water

Article

A SMA-SVM-Based Prediction Model for the Tailings Discharge Volume After Tailings Dam Failure

Gaolin Liu [1], Bing Zhao [1,2,*], Xiangyun Kong [2,3], Yingming Xin [1], Mingqiang Wang [1] and Yonggang Zhang [4,*]

1. Faculty of Land Resources Engineering, Kunming University of Science and Technology, Kunming 650093, China; liugaolin7204@163.com (G.L.)
2. Yunnan International Technology Transfer Center for Mineral Resources Development and Solid Waste Resource Utilization, Kunming 650600, China
3. Faculty of Civil Engineering and Mechanics, Kunming University of Science and Technology, Kunming 650500, China
4. Engineering Research Institute, China Construction Eighth Engineering Division Corp., Ltd., Shanghai 200122, China

* Correspondence: b912681848z@163.com (B.Z.); demonzhangyg@163.com (Y.Z.)

Abstract: Tailings ponds can recycle water resources through the water recirculation system by clarifying and purifying the wastewater discharged from the mining production process. Due to factors such as flooding and heavy rainfall, once a tailings dams burst, the spread of heavy metals in the tailings causes underground and surface water pollution, endangering the lives and properties of people downstream. To effectively assess the potential impact of tailings dams bursting, many problems such as the difficulty of taking values in predicting the volume of silt penetration through empirical formulae, model testing, and numerical simulation need to be solved. In this study, 65 engineering cases were collected to develop a sample dataset containing dam height and storage capacity. The Support Vector Machine (SVM) algorithm was used to develop a nonlinear regression model for tailings discharge volume after tailings dam failure. In addition, the model penalty parameter C and kernel function g were optimized using the powerful global search capability of the Slime Mold Algorithm (SMA) to develop an SMA–SVM prediction model for tailings discharge volume. The results indicate that the volume of tailings discharged increases nonlinearly with increasing dam height and tailings storage capacity. The SMA-SVM model showed higher prediction accuracy compared to the predictions made by the Random Forest (RF), Radial Basis Function (RBF), and Least Squares SVM (LS-SVM) algorithms. The average absolute error in tailings discharge volume compared to actual values was 30,000 m^3, with an average relative error of less than 25%. This is very close to practical engineering scenarios. The ability of the SMA-SVM optimization algorithm to produce predictions with minimal error relative to actual values was further confirmed by the combination of numerical simulations. In addition, the numerical simulations revealed the flow characteristics and inundation area of the discharged sediment during tailings dam failure, and the research results can provide reference for water resource protection and downstream safety prevention and control of tailings ponds.

Keywords: support vector machine; slime mold algorithm; numerical simulation; flow characteristics

check for updates

Received: 19 December 2024
Revised: 5 February 2025
Accepted: 13 February 2025
Published: 19 February 2025

Citation: Liu, G.; Zhao, B.; Kong, X.; Xin, Y.; Wang, M.; Zhang, Y. A SMA-SVM-Based Prediction Model for the Tailings Discharge Volume After Tailings Dam Failure. *Water* **2025**, *17*, 604. https://doi.org/10.3390/w17040604

1. Introduction

By 2023, there were more than 6000 existing tailings dams in China, the largest number in the world. These tailings dams serve as critical infrastructure for mining operations and are subject to flooding, rainfall, and tailing. They also pose a significant safety threat to downstream residents. The impact on downstream water resources and human life and property can be devastating if a tailings dam fails [1]. In recent years, tailings dam failures have become increasingly common. For example, in 2008, the Kingston Fossil Plant tailings dam in the United States failed, affecting an area of 1.6 km^2 and releasing 5.4 million m^3 of tailings [2]. In 2010, excessive saturation and inadequate flood storage capacity led to the failure of a tailings dam in Kolontár, Hungary. This incident affected an area 8 km downstream and resulted in 10 deaths and 120 injuries [3]. In 2019, a tailings dam failure at the Córrego do Feijão iron mine in Brumadinho, Brazil, caused nearly 300 deaths [4]. In 2020, a tailings dam leak at Yichun Luming Mining Co., Ltd. in Yichun, Heilongjiang Province, released more than 2 million m^3 of tailings. This resulted in excessive concentrations of molybdenum along a stretch of about 340 km of the river that runs from the Yijimi River to the Hulan River [5]. Many researchers have studied the effects of tailings dam failures in order to develop appropriate preventive measures to reduce the loss of life and property, given the immense destructive potential of tailings dam failures.

Investigative analyses show that factors such as the storage capacity and the height of the dam of the tailings reservoir are critical factors responsible for the tailings discharge volume and the impact distance [6,7]. Xu et al. summarized empirical formulae for dam breach flow rate and impact distance, on the basis of which Chen et al. [8,9] optimized these empirical formulae by considering the properties of tailings and their deformation during flow using the MAC staggered grid. However, the theoretical calculations are generally based on the worst-case scenario and do not take into account the influence of the topography of the downstream area. In order to accurately reflect the actual conditions of tailings dam failures and predict their effects, Yin [10] conducted large-scale model tests to investigate the flow characteristics of tailings dam failures for the first time. Du et al. [11–16] conducted dam-failure experiments to obtain information on peak flow rates and impact areas. They also proposed prediction methods and preventive measures for dam-failure effects. Zhang et al. [17] set up a 42-meter-long large water tank to conduct physical model tests with different particle sizes and dam-failure shapes. The results show that any dam material could experience instability and failure, with finer debris travelling further during the dam-failure process. Zhao et al. [18] derived a flow regime equation to support the prediction of dam-failure impact ranges by investigating the impact velocity and flow characteristics of water–sand mixtures under seismic conditions. However, due to site and technical constraints, similar dam-failure model tests can be affected by issues such as mechanical and material similarity, which can bias the results [19–21]. Therefore, in order to address engineering issues, numerical simulations have become essential to create virtual topographic conditions. The results indicate that any dam material could experience instability and failure, with finer tailings traveling farther during the dam breach process. Zhao et al. [18] investigated the impact velocity and flow characteristics of water–sand mixtures under seismic conditions and derived a flow regime equation that supports predictions of dam breach impact ranges. However, due to site and technical constraints, similar model tests on dam failures may be affected by issues such as mechanical and material similarity, which can distort results [19–21]. Consequently, numerical simulations to construct virtual topographical conditions have become essential for addressing engineering issues. For instance, the Volume of Fluid method is employed to establish mechanical models that capture fluid surfaces during dam failure [22,23]. This method simplifies the post-failure sediment flow into cubic particles through grid division. Moreover, particle flow discrete element–finite difference simulations have emerged as a primary method for tailings dam modeling [24–26]. Nevertheless, the selection of constitutive relations and parameters during numerical simulations can lead to significant deviations between simulated and actual results for some projects [27,28], All of the above studies provide reference and basis for tailings pond failure prediction; however, understanding of how to realize rapid prediction of sand breakage volume based on a large number of data statistics is still lacking.

With the advancement of artificial intelligence, algorithms such as Random Forest, Neural Networks, and SVM are increasingly applied in engineering fields [29–31]. To address the issue of low predictive accuracy caused by parameter selection in theoretical calculations, model tests, and numerical simulations, a sample dataset of tailings dam-failure cases was developed based on the principles of machine learning. By training these samples, a nonlinear mapping relationship was established to predict tailings discharge volume. Based on this, this paper selects the mucilage optimization support vector machine algorithm based on the comparison of various optimization algorithms, constructs the mucilage optimization support vector machine model (SMA–SVM), and predicts the amount of sand breakage after the tailing pond breaks with the tailing pond capacity and the dam height as the characteristic factors. The predictions were compared with results from Random Forest, Radial Basis Function Neural Networks, Least Squares SVM, large-scale physical model tests, and numerical simulations, which validated the predictive accuracy of the SMA optimization algorithm. The flow characteristics of tailings after dam failure were analyzed. The findings provide valuable references for dam-failure prevention of tailings reservoirs.

2. Impact Assessment Model for Tailings Dam Failure

2.1. SVM Algorithm for Assessing the Impact of Tailings Dam Failure

SVM is a classical classification method in machine learning and has been extensively employed for both linear and nonlinear regression. The assessment of the impact of tailings dam failure can be viewed as a complex nonlinear regression problem [32]. Considering the primary factors influencing dam failure, the input variables are the dam height of the tailings reservoir (x_1) and its storage capacity (x_2), while the output is the tailings discharge volume (f_1). According to the principle of the SVM, input samples x_i were mapped into a high-dimensional feature space, where the fitting relationship between inputs and outputs is expressed as Equation (1).

$$f_{(x)} = \omega\phi(x_i) + b \tag{1}$$

where x_i represents the features of the input samples (dam height and storage capacity); ω denotes the weights of each feature; $\phi(x_i)$ is the nonlinear mapping function; and b is the threshold.

