smart cities

MDPI

Article

Machine-Learning-Based Prediction of Gushing-Induced Ground Disturbance Around Shield Tunnels

Xiao-Chuang Xie [1], Zhao-Geng Chen [1] and Yu-Xin Zhang [2,*]

[1] Department of Geotechnical Engineering, College of Civil Engineering, Tongji University, Shanghai 200092, China
[2] Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Hong Kong 999077, China
* Correspondence: yx.zhang@polyu.edu.hk

Highlights

What are the main findings?

- A validated two-phase Material Point Method (MPM) database of 39,810 samples is established to capture gushing-induced ground disturbance around shield tunnels under diverse hydro-mechanical conditions.
- Among five machine-learning algorithms (MLP, RF, XGBoost, SVR, Ridge), nonlinear models, particularly MLP and RF, consistently achieve the highest predictive accuracy and the most favourable error distributions across all three disturbance descriptors (maximum ground settlement, flow-zone width, and flow-zone centroid angle).
- SHAP-based interpretation reveals that soil gushing mass (SGM) dominates settlement and flow-zone width, while gushing location primarily governs flow-zone geometry, confirming that the learned relationships are physically meaningful.

What are the implications of the main findings?

- The proposed ML surrogate framework replaces computationally expensive numerical simulations with rapid, interpretable predictions, enabling real-time or near-real-time hazard assessment of tunnel gushing incidents, a critical capability for smart city underground infrastructure management.
- The physically interpretable SHAP analysis supports data-driven decision-making for preventive maintenance and risk-informed evaluation of shield tunnels, directly contributing to the resilience and safety of smart urban underground systems.
- Shield tunnels form a core component of urban underground transportation networks in modern smart cities. The integration of AI-based surrogate modelling with physics-based simulation addresses a key challenge in smart city infrastructure: achieving rapid, data-driven situational awareness and risk assessment without sacrificing physical interpretability. The framework proposed in this study can be embedded into smart monitoring and early-warning systems for underground infrastructure, thereby supporting sustainable and resilient urban development.

Check for updates

Academic Editors: Pierluigi Siano, Rachid Belaroussi, Nuno Datia and Javier J. Sanchez-Medina

Received: 16 April 2026
Revised: 2 June 2026
Accepted: 11 June 2026
Published: 13 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Abstract

Water-soil gushing caused by tunnel leakage can induce severe ground disturbance and threaten the safety of shield tunnels, yet rapid prediction remains difficult because high-fidelity numerical simulations are computationally expensive. This study develops an interpretable machine-learning framework for predicting gushing-induced ground disturbance around shield tunnels based on a validated two-phase Material Point Method database. Six governing variables are considered, including the tunnel depth ratio, gushing

location, soil friction angle, Young's modulus, intrinsic permeability, and soil gushing mass. Three representative response variables were selected, namely the maximum ground settlement, flow-zone width, and flow-zone centroid angle. Five algorithms, including MLP, RF, XGBoost, SVR, and Ridge, were established and compared, with hyperparameters optimised using Optuna. The results show that nonlinear models consistently outperform the linear baseline, among which MLP, RF, and XGBoost achieve the best overall accuracy and robustness. Error-distribution analysis further indicates that MLP and RF yield the highest proportion of low-error predictions. SHAP interpretation shows that SGM is the dominant factor governing maximum settlement and flow-zone width, whereas gushing location primarily controls the flow-zone centroid angle. The proposed framework provides an efficient and physically interpretable surrogate for rapid hazard assessment of gushing-induced ground disturbance in shield tunnelling.

Keywords: shield tunnel; water-soil gushing; ground disturbance; material point method; machine learning

1. Introduction

Shield tunnels are widely used in urban underground transportation systems because of their high construction efficiency and limited disturbance to surface activities. However, as groundwater fluctuations intensify under the combined effects of extreme weather and construction activities, shield tunnels are increasingly exposed to leakage-related hazards because their prefabricated lining system contains numerous segmental joints. Large transverse convergence may lead to joint opening or segment misalignment [1,2]. When such deformation exceeds the allowable limit, groundwater may transport surrounding soil into the tunnel, thereby triggering water–soil gushing accidents [3–5]. The associated soil loss can induce severe ground disturbance and even large-scale collapse, thereby threatening tunnel serviceability and overburden strata (Figure 1). For instance, in 2016, a water-leakage accident occurred during construction of the Fukuoka Subway Nanakuma Line in Japan. Continuous inflow of sandy soil into the tunnel triggered a sudden large-scale ground collapse, forming a sinkhole of about 30 × 27 × 15 m; in 2024, a similar water-and-sand gushing incident occurred between Kaiyuanmen Station and Tumen Station on Xi'an Metro Line 8, causing extensive surface collapse; in 2025, a metro line in Bangkok also experienced a water–sand inrush, which abruptly generated a large sinkhole on Samsen Road, and the void was estimated to extend nearly 30 m across and up to 50 m deep. As illustrated in Figure 1b, water–soil inflow may first disturb the surrounding ground and form internal voids above the tunnel crown; continued soil gushing then aggravates upward propagation of the disturbed zone, eventually resulting in progressive soil collapse and sudden surface settlement. These typical accident manifestations indicate that water–soil gushing is not merely a local seepage defect, but a coupled stratum–structure hazard with potentially catastrophic consequences. Efficient prediction of the gushing-induced ground disturbance is therefore of clear practical significance for hazard assessment and preventive maintenance [6].

