Article

Using a data driven approach to predict waves generated by gravity driven mass flows

Zhenzhu Meng [1], Yating Hu [1,2], Christophe Ancey [1]

[1] ENAC/IIC/LHE, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[2] College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing, China
* Correspondence: zhenzhu.meng@epfl.ch; Tel.: +41-2169-35916

Version December 22, 2019 submitted to Water

Abstract: When colossal gravity-driven mass flows enter a body of water, they may generate waves whose effects can have destructive consequences on coastal areas. A number of empirical equations—in the form of power functions of several dimensionless groups—have been developed to predict wave characteristics. However, in some complex cases (for instance, when the mass striking the water is made up of varied slide materials), fitting an empirical equation with a fixed form to the experimental data may be problematic. In contrast to previous empirical equations that specified the mathematical operators in advance, we developed a purely data-driven approach which relies on datasets and does not need any assumptions about functional form or physical constraints. Experiments were carried out using Carbopol Ultrez 10 (a viscoplastic polymeric gel) and polymer–water balls. We selected an artificial neural network model as an example of a data-driven approach to predicting wave characteristics. We first validated the model by comparing it with best-fit empirical equations. Then, we applied the proposed model to two scenarios which run into difficulty when modeled using those empirical equations: (i) predicting wave features from landslide parameters at their initial stage (with the mass beginning to move down the slope) rather than from the parameters at impact; and (ii) predicting waves generated by different slide materials, specifically, viscoplastic slides, granular slides, and viscoplastic–granular mixtures. The method proposed here can easily be updated when new parameters or constraints are introduced into the model.

Keywords: viscoplastic slide; granular slide; landslide-generated waves; data-driven approach; artificial neural network approach; empirical equation.

1. Introduction

When colossal gravity-driven mass flows enter a body of water, such as a sea, a lake, or a reservoir, they sometimes generate waves. These events are particularly relevant in coastal areas and mountainous countries. Such waves occurred, for example, in Papua New Guinea in 1998 [1] and in Vajont, Italy, in 1963 [2]. Predicting the characteristics of waves induced by landslides is of great importance for risk management in coastal areas [3].

Researchers have conducted experiments using physical models to mimic the physical processes of these events. They have simplified water geometry by using 2D or 3D flumes and idealized the sliding masses as rigid blocks [4–6], granular solids [7–13], or viscoplastic fluids [14,15]. Based on reliable experimental data, a number of empirical or semi-empirical equations has been established, either by combining regression techniques with dimensional analysis [16–18] or by a scaling analysis of governing equations [19,20]. All the equations to date have expressed wave characteristics as power functions of several of the slide parameters on impact.

One significant issue has emerged from previous research: on many occasions, empirical equations have fit well with their own experimental data, but they then exhibited large deviations from the datasets obtained by other teams, especially when different slide materials were involved [15,21–23]. Until now, none of the existing equations has been adaptable to experiments conducted with different categories of slide materials. Indeed, the performances of the different equations on a given dataset remain uncertain—a reflection of the limitations of empirical equations of a given functional form. Applying empirical equations is particularly tricky when attempted using a more informative experimental dataset. For instance, previous experimental studies merely used a single slide material (granular materials, viscoplastic materials, or solid blocks); however, many landslides in the real world are made up of mixtures of two or more materials. Taking the viscoplastic–granular mixture as an example, the representative parameters of these two materials are the grain diameter and yield stress, respectively. Due to the current lack of understanding about how these two materials affect the underlying physics of the slide–water interaction, integrating these two parameters into one equation might be problematic if we have a presumed functional form for that equation in advance.

Another key issue is that all the existing empirical equations express wave characteristics from the parameters relating to the sliding masses on impact; none uses the parameters related to the initial stage (i.e., when the mass is still on the slope and starts moving). Putting the emphasis on the parameters on impact makes it easier to control the variables and to provide a quantitative analysis, however, for engineering, there is a need to predict wave characteristics before the sliding has occurred. For example, in May 2009, a slight slope failure occurred on the Guopu bank of the Laxiwa reservoir, in China. Based on monitoring data, a faulted rock mass with an approximate volume of 3×10^7 m^3 showed signs of general displacement [24]. Although there is a very small probability, should the mass drop into the reservoir, it would generate surge waves which may well destroy the nearby arch dam ([25]). In this situation, estimating the characteristics of the potential waves from information on the potential landslide (which is still at rest on the slope) is more than warranted. However, a theoretical study describing the whole process (from the initial stage to wave generation) is lacking, which increases the difficulty in providing physical constraints on the mathematical operators of prediction equations.

Using an approach that did not assume the functional form of the equation in advance and relied strictly on the data alone, would be preferable for dealing with both of the above issues. To the best of our knowledge, there are no previous examples of this in the literature. To overcome the limitations of empirical equations, the present study presents a data-driven method, known as an artificial neural network (ANN) method, which has been successfully employed in other fields to cope with complicated parameters in experimental data processing and to develop highly accurate predictive models [26–31]. In contrast to previous empirical equations, in which mathematical dependence was fixed in advance, the present study provides a new approach in which both the explanatory and explained variables in the data ultimately define their internal relationship without any prior assumptions about the equation's functional form or physical constraints. Moreover, the model can be easily calibrated when new data or parameters become available, which makes it powerful in solving complex problems [32]. Using the ANN method, we (*a*) estimate the wave characteristics from the parameters of a mass at the initial stage, when it is at rest and starts moving down the slope, and (*b*) predict the wave characteristics generated by different slide mass materials (specifically, viscoplastic slides, granular slides, and mixtures of them), all within one model.

2. Experiments

2.1. Physical model

Figure 1 illustrates a physical model of a mass flow moving down a slope and intruding into a body of water. The whole process can be divided into three stages: in stage I, the slide is at rest, in the container box, and then starts moving; in stage II it moves down the slope and reaches the shoreline; in stage III, it enters the body of water and generates waves. We consider a slope with an inclination of θ entering a horizontal flume filled with water. The still-water depth is denoted by h_0, and the water density is denoted by ρ_w. We defined two coordinate systems. The first coordinate system (x, y) is defined with its origin located at the shoreline, with the x-axis proceeding out across the water, stream-wise, and the y-axis going directly upward. The second coordinate system (s, l) is defined with the l-axis being along the slope and the s-axis being perpendicular to the slope. A slide mass, with a volume of V_I and density of ρ_s, is released at a distance l_s from the shoreline. The slide's initial shape is idealized as a rectangle with a height of s_0 and length of l_0. When the sliding mass moves down the slope, its thickness $s(l,t)$ and depth average velocity $v_s(l,t)$ vary as a function of l and t, respectively. The volume of the immersed slide is denoted by V_s. The free water surface $\eta(x,t)$ depends on the horizontal coordinate x and time t. The wave created by the incursion of the sliding mass is evaluated quantitatively by its height h and amplitude a. The gravity acceleration is denoted by g.