Following the principle of structural risk minimization, slack variables ξ and ξ^* were introduced to obtain the optimal coefficients ω and b. Consequently, the nonlinear regression was transformed into a solvable optimization problem, as presented in Equation (2). The corresponding constraints are represented in Equation (3).

$$\min_{y,\xi_i b}\left[\frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)\right] \tag{2}$$

$$\begin{cases} \omega\phi(x_i) + b - y_i \leq \varepsilon + \xi_i \\ y_i - \omega\phi(x_i) - b \leq \varepsilon + \xi_i^n \\ \xi_i \geq 0, \xi_i^n \geq 0 \\ i = 1, 2, \cdots n \end{cases} \tag{3}$$

where C is the penalty parameter, and ξ_i, ξ_i^n represents the training error.

To transform the samples into completely separable samples in the feature space, a Lagrangian function was introduced to convert the original problem into a new dual problem for solving, and the predictive model was obtained, as shown in Equation (4).

$$f(x) = \sum_{i=1}^{n} \alpha_1 K(x_i, x) + b \tag{4}$$

where n is the number of samples; α_1 the Lagrange multiplier; and $K(x_i,x)$ is the kernel function, whose choice determines the structure and complexity of the high-dimensional feature space. Given the complexity of predicting the impacts of tailings dam failures [33], this study selected the Gaussian kernel function to map input tailings dam features into high-dimensional space, as shown in Equation (5).

$$K(x_i, x) = \exp\left\{ \frac{-\|x_i - x\|^2}{2g^2} \right\} \tag{5}$$

where g represents the bandwidth of the Gaussian kernel.

Based on the principles of the SVM outlined above, the nonlinear training process for predicting tailings discharge volume and influence distance after tailings dam failure is illustrated in Figure 1.

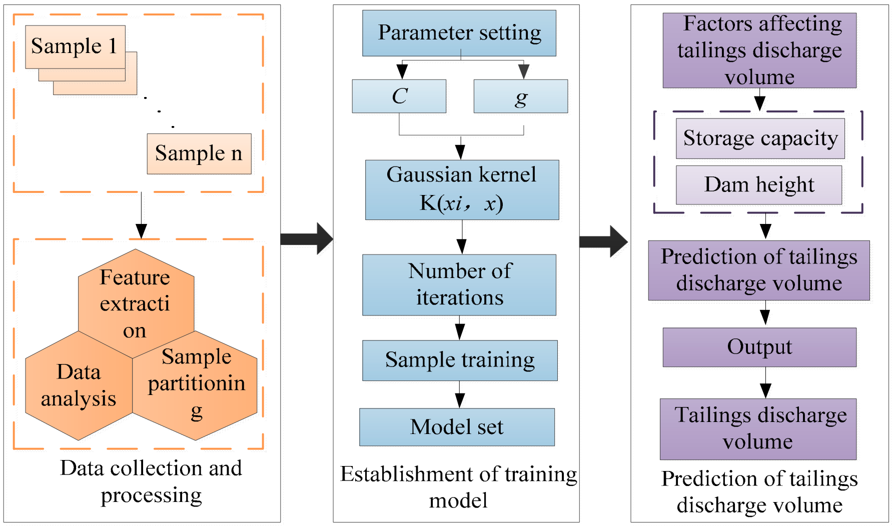

Figure 1. Principle of the SVM Algorithm.

2.2. Principle of the SMA Optimization Algorithm

Inspired by the foraging behavior of slime molds, the Slime Mold Algorithm (SMA) is a heuristic optimization algorithm. Shimin et al. [34] summarized this behavior into three stages: food seeking, food approach, and food capture. Slime molds develop a gradually expanding fan shape and extend tentacle-like structures during the foraging process. After several rounds of development, an interconnected vein-like structure is finally formed, which allows for a comprehensive search for food.

Slime molds update their paths based on optimal theory during the foraging process. When potential food sources vary in quality, they select sources of high concentration and choose the optimal paths to connect these food sources, thereby constructing robust pathways. Research suggests that the food source along the foraging path is inversely proportional to the speed of decision making, with faster decision making reducing the likelihood of finding food [35]. This foraging process enables molds to weigh optimal weighting against potential risk, ultimately determining the best weighting for each feature of the sample.

When slime molds encounter high-concentration food sources during their division process, they do not gather in that region for a long time; instead, they continue to divide and search for other food sources. If a new food source has a lower concentration than others, the slime molds quickly move away [36,37]. Therefore, during the process of surrounding food, slime molds form connection channels of varying sizes. The mathematical model for the SMA is established as follows.

(1) Approaching food: Slime molds move closer to food based on airborne scents; the higher the food concentration, the larger the mass of the slime molds. When the concentration is too low, they search in other directions, as expressed in Equation (6).

$$X(t+1) = \begin{cases} X_b(t) + V_b(W \cdot X_A(t) - X_B(t), r<p \\ V_c \cdot X(t), r \geq p \end{cases} \tag{6}$$

where t is the current iteration count; X_b is the optimal individual position for fitness; V_b, and V_c are control parameters within $[-a, a]$ and $[-1, 1]$, respectively; W is the fitness; $X_A(t)$ and $X_B(t)$ are two random individual positions; and r is a random number within $[-0, 1]$.

(2) Surrounding food: After surrounding high-concentration food, slime molds continue to separate and search for even higher concentrations. The position iteration at this stage is represented in Equation (7).

$$X(t+1) = \begin{cases} r \cdot (B_u - B_l) + B_l, r \in [0,1] \\ X_b(t) + V_b(W \cdot X_A(t) - X_B(t), r<p \\ V_c \cdot X(t), r \geq p \end{cases} \tag{7}$$

where B_u and B_l are the upper and lower limits of the search range.

(3) Capturing food: During the processes of approaching and surrounding food, the control parameters V_b and V_c oscillate between $[-a, a]$ and $[-1, 1]$ during iterations. When these parameters approach 0, food is captured. The data search process of the SMA is illustrated in Figure 2.

2.3. Construction of the SMA–SVM Algorithm

The selection of parameters in machine learning models directly affects learning effectiveness and efficiency, which makes it a crucial process for enhancing model performance [38,39]. The essence of SVM parameter optimization lies in global searching during the sample learning process. To achieve the optimal performance of the learning model, efficient search strategies must be employed in the parameter space to identify a set of approximately globally optimal parameter combinations during the searching process [40]. According to the basic principles of the SVM learning algorithm, the penalty parameter C and the Gaussian kernel bandwidth g are crucial parameters that need optimization during model training. The penalty parameter C indicates the degree of tolerance of the machine to the interference samples when constructing the optimal hyperplane, while the Gaussian kernel bandwidth g directly determines the structure of the kernel function mapping and affects the radius of samples selected as support vectors by the model. Therefore, both C and g directly impact predictive accuracy. The algorithm is prone to fall into the local optimum when the mapping relationship is complex.

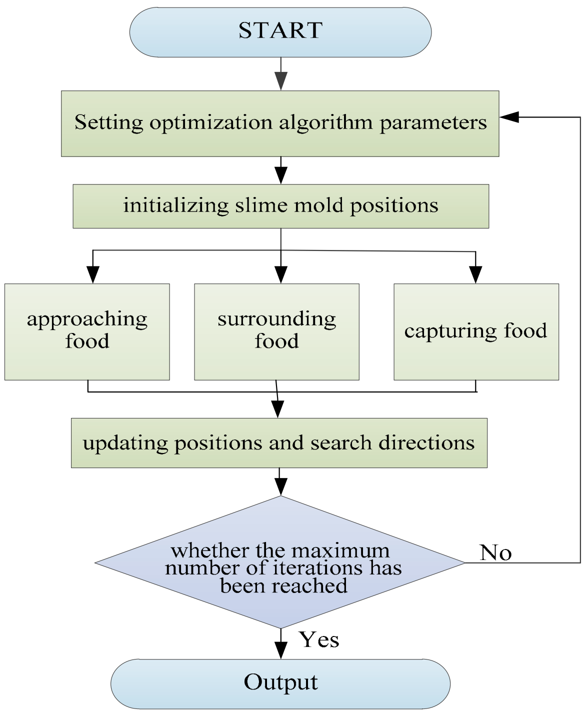

Figure 2. SMA optimization process.

In SVM training, *C* and g tend to be selected from a relatively wide range, typically [0.001, 1000]. Research has shown that overfitting or difficulty in capturing the sample data can occur when empirical methods are used to determine the values of *C* and *g* [41,42]. The introduction of the SMA algorithm uses the characteristics of the approaching and surrounding food to perform a global search on the samples, which allows it to escape local optima to determine the values of *C* and *g*. Algorithm 1 presents the main code for the SMA optimization process, while the principle of SMA optimization is illustrated in Figure 3.