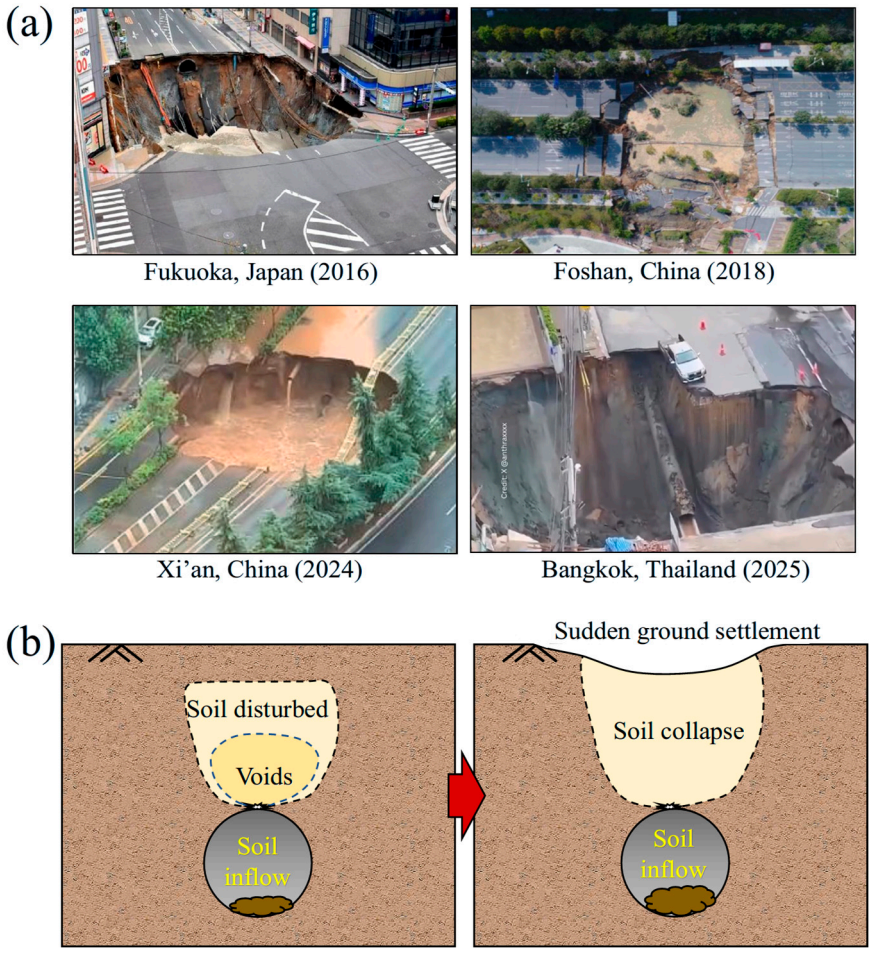

Figure 1. Water–soil gushing hazards in shield tunnels: (**a**) representative accident cases; (**b**) illustration of progressive soil collapse.

Considerable effort has been devoted to understanding leakage- and gushing-induced disturbance around tunnels through field investigation, laboratory testing, and numerical simulation. Previous studies have shown that the evolution of water–soil gushing is governed by strongly coupled hydro-mechanical processes, including seepage-driven particle migration, progressive formation of flow channels, stiffness degradation of the surrounding ground, and stress redistribution between the stratum and tunnel lining [7–9]. Physical model tests and case studies have provided valuable insight into the development of cavities, disturbed zones, and settlement troughs, and have clarified the influence of burial depth, leakage position, soil type, and hydraulic conditions. Karoui et al. (2018) [10] investigated the settlement behaviour of deep soils induced by tunnel sand leakage under both constant and varying hydraulic heads. Liang et al. (2025) [11] studied the collapse behaviour of composite strata with upper silt and lower sand under tunnel gushing by means of physical model tests. Nevertheless, laboratory tests are constrained by scale effects, simplified boundary conditions, and limited parameter combinations. Numerical simulation has therefore become a principal tool for investigating the mechanism and consequences of gushing hazards. Zheng et al. (2024) [12] examined the deformation and failure of shield tunnels induced by contact loss using the Coupled Eulerian–Lagrangian method. Zhang et al. (2026) [13] established a closed-loop governing framework for the transition from suffusion to leakage in seepage erosion based on CFD-DEM simulations and particle-scale upscaling. Other large-deformation methods, such as smoothed particle hydrodynamics (SPH) and coupled DEM–SPH models [14] and large-deformation finite-element (LDFE) analysis [15], have likewise been applied to landslide surge waves

and tunnelling-induced sinkholes. In this study, the Material Point Method (MPM) is preferred because its material points retain history-dependent state variables (effective stress, pore pressure, and porosity) while the background grid avoids severe mesh distortion and naturally enforces the soil–structure contact. It has been widely applied to a range of geotechnical problems, including landslides [16], tunnelling [17], and soil–structure interaction [18]. These MPM-based studies have demonstrated its capability in reproducing the staged development of water–soil gushing and capturing key response characteristics such as ground settlement and subsurface ground movement.

Despite these advances, numerical simulation is still computationally demanding, especially when a broad influence parameter space must be explored. In practical engineering, decision-makers often require rapid estimates of key response indices under varying conditions, rather than a small number of detailed forward simulations. Repeated simulations for these purposes are costly and are therefore not well suited to emergency evaluation or large-scale parametric assessment. This limitation motivates the introduction of machine-learning surrogate models [19,20]. For instance, Zhang et al. (2025) [20] applied machine-learning models to predict the maximum seismic response of pile-supported structures in liquefiable soils and showed that XGBoost achieved the best overall predictive performance. Compared with empirical correlations and simplified analytical approaches, they are more capable of capturing strong nonlinearity and multi-parameter interaction, both of which are intrinsic to gushing-induced ground disturbance. In recent years, such data-driven approaches have shown considerable promise in underground engineering applications, including deformation prediction [21], settlement assessment [19], and parameter inversion [22].