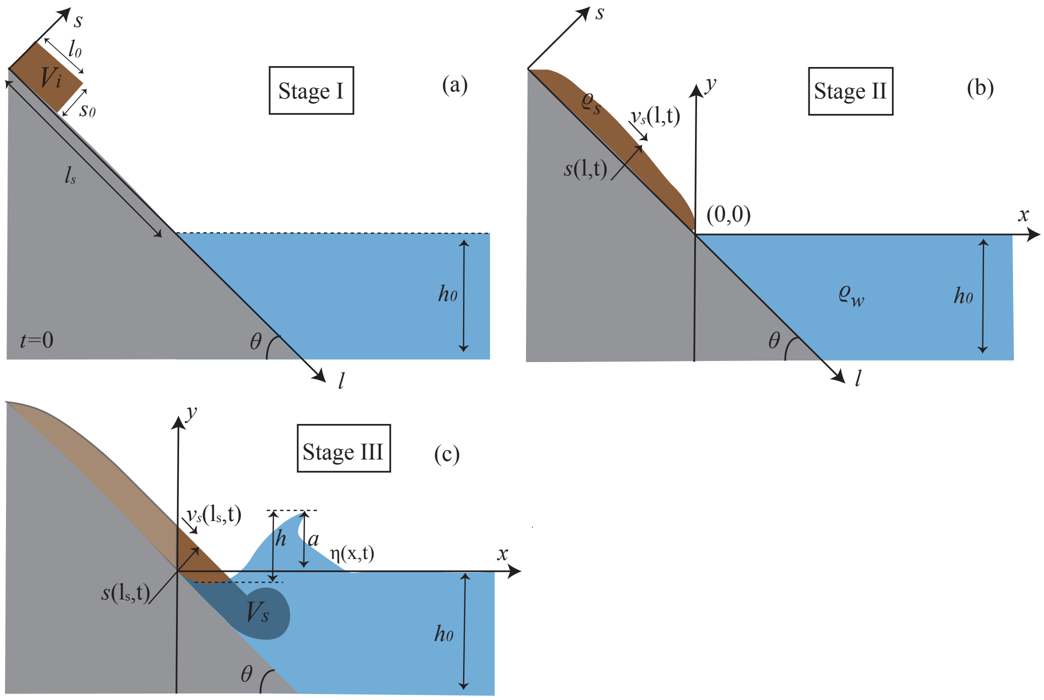

Figure 1. Two dimensional physical model of a landslide generating waves: (a) the slide material is at rest and then starts moving (stage I), (b) the slide material moves down the slope till it reaches the shoreline (stage II), and (c) the slide material intrudes into the body of water and generates waves (stage III).

2.2. Experimental method

Experiments were conducted in a two-dimensional flume at the Swiss Federal Institute of Technology Lausanne (see Figure 2). The experimental facility was devised to mimic snow avalanches penetrating mountain lakes (further information see [20]). The scale factor between the real world and this facility was approximately $r \sim 100$. The flume consisted of two parts. The first part was a 1.5 m long and 0.12 m wide chute, and it could be tilted at an angle θ ranging from 30° to 50°. Its bottom was lined with sandpaper to provide consistent basal friction and its slide walls were made of PVC. The second part was a water-filled, transparent glass flume, 2.5 m long, 0.4 m deep, and 0.12 m wide. The slide mass material was initially contained in a box located at the chute entrance, closed off by a locked gate 0.4 m high and 0.12 m wide. The gate was pneumatically activated and could be opened in less than 0.1 s to release the material from the box. The distance from the gate to the shoreline could be varied from 0.5 m to 1.0 m. Once the slide mass material was released, it accelerated energetically, under gravity, and reached velocities as high as 2.5 m/s. Each experiment's initial settings,including slide mass volume V_i, initial slide length l_0, initial slide height s_0, slope length l_s, still-water depth h_0, and slope angle θ, were recorded before the slide mass material was released.

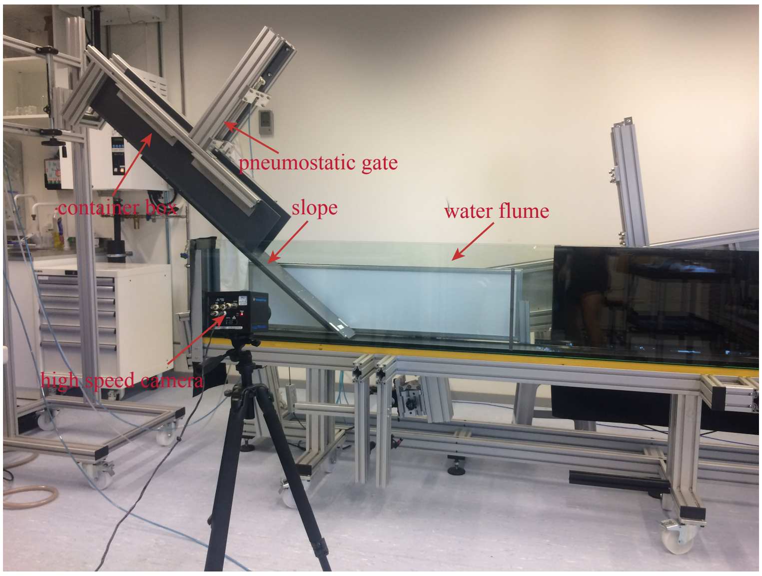

Figure 2. The experimental facility.

108 We selected Carbopol Ultrez 10 viscoplastic material to mimic cohesive landslides, whose
109 rheological behavior can be described using the Herschel–Bulkley model:

$$\tau = \tau_c + K\dot{\gamma}^n \quad (1)$$

110 where τ_c is the yield stress, $\dot{\gamma}$ is the shear rate, K is the slide mass consistency, and n is a power-law
111 index that reflects shear thinning (or shear thickening when n > 1). The rheological measurements
112 of Carbopol were conducted using a Bohlin Gemini rheometer equipped with striated parallel plates
113 (40 mm diameter; 1 mm gap size). The values of τ_c, K and n in the Herschel–Bulkley equation were
114 fitted to the rheological measurements. Table 1 shows how the rheological parameters of Carbopol
115 depend on its concentration C and the proportion of NaOH to Ultrez 10 in the composite. See [33] for
116 the Carbopol Ultrez 10 preparation procedure.

Table 1. Rheological characteristics of the Carbopol used in the present study.

C [%]	Ultrez 10 [g]	NaOH [g]	H_2O [L]	τ_c [Pa]	K [Pa · s^n]	n [-]
1.5	45	18.0	30	38	10.3	0.289
1.6	50	20.7	30	43	12.3	0.293
1.7	53	22.0	30	49	14.4	0.295
1.8	55	22.8	30	53	16.2	0.315
1.9	58	24.0	30	55	17.1	0.321
2.0	60	24.9	30	58	18.9	0.330
2.2	65	26.9	30	60	19.8	0.333
2.3	68	28.2	30	65	23.2	0.339
2.4	70	29.0	30	68	24.6	0.348
2.5	75	31.0	30	74	29.1	0.364
2.7	80	33.2	30	78	32.1	0.388
2.8	85	35.0	30	80	35.8	0.390
3.0	90	37.3	30	85	42.1	0.392

117 We used polymer–water balls to represent granular avalanches. These were produced by soaking
118 dry, water-absorbent beads in water for 4–5 hours. Both Carbopol and the polymer–water balls have a
119 density very close to that of water (1000 kg·m^{-3}), which is also similar to that of the ice (910 kg·m^3)
120 mobilized in snow or ice avalanches. Taking advantage of the similar densities of Carbopol and
121 polymer–water balls, we were able to investigate how mixtures of cohesive and granular materials
122 generated waves without having to consider the effects of the densities of the varying proportions

of each material in the mixtures. Due to the difficulties in finding materials with matching higher densities, the question of how density and mixture proportions interact during wave formation could not be investigated in the current study.

A high-speed camera was placed in front of the shoreline, with its optical axis perpendicular to the sidewall. The camera collected images at a frequency of 200 frames per second, acquiring 600 $\times$ 800-pixel images, corresponding to an observation window of 48 $\times$ 64 cm. We used a 0.2 $\times$ 0.4 m^2 mesh grid to calibrate the raw images and determine the size conversion factor. For each image, we measured (a) the free-water surface when the leading wave reached its maximum height, which helped to deduce the wave amplitudes a_m and h_m, (b) the velocity v_s and thickness s of the sliding mass upon impact, and (c) the volume of the underwater part of the sliding mass V_s.