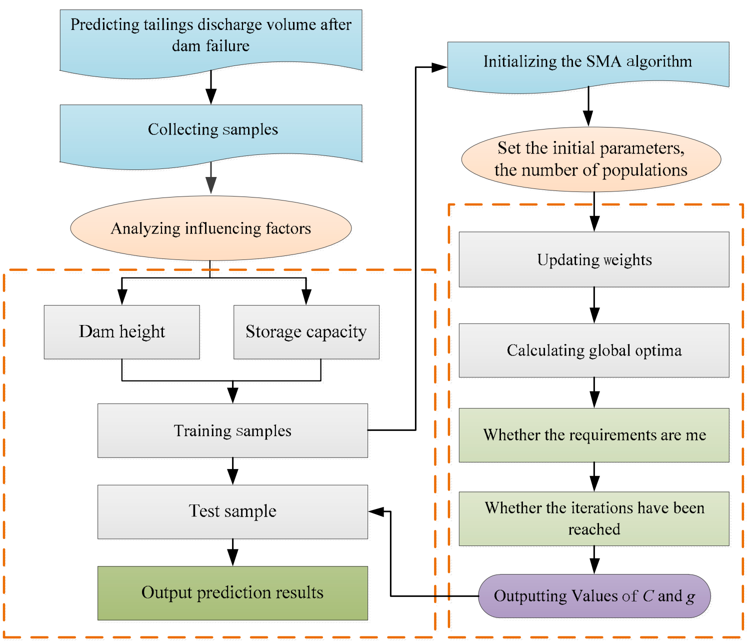

Figure 3. Flowchart of the SMA-optimized SVM algorithm.

Algorithm 1 Slime Mold Optimization Algorithm

1. For each amoeba
2. Begin amoeba at a uniformly distributed random location within the search space.
3. Evaluate the objective function at the current location.
4. Store the objective function value and location as the amoeba's current location, current objective function value, personal best objective function value, and personal best location.
5. Set the amoeba's state to VEGETATIVE.
6. End for
7. Archive the best objective function value from the amoebae.
8. Output optimum

3. Analysis of Predictive Results for Tailings-Dam-Failure Impact Using SMA–SVM

3.1. Prediction Using the SMA–SVM Model

The dam height and tailings storage capacity were used as sample inputs x_i, while the predicted target tailings discharge volume was used as model outputs $f_{(xi)}$, based on the principles of the SVM algorithm and the key factors influencing dam failure. As detailed in Table 1, relevant information from 65 dam-failure cases summarized by Rico et al. was collected to construct a sample dataset [43–45], including 56 training samples and 9 testing samples. During the optimization process of the SMA–SVM algorithm, the sma SVMcg ForRegress program was first used to train the samples and identify the optimal SVM parameters. After SMA optimization, the best penalty parameter C and the best width of the kernel function g were determined and then entered into the test set for the prediction.

Table 1. Collection of tailings dam-failure case samples.

No	Name of the Dam	Impoundment Volume /10^4 m^3	Dam Height /m	Released Volume /10^4 m^3
1	Brumadinho, Mina Corrego'do Feijao, Minas Gerais, Brazil (Vale)	1200	87	957
2	Louyang Xiangjiang Wanji Aluminum, China	200	45	200
3	Fundao-Santarem (Germano), Minas Gerais, Brazil (Samarco = Vale &BHP)	5640	110	4370
4	Imperial Metals, Mt. Polley, British Columbia, Canada	7400	40	2360
5	Ajka Alumina Plant, Kolontár, Hungary (MAL Magyar)	3000	22	100
6	AguaTranque No. 5Nogales, V Region, Valparaíso, Chile	8	16	3
7	Mineracao RioPomba Cataguases, Mirai, MinasGerais, Brazil, Mineraçao (IndustriasQuimicas Cataguases)	380	35	200
8	Tailings Dam, USA	50	43	17
9	Baia Mare, Romania	80	7	10
10	Los Frailes, nearSeville, Spain (Boliden Ltd.)	1500	27	680
11	Omai Mine, Tailings dam No 1, 2, Guyana (Cambior)	525	44	420
12	Middle Arm, Launceston, Tasmania	2.5	4	0.5
13	Merriespruit, nearVirginia, SouthAfrica (Harmony) No 4A Tailings Complex	704	31	60
14	Maritsa Istok 1, Bulgaria	5200	15	50
15	Ajka Alumina Plant, Kolontár, Hungary	450	25	4.3
16	Unidentified, Hernando, County, Florida, USA #2	330	12	0.5
17	Story's Creek, Tasmania	3	17	0.01

Table 1. *Cont.*

No	Name of the Dam	Impoundment Volume /10^4 m^3	Dam Height /m	Released Volume /10^4 m^3
18	Niujiaolong tailings pond, China	110	40	73
19	Bonsal, North Carolina, USA	3.8	6	1.1
20	Cerro Negro No. (4 of 5)	200	40	50
21	Niujiaolong, Hunan (Shizhuyuan Non ferrous Metals Co.)	110	40	73.1
22	Olinghouse, Nevada, USA	12	5	2.5
23	Sipalay, Phillippines, No. 3 Tailings Pond (Maricalum Mining Corp)	2200	25	1500
24	Balka Chuficheva, Russia	2700	25	350
25	Tyrone, New Mexico (Phelps Dodge)	250	66	200
26	Churchrock, New Mexico, United Nuclear	37	11	37
27	Arcturus, Zimbabwe	68	25	3.9
28	Mochikoshi No. 2, Japan (2 of 2)	48	19	0.9
29	Mochikoshi No. 1, Japan (1 of 2)	48	28	8
30	Zlevoto No. 4, Yugoslavia	100	25	30
31	Bafokeng, South Africa	1300	20	300
32	Silver King, Idaho, USA	3.7	9	1.4
33	(unidentified), Southwestern USA	50	43	17
34	Buffalo Creek, West Virginia, USA (Pittson Coal Co.)	50	16	50
35	Cities Service, Fort Meade, Florida, phosphate	1234	15	900
36	Mufulira, Zambia (Roan Consolidated Mines)	100	50	6.8
37	Mir mine, Sgurigrad, Bulgaria	152	45	45
38	El Cobre Old Dam	425	35	190
39	Los Maquis No. 3	4.3	15	2.1
40	Castano Viejo Mine, San Juan, Argentina	2.7	9	1.7
41	Louisville, USA	91	31	66.7
42	Huogudu, Yunnan Tin Group Co.,Yunnan	542	19	330
43	Jupille, Belgium	55	46	13.6
44	Los Cedros, Tlalpujahua, Michoac'an, M'exico	1148	35	250
45	Barahona, Chile	2000	61	280
46	Aitik mine, near Gallivare, Sweden (Boliden Ltd.)	1500	15	180
47	Sgurigrad, Bulgaria	152	45	22
48	Las PalmasPencahue, VIRegion, Maule, Chile (COMINOR)	22	15	17
49	Brunita Mine, Caragena, Spain (SMM Penaroya)	108	25	7
50	Stancil, Maryland, USA	7.4	9	3.8
51	Deneen Mica Yancey County, North Carolina, USA	30	18	3.8
52	Veta de Agua No. 1, Chile	70	24	28
53	Madjarevo, Bulgaria	300	40	25
54	El Cobre New Dam	35	19	35
55	Taoshi, Linfen City, Xiangfen county, Shanxi province, China (TahsanMining Co.)	29	50.7	19
56	Dan River Steam Station, North Carolina (Duke Energy)	155	12	33.4
57	Sotkamo, Kainuu Province, Finland (Talvivaara)	540	25	24
58	Gypsum Tailings Dam (Texas, USA)	636	16	13
59	Prestavel Mine-Stava, North Italy, 2, 3 (Prealpi Mineraria)	40	29.5	18
60	La Luciana, Reocín (Santander), Cantabria, Spain	125	24	10
61	Hokkaido, Japan	30	12	9
62	Hector Mine Pit Pond, MN, USA	18.5	17	12.3
63	Cerro Negro No. (3 of 5)	50	20	8.5
64	Mike Horse, Montana, USA (Asarco)	75	18	15
65	Bellavista, Chile	45	20	7

Tailings discharge volume generally increased nonlinearly with increasing dam height and storage capacity, and the degree of linear fit of a single variable to tailings discharge volume was relatively low, as shown in the tailings dam-failure cases in Table 1. Within the sample dataset, some dam-failure samples had a tailings discharge volume/storage capacity ratio significantly higher than the remaining samples (e.g., samples 2 and 34), which could affect SVM classification performance and lead to larger prediction errors.

When constructing the sample dataset based on Table 1, the value range of the storage capacity in the sample dataset is [2.7, 7400], which differs significantly from those of other parameters. Therefore, normalization was applied to the storage capacity and tailings discharge volume to ensure prediction accuracy. A total of 85% of the samples were used as training samples, while the remaining samples were used as testing samples for prediction. The predictions for the test samples were performed using Random Forest (RF), Radial Basis Function (RBF), Least Squares Support Vector Machine (LS-SVM), and SMA-SVM algorithms. The optimized values of C and g from the SMA-SVM algorithm are listed in Table 2, the fitness curve during the optimization of the SMA algorithm is shown in Figure 4, and the comparison of calculated results with actual values is presented in Table 3.

Table 2. Output results for optimized values of C and g.