However, several gaps still remain in the prediction of gushing-induced ground disturbance around shield tunnels. First, most existing studies have focused on numerical simulation of leakage-induced deformation [23], while an integrated framework linking gushing conditions to both surface and subsurface disturbance characteristics is still lacking. Second, the hydro-mechanical response induced by water-soil gushing is strongly nonlinear and governed by coupled effects of gushing location, soil properties, and so on, yet the applicability and comparative performance of different machine-learning models for this problem have not been systematically clarified. Third, although interpretability is essential for engineering adoption, limited attention has been paid to identifying the governing factors behind different disturbance descriptors and verifying whether the learned relationships are physically meaningful rather than purely statistical.

To address these issues, this study develops an interpretable machine-learning surrogate framework for predicting gushing-induced ground disturbance around shield tunnels, based on a validated two-phase MPM database. The framework links six governing variables to the key surface and subsurface descriptors of disturbance, and combines multi-model comparison with SHAP-based interpretation to achieve both predictive accuracy and physical transparency. The main novelties of this study are summarized as follows. First, in contrast to previous numerical studies on water–soil gushing, which mostly rely on a limited number of computationally expensive forward simulations, a validated two-phase MPM model is used to construct a systematic physics-based database that links gushing conditions to both surface and subsurface disturbance characteristics. Second, soil gushing mass (SGM) is introduced as a physically meaningful intensity variable to quantify the development stage of gushing; compared with empirical correlations that rely mainly on elapsed time or a single volume-loss measure, this enables the strongly nonlinear and coupled hydro-mechanical response to be represented more directly. Third, unlike most existing machine-learning studies on tunnel-induced ground response, which predict surface settlement alone using a single algorithm, five machine-learning models are systematically

compared for three complementary descriptors, including maximum ground settlement, flow-zone width, and flow-zone centroid angle, so that the intensity, extent, and geometric pattern of disturbance are predicted simultaneously. Fourth, SHAP-based interpretation is employed to quantify the contribution of each governing variable and to verify that the learned input–output relationships are physically meaningful rather than purely statistical.

2. MPM Modeling of the Water–Soil Gushing

Although water–soil gushing events in shield tunnels have been observed in practice, acquiring systematic field data for data-driven modelling is still difficult. Existing case records are commonly incomplete with important hydro-mechanical parameters. In addition, the limited number of well-documented cases hinders the direct construction of reliable prediction models. Therefore, a synthetic database is generated in this study using a developed two-phase coupled MPM framework.

2.1. Description of Two-Phase MPM

Various MPM formulations have been developed for fully coupled dynamic analysis of two-phase porous media [24]. In general, these formulations can be classified into two categories according to the representation of the saturated medium: (1) a single-point MPM formulation, in which one set of material points is used to describe the coupled solid-fluid mixture, as illustrated in Figure 2a; and (2) a two-point MPM formulation, in which separate sets of material points are assigned to the solid and fluid phases. Although the latter can provide a clearer phase-wise description, it usually involves higher computational cost and requires additional treatment of the solid–fluid interaction. For this reason, the single-point MPM formulation has been more widely adopted in coupled dynamic analyses [25] and is therefore employed in the present study.

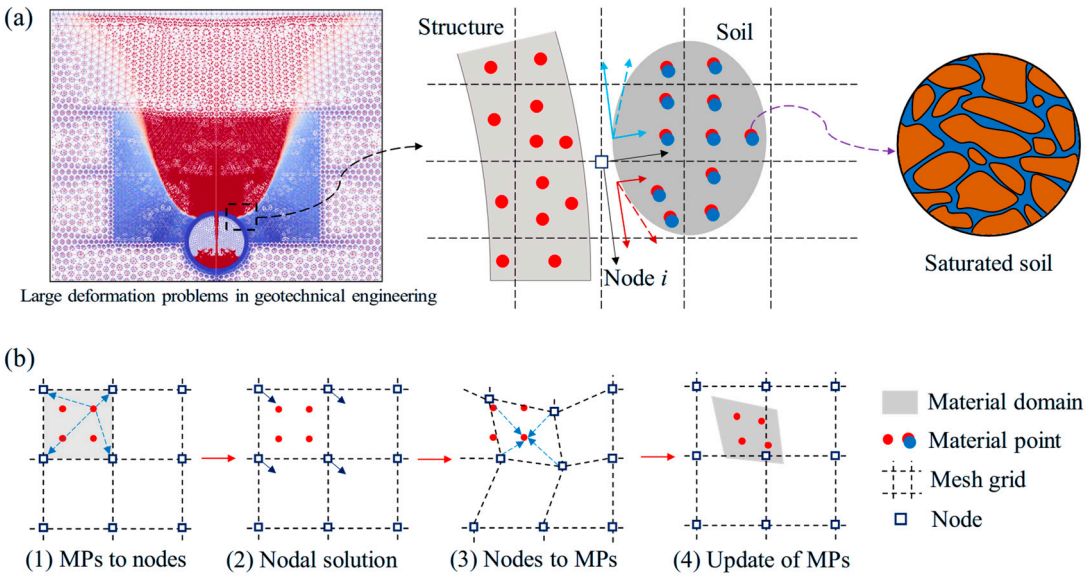

Figure 2. Illustration of two-phase MPM for simulating large deformation problems: (**a**) coupled MPM implementation with a contact algorithm for saturated soil–structure interfaces; (**b**) computational cycle of the MPM.