3. The artificial neural network method

The ANN method is inspired by how the human brain processes information, and it is constructed from interconnected processing elements called neurons [34] (see Figure 3). ANNs are receiving ever greater attention because of their ability to express complex functions in a flexible form. A typical ANN model consists of three main parts: learning rules, network architecture, and an activation function. The network structure is formed of several layers: one input layer, one output layer, and one or several hidden layers, with each layer containing several neurons. Each of the neurons in a layer is connected to neurons of the adjacent layers via coefficients called weightings.

From a mathematical perspective, the principle of neural networks involves the composition of non-linear functions. Starting with a linear model, considering a dataset z and a vector of inputs x, a linear model for the output $\hat{z}(x)$ can be constructed considering $\hat{z}(x) = Wx + \beta$, where the weighting matrix W and the bias vector β are obtained by solving an optimization problem that minimizes the overall difference between z and $\hat{z}$. This process is called $model training$. Such a simple model may lack the flexibility to represent complex functional mapping and, therefore, intermediate variables (layers) y are introduced: $y = \sigma(W^{(1)}x + \beta^{(1)})$ and $z = W^{(2)}y + \beta^{(2)}$, where σ is a user-specified activation function, like the hyperbolic tangent. The composition of several intermediate layers results in a neural network capable of efficiently representing arbitrarily complex function forms.

In this study, we selected a one-hidden-layer network, as an example, and adopted a back-propagation algorithm to train the network [35]. Establishing an ANN model consists of three steps: (i) preparing the required data for training the network; (ii) evaluating neural networks with different structures and choosing the optimal one; and (iii) testing the neural network's performance using data which have not been used previously for training the network.

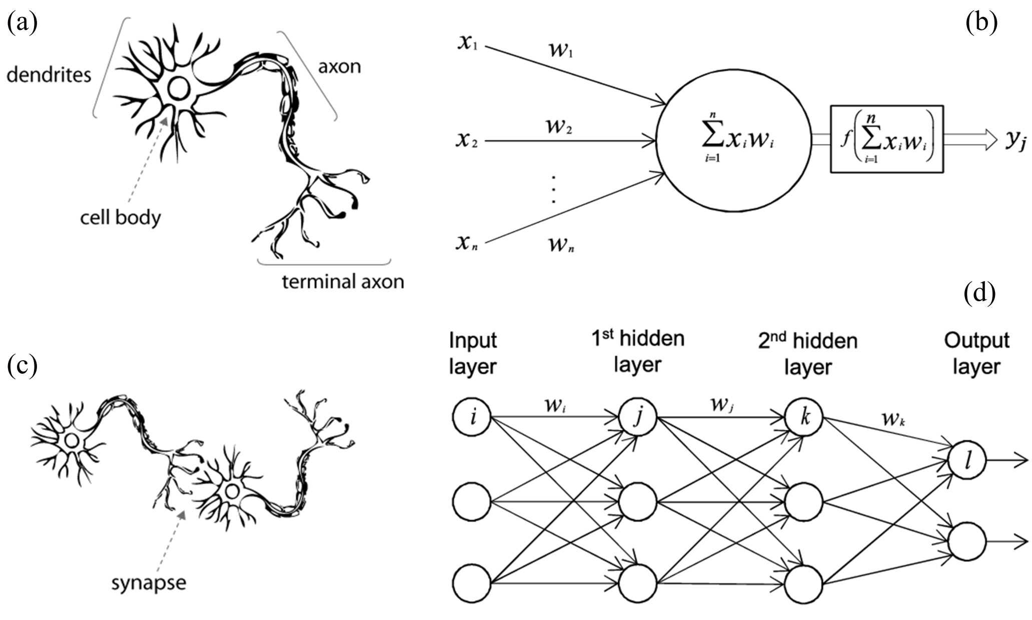

Figure 3. A biological neuron in comparison to an artificial neural network: (a) human neuron; (b) artificial neuron; (c) biological synapse; and (d) ANN synapses [36].

The back-propagation artificial neural network algorithm (BP-ANN) consists of two paths: the feed-forwards and the feed-backwards paths. The feed-forwards path is expressed by equations (2) and (3).

$$y_i = F\left(X_j\right) = F\left(W_{oj} + \sum_{i=1}^{I} W_{ij} x_i\right) \tag{2}$$

$$Z_k = F\left(Y_k\right) = F\left(W_{ok} + \sum_{j=1}^{J} W_{jk} y_i\right) \tag{3}$$

where x_i, y_j, and Z_k represent the input, hidden, and output layers, respectively, W_{oj} and W_{ok} are the bias weights for setting the threshold values, X_j and Y_k temporarily represent computing results before using the activation function, and F is the activation function applied in the hidden and output layers. For the activation function, we chose the sigmoid function, which ranges between 0 and 1 (see equation (4)).The activation function is defined on each layer's neurons and is applied to the sum of the weighted inputs and to each neuron's bias to generate the neuron output.

$$F(a) = \frac{e^a}{e^a + 1} \quad (a = X_j, Y_k) \tag{4}$$

Equation (5) displays the residual function for residual back-propagation training.

$$E = \frac{1}{2}\sum_{k=1}^{K} e_k^2 = \frac{1}{2}\sum_{k=1}^{K} (t_k - z_k)^2 \tag{5}$$

where t_k is the predefined target value and e_k is the residual of each output node. E is the residual between the expected and actual output values. We used a gradient-descent strategy to adjust the weightings, aiming to obtain a minimum E. Equations (6) to (9) express the weightings between the hidden and output layers.

$$\frac{\partial E}{\partial w_{jk}} = -e_k \frac{\partial F(Y_k)}{Y_k} y_j = -\delta_k y_j \tag{6}$$

and hence

$$\delta_k = e_k F'(Y_k) = (t_k - z_k)F'(Y_k) \quad (7)$$

Therefore, the weighting adjustments in the hidden and output link Δw_{jk} can be expressed by equation (8).

$$\Delta w_{jk} = \eta \times y_j \times \delta_k \quad (8)$$

where η is the learning rate ranging between 0 and 1. With a lower learning rate, the network model will take longer time to converge. Conversely, a higher learning rate may lead to a widely oscillating network. In addition, maintaining a consistent learning rate across the model is preferable. The new weighting w_{jk} is updated by the following equation (9), where n is the number of iterations.

$$w_{jk}(n+1) = w_{jk}(n) + \Delta w_{jk}(n) \quad (9)$$

Similarly, the error gradient in the links between the input and hidden layers can be derived from the partial derivative with respect to w_{ij}.

$$\frac{\partial E}{\partial w_{ij}} = \left(\sum_{k=1}^{K} \frac{\partial E}{\partial z_k} \frac{\partial z}{\partial Y_k} \frac{Y_k}{y_j} \right) \times \frac{\partial y_i}{\partial X_j} \times \frac{\partial X_j}{\partial w_{ij}} = -\Delta j x_i \quad (10)$$

where

$$\Delta j = F'(X_j) \sum_{k=1}^{K} \delta_k w_{jk} \quad (11)$$

The new weighting makes (and dominates) the link between the input and hidden layer δw_{jk} can be updated as:

$$\delta w_{ij} = \eta \times x_i \times \delta_j \quad (12)$$

$$w_{ij}(n+1) = w_{ij}(n) + \delta w_{ij}(n) \quad (13)$$

All the input data were normalized in the range between 0 and 1 using the following equation:

$$Y = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (14)$$

where X is the raw data and Y is the normalized data. The initial parameter settings are shown in Table (2).

Table 2. Initial settings for the parameters in the ANN model.

