1}{\sqrt{2\pi}y} \exp\left[-\frac{1}{2}(\ln y)^2\right] \quad (9)$$

then the four statistical moments of this distribution can be calculated analytically:

$$\langle y \rangle = \frac{1}{\sqrt{2\pi}} \int_0^\infty \exp\left[-\frac{(\ln y)^2}{2}\right] dy = \sqrt{e} \cong 1.65, \quad (10)$$

$$\sigma^2 = \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{(y - \langle y \rangle)^2}{y} \exp\left[-\frac{(\ln y)^2}{2}\right] dy = e^2 - e \cong 4.67, \sigma \cong 2.16, \quad (11)$$

$$Sk = \frac{1}{\sqrt{2\pi}\sigma^3} \int_0^\infty \frac{(y - \langle y \rangle)^3}{y} \exp\left[-\frac{(\ln y)^2}{2}\right] dy = \frac{e^{9/2} - 3e^{5/2} + 2e^{3/2}}{(e^2 - e)^{3/2}} = \sqrt{e-1}(2+e) \cong 6.18, \quad (12)$$

$$Ku = \frac{1}{\sqrt{2\pi}\sigma^4} \int_0^\infty \frac{(y - \langle y \rangle)^4}{y} \exp\left[-\frac{(\ln y)^2}{2}\right] dy = e^4 + 2e^3 + 3e^2 \cong 113.9. \quad (13)$$

Here $\langle y \rangle$ is the average, σ is standard deviation, Sk is skewness and Ku is kurtosis of variable y .

From these estimates follows that the lognormal distribution is very different from the normal (Gaussian) distribution, see Figure 2, which also shows a segment of the Gaussian distribution with the same mean value. In the range of large values lognormal distribution decreases slower than the normal one, demonstrating the decisive contribution of high wave runup heights to the statistical moments.

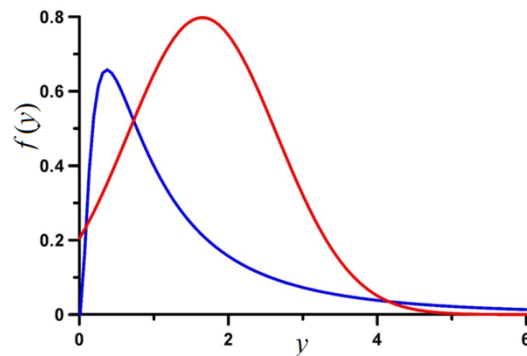


Figure 2. Dimensionless lognormal (blue) and Gaussian (red) probability density functions.

However, in practice, it is quite difficult to use the probability density functions for analysis of real tsunamis due to the large scatter of measurements and data of observations. It is much more convenient to use the integral distribution function, which is smoother:

$$F(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty \frac{1}{z} \exp\left[-\frac{1}{2}(\ln z)^2\right] dz \quad (14)$$

In practice, distribution functions are built in decimal logarithms, not natural ones. This leads to the following modification of the integral distribution function and its argument:

$$F(y) = \frac{1}{\sqrt{2\pi \ln 10}} \int_y^\infty \frac{1}{z} \exp\left[-\frac{1}{2}(\log z)^2\right] dz \tag{15}$$

$$y = \frac{H^{1/\sigma_{ln}}}{\exp(a/\sigma_{ln})}. \tag{16}$$

The last Equations (15) and (16) are used below for analysis of field survey data.

3. Tsunami Observations

In Figure 3 the data of 11 tsunamis from 1990s are plotted on a single graph. The data include: Flores Island tsunami, Indonesia, 12 December 1992; tsunami at the East Korean Coast, 12 July 1993; Hokkaido tsunami, Japan, 12 July 1993; Java tsunami, Indonesia, 2 June 1994; tsunami in Kuril Islands, Russia, 4 October 1994; Mindoro Island tsunami, Philippines, 14 November 1994; Chile tsunami, 30 July 1995; Sulawesi tsunami, Indonesia, 1 January 1996; Western Irian Jaya tsunami, Indonesia, 17 February 1996; tsunami in Peru, 21 February 1996; and Papua New Guinea tsunami, 17 July 1998. These data, which include both seismic and landslide induced tsunamis, are discussed in [5]. Most of them are tsunamis of moderate magnitude. It can be seen that all these data nicely follow the lognormal distribution.