Optimization Details	Empirical Values [7]	Optimized Values
Kernel function width g	[0, 5.6]	1.066
Penalty parameter C	[0, 200]	12.991

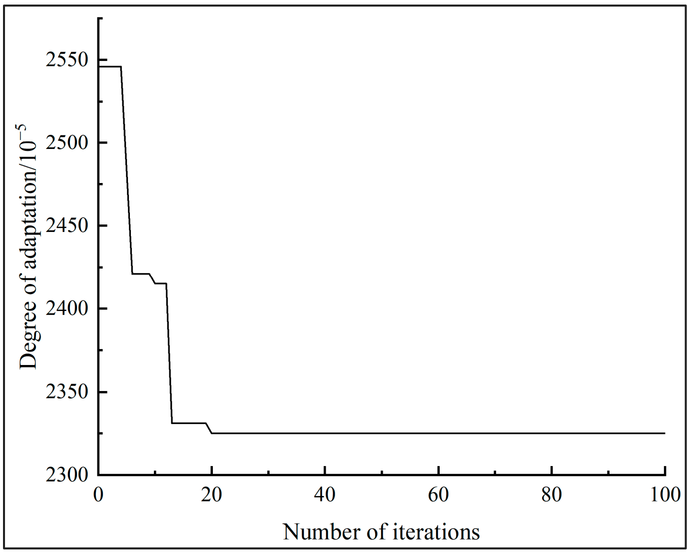

Figure 4. Adaptation curve.

Table 3. Prediction results from different algorithms (10^4 m^3).

Sample ID	Real Values	RF	RBF	LS-SVM	SMA-SVM
Test sample 1	24	342.0	225.3	109.4	21.8
Test sample 2	13	277.5	267.4	104.1	11.0
Test sample 3	18	29.4	17.5	32.2	16.5
Test sample 4	10	29.0	61.2	30.3	11.8
Test sample 5	9	18.1	14.3	13.3	5.0
Test sample 6	12.3	6.0	14.4	5.5	6.2
Test sample 7	8.5	30.1	30.0	11.1	7.8
Test sample 8	15	19.3	41.6	14.5	7.0
Test sample 9	7	24.2	27.6	10.3	7.7

3.2. Analysis and Discussion

As depicted in Table 3, the RF, RBF, LS-SVM, and SMA-SVM algorithms could all predict the impact of dam failures. However, the predictions from the other algorithms exhibited greater variability compared to the SMA-SVM algorithm, while Test1 and Test2 yielded results significantly higher than the actual values. To facilitate analysis and statistics, their impacts were not considered. This study employed the Relative Error (RE) (Equation (8)) and Mean Absolute Error (MAE) (Equation (9)) to analyze the prediction results [46,47] to verify the applicability of the SMA-SVM model in predicting tailings-dam-failure impacts. The absolute errors between the predicted and actual values are presented in Figure 5.

$$RE = \left| y_i - f_{(xi)} \right| \tag{8}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left| y_i - f_{(xi)} \right| \tag{9}$$

where RE represents the relative error; MAE denotes the mean absolute error that reflects the absolute deviation between multiple predicted results and actual values; ρ is the correlation coefficient, where values closer to 1 indicate predictions closer to actual values; n is the number of test samples (1 to 9); y_i is the actual value of the i-th test sample; $\overline{y}$ is the average value of the test samples; $f_{(xi)}$ is the predicted value of the i-th test sample; and $\overline{f}$ is the average value of the predicted values.

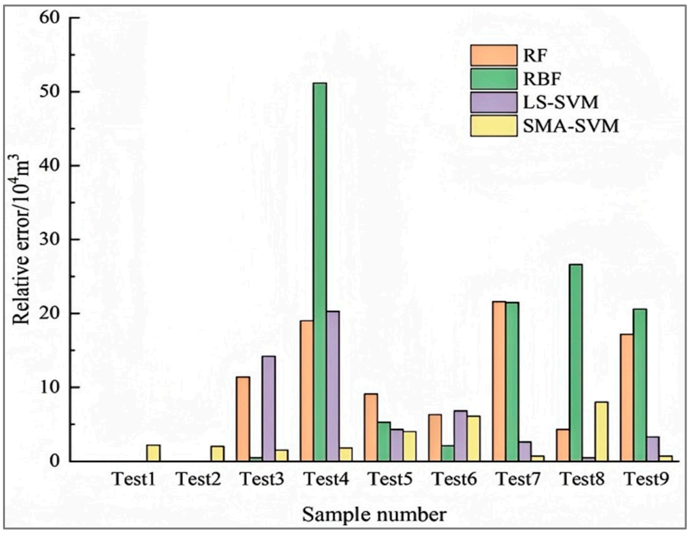

Figure 5. Absolute error of prediction results for tailings discharge volume.

As illustrated in Figure 4, the discrepancies between the predicted results from the SMA-SVM model and the actual values were generally within 60,000 m^3, and predicted values were significantly lower than the storage capacity of the tailings reservoir (3 to 15,500 m^3). Except for Test1 and Test2, the errors of other algorithms were all within 600,000 m^3. This indicates that machine learning can be used to predict the tailings discharge volume after tailings dam failure, but the SMA-SVM model exhibited higher prediction accuracy. The predicted results and mean absolute errors for each sample are plotted in Figure 6, while the relative errors from the SMA-SVM algorithm predictions are illustrated in Figure 7.

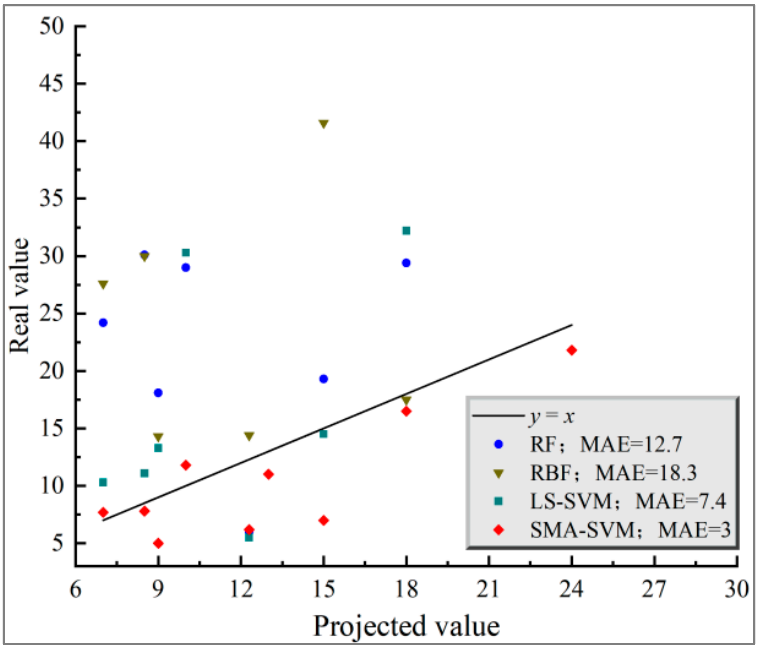

Figure 6. Predicted results from different algorithms.

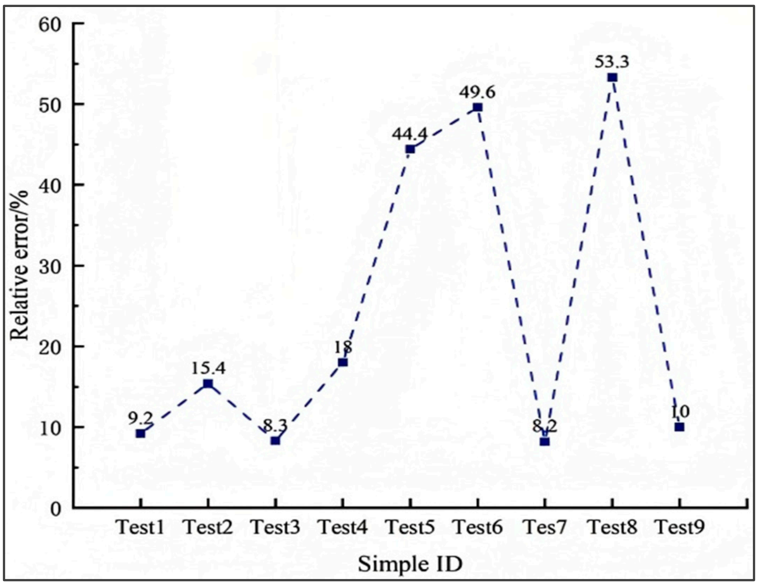

Figure 7. Error analysis of prediction results from SMA-SVM algorithm.

The above analyses confirmed that, with an average absolute error of 30,000 m^3 and an average relative error of less than 25% for tailings discharge volume predictions, the SMA-SVM algorithm provides higher prediction accuracy than other algorithms. Compared to the other algorithms, the maximum improvement in the mean absolute error for the prediction of the tailings discharge volume was 97,000 m^3. This was attributed to the high sensitivity of the RF algorithm to the decision tree parameters. This resulted in poor handling of small sample data due to the randomness in the computational process. The objective of training RBF neural networks is to derive the minimum error function, but the use of the gradient descent strategy often results in local optima. As a result, the RF and RBF algorithms produced relatively high prediction deviations from actual values. SVM is suitable for nonlinear predictive regression. The SMA optimization algorithm provides optimal values of C and g through global search to ensure high prediction accuracy for the model, making it more applicable than the LS-SVM algorithm. However, in addition to the effects of storage capacity and dam height on the tailings flow rate, the method of dam construction also affects the tailings flow rate, leading to certain differences between the prediction results and the actual results.