Parameters	Initial setting
Initial weightings	0.2–0.5
Learning rate	0.1
Maximum number of epochs	200
Objective mean square error	0.00001
Training function	traingdx
Momentum parameters	0.9
Activation function	Sigmoid function

4. Results

In section 4.1, we validated the ANN method by comparing its prediction accuracy against empirical equations, using the experimental data generated by the viscoplastic flow. In section 4.2, we predicted the wave characteristics from the slide mass features at rest and as it started moving (stage I in Figure 1). In section 4.3, we developed an ANN model which aimed to cope with the parameters of a landslide with complex properties, specifically, a mixture of cohesive and granular slide mass materials.

Each model's performance was evaluated by its coefficient of determination (R^2), mean square error (MSE), and its sum of squares due to error (SSE), which are expressed as follows:

$$R^2 = 1 - \sum_{i=1}^{n} \left(\frac{\left(y_{pred,i} - y_{exp,i} \right)^2}{\left(y_{pred,i} - y_{exp,avg} \right)^2} \right) \tag{15}$$

$$\mathrm{MSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(\mathrm{y_{pred,i}} - \mathrm{y_{exp,i}} \right)^2}{\mathrm{n}}} \tag{16}$$

$$\mathrm{SSE} = \sum_{i=1}^{n} (\mathrm{y_{exp,i}} - \mathrm{y_{pre,i}}) \tag{17}$$

where n is the number of series of experimental data, $y_{pred,i}$ and $y_{exp,i}$ are the predicted and observed data, respectively, and $y_{exp,avg}$ is the average of observed data.

4.1. Model validation

Most commonly used empirical equations involve the following common dimensional parameters:

$$\eta(x,t) = \eta(h_0, s, v_s, g, V_s, \theta, t, \rho_w, \rho_s) \tag{18}$$

where $\eta(x,t)$ is the free water surface, with x denoting the horizontal coordinate and t denoting time, h_0 is the still-water depth, s is the slide mass thickness, v_s denotes the slide mass velocity on impact, V_s denotes the immersed slide mass volume, θ is the slope angle, and ρ_w and ρ_s are the densities of the water and the slide mass, respectively.

Based on a dimensional analysis or a scale analysis, the scaled wave characteristics can be expressed as a function of several dimensionless groups:

$$X_n = \delta \prod_{i=1}^{N} \Pi_i^{\beta_i} \tag{19}$$

where X represents the scaled wave characteristics (e.g., the scaled maximum wave amplitude, wave height, wave length, impact radius, wave period); Π_i indicates the explanatory variables selected, and where N is the number of explanatory variables.

The predicting equations developed by Zitti *etal*. [20] were the best fit with our experimental data (see equation (20)).

$$X_{1,2} = \delta \Pi_1^{\beta_1} \Pi_2^{\beta_2} \Pi_3^{\beta_3} \tag{20}$$

where $X_{1,2} = H_m, A_m$, and $\Pi_1 = \dfrac{v_s}{\sqrt{g h_0}}$ is the slide mass Froude number, $\Pi_2 = \dfrac{s}{h_0}$ is the scaled slide mass thickness, and $\Pi_3 = \dfrac{\rho_s V_s}{B h_0^2}$ is the scaled impacted slide mass.

The coefficients of explanatory variables δ and $\beta_{1,2,3}$ were acquired by fitting the experimental data based on a linear regression technique. The empirical equations of A_m and H_m for the present study were:

$$A_m = 1.2973 \Pi_1^{0.6170} \Pi_2^{0.1626} \Pi_3^{0.6406} \tag{21}$$

$$H_m = 1.4368 \Pi_1^{0.9700} \Pi_2^{0.0768} \Pi_3^{0.6076} \tag{22}$$

Using the same database and explanatory variables as Equation (21), we modeled the experimental data using our ANN method. Thus, the three neurons in the input layer and the two neurons in the output layer were:

- 3 inputs: Π_1, Π_2, and Π_3
- 2 outputs: A_m and H_m

Of the 291 samples of Carbopol mass slides in the experimental database, 80 % (233 samples) were selected as training data for model construction and 20 % (58 samples) were saved as test data for model validation, providing an independent measure of ANN performance after training. Samples for each group were selected randomly.

We used a basic three-layer network structure, namely, one input layer, one hidden layer, and one output layer. To select the optimal number of neurons in the hidden layer, we set a random number of neurons and ran the program, determining their performance by the coefficient of determination R^2. Each run was repeated 5 times and R^2 was calculated by eliminating the maximum and minimum coefficients of determination and averaging the results of the remaining three tests. As shown in Figure 4, the R^2 of both H_m and A_m reached their maximum values when the hidden layer contained 5 neurons. Thus, the optimum network for the present study was a 3–5–2 structure (input–hidden–output).

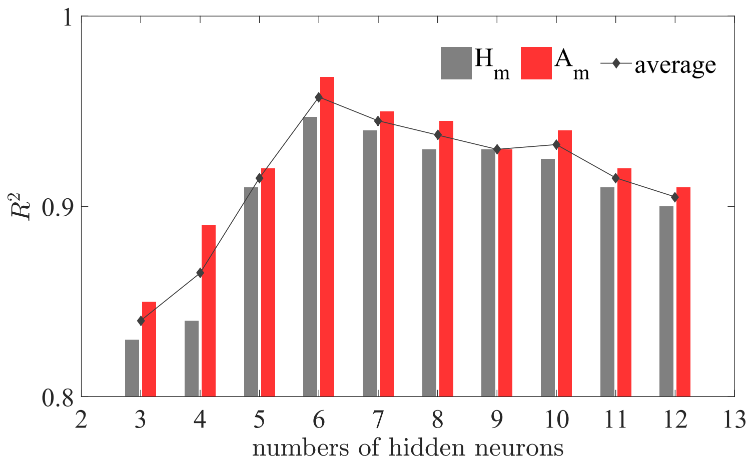

Figure 4. Variation of R^2 versus the number of neurons in the hidden layer.

Model training was constrained by the following indicators: the maximum epoch number was initially set to 100; the objective MSE was set to 1×10^{-4}; the minimum gradient was set to 1×10^{-5}; the maximum number of validation fails, which represents the number of successive iterations that the validation performance fails to decrease, was initially set to 6. Training would stop once one of the indicators mentioned above reached its initial value; for instance, in the present study, training stopped when the number of validation fails reached 6. Figure 5 illustrates the evolution of these indicators (i.e., gradient, validation fails, and MSE) at each epoch until the training is stopped.

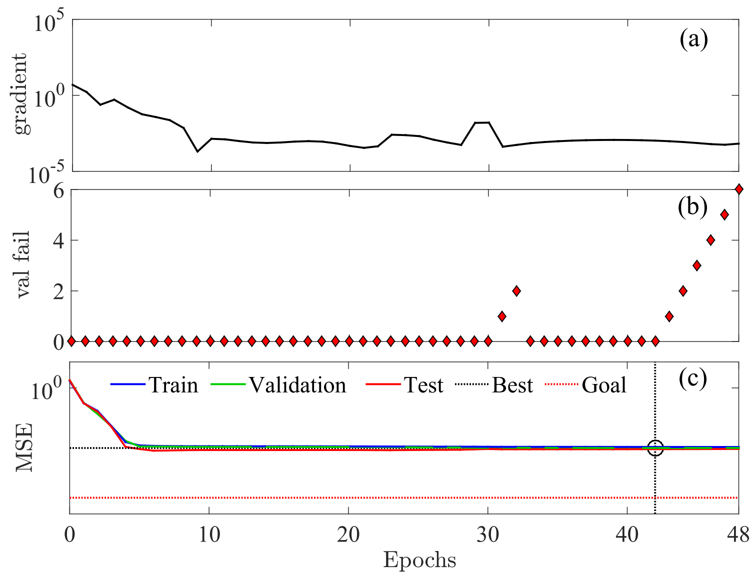

Figure 5. Variations in (a) the gradient, (b) the number of validation fails, and (c) MSE, against epochs.