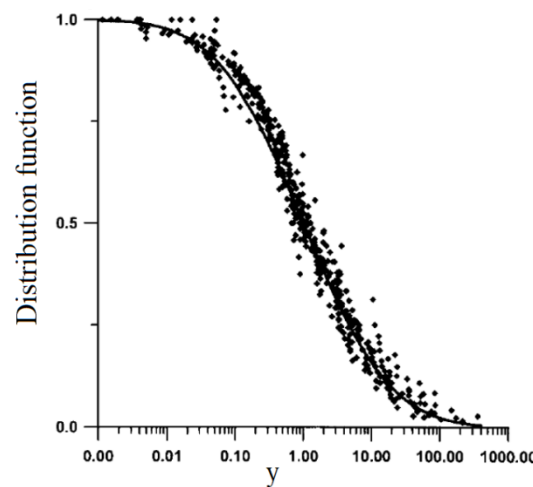


Figure 3. Dimensionless integral probability distribution function for tsunami runup heights from 11 different tsunamis of 1990s. Solid line represents theoretical curve, while dots denote the field survey data of runup heights of 11 tsunamis from 1990s [5].

The 2004 Indian Ocean tsunami was the largest ever recorded in several countries. Analysis of its runup height distribution function was made in [6] and is shown in Figure 4. As we can see, the observation data of this exceptionally large tsunami are also well approximated by the lognormal distribution.

Similar results were obtained also for runup heights of the 2011 Japan tsunami (Figure 5). This dataset was compiled from several field surveys run by different scientific groups, references to which can be found in [7]. Most of these data with a high spatial resolution were obtained in the region of very high tsunami runups and were correlated. This led to the violation of the ‘randomness’ criterion of the wave propagation paths, which is the basic assumption underlying the lognormal distribution. The importance of the ‘valid’ spatial resolution of the data is discussed in [7]. In Figure 5 one can see the distribution functions of the 2011 tsunami using a different spatial resolution of the observations. It can

be seen that an increase in a spatial scale along the coast in several kilometers leads to a better fit of the measurement data to the lognormal distribution.

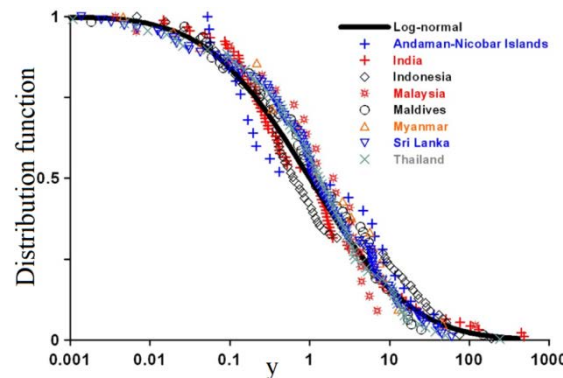


Figure 4. Runup height integral probability distribution functions of the 2004 tsunami manifestations in different countries of the Indian Ocean. Solid line represents theoretical curve, while different markers denote the field survey data of runup heights [6].

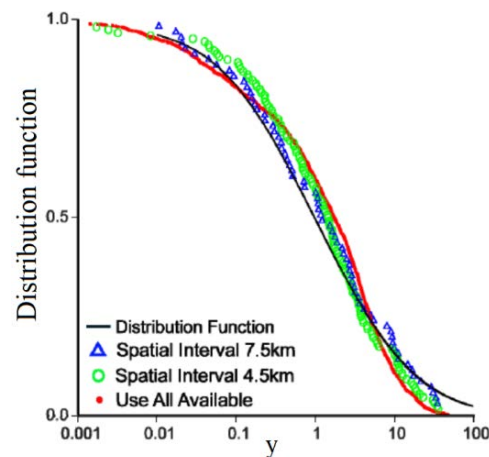


Figure 5. Runup height integral probability distribution functions of the 2011 tsunami along Japanese coastline with different spatial resolution. Solid line represents theoretical curve, while different markers denote the field survey data of runup heights [7].

4. Palu-Sulawesi 2018 Tsunami

Another interesting example is the Palu–Sulawesi tsunami, which occurred on 28 September 2018. Its maximum runup height reached 9.1 m. The tsunami was generated by a strike-slip earthquake with magnitude 7.5, which subsequently caused several tsunami-genic submarine and subaerial landslides; [16–19] suggested that the observed tsunami was initiated predominantly by submarine/subaerial landslides and the contribution of earthquake tsunami is <20% of the maximum wave height.

Thus, the observational data are divided into sets of measurements from many individual tsunami events. Moreover, the final dataset is compiled not from their linear superposition, but with a cumulative effect since the tsunami runup height in a particular location depends on the sum of these events with unknown weights. It is unclear whether the lognormal distribution is suitable for such heterogeneous sources.