In summary, the prediction accuracy of the SMA-SVM algorithm is higher than other algorithms, such as the RF algorithm and RBF neural network, in predicting the amount of sand breakage in tailing ponds, so the method can be used for predicting the impact of dam breakage in tailing ponds, such as quickly estimating the impact range of dam breakage according to the amount of sand breakage and the flow characteristics of the discharged tailing slurry during the breakage, which can be used to provide a reference for the preparation of emergency response plans for this kind of accident, and for the implementation of protection design and other projects. However, the accuracy of machine learning is greatly affected by the number of samples, in order to avoid overfitting and underfitting phenomenon, in future research, we can continue to collect similar engineering cases, increase the number of samples and characteristic factors, so as to realize more accurate prediction.

4. Model Test for Tailings Dam Failure and Validation of the SMA-SVM Algorithm

To validate the SMA-SVM algorithm, this study selects a case of dam failure from a valley-type tailings reservoir in Guangdong Province as the background [48]. The initial elevation of the tailings reservoir's dam base is 709.5 m, with the final designed elevation of the stockpile dam top being 830 m. The total dam height is 120.5 m, and the initial dam height is 40.5 m. The storage capacity below the initial dam top elevation of the tailings reservoir is 1,028,000 m^3, while the total capacity is 1.6 million m^3. The tailings reservoir is classified as a Class II reservoir. The initial dam is a permeable rockfill dam, built at the entrance of a narrow and steep valley downstream. The elevation of the dam top is 755 m, with a dam height of 45.5 m and a dam top width of 7.5 m. The length of the dam top axis is 110 m. The upstream and downstream slopes of the dam are 1:1.6 and 1:2, respectively. In 2010, influenced by Typhoon Fanapi, the tailings reservoir experienced an overtopping dam failure, resulting in a tailings discharge volume of approximately 200,000 m^3. Before the tailings discharge, the dam height of the tailings reservoir was 45.5 m, and the storage capacity was about 300,000 m^3. A series of three-dimensional models were established based on the above information about the tailings reservoir. Different numerical simulation scenarios were designed, as shown in Table 4.

Table 4. Numerical simulation scenarios for tailings dam failure.

Simulation Scenario	Dam Height	Storage Capacity
Scenario 1	45.5	30
Scenario 2	45.5	120.8
Scenario 3	120.5	1561

Based on the dam failure simulation scenarios outlined in Table 4, three-dimensional simulations are conducted using FLOW-3D v 11.2 numerical simulation software to obtain the tailings discharge volume after tailings dam failure. During the simulation, five monitoring points (M1 to M5) were established along the flow path of the sediment to collect data on outflow velocity and inundation depth. The three-dimensional numerical model is presented in Figure 8, while the model parameters are listed in Table 5.

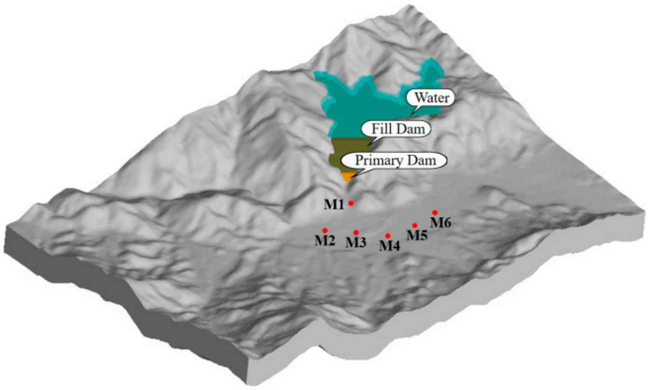

Figure 8. Numerical simulation model of overtopping dam failure (Dam height: 120.5 m).

Table 5. Parameters for numerical simulation.

Simulation Parameters	Value	Simulation Parameters	Value
Diameter/mm	0.0385	Entrainment Coefficient	0.18
Density/kg·m^{-3}	2020	Bed Load Coefficient	8
Critical Shields Number	0.05	Angle of Repose (Degrees)/°	32

Tailings dam failure is a complex and variable process characterized by severe free surface deformations due to erosion, embankment collapse, splashing, and coalescence. The investigation cannot be limited to a single fluid model. FLOW-3D uses the full Navier–Stokes equations together with the continuity equation as the governing equations for fluid motion. It uses the Volume of Fluid (VOF) method to calculate free flows on the fluid surface and combines it with the Fractional Area/Volume Ratios (FAVOR) grid processing approach to optimize the computational grid. Its uniqueness lies in the separate definition of controllable fluids, so that the geometry of the computational grid is independent of the model geometry. Its governing equations are

$$\frac{\partial(uA_x)}{\partial x} + \frac{\partial(uA_y)}{\partial y} + \frac{\partial(uA_z)}{\partial z} = 0 \tag{10}$$

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{1}{V_F}\left(uA_x\frac{\partial u}{\partial x} + vA_y\frac{\partial u}{\partial y} + wA_z\frac{\partial u}{\partial z}\right) = -\frac{1}{\rho}\frac{\partial P}{\partial x} + G_x + f_x \\ \frac{\partial v}{\partial t} + \frac{1}{V_F}\left(uA_x\frac{\partial v}{\partial x} + vA_y\frac{\partial v}{\partial y} + wA_z\frac{\partial v}{\partial z}\right) = -\frac{1}{\rho}\frac{\partial P}{\partial y} + G_y + f_y \\ \frac{\partial w}{\partial t} + \frac{1}{V_F}\left(uA_x\frac{\partial w}{\partial x} + vA_y\frac{\partial w}{\partial y} + wA_z\frac{\partial w}{\partial z}\right) = -\frac{1}{\rho}\frac{\partial P}{\partial z} + G_z + f_z \end{cases} \tag{11}$$

where ρ represents the fluid density; P is the pressure acting on the fluid micro-element; u, v, and w are the corresponding velocity components in the x, y, and z directions; V_F denotes the flowable volume fraction; A_x, A_y, and A_z represent the flowable area fractions in the x, y, and z directions; f_x, f_y, and f_z represent the viscous force accelerations in the x, y, and z directions; G_x, G_y, and G_z represent the gravitational accelerations of the fluid in the three directions.

To truly reflect the flow characteristics during dam failure, this study employs a sediment module to simulate the movement of particles throughout the dam-failure process. In the calculations, the movement of tailings and their interaction with the water flow are represented by the suspended sediment equation. The ability of water flow to carry sediment is highly correlated with flow velocity. As water flow velocity increases, its capacity to carry sediment also increases. Once the water flow velocity reaches a threshold, the deposited tailings are entrained by the water flow, which facilitates them to transition from bed load to suspended load. These two states change alternately with the variations in water flow's entrainment capability. According to the aforementioned theory, the simulation results for various scenarios are presented in Figure 9.

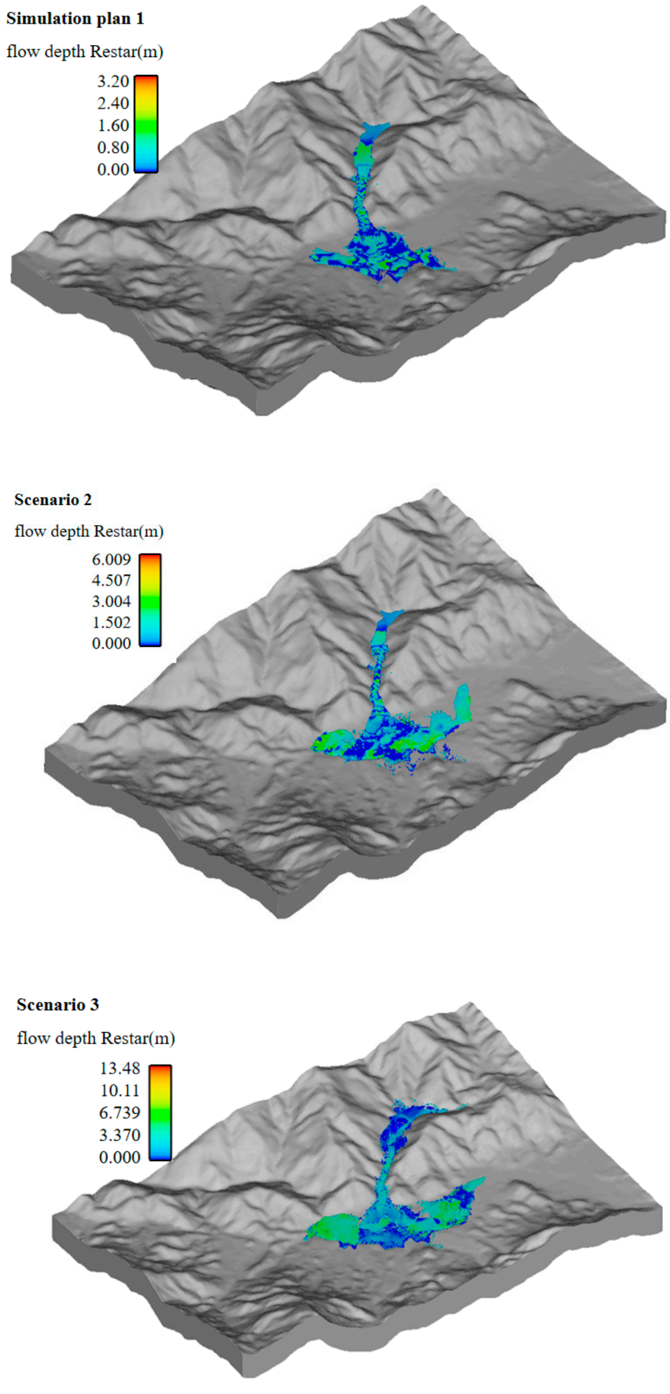

Figure 9. Influence range of tailings dam failure for different scenarios.