In Figure 5 (c), the MSEs of the training data and the test data were counted separately. The curves of the evolution of the MSE for these three data series were very close, indicating the model's high level of adaptability. The best validation performance was an MSE = 0.00025337 at epoch 43, and the training terminated at epoch 48 as the number of validation fails reached 6. The gradient = 0.0011736 at epoch 48. Figure 6 displays a histogram of the residuals between the predicted A_m and the observed A_m. The probability density of the residuals approximately follows a Gaussian distribution.

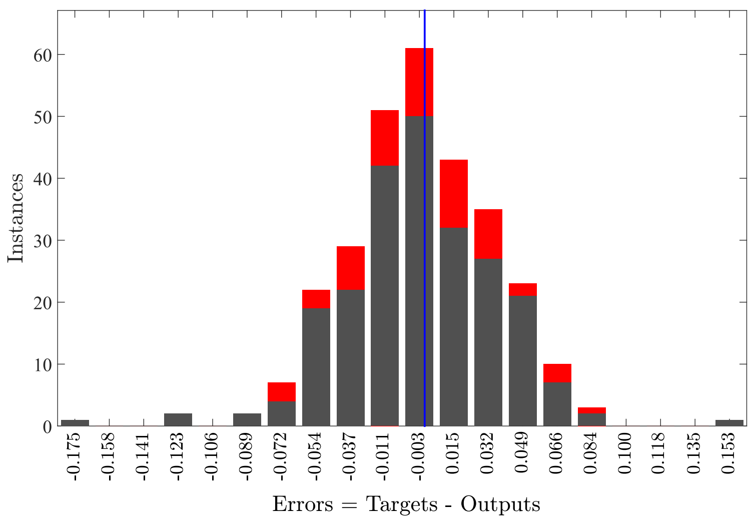

Figure 6. Error histogram of A_m with 20 bins. Red part denotes test data and grey part denotes training data.

Figure 7 displays the observed A_m and H_m versus the predicted data modeled using the ANN model and the empirical equations. The R^2 of A_m and H_m in the ANN model were 0.9682 and 0.9479, respectively; the R^2 of A_m and H_m predicted by the empirical equations were 0.9214 and 0.9062, respectively. In view of this prediction accuracy, the ANN model outperformed the best-fitting empirical equation. In addition, the R^2 of A_m was always slightly higher than that of H_m, in both models, which may result from measurement errors in the experiments.

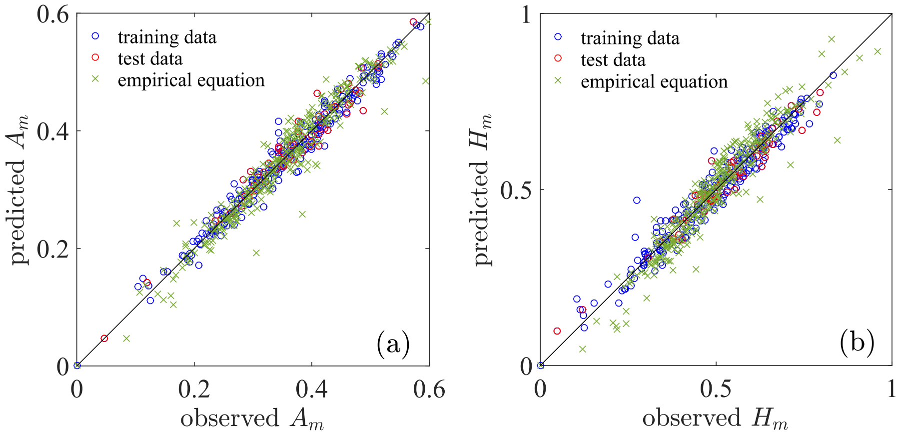

Figure 7. Q-Q plot of observed and predicted (a) A_m and (b) H_m, for the empirical equations and the ANN model. Training data and test data in the ANN model are displayed separately.

4.2. Prediction of wave characteristics from initial mass slide parameters

Previously, empirical or semi-empirical equations determined wave characteristics from the mass slide features on impact (illustrated as stage II in Figure 1), and these equations were established in the form of the power-law equations of several dimensionless groups (see Equation (21)). None predicted the wave characteristics from the slide mass features at the initial stage (illustrated as stage I in Figure 1).

A theoretical study focusing on a description of the whole process is still lacking, which makes it difficult to provide physical constraints on the mathematical operator of the prediction equation. In this case, assuming a functional form for the prediction equation in advance might be problematic. Therefore, a data-driven approach that relies strictly on the data rather than on a fixed form equation is preferable, and the ANN method thus fits this requirement. The process involves the following parameters:

$$\eta(x,t) = \eta(\tau_c, K, n, l_0, s_0, l_s, V_I, h_0, \theta, \rho_w, \rho_s, t, g) \tag{23}$$

where τ_c is the yield stress, K is the consistency, n is the power-law index, l_0 and s_0 are the length and height of the slide mass in its container box, l_s is the slope length, V_I is the initial slide mass volume, h_0 is the still-water depth, θ is the slope angle, ρ_w is the water density, ρ_s is the slide mass density, t is time, and g is the gravity acceleration.

The slide mass's rheological parameters include τ_c, K, and n. Although they have little effect on the slide mass–water interaction and wave formation [15], they have great effects on the slide mass flowing down the slope. The Pearson correlation coefficients between each pair of these three parameters were all above 0.9 (see Table 3), indicating that all three parameters correlated highly. We therefore selected the yield stress τ_c, namely the stress at which the material starts yielding, to represent the rheological parameters.

Table 3. The Pearson correlation coefficients between τ_c, K, and n.

	τ_c	K	n
τ_c	1	0.9739	0.9604
K	0.9739	1	0.9633
n	0.9604	0.9633	1

Figure 8 provides a first insight into how the wave characteristics depend on the rheological properties of the slide mass and on its parameters at the initial stage. It shows experimental data with the yield stress set at τ_c = 41 Pa, 62 Pa, and 80 Pa. Overall, the maximum wave amplitude a_m increased with rises in the yield stress τ_c and the initial slide mass m_I, and with decreases in the slope length l_s.

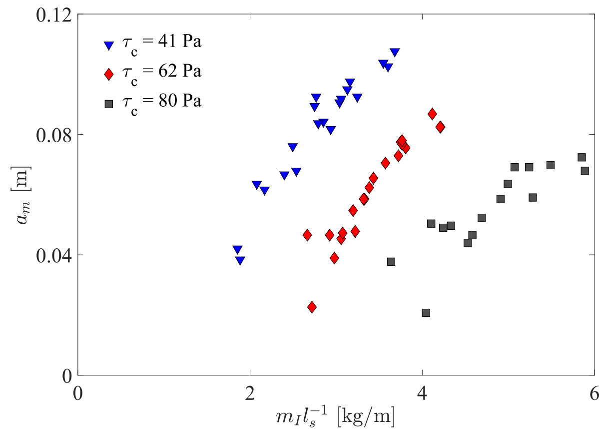

Figure 8. Variations in wave amplitude a_m against $m_I l_s^{-1}$, with the water depth h_0 = 0.2 m and slope angle θ = 45°.