The 55 runup height measurements from the field survey conducted by [18] were used to plot the runup height distribution function for the 2018 Palu–Sulawesi tsunami (Figure 6). It can be seen that even in this complex multi-source tsunami generation the data still follow closely the lognormal distribution.

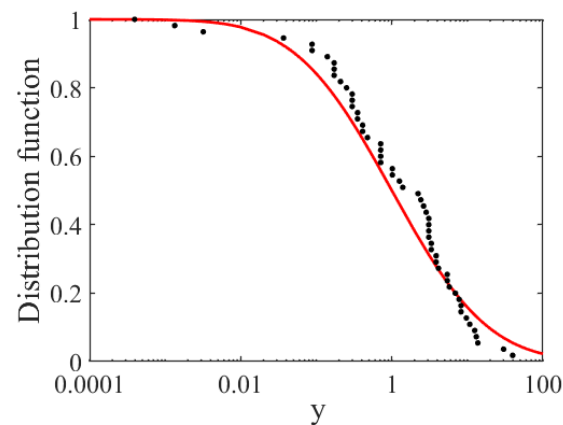


Figure 6. Runup height integral probability distribution functions of the 2018 Palu-Sulawesi tsunami. Solid line represents theoretical curve, while different markers denote the field survey data of runup heights.

5. Discussion

In this paper we discuss the application of the lognormal distribution to the data of tsunami wave runups at the coast. This topic was of a special interest of Professor Byung Ho Choi for several decades. Here we discuss the applicability of lognormal distribution for description of tsunami runup heights, calculate exactly the four statistical moments of the lognormal distribution and apply it to several different historic tsunami datasets. It has worked very well to describe the available tsunami data with a spatial resolution of several kilometers. For more detailed data, the “randomness” condition, lying in the basis of applicability of lognormal distribution is violated, so that the propagation paths cannot be considered independent and lognormal distribution does not work so well.

We also apply lognormal distribution to the 2018 Palu–Sulawesi tsunami, which was induced by the 7.5 earthquake and several subsequent landslides. It is shown that even in this complex multi-source case, the tsunami runup heights still nicely follow the lognormal distribution.

We note that the real probability distribution is unaware of the assumptions of the theoretical model. However, it confirms the main ideas of the model. First of all, it is the inhomogeneity of the bathymetry, which leads to a lognormal distribution of wave heights along the coast. Regardless of its origin (earthquake or landslide), a tsunami propagates over a basin of non-uniform bathymetry. Of course, the distribution parameters should depend on the characteristics of the source. However, in practice, when measuring wave heights, we do not know the exact cause of it, which is often the superposition (linear or non-linear) of the heights of partial events (earthquake or landslide, or several landslides). It averages the observed characteristics of the distribution, making it single-peak. Deviations from the lognormal distribution are most likely due to the multiple sources of the event, but it is not possible to identify this from the available data. To understand how each tsunami source affects the total distribution function, numerical modelling of tsunami from each tsunami source is needed.

All this is important for the evaluation of tsunami magnitude, which is defined by the averaged tsunami runup height on the coast. Knowing the distribution function, we can estimate the probability of missing measurements (this can happen due to the limited access to certain locations), with an extreme wave height, and estimate variations in the tsunami magnitude.

Author Contributions: Conceptualization, E.P. and I.D.; methodology, E.P. and I.D.; analysis, I.D. and A.Z.; writing—original draft preparation, I.D.; writing—review and editing, E.P, I.D. and A.Z. All authors have read and agreed to the published version of the manuscript.

Funding: EP acknowledges support of the Laboratory of nonlinear hydrophysics and natural hazards of V.I. Il'ichev Pacific Oceanological Institute, Far Eastern Branch Russian Academy of Sciences, project of Ministry of science and education of Russia, agreement No 075–15-2022–1127 from 1 July 2022.