As illustrated in Figure 9, numerical simulations revealed that the tailings flow generated by the dam failure of the tailings reservoir is the process in which floodwaters overtopped and scoured the dam surface. During this process, the tailings moved downstream in a slurry state with the water flow and spread to both sides at the valley exit. The dam-failure process ceased when the shear stress increment of the outflow substrate was less than the critical shear stress required to mobilize the tailings. Analyses of the fluid surface area and Facked sediment mass data from the simulation results demonstrate that as the storage capacity and dam height of the tailings reservoir increased, the tailings discharge volume and inundation range after overtopping dam failure exhibited a nonlinear increase. The dam height and storage capacity data from simulation scenarios 1 to 3 were put into the SMA-SVM algorithm to predict the tailings discharge volume. The comparison between the predicted and simulated results is presented in Table 6. To capture the flow characteristics of the discharged tailings, simulation results were collected every 10 s during the simulation process. The flow velocity variation curves for monitoring points M1 to M6 at different times are plotted in Figures 10–15.

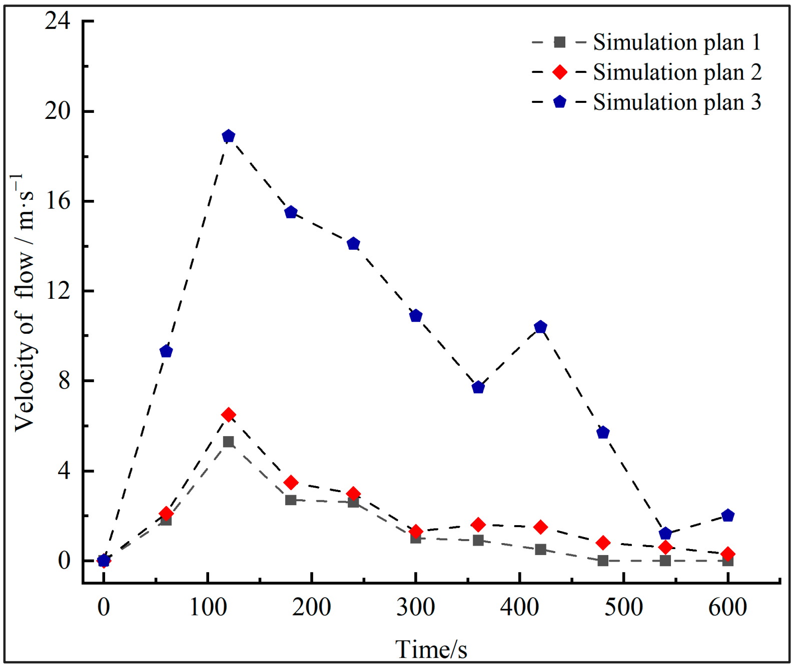

Figure 10. Flow velocity at monitoring point M1 for different dam-failure scenarios.

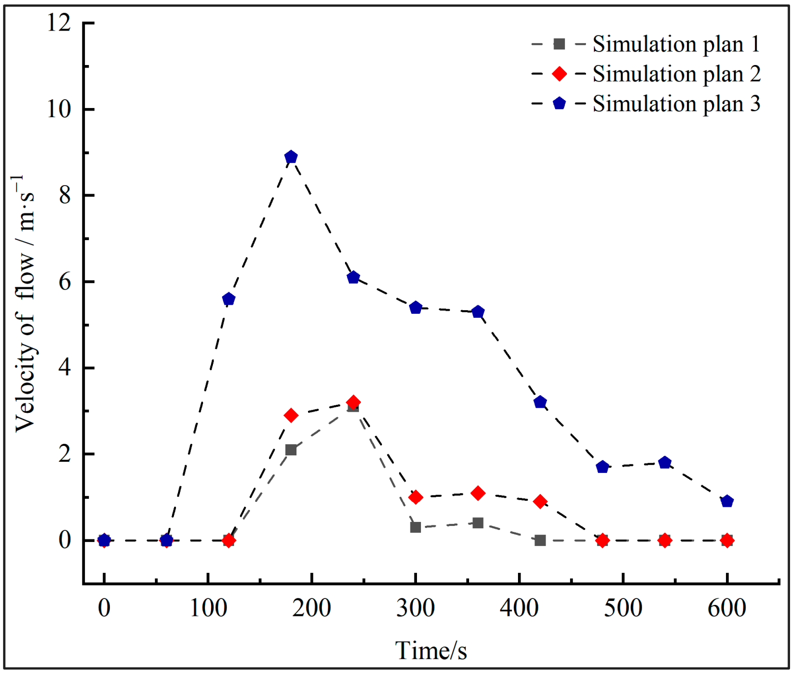

Figure 11. Flow velocity at monitoring point M2 for different dam-failure scenarios.

Table 6. Ratio of numerical simulation results to SMA-SVM prediction results.

Simulation Scenario	Simulated Tailings Discharge Volume/10^4 m^3	Predicted Tailings Discharge Volume by SMA-SVM Algorithm/10^4 m^3
Scenario 1	20.0	15.2
Scenario 2	36.7	18.1
Scenario 3	487.6	427.5

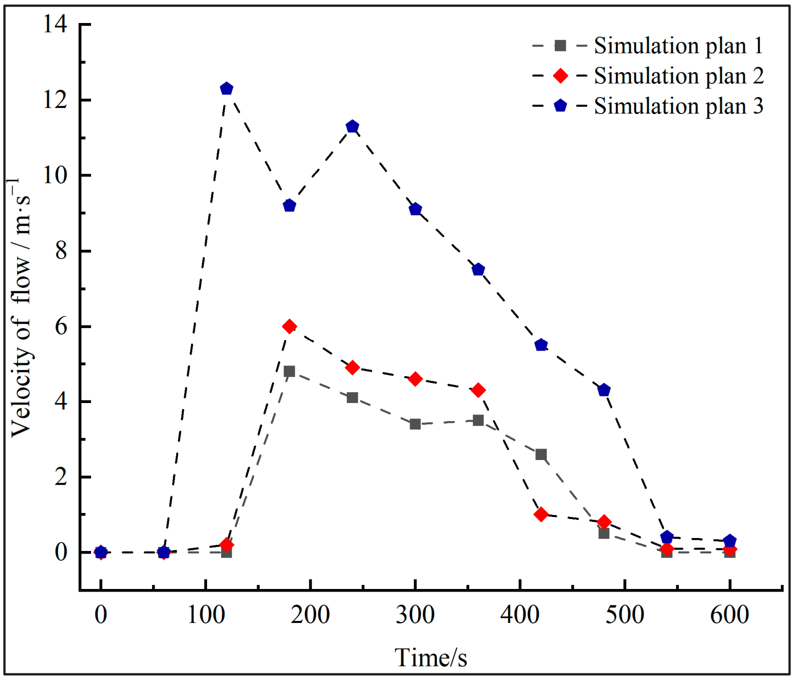

Figure 12. Flow velocity at monitoring point M3 for different dam-failure scenarios.

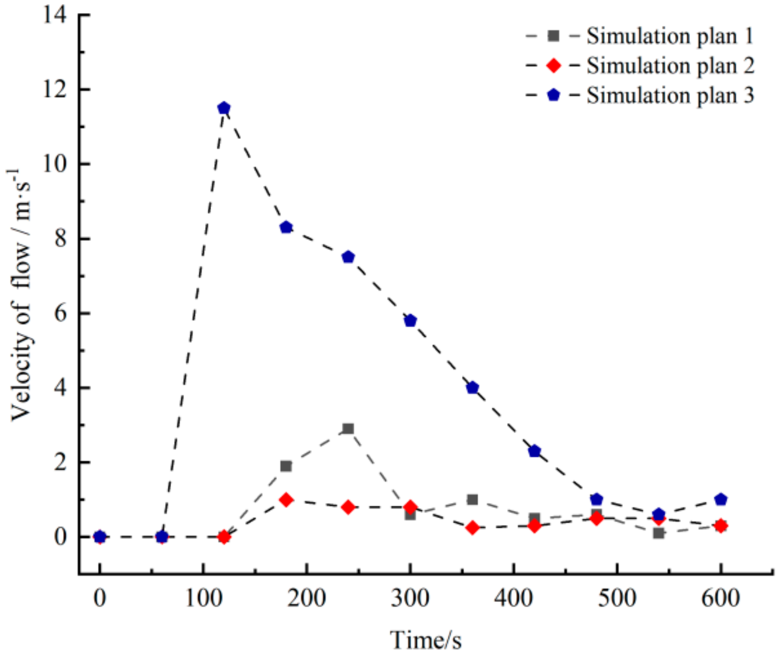

Figure 13. Flow velocity at monitoring point M4 for different dam-failure scenarios.