$\epsilon = \dfrac{l_*}{h_*}$ and $\varsigma = \dfrac{s_*}{h_*}$ are aspect ratios for the l-axis to the y-axis, and for the s-axis to the y-axis, respectively. The natural choice for defining the typical scale introduced by these ratios was to take the dimensions of the reservoir: $l_* = l_0$, $h_* = h_0$, and $s_* = s_0$. The Bingham number can be expressed as $Bi = \dfrac{\tau_c}{K(v_*/s_*)}$, which is a dimensionless yield stress (relative to the viscous forces). We assumed that the viscoplastic flow reached a near-equilibrium regime, where viscous forces balanced gravity acceleration, and the velocity scale was then $v_* = (\rho g \sin\theta/K)^{1/n} s_*^{1+1/n}$. The Bingham number then became $Bi = \dfrac{\tau_c}{\rho g s_0 \sin\theta}$.

The dimensions involved in equation (23) are length [L], mass [M], and time [T]. We chose three scaling parameters: water density ρ_w, still-water depth h_0, and gravitational acceleration g. Thus, the dimensionless form could be expressed as:

$$\eta' = \frac{\eta(x,t)}{h_0} = \eta'\left(\frac{\tau_c}{\rho g s_0 \sin\theta}, \frac{l_0}{h_0}, \frac{s_0}{h_0}, \frac{l_s}{l_0}, \theta, \frac{\rho_s}{\rho_w}\right) \tag{24}$$

where the η' is the scaled free-water surface elevation. As with Section 4.1, we selected the scaled maximum wave amplitude A_m and height H_m to represent the water surface elevation. As the slide mass density ρ_s and water density ρ_w were constant throughout our experiments, $\dfrac{\rho_s}{\rho_w}$ can be eliminated. There were therefore five neurons in the input layer and two neurons in the output layer:

- 5 inputs: Bi, ϵ, ς, $\dfrac{l_s}{l_0}$, and θ
- 2 outputs: A_m and H_m

The modeling method used was the same as in Section 4.1. First, based on the optimal number of hidden neurons determined, a 5–10–2 network structure was developed; then, the experimental data were divided into training data and test data; finally, the ANN model was trained using the training data and validated using the test data. The R^2, MSE, and SSE of A_m were 0.8983, 0.00089, and 0.2591, respectively. The H^2, MSE, and SSE of A_m were 0.8497, 0.00295, and 0.8483, respectively. Because the $R^2 > 0.8$, the present model is validated. Yet compared with the scenario that predicted wave characteristics from the slide mass parameters on impact, the prediction accuracy of the ANN method in the present scenario was lower. The more complicated the physical process is, the more information could be lost in prediction.

4.3. Waves generated by viscoplastic–granular mixtures

Recent studies have mimicked landslides in the real world by using a single slide mass material, including granular slides, viscoplastic materials, or solid blocks. However, many landslides in the natural world are mixtures of granular and viscoplastic materials, and these have not been studied either experimentally or theoretically. In the present study, we conducted experiments using mixtures of polymer–water balls and Carbopol, with the percentage of Carbopol in volume varying symmetrically (0 %, 20 %, 50 %, 80 % and 100 %). Figure 9 shows raw images, captured by a high-speed camera, of Carbopol, polymer–water balls, and mixtures of them, entering the body of water. These represented landslides with different degrees of cohesion.

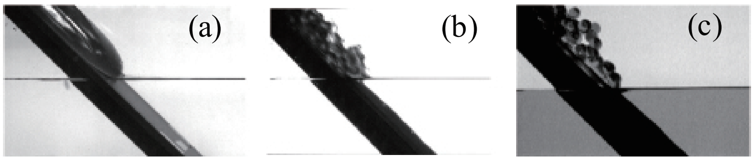

Figure 9. Raw images of landslides intruding into a body of water, as recorded by a high-speed camera: (a) Carbopol, (b) mixture of 50 % Carbopol and 50 % polymer–water balls, and (c) polymer–water balls.

As shown in Figure 10, higher waves could be generated with higher proportions of Carbopol in the mixture, which implies that the slide mass material's composition influenced wave generation. Here, to provide identical criteria for all slide mass materials, we quantified the slide mass properties using a universal dimensionless group named the *Impulse product parameter* P, which was proposed by [9]:

$$\mathrm{P} = \Pi_1 \Pi_2^{1/2} \Pi_3^{1/4} \cos(6/7\theta)^{1/2} \tag{25}$$

where Π_1, Π_2, and Π_3 denote the same parameters as Equation (20).

One issue which should be noted is that the properties of granular slides are usually represented by their grain diameters, whereas the rheological behavior of viscoplastic materials is commonly described using yield stress. It is difficult to integrate these two parameters into one equation in the form of a power-law equation. To overcome this limitation and provide a compatible model for these parameters, we applied the ANN method so as to avoid assuming the functional form of a prediction equation.

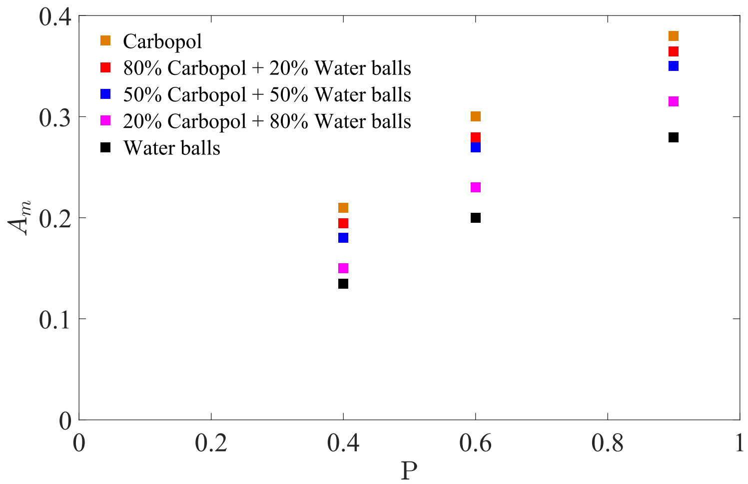

Figure 10. Effects of slide mass material composition on the scaled maximum wave amplitude A_m.

As underlined above, the dimensionless parameters in modeling experiments with a single material commonly involve the slide Froude number Π_1, relative slide mass Π_2, and the relative slide thickness Π_3. To quantify the properties of mixed viscoplastic and granular slides, we introduced the following dimensionless groups: the Bingham number Bi$= \dfrac{\tau_c}{\rho_s g s_0 \sin\theta}$, which represents the rheological properties of a cohesive material; the scaled diameter of the granular slide mass $D_s = \frac{d_g}{h_0}$, where d_g is the diameter of a granular particle and h_0 is the still-water depth; the volume ratio of the viscoplastic material in the mixture $R_V = \frac{V_s}{V_g + V_s}$, where V_s is the volume of the viscoplastic slide mass and V_g is the volume of the granular slides; and the density ratio between the two materials $R_\rho = \frac{\rho_s}{\rho_g}$, which is a constant in present study.

Hence, the input layer contained 6 neurons $\{\Pi_1, \Pi_2, \Pi_3, \mathrm{Bi}, D_s, \text{and } R_V\}$, and the output layer, as in the above sections, contained $\{A_m \text{ and } H_m\}$. Using the same method presented in Section 4.1, the number of hidden neurons was determined, and the network's optimum structure was 6–8–2. The R^2, MSE, and SSE of A_m were 0.9325, 0.0072, and 0.2172, respectively. The R^2, MSE, and SSE of H_m were 0.9173, 0.00178, and 0.6154, respectively. The prediction performance for waves generated by viscoplastic–granular mixtures was quite good.