Similar to FEM-based poromechanical formulations, fully coupled two-phase MPM formulations can be classified by their primary unknowns, namely the solid velocity v, the fluid velocity w, and the pore pressure p. When the relative fluid acceleration is neglected, the v-p formulation [26] is straightforward to implement but may be inadequate for rapid transient responses. In contrast, the v-w formulation [27] retains soil–water acceleration

contributions and is therefore better suited to a broad range of dynamic problems. The hydro-mechanical problem is formulated in terms of effective stress and Darcy flow. The discretized equations are derived from the momentum balance equation of the water phase and the soil-water mixture. The mixture momentum balance reads:

$$(1-n)\rho_s\dot{\boldsymbol{v}} + n\rho_w\dot{\boldsymbol{w}} = \nabla\boldsymbol{\sigma} + \bar{\rho}\boldsymbol{g} \tag{1}$$

where $\boldsymbol{\sigma}$ is the effective stress tensor of the soil skeleton; n represents porosity; ρ_s and ρ_w are the densities of soil and water, respectively; $\bar{\rho}$ is the bulk density of the mixture, defined as $\bar{\rho} = (1-n)\rho_s + n\rho_w$; $\boldsymbol{g}$ denotes the gravitational acceleration vector.

The liquid-phase momentum balance is described by Equation (2):

$$\rho_w\dot{\boldsymbol{w}} = \nabla p_w + \rho_w\boldsymbol{g} - \frac{n\gamma_w}{k_w}(\boldsymbol{w}-\boldsymbol{v}) \tag{2}$$

where p_w denotes the pore pressure; γ_w represents the unit weight of the water; k_w is the Darcy permeability coefficient. Additional information is available in previous studies [25,27].

The MPM framework can readily enforce no-slip contact between different bodies, since the velocities of all bodies are defined on a shared background grid. To model the interaction between tunnel structures and fully saturated soil, a contact algorithm based on nodal velocities is adopted in this study. This algorithm was originally proposed by Bardenhagen et al. (2001) [28] for single-phase materials and was subsequently extended to two-phase problems by Ceccato et al. (2016) [25], where full implementation details can be found.

Figure 2b illustrates the standard MPM procedure within one time increment. The soil domain is represented by a set of Lagrangian material points (MPs), which store mass, momentum, stress, pore pressure, porosity, and other internal variables, while a background Eulerian grid is introduced at each step as the interpolation domain. Within each increment, the computation consists of four main steps: (1) projection of MP information to the grid nodes; (2) solution of the governing equations on the grid to obtain nodal kinematic variables; (3) interpolation of nodal results back to the MPs to update displacements, velocities, strains, effective stresses and pore pressures; and (4) updating of MPs' positions followed by resetting of the background grid to its initial configuration. In this way, mesh distortion is avoided while the history-dependent state variables remain carried by the MPs. Based on this framework, an explicit single-point MPM with a $\boldsymbol{v}$-$\boldsymbol{w}$ formulation is adopted to simulate hydrodynamic gushing-induced ground collapse in this study. The formulation is implemented in the Anura3D MPM software [29].

2.2. Validation of MPM Approach

To provide confidence in the use of MPM results as a basis for data-driven analysis, the two-phase MPM framework is first verified against field observations from a Shanghai case history [5]. Figure 3a shows that the tunnels are buried under 23.1 m of overburden, and the tunnel outer diameter is 6.6 m. The alignment is located within a highly permeable sandy silt stratum subjected to elevated pore-water pressure. During construction, accidental leakage from the freezing pipes triggered rapid soil inflow through the gushing channel, which led to pronounced ground settlement.

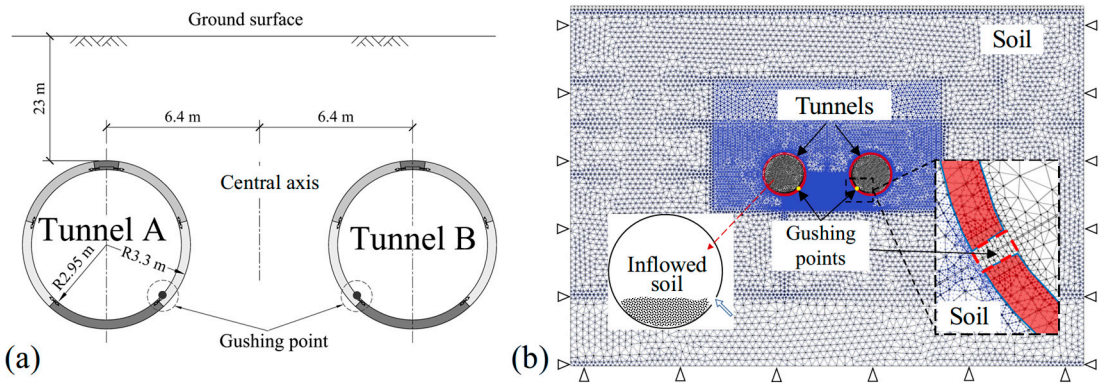

Figure 3. Geometry of case study: (**a**) cross-section, and (**b**) simulation model.

As shown in Figure 3b, a simulation model is constructed to reproduce the gushing process. It should be noted that, for computational efficiency, the tunnel gushing problem is analysed under plane-strain conditions. A drained assumption is adopted during the gushing process because excess pore-water pressure is expected to dissipate rapidly in the sandy ground. The geotechnical parameters obtained from the site investigation are listed in Table 1. Water–soil gushing is assumed to start from an active leakage channel. Based on post-incident inspection of the damaged lining rings, which identified a failed freezing borehole about 0.2 m wide at the tunnel invert, the channel width is set to 0.2 m. It should be noted that this prescribed channel represents the net hydraulic effect of leakage rather than the failure process of the lining waterproof system itself. In practice, the segmental joints, gaskets, and grouting layer constitute a complex waterproof system whose failure mode governs the initiation position, opening size, and rate of water–soil gushing [30]; explicitly incorporating variables that characterise the waterproof-system state is left for future work. To maintain an adequate geometric description and stable traction transmission at the soil-lining interface in the vicinity of the gushing channel, local mesh refinement is introduced. In this region, the mesh size is reduced to 0.05 m across the channel width, such that the 0.2 m-wide channel is resolved by four elements, with each element containing 12 material points. Based on preliminary mesh sensitivity analysis and local refinement tests, the adopted discretisation provides a satisfactory balance between accuracy and computational efficiency. As boundary conditions, horizontal displacements are restrained at the left and right boundaries, while both horizontal and vertical displacements are fixed at the bottom boundary.