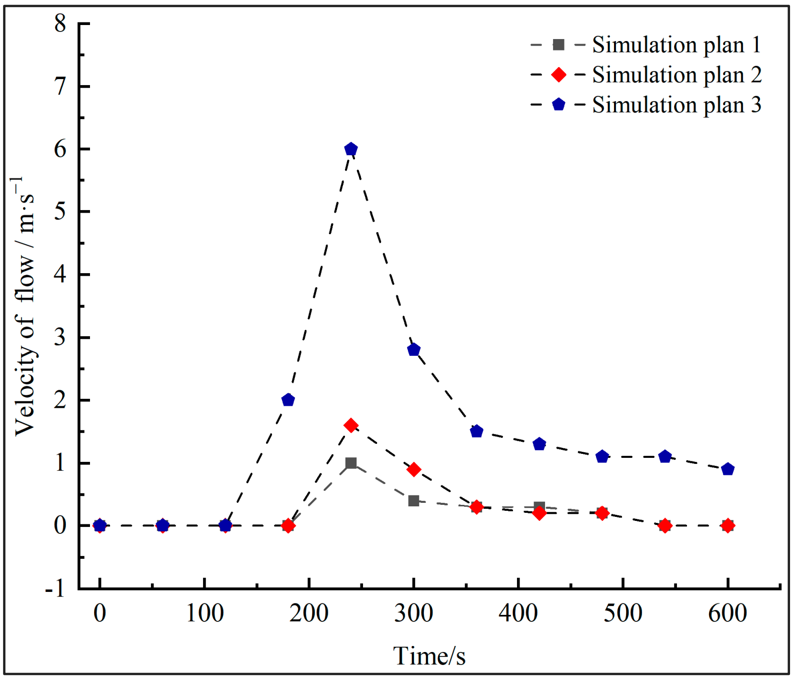

Figure 14. Flow velocity at monitoring point M5 for different dam-failure scenarios.

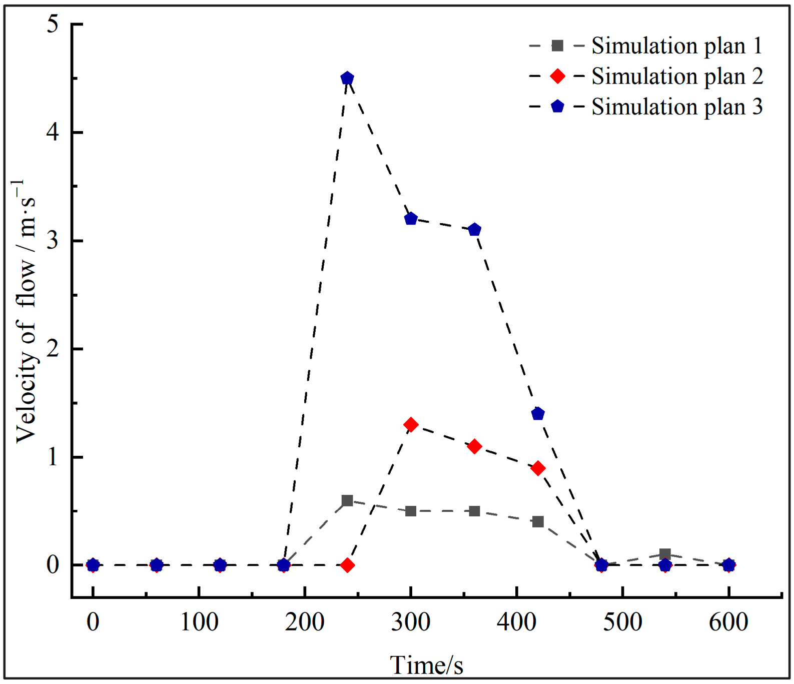

Figure 15. Flow velocity at monitoring point M6 for different dam-failure scenarios.

Analyses of the flow characteristics at different monitoring points over time revealed that the flow velocity at different monitoring points decreases with the increase in the flow path at the same moment. After the tailings dam failed, the discharged slurry reaches a peak velocity as it flowed through the monitoring point, and the time to reach this peak shortened as the tailings discharge volume and dam height increased. Throughout the dam-failure process, the influence range of the discharged slurry continuously expanded. The flow must overcome surface resistance contributing to a gradual decrease in velocity, and the velocity curve can be approximated as triangular. Once the kinetic energy carried by the slurry was exhausted, the flow ceased, and the influence range of the dam failure tended to stabilize.

5. Conclusions

(1) Combined with the collected engineering cases, it can be seen that by the joint influence of reservoir capacity, dam height, and other aspects, the amount of sand breakage of tailings pond dam failure increases with the increase in dam height and reservoir capacity, and the correlation between a single variable and the amount of sand breakage is relatively low, and through a large number of cases of dam failure to construct the training samples, the use of machine learning can solve the problem of the impact of dam-failure prediction of the small number of parameters and the long computation time.

(2) There is a nonlinear relationship between the impact of tailing pond dam breach and the existing conditions of tailing pond. The SVM algorithm cannot fully reflect the correlation of global samples when performing nonlinear regression due to its weak search ability, and the SMA algorithm gives the optimal penalty parameter C and the kernel function parameter g through global search, which can improve the algorithm accuracy.

(3) The SMA-SVM optimization algorithm is established to predict the tailing sand discharge volume after dam failure, and the results show that, compared with the RF algorithm, RBF algorithm, and LSSVM algorithm, the prediction results of the SMA-SVM optimization algorithm are closer to the real value, and the average absolute relative error of the sand discharge volume is about 30,000 m^3 The prediction results have a high degree of agreement with the actual value, and the errors with the results of numerical simulation are 48,000 m^3, 186,000 m^3, and 186,000 m^3, and 601,000 m^3, respectively, and the research results can be used as an effective analysis tool for the study of tailings pond dam failure.

(4) There are still deficiencies in the model, in addition to the main impact of the tailings pond dam failure by the reservoir capacity, dam height, rainfall, and dam construction method, as well as other factors. These deficiencies represent one of the influencing factors on the amount of sand breakage; however, this paper does not consider the impact of other factors, so the SMA-SVM optimization algorithm proposed in the paper does not have universality, and in future research, it is necessary to carry out a more comprehensive analysis and assessment of these factors in order to analyze the sand breakage of the tailings released after the tailings pond dam failure more accurately.

Author Contributions: Data curation, G.L., B.Z., Y.X. and M.W.; Funding acquisition, X.K.; Writing—original draft, G.L., B.Z. and Y.Z.; Writing—review and editing, G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Yonggang Zhang were employed by the company China Construction Eighth Engineering Division Corp., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflicts of interest.

References

1. Lin, Z.; Wang, H.; Yang, J. Numerical Simulation Study on Different Forms of Dam Failure in Tailings Ponds. *China Mine Eng.* **2023**, *52*, 12–16+21. [CrossRef]