5. Discussion

5.1. Model adaptability

In Sections 4.2 and 4.3, we presented two applications which were difficult to model using empirical equations with a fixed functional form:

- One application was predicting wave characteristics from slide mass features at the initial stage I. As a theoretical study analyzing this whole process is still lacking, it is difficult to provide physical constraints for the mathematical operator of the empirical equation. In this case, assuming a functional form for the predictive equation in advance might be problematic.
- Another application was predicting waves generated by viscoplastic–granular mixtures. Previous studies had mimicked real-world landslides by using single slide mass materials, however, many landslides in the natural world are mixtures of granular and viscoplastic materials, and these have not been studied either experimentally or theoretically. The properties of granular slides are usually represented by their grain diameters, whereas the rheological behaviors of viscoplastic materials are commonly described using yield stress. It is difficult to integrate these two parameters into one equation in the form of a power-law equation.

Both these scenarios can easily be adapted using the ANN method's high prediction accuracy (see Table 4). This clearly demonstrates the advantage of using a purely data-driven method in terms of model adaptability (and this is not limited to an ANN method). In contrast to equations with fixed formulae, the ANN method has no external constraints, making it a scalable open system. In addition, it has the ability to self-update and is highly adaptable when new parameters become available or fresh constraints appear (they are not limited to the two scenarios presented in this study). With more informative, richer datasets, stronger correlations can be built from the input layer to the output layer.

Table 4. The R^2, MSE, and SSE values of the models described.

	empirical equations		ANN model (3–6–2)		ANN model (5–10–2)		ANN model (6–8–2)	
	A_m	H_m	A_m	H_m	A_m	H_m	A_m	H_m
R^2	0.9214	0.9062	0.9682	0.9479	0.8983	0.8497	0.9325	0.9173
MSE	0.00081	0.00197	0.00025	0.00107	0.00089	0.00295	0.00072	0.00178
SSE	0.2571	0.6266	0.0865	0.3088	0.2591	0.8483	0.2172	0.6154

5.2. Prediction accuracy

Table 4 displays the coefficient of determination R^2, mean square error (MSE), and sum of squares due to error (SSE) values for each of the models presented in Section 4. The following features are worth noting:

- Compared with the empirical equations based on regression techniques, the ANN model gave more precise predictions. Using the same explanatory variables, the coefficient of determination R^2 improved from 0.9214 to 0.9682 for A_m, and from 0.9062 to 0.9479 for H_m. Of course, the improvement in prediction accuracy is not large.
- The prediction precision for A_m was greater than for H_m in predictions made with empirical equations and with the ANN models. This may be because the experimental measurement errors of wave height h_m were larger than those for wave amplitude a_m. Prediction precision not only depends on the prediction performance of the model selected, but it also relies on experimental accuracy.
- The predictions of wave features from the parameters at impact were better than the predictions from the parameters at the initial stage. Also, prediction precision decreased when the dataset involved combinations of different slide mass materials. Thus, prediction precision decreased as experimental complexity increased and more parameters were involved.

5.3. Multicollinearity

Multicollinearity is a phenomenon where one explanatory variable in a multiple regression model can be linearly predicted from the others with a substantial degree of accuracy. This may lead to the problem that the multiple regression's coefficient estimates change erratically in response to small changes in the model. The natural logarithmic form of empirical equation (Equation (20)) can be written as:

$$\ln X = \ln \delta + \alpha \ln \Pi_1 + \beta \ln \Pi_2 + \gamma \ln \Pi_3 \tag{26}$$

The coefficients $\ln \delta$, α, β, and γ were estimated using the least squares (linear regression) method based on experimental data. As length [L] was scaled by the still-water depth h_0, h_0 appears in the three aggregated parameters Π_1, Π_2 ,and Π_3, and specifically, they are correlated with $h_0^{-1/2}$, h_0^{-1}, and h_0^{-2}, respectively. The high correlations among explanatory variables may result in multicollinearity during the linear regression. However, to date, none of the studies using empirical equations has discussed multicollinearity.

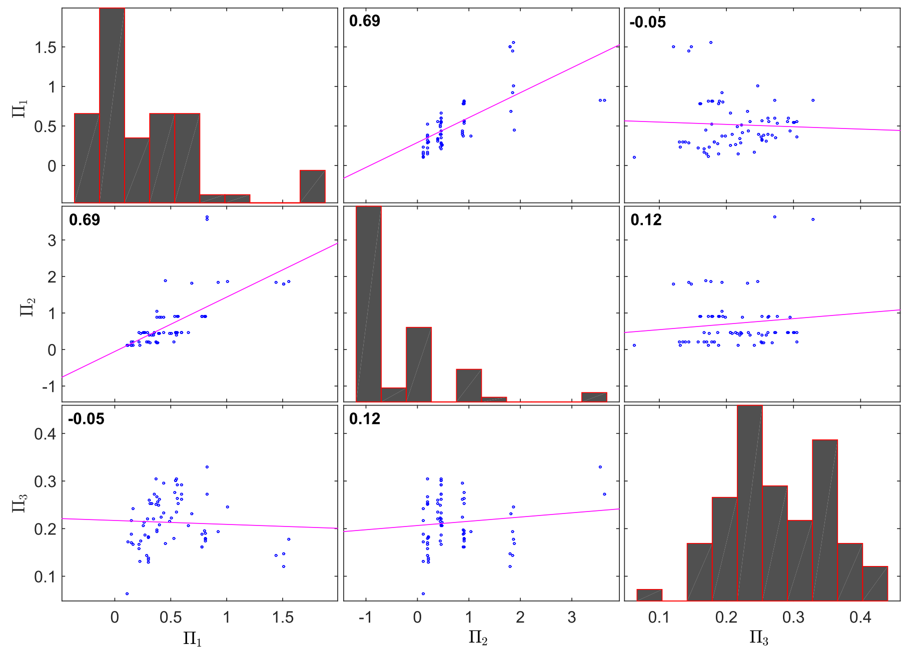

Figure 11. Correlation matrix of explanatory variables Π_1, Π_2, and Π_3 in Equation (20).

To estimate the correlations between each pair of explanatory variables, we calculated their Pearson correlation coefficients. As illustrated in Figure 11, the r between Π_1 and Π_2 is relatively high ($r = 0.69$), however, it is still under the upper limit of 0.8. We may therefore consider that the multicollinearity lies within an acceptable range.

5.4. Limitations

The present study explored the possibility of extracting models purely from data, however, data-driven models may suffer from a lack of interpretability, e.g., the difficulty in explaining causal relationships between the data, the discrepancy, and the corresponding prediction. The use of deep learning strategies and vast amounts of data in the inference process exacerbate this issue.

6. Conclusion

This study applied an artificial neural network (ANN) method—one of the most commonly used machine learning methods—to predict the characteristics of waves generated by gravity-driven slide masses. Laboratory experiments were conducted using a viscoplastic material (Carbopol), a granular material (polymer–water balls), and mixtures of them. After validating the ANN model by comparing its prediction accuracy with that of empirical equations, we applied the model to two scenarios: (i) predicting wave characteristics from the parameters of landslides initially at rest on the slope and (ii) integrating the parameters of different categories of slide mass material into one model, i.e., a Bingham number for the viscoplastic material and the grain diameter for the granular material. In the first scenario, the R^2 for H_m and A_m were 0.8983 and 0.8497, respectively, and in the second scenario, the R^2 for H_m and A_m were 0.9325 and 0.9173, respectively. As a purely data-driven method, this ANN method was easy to adapt when new parameters were included or fresh constraints occurred.

Author Contributions: conceptualization, C.A. and Z.M.; experiments, Z.M. and Y.H.; methodology, Z.M. and Y.H.; validation, Z.M.; formal analysis, Z.M.; writing–original draft preparation, Z.M.; writing–review and editing, Z.M. and C.A.; supervision, C.A.; project administration, C.A.; funding acquisition, C.A.