Table 1. Soil parameters for the case study simulation.

Parameter	Value	Unit
Elastic modulus, E	40,000	kPa
Poisson ratio, υ	0.3	–
Cohesion, c	0	kPa
Friction angle, ϕ	25	°
Density of water, ρ_w	1000	kg/m^3
Liquid viscosity, m	1.002×10^{-6}	kPa·s
Liquid bulk modulus, K_w	80,000	kPa
Porosity, n	0.45	–

Figure 4 illustrates the ground-surface settlement induced by the gushing event. During the incident, a large volume of soil is carried into the tunnel, and the cumulative inflow mass can therefore be used as a direct indicator of gushing intensity. To validate the numerical model against field observations, settlement curves corresponding to identical

inflowed soil masses are plotted: simulated results appear as continuous lines, whereas the field measurements are plotted as filled markers. The settlement profiles in Figure 4 correspond to a cumulative inflowed soil mass of about 4591 kg/m, which is equivalent to a gushed-soil volume of approximately 1.53 m^3 per meter of tunnel length. Because the simulated and field settlements are compared at the same inflowed soil mass, the calculated gushing amount is consistent with the field-reported value [5]. The close agreement between the two sets of results confirms that the model can accurately reproduce the large soil deformation induced by the gushing event. Therefore, the two-phase MPM approach adopted in this study can effectively reproduce the large soil deformation induced by engineering-scale gushing incidents, and the validated methodology can be readily applied in the subsequent database construction for machine-learning modelling.

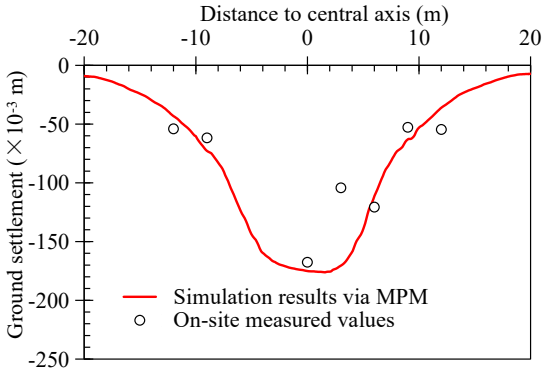

Figure 4. Comparison of simulated and field-measured ground settlement at the same cumulative inflowed soil mass.

Figure 5 indicates that the cumulative soil mass entering the twin tunnels grows in an almost linear manner with time. In the numerical simulations, this quantity is obtained by integrating the mass of soil that has migrated into the tunnel cavities, and is reported in kg/m, i.e., the inflowed soil mass per unit longitudinal tunnel length. Nevertheless, the gushing time alone is not an appropriate basis for cross-case comparison. Under different combinations of influence factors, gushing may progress at substantially different rates, meaning that the same physical time can correspond to very different stages of instability. In contrast, the cumulative gushed soil mass more directly reflects how far the accident has developed because it captures both the intensity of soil loss and its accumulated effect on the surrounding ground. For this reason, the soil gushing mass (*SGM*) is adopted in the following discussion as the key variable for tracking gushing severity and evolution stage.

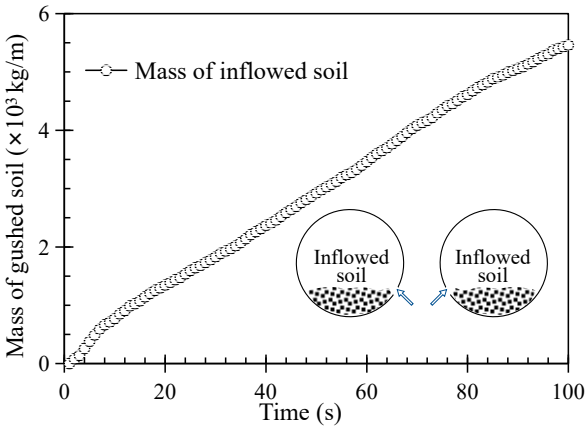

Figure 5. Development of the soil gushing mass.

As the gushing develops, pronounced soil disturbance is induced around shield tunnels, as illustrated in Figure 6. Owing to symmetry, only one half of the model domain is shown, and the horizontal coordinate is measured from the central axis (marked in Figure 3a). Figure 6a shows that a broad zone around the tunnels undergoes noticeable deformation. At the same time, a flow zone close to the gushing point continues to evolve and propagate upward. The displacement within the flow zone exceeds 0.05 m, which also corresponds to the threshold for first-level engineering impact on surrounding ground in China [31]. As further shown in Figure 6b, the shear-slip boundary develops in close agreement with the outer contour of the flow zone. This correspondence suggests that the geometry of the flow zone can serve as an effective indicator of the internal development of the gushing process. Taken together, these results show that gushing-induced disturbance is manifested not only by the surface response, but also by the geometric evolution of the subsurface flow zone.