2. Deonarine, A.; Hsu-Kim, H.; Ruhl, L.; Ruhl, L.; Vengosh, A. Riverine mercury contamination after the 2008 coal ash spill at the Kingston Fossil Plant, TN. In *Geochimica et Cosmochimica Acta*; Pergamon-Elsevier Science Ltd.: Oxford, UK, 2010; Volume 74, p. A223.
3. Kovács, T.; Sas, Z.; Jobbágy, V.; Csordás, A.; Szeiler, G.; Somlai, J. Radiological aspects of red mud disaster in Hungary. *Acta Geophys.* **2013**, *61*, 1026–1037. [CrossRef]
4. Quintarelli, J.M.; da Silva Junior, G.C.; Viglio, E.P. A Note on the Influence of the Mine Tailings Released in the Córrego do Feijão Mine Disaster on the Water Bodies of Brumadinho, Minas Gerais, Brazil. *Mine Water Environ.* **2023**, *42*, 187–199. [CrossRef]
5. Ma, Y. A Study on the Problems and Countermeasures of Emergency Management of Environmental Emergencies in Heilongjiang Province. Master's Thesis, Changchun University of Technology, Changchun, China, 2022. [CrossRef]
6. Wang, G. *Case Study of Tailings Pond Failure Accidents*; Metallurgical Industry Press: Beijing, China, 2022.
7. Chen, C. Research and Application of Risk Assessment and Risk Reduction for Tailings Dam Failure. Ph.D. Thesis, University of Science and Technology Beijing, Beijing, China, 2023.
8. Xu, Y.; Qiao, Y. Several problems in the calculation of dam break flow. *Yellow River* **1983**, *06*, 30–33.
9. Chen, Q.; Sun, J. Calculation and simulation of dam break sand flow in mine tailings reservoir. *J. Hohai Univ.* **1995**, *05*, 99–105.
10. Yin, G.; Jing, X.; Zhang, W.; Wei, Z.; Xiao, Q. Experiments of debris flow disaster of tailings dam-break. *Disaster Adv.* **2011**, *4*, 4–9.
11. Du, C.; Niu, B.; Yi, F.; Jiang, X.; Liang, L. Impact of overtopping range of overtop dam break of tailings pond under actual terrain conditions. *PLoS ONE* **2023**, *18*, e0295056. [CrossRef] [PubMed]
12. Du, C.; Tao, H.; Yi, F.; Yi, F.; Jiang, X.Q.; Liang, L.D. Experimental study on overtop** dam-break of a tailing reservoir under extreme conditions. *Environ. Sci. Pollut. Res.* **2023**, *18*, 1–17.
13. Lin, H.; Luo, J.; Zhou, C.; Tao, Z. Research on model material selection based on inception similarity in impact analysis of flood overtopping on tailings dam. *Environ. Earth Sci.* **2023**, *82*, 1–14. [CrossRef]
14. Luo, C.; Li, D.; Xu, B. Study on the Evolution of Tailings Dam Break Disaster under Complex Environment. *Sustainability* **2023**, *15*, 14728. [CrossRef]
15. Jing, X.; Chen, Y.; Xie, D.; Xie, D.; Williams, D.J.; Wu, S.W.; Wang, W.S.; Yin, T.W. The effect of grain size on the hydrodynamics of mudflow surge from a tailings dam-break. *Appl. Sci.* **2019**, *9*, 2474. [CrossRef]
16. Wu, S.; Yan, J.; Cai, H. Experimental study on dam break mode of tailing piping and dischargeimpact characteristics. *J. Geotech. Eng.* **2021**, *43*, 2134–2141.
17. Zhao, H.; Yao, L.; Wang, B.L.; Ali, S. Surge Waves Under Earthquake and Clastic Flow Landslide. *J. Earthq. Tsunami* **2018**, *12*, 1850009. [CrossRef]
18. Zhang, F.; Ti, Z.; Zhang, A. Simulation of impact stability test model in underwater coal mining of dam. *J. Coast. Res.* **2018**, *83*, 642–645. [CrossRef]
19. Jing, X.; Yin, G.; Wei, Z.; Huang, G. Model test and numerical simulation of tailing dam safety forecasting. *J. Chongqing Univ.* **2009**, *32*, 308–313.
20. Zhang, L.T.; Qi, Q.; Li, Q.; Zhang, S.X. 3D numerical simulation on the evolution of debris flow from dam break of tailing pond. *Int. J. Earth Sci. Eng.* **2014**, *7*, 111–115.
21. Chang, M.; Qin, W.; Wang, H.; Wang, H.B.; Wang, C.T.; Zhang, X.L. Study on the evolution of a flooded tailings pond and its post-failure effects. *Water* **2022**, *15*, 173. [CrossRef]
22. Shuai, X.; Liu, X.; Wan, J.; Wang, J.; Zhang, Y.; Zhou, Y. Numerical Analysis of Tailings Dam Break Based on GDEM-GAVA. *Geotech. Geol. Eng.* **2024**, *42*, 2833–2843. [CrossRef]
23. Zhuang, Y.; Jin, K.; Cheng, Q.; Xing, A.; Luo, H. Experimental and numerical investigations of a catastrophic tailings dam break in Daye, Hubei, China. *Bull. Eng. Geol. Environ.* **2022**, *81*, 1–16. [CrossRef]
24. Gao, T.; Sun, W.; Li, Z.; Li, Z.Y.; Fan, K.; Jiang, M.G.; Cheng, H.Y. Study on shear characteristics and failure mechanism of inclined layered backfill in mining solid waste utilization. *Minerals* **2022**, *12*, 1540. [CrossRef]
25. Huang, G.; Yang, Y.; Liu, M.; Liu, M.Y.; Zhang, J.H.; Liu, F.P.; Mwangi, A.D.; Ye, H.W. Numerical Study on Mechanical Properties of the Freezing–Thawing Cycle of Tailings Based on Particle Discrete Element Method. *Minerals* **2022**, *12*, 904. [CrossRef]
26. Wang, X.; Sun, J.; Zhou, S.; Zhou, J. Numerical simulation study on movement characteristics of tailings dam break. *J. Eng. Geol.* **2019**, *27*, 144–151.
27. Tian, Z.; Yang, H.; Wang, W.; Cao, D.S. Numerical analysis of sand bed degrading and sediment transport rate under tailings dam break. *Front. Earth Sci.* **2021**, *9*, 686277. [CrossRef]
28. Clarkson, L.; Williams, D. An overview of conventional tailings dam geotechnical failure mechanisms Mining. *Metall. Explor.* **2021**, *38*, 1305–1328.
29. Wen, T. Prediction Model for Open-pit Coal Mine Slope Stability Based on Random Forest. Prediction Model for Open-pit Coal Mine Slope Stability Based on Random Forest. *Sci. Technol. Rev.* **2014**, *32*, 105–109.

30. Liang, H.; Zhang, H. Identification of slope stability based on the contrast of BP neural network and SVM. In Proceedings of the 2010 3rd International Conference on Computer Science and Information Technology, Chengdu, China, 9–11 July 2010; Volume 9, pp. 347–350.
31. Jin, J.X.; Yuan, S.; Cui, H.; Xiao, X.C.; Jia, B.X. A threshold model of tailings sand liquefaction based on PSO-SVM. *Sustainability* **2022**, *14*, 2720. [CrossRef]
32. Aburomman, A.A.; Reaz, M.B.I. A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection systems. *Inf. Sci.* **2017**, *414*, 225–246. [CrossRef]
33. Deif, M.A.; Solyman, A.A.A.; Alsharif, M.H.; Jung, S.; Hwang, E. A hybrid multi-objective optimizer-based SVM model for enhancing numerical weather prediction: A study for the Seoul metropolitan area. *Sustainability* **2021**, *14*, 296. [CrossRef]
34. Monismith, D.R.; Mayfield, B.E. Slime Mold as a model for numerical optimization. In Proceedings of the 2008 IEEE Swarm Intelligence Symposium, St. Louis, MO, USA, 21–23 September 2008. [CrossRef]
35. Latty, T.; Beekman, M. Speed–accuracy trade-offs during foraging decisions in the acellular slime mould *Physarum polycephalum*. *Proc. R. Soc. B Biol. Sci.* **2011**, *278*, 539–545. [CrossRef] [PubMed]
36. Latty, T.; Beekman, M. Food quality affects search strategy in the acellular slime mould. *Physarum polycephalum. Behav. Ecol.* **2009**, *20*, 1160–1167. [CrossRef]
37. Ghannadi, P.; Kourehli, S.S. Efficiency of the slime mold algorithm for damage detection of large-scale structures. *Struct. Des. Tall Spec. Build.* **2022**, *31*, e1967. [CrossRef]
38. Xiao, Y.; Wang, H.; Xu, W. Parameter Selection of Gaussian Kernel for One-Class SVM. *IEEE Trans. Cybern.* **2015**, *45*, 927–939. [CrossRef] [PubMed]
39. Zhou, W.; Xiang, Y.; Chen, X.; Wu, J.; Zhang, J. Robust Network Traffic Classification. In *IEEE/ACM Transactions on Networking: A Joint Publication of the IEEE Communications Soceity, the IEEE Computer Society, and the ACM with Its Special Interest Group on Data Communication*; Institute of Electrical and Electronics Engineers, and the Association for Computing Machinery: New York, NY, USA, 2015; Volume 1.
40. Fofana, A.; Haas, O.; Ersanilli, V.; Burnham, K.; Mahtani, J.; Woolley, C.; Vithanage, K. Multi-objective genetic algorithm for an automatic transmission gear shift map. *IFAC-Pap.* **2016**, *49*, 123–128. [CrossRef]
41. Jin, C.; Ji, Z.; Gong, L.; An, X.; Zhou, Y. Evaluation model of rockburst intensity of diversion tunnel based on WOA-SVM. *China Saf. Sci. J.* **2023**, *33*, 41–48. [CrossRef]
42. Wei, Y. Research on the Application of Heuristicslime Mould Algorithm. Master's thesis, Guangxi University for Nationalities, Nanning, China, 2022. [CrossRef]
43. Rico, M.; Benito, G.; Diez-Herrero, A. Floods from tailings dam failures. *J. Hazard. Mater.* **2008**, *154*, 79–87. [CrossRef] [PubMed]
44. Piciullo, L.; Storrøsten, E.B.; Liu, Z.; Nadim, F.; Lacasse, S. A new look at the statistics of tailings dam failures. *Eng. Geol.* **2022**, *303*, 106657. [CrossRef]
45. Concha Larrauri, P.; Lall, U. Tailings dams failures: Updated statistical model for discharge volume and runout. *Environments* **2018**, *5*, 28. [CrossRef]
46. Gao, W.; Ibaraki, S.; Donmez, M.A.; Kono, D.; Mayer, J.R.R.; Chen, Y.L.; Szipka, K.; Archenti, A.; Linares, J.M.; Suzuki, N. Machine tool calibration: Measurement, modeling, and compensation of machine tool errors. *Int. J. Mach. Tools Manuf.* **2023**, *187*, 104017. [CrossRef]
47. Sun, H.; Sun, J.; Zhao, K.; Wang, L.C.; Wang, K. Data-driven ICA-Bi-LSTM-combined lithium battery SOH estimation. *Math. Probl. Eng.* **2022**, *2022*, 9645892. [CrossRef]
48. Yang, Y.; Zhou, X. Numerical Simulation of Tailings Mortar Flowing from Dam Break by Variable Concentration Two-phase Flow Algorithm. *IOP Conf. Ser. Earth Environ. Sci.* **2024**, *1331*, 012027. [CrossRef]