Institutional Review Board Statement: This study did not require ethical approval.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data used are available through published sources.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Selva, J.; Lorito, S.; Volpe, M.; Romano, F.; Tonini, R.; Perfetti, P.; Bernardi, F.; Taroni, M.; Scala, A.; Babeyko, A.; et al. Probabilistic tsunami forecasting for early warning. *Nat. Commun.* **2021**, *12*, 5677. [[CrossRef](#)] [[PubMed](#)]
2. Van Dorn, W.G. Tsunamis. In *Advances in Hydrosience*; Chow, V.T., Ed.; Academic Press: London, UK, 1965; Volume 2, pp. 1–48.
3. Kajiura, K. Some statistics related to observed tsunami heights along the coast of Japan. In *Tsunamis Their Science and Engineering*; Terra Scientific Publishing Company: Tokyo, Japan, 1983; pp. 131–145.
4. Go, C.N. Statistical properties of tsunami runup heights on the coasts of the Kuril Isles and Japan. In *Tsunami along the Coast*; Choi, B.H., Pelinovsky, E., Eds.; Hanrimwon Publishing Co.: Seoul, Korea, 2016; pp. 155–180.
5. Choi, B.H.; Pelinovsky, E.; Riabov, I.; Hong, S.J. Distribution functions of tsunami wave heights. *Nat. Hazards* **2002**, *25*, 1–21. [[CrossRef](#)]
6. Choi, B.H.; Hong, S.J.; Pelinovsky, E. Distribution of runup heights of the December 26, 2004 tsunami in the Indian Ocean. *Geophys. Res. Lett.* **2006**, *33*, L13601. [[CrossRef](#)]
7. Choi, B.H.; Min, B.I.; Pelinovsky, E.; Tsuji, Y.; Kim, K.O. Comparable analysis of the distribution functions of runup heights of the 1896, 1933 and 2011 Japanese Tsunamis in the Sanriku Area. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 1463–1467. [[CrossRef](#)]
8. Wahyuni, D.; Adnan, A.; Yendra, R.; Desvina, A.P.; Muhaijir, M.N.; Arrafie Abduh, M. Best Fit Probability Model for Runup Height tsunami in Aceh Using Some Mixture Distribution. *Appl. Math. Sci.* **2018**, *12*, 1191–1199. [[CrossRef](#)]
9. Kim, D.; Cho, Y.S.; Kim, Y.C. On the spatial pattern of the distribution of the tsunami run-up heights. *Stoch. Hydrol. Hydraul.* **2012**, *27*, 1333–1346. [[CrossRef](#)]
10. Kim, D.; Kim, B.J.; Lee, S.-O.; Cho, Y.-S. Best-fit distribution and log-normality for tsunami heights along coastal lines. *Toch. Hydrol. Hydraul.* **2013**, *28*, 881–893. [[CrossRef](#)]
11. Kim, D.; Choi, Y.S. Best-fit distribution and log-normality for tsunami runup heights. In *Tsunami along the Coast*; Choi, B.H., Pelinovsky, E., Eds.; Hanrimwon Publishing Co.: Seoul, Korea, 2016; pp. 103–124.
12. Song, M.-J.; Sho, Y.-S. Modeling Maximum Tsunami Heights Using Bayesian Neural Networks. *Atmosphere* **2020**, *11*, 1266. [[CrossRef](#)]
13. Burroughs, S.M.; Tebbens, S.F. Power-law scaling and probabilistic forecasting of tsunami runup heights. *Pure Appl. Geophys.* **2005**, *162*, 331–342. [[CrossRef](#)]
14. Kaistrenko, V.M. Tsunami recurrence versus tsunami height distribution along the coast. *Pure Appl. Geophys.* **2011**, *168*, 2065–2069. [[CrossRef](#)]
15. Satake, K. Effects of Bathymetry on Tsunami Propagation: Application of Ray Tracing to Tsunamis. *Pure Appl. Geophys.* **1988**, *126*, 27–36. [[CrossRef](#)]
16. Arikawa, T.; Muhari, A.; Okumura, Y.; Dohi, Y.; Afriyanto, B.; Sujatmiko, K.A.; Imamura, F.; Fisheries, J.M.O.M.A.A.; Bandung, W.J.I.T. Coastal Subsidence Induced Several Tsunamis during the 2018 Sulawesi Earthquake. *J. Disaster Res.* **2018**, *13*, 1–3. [[CrossRef](#)]
17. Heidarzadeh, M.; Muhari, A.; Wijanarto, A.B. Insights on the source of the 28 September 2018 Sulawesi tsunami, Indonesia based on spectral analysis and numerical simulations. *Pure Appl. Geophys.* **2018**, *176*, 25–43. [[CrossRef](#)]
18. Omira, R.; Dogan, G.G.; Hidayat, R.; Husrin, S.; Prasetya, G.; Annunziato, A.; Proietti, C.; Probst, P.; Paparo, M.A.; Wronna, M.; et al. The September 28th, 2018, Tsunami in Palu-Sulawesi, Indonesia: A Post-Event Field Survey. *Pure Appl. Geophys.* **2019**, *176*, 1379–1395. [[CrossRef](#)]
19. Sassa, S.; Tagakawa, T. Liquefied gravity flow-induced tsunami: First evidence and comparison from the 2018 Indonesia Sulawesi earthquake and tsunami disasters. *Landslides* **2019**, *16*, 195–200. [[CrossRef](#)]