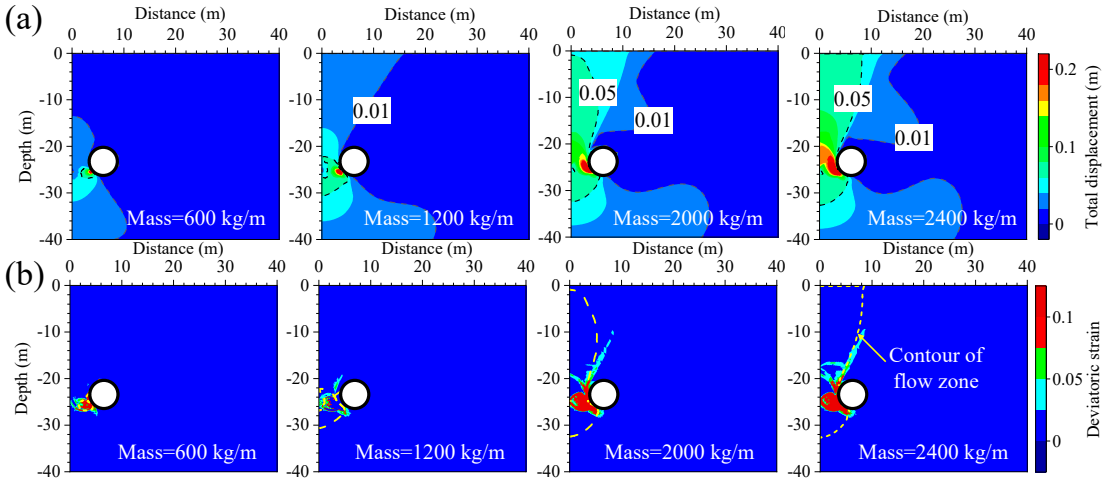

Figure 6. Soil disturbance due to gushing: (**a**) soil deformation; (**b**) deviatoric strain.

3. Machine Learning

Given the high computational cost of the numerical model for large-scale parametric analysis and rapid prediction, machine learning is adopted to establish efficient surrogate models for gushing-induced ground disturbance.

3.1. Candidate Algorithms

3.1.1. Multi-Layer Perceptron

A multi-layer perceptron (MLP), as a class of artificial neural networks, comprises an input layer, one or more hidden layers, and an output layer, as illustrated in Figure 7a. By iteratively updating the network parameters during training, the MLP establishes a nonlinear function f: $\mathbb{R}^m \rightarrow \mathbb{R}^o$, in which m and o are the input and output dimensions, respectively. For a single-hidden-layer MLP, the output of the j-th hidden neuron, h_j, is expressed as:

$$h_j = \sigma\left(\sum_{i=1}^{m} \omega_{j,i} x_i + \theta_j\right) \quad (3)$$

where x_i is the i-th input variable; $s(\cdot)$ represents the activation function in the hidden layer; $w_{j,i}$ and θ_j denote the connection weight and bias associated with the j-th hidden neuron, respectively. The predicted value of the k-th output neuron, y_k, is given by:

$$y_k = \phi\left(\sum_{j=1}^{N_h} \omega_{k,j} h_j + \theta_k\right) \quad (4)$$

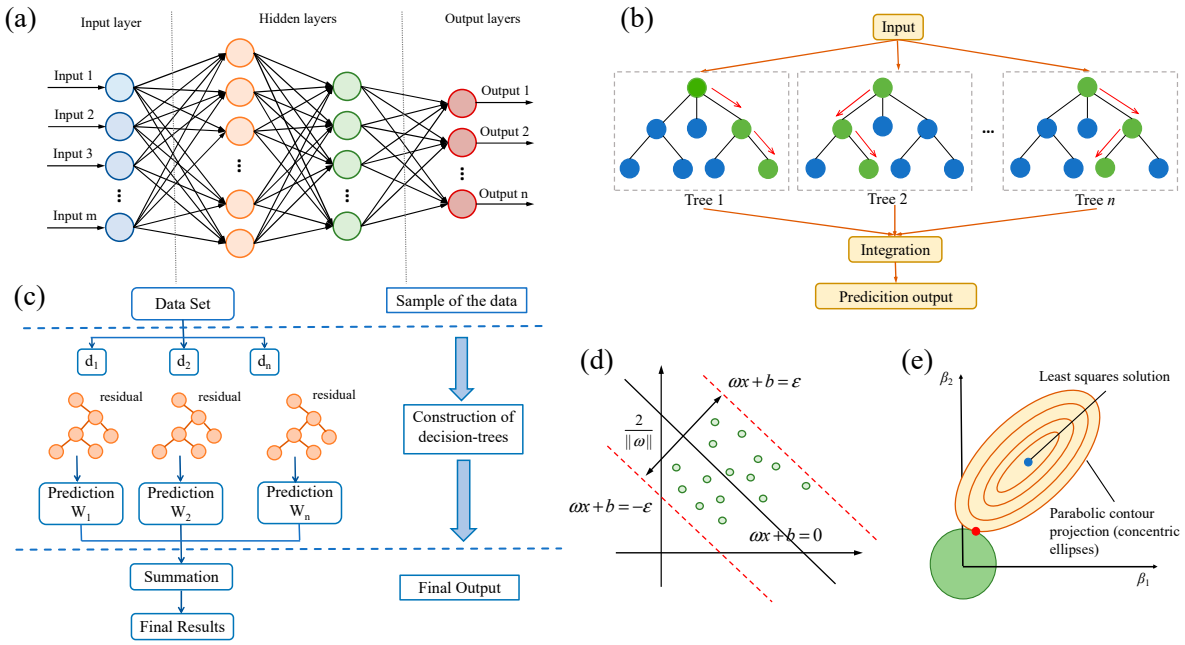

Figure 7. Schematic diagrams of candidate algorithms: (**a**) MLP; (**b**) RF; (**c**) XGBoost; (**d**) SVR; (**e**) Ridge.