Funding: This research was funded by the EPFLs Civil Engineering Department and the Swiss National Science Foundation (Grant No. 200021 146271/1 for a project called *Physics of Basal Entrainment*). Z.M. acknowledge the support of the China Scholarship Council (Grant No. 201506710074).

Acknowledgments: Part of the preliminary tests were conducted by EPFL student Jeremy Bussat. We are grateful to our colleagues Bob de Graffenried for his help.

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

References

1. Synolakis, C.E.; Bardet, J.P.; Borrero, J.C.; Davies, H.L.; Okal, E.A.; Silver, E.A.; Sweet, S.; Tappin, D.R. The slump origin of the 1998 Papua New Guinea tsunami. Proceedings of the royal society of london A: Mathematical, physical and engineering sciences. The Royal Society, 2002, Vol. 458, pp. 763–789.
2. Muller, L. The rock slide in the Vajont Valley. *Rock Mechanics and Engineering Geology* **1964**, *2*, 148–212.
3. Fuchs, H.; Pfister, M.; Boes, R.; Perzlmaier, S.; Reindl, R.; Zuerich, E.; und Glaziologie, H. Impulse waves due to avalanche impact into Kuehtai reservoir; Impulswellen infolge Lawineneinstoss in den Speicher Kuehtai **2011**.
4. Wiegel, R.L.; Noda, E.K.; Kuba, E.M.; Gee, D.M.; Tornberg, G.F. Water waves generated by landslides in reservoirs. *Journal of the Waterways, Harbors and Coastal Engineering Division* **1970**, *96*, 307–333.
5. SLINGERLAND, R.L.; Voight, B. Occurrences, properties, and predictive models of landslide-generated water waves. In *Developments in geotechnical engineering*; Elsevier, 1979; Vol. 14, pp. 317–394.
6. Watts, P. Tsunami features of solid block underwater landslides. *Journal of waterway, port, coastal, and ocean engineering* **2000**, *126*, 144–152.
7. Huber, A. Impulse waves in Swiss lakes as a result of rock avalanches and bank slides, Experimental results for the prediction of the characteristic numbers of these waves. Trans. 14th Int. Congress Large Dams, 1982, Vol. 3, pp. 455–476.

8. Fritz, H.M. Initial phase of landslide generated impulse waves. PhD thesis, ETH Zurich, 2002.
9. Heller, V.; Hager, W.H. Impulse product parameter in landslide generated impulse waves. *Journal of waterway, port, coastal, and ocean engineering* **2010**, *136*, 145–155.
10. Heller, V.; Moalemi, M.; Kinnear, R.D.; Adams, R.A. Geometrical effects on landslide-generated tsunamis. *Journal of Waterway, Port, Coastal, and Ocean Engineering* **2011**, *138*, 286–298.
11. Heller, V.; Spinneken, J. 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* **2015**, *104*, 113–134.
12. Heller, V.; Bruggemann, M.; Spinneken, J.; Rogers, B.D. Composite modelling of subaerial landslide–tsunamis in different water body geometries and novel insight into slide and wave kinematics. *Coastal Engineering* **2016**, *109*, 20–41.
13. Miller, G.S.; Andy Take, W.; Mulligan, R.P.; McDougall, S. Tsunamis generated by long and thin granular landslides in a large flume. *Journal of Geophysical Research: Oceans* **2017**, *122*, 653–668.
14. Meng, Z. Experimental study on impulse waves generated by a viscoplastic material at laboratory scale. *Landslides* **2018**, *15*, 1173–1182.
15. Meng, Z.; Ancey, C. The effects of slide cohesion on impulse-wave formation. *Experiments in Fluids* **2019**, *60*, 151.
16. Heller, V. Landslide generated impulse waves: Prediction of near field characteristics. PhD thesis, ETH Zurich, 2007.
17. Mohammed, F.; Fritz, H.M. Physical modeling of tsunamis generated by three-dimensional deformable granular landslides. *Journal of Geophysical Research: Oceans* **2012**, *117*.
18. Zitti, G.; Ancey, C.; Postacchini, M.; Brocchini, M. Impulse waves generated by snow avalanches falling into lakes. Proceedings of 36th IAHR World Congress, The Hague, Netherlands, 2015. IAHR, 2015, number EPFL-CONF-215899.
19. Walder, J.S.; Watts, P.; Sorensen, O.E.; Janssen, K. Tsunamis generated by subaerial mass flows. *Journal of Geophysical Research: Solid Earth* **2003**, *108*.
20. Zitti, G.; Ancey, C.; Postacchini, M.; Brocchini, M. Impulse waves generated by snow avalanches: Momentum and energy transfer to a water body. *Journal of Geophysical Research: Earth Surface* **2016**, *121*, 2399–2423.
21. Zweifel, A. Impulswellen: Effekte der Rutschdichte und der Wassertiefe. PhD thesis, ETH Zurich, 2004.
22. Lindstrøm, E.K. Waves generated by subaerial slides with various porosities. *Coastal Engineering* **2016**, *116*, 170–179.
23. Meng, Z.; Ancey, C. Experimental comparison of tsunami generated by viscoplastic and granular slides. EGU General Assembly Conference Abstracts, 2018, Vol. 20, p. 6170.
24. Su, H.; Li, J.; Cao, J.; Wen, Z. Macro-comprehensive evaluation method of high rock slope stability in hydropower projects. *Stochastic environmental research and risk assessment* **2014**, *28*, 213–224.
25. Liu, Y.; Wang, X.; Wu, Z.; He, Z.; Yang, Q. Simulation of landslide-induced surges and analysis of impact on dam based on stability evaluation of reservoir bank slope. *Landslides* **2018**, *15*, 2031–2045.
26. Abraham, A. Artificial neural networks. *handbook of measuring system design* **2005**.
27. Yegnanarayana, B. *Artificial neural networks*; PHI Learning Pvt. Ltd., 2009.
28. Kim, D.H.; Park, W.S. Neural network for design and reliability analysis of rubble mound breakwaters. *Ocean engineering* **2005**, *32*, 1332–1349.
29. Lee, A.; Geem, Z.W.; Suh, K.D. Determination of optimal initial weights of an artificial neural network by using the harmony search algorithm: application to breakwater armor stones. *Applied Sciences* **2016**, *6*, 164.
30. Armaghani, D.J.; Mohamad, E.T.; Hajihassani, M.; Abad, S.A.N.K.; Marto, A.; Moghaddam, M. Evaluation and prediction of flyrock resulting from blasting operations using empirical and computational methods. *Engineering with Computers* **2016**, *32*, 109–121.
31. Gedik, N. Least Squares Support Vector Mechanics to Predict the Stability Number of Rubble-Mound Breakwaters. *Water* **2018**, *10*, 1452.
32. Dou, J.; Yamagishi, H.; Pourghasemi, H.R.; Yunus, A.P.; Song, X.; Xu, Y.; Zhu, Z. An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan. *Natural Hazards* **2015**, *78*, 1749–1776.
33. Cochard, S. Measurements of time-dependent free-surface viscoplastic flows down steep slopes. PhD thesis, EPFL Lausanne, 2007.

34. Liu, J.; Chang, H.; Hsu, T.; Ruan, X. Prediction of the flow stress of high-speed steel during hot deformation using a BP artificial neural network. *Journal of materials processing technology* **2000**, *103*, 200–205.
35. Rumelhart, D.E.; McClelland, J.L.; Group, P.R.; others. *Parallel distributed processing*; Vol. 1, MIT press Cambridge, MA, 1987.
36. Suzuki, K. *Artificial Neural Networks-Architectures and Applications*; 2013.