Here, N_h is the number of hidden neurons; $w_{k,j}$ denotes the weight from hidden neuron j to output neuron k; θ_k is the bias of the output layer; and $f(\cdot)$ represents the activation function at the output layer. Common activation functions include Identity ($f(x) = x$), Tanh ($f(x) = \tanh(x)$), Logistic ($f(x) = 1/(1 + e^{-x})$), and ReLU ($f(x) = \max\{0, x\}$). Here, ReLU was employed. The model performance was further enhanced by tuning the network architecture, including the number of hidden layers, the number of neurons per layer, and the associated weight and bias parameters.

3.1.2. Random Forest

Random forest (RF) is a supervised ensemble-learning method that can be applied to both classification and regression tasks. It builds a collection of decision trees [32] based on bootstrap samples and random subsets of input features, which helps suppress overfitting and improve prediction reliability. As illustrated in Figure 7b, the final output is produced by combining the predictions of individual trees, usually by majority voting or mean averaging:

$$y = \frac{1}{N}\sum_{i=1}^{N} f_i(\mathbf{X}) \tag{5}$$

where y represents the prediction output; $\mathbf{X} = [x_1, x_2, \ldots, x_m]^T$; and N denotes the number of decision trees in the forest.

3.1.3. XGBoost

XGBoost, developed by Chen and Guestrin (2016) [33], is an efficient tree-based ensemble learning algorithm that has been widely applied in regression and classification tasks. Owing to its high predictive accuracy, computational efficiency, and built-in regularisation, it is particularly suitable for engineering prediction problems. Methodologically, XGBoost is an implementation of gradient boosting decision trees (GBDT) [34], in which decision trees are added sequentially and each new tree is trained to fit the residuals, i.e., the negative gradients of the loss function. In this study, XGBoost was adopted instead of a conventional GBDT framework because it provides a more efficient and robust solution for large-scale

modelling and repeated hyperparameter optimisation. A key advantage of XGBoost is that it incorporates regularisation into the additive tree model to control model complexity and reduce overfitting. The regularised objective function is expressed as:

$$\mathcal{L}(\phi) = \sum_i l(y_i, \hat{y}_i^{(t-1)}) + \sum_k \Omega(f_k) \tag{6}$$

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{7}$$

where $l(\cdot)$ denotes the training loss function; W$(\cdot)$ is the regularisation term; T represents the number of leaves in a tree; and w_j is the score of leaf j. The regularisation strength is controlled by the parameters g and l.

At the t-th boosting iteration, the objective function is formulated as:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{8}$$

By applying a second-order Taylor expansion of the loss around $\hat{y}_i^{(t-1)}$, the objective can be rewritten in a leaf-wise summation form:

$$\left.\begin{aligned} \widetilde{\mathcal{L}}^{(t)} &= \sum_{j=1}^{T}\left[(\sum_{i\in I_j} g_i)\omega_j + \frac{1}{2}(\sum_{i\in I_j} h_i + \lambda)\omega_j^2\right] + \gamma T \\ g_i &= \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \\ h_i &= \partial^2_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \\ I_j &= \{i | q(X_i) = j\} \end{aligned}\right\} \tag{9}$$

where I_j denotes the set of samples assigned to leaf j, and $q(\cdot)$ is the leaf index function. In addition, XGBoost adopts an efficient greedy strategy to search for tree splits, which, together with the regularized formulation above, contributes to its strong generalization ability and high training efficiency in practice.

3.1.4. Support Vector Regression

SVR is adopted in this study as a kernel-based learning approach for regression prediction [35]. Built on the structural risk minimization principle, SVR seeks a regression function:

$$f(x) = \sum_{i=1}^{m} \omega_i K_i(x) + b \tag{10}$$

which achieves a trade-off between fitting accuracy and model complexity by introducing an e-insensitive tolerance in the optimization:

$$\left.\begin{aligned} &\min(\|\boldsymbol{\omega}\|^2/2) \\ &\text{s.t.}\left|\sum_{i=1}^{m} \omega_i K_i(x) + b - y_i\right| \le \varepsilon \end{aligned}\right\} \tag{11}$$

where $\boldsymbol{w}$ and b denote the weight vector and bias term, respectively; m is the number of samples; $K(\cdot)$ is the selected kernel function, w_i is the coefficient of the ith sample; y_i is the corresponding target value; and e represents the tolerance parameter. SVR is chosen here because it can capture the nonlinear coupling between the input factors and the gushing-induced ground responses while maintaining stable generalization through regularization. By introducing kernel functions in place of inner products, the input variables are implicitly

projected into a higher-dimensional feature space, where linear regression can be carried out (Figure 7d). In this study, the radial basis function (RBF) kernel is adopted:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (12)$$

where g is controlling the kernel width and thus governs the flexibility of the SVR model.

3.2. Ridge Regression Baseline

In addition to the nonlinear ML models in Section 3.1, a linear regression baseline based on Ridge regression [36] is introduced for comparison. Ridge regression augments ordinary least squares with an L2 regularisation term, which penalises large coefficients and thus improves numerical stability and generalisation, particularly when input variables exhibit multicollinearity. Given its linear functional form, Ridge provides an interpretable and computationally efficient benchmark; however, its capacity to capture complex nonlinear input–output relationships is inherently limited. Therefore, the performance gap between Ridge and the proposed ML models can be used to quantify the added value of nonlinear learning for predicting gushing-induced ground disturbance.

3.3. Hyper-Parameter Tuning

The Appropriate hyper-parameters are essential for achieving reliable predictive performance of classical machine-learning models. In this study, a unified hyper-parameter optimisation framework based on Optuna [37] is adopted for all surrogate models, as shown in Figure 8. Optuna employs the Tree-structured Parzen Estimator (TPE) sampler, a Bayesian optimisation method that updates a probabilistic surrogate of the objective using accumulated trial results and proposes new hyper-parameter configurations with higher expected improvement. This strategy is well suited to mixed search spaces that commonly arise in classical machine-learning models. The optimisation objective is defined as the validation mean squared error (MSE), and the optimal hyper-parameter set is subsequently used for model retraining and final testing. The universal workflow of Optuna used in this study can be summarised in five steps:

(1) Step 1: Define the objective function (validation MSE), the hyper-parameter search space, and the maximum number of trials N_t. The dataset is first divided at the group level into a training set and an independent testing set, and a fixed random seed is specified for reproducibility.

(2) Step 2: For each trial, Optuna suggests a candidate hyper-parameter set l_i, trains the corresponding model on the training subset, and evaluates its performance on the validation subset.

(3) Step 3: Compute the trial loss L_i and update the current best objective value as its fitness T_L:

$$T_L = \min(L_1, L_2, \ldots, L_i) \quad (13)$$

(4) Step 4: Repeat Steps (2)–(3) until N_t is reached. For models equipped with built-in early stopping (e.g., XGB boosting rounds or MLP training), early stopping is activated during training to reduce overfitting and unnecessary computations.

(5) Step 5: Select λ^* corresponding to the minimum validation loss and retrain the model using the optimal hyper-parameters, followed by final evaluation on the independent test set. The trial history is recorded for further analysis (e.g., optimisation curves and best-trial statistics).

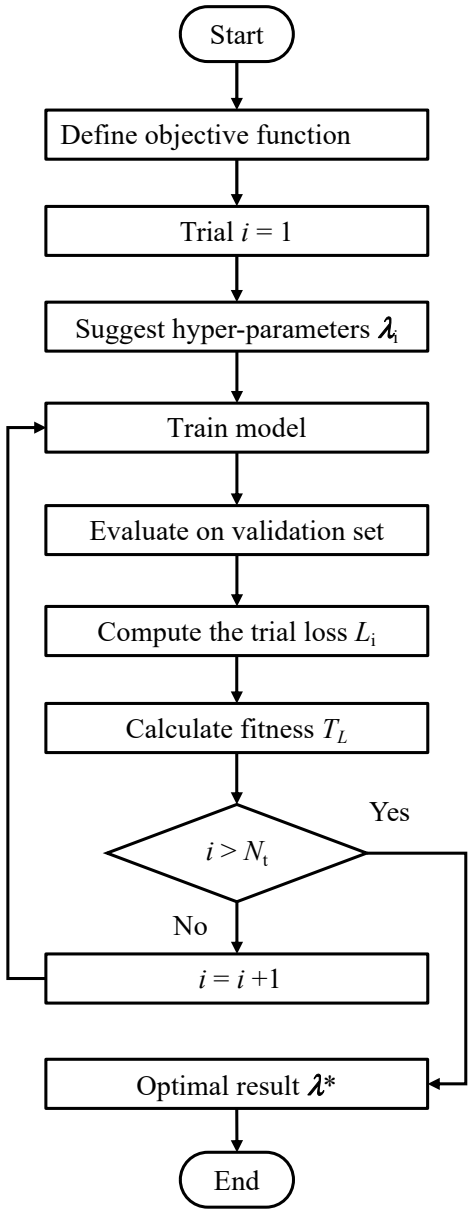

Figure 8. Universal workflow of Optuna.

With the assistance of Optuna, we develop a hybrid algorithm termed "Optuna-ML" to calibrate the hyper-parameters of the adopted machine-learning models for predicting gushing-induced ground disturbance. For each Optuna-ML algorithm, the key hyper-parameters to be tuned are specified together with their data types and searching ranges. The selection of these critical hyper-parameters and their bounds is determined according to modelling experience and widely used practices in previous machine-learning studies.

Before running the optimisation, several Optuna settings are predefined to ensure a consistent and reproducible tuning process. Specifically, the sampler is set to a TPE-based Bayesian optimiser with a fixed random seed, and the maximum number of trials is limited to N_t as the termination condition. During optimisation, Optuna iteratively proposes candidate hyper-parameter sets, trains the corresponding model on the training subset, and evaluates its performance on the validation subset using the MSE as the objective value. Once the trial budget is exhausted, the best trial is selected as the optimal result and is used for subsequent model retraining and testing.

3.4. Auxiliary Methods

3.4.1. Group k-Fold Cross-Validation

To enhance the generalisation ability of the prediction model while avoiding information leakage caused by correlated samples, Group k-fold cross-validation (GroupKFold) is adopted on the model-development set in this study. Unlike standard k-fold cross-validation [38], which randomly splits individual samples, GroupKFold partitions the data according to a predefined group label. The model-development set is divided into k folds at the group level, such that samples belonging to the same group are always assigned to the same fold. In each iteration, $k-1$ folds are used for training and the remaining fold is used for validation, and the procedure repeats k times so that each fold serves as the validation set once. The overall cross-validation performance is taken as the average of the fold-wise errors, expressed as:

$$T_L = \frac{1}{k}\sum_{i=1}^{k} \mathrm{MSE}_i \tag{14}$$

where MSE_i denotes the validation error of the i-th fold. By enforcing group-wise separation, GroupKFold provides a more realistic assessment of model performance when samples within the same group share similar boundary conditions or physical settings, which is consistent with the data structure in this work.

3.4.2. Evaluation Metrics

To quantitatively evaluate the predictive performance of the proposed models, four evaluation indices are adopted in this study, including the root mean square error (RMSE), mean absolute error (MAE), symmetric mean absolute percentage error (sMAPE), and goodness of fit (R^2). Their computational definitions are given as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i}^{m}(y_i - \hat{y}_i)^2} \tag{15}$$

$$\mathrm{MAE} = \frac{1}{m}\sum_{i}^{m}|y_i - \hat{y}_i| \tag{16}$$

$$\mathrm{sMAPE} = \frac{100\%}{m}\sum_{i}^{m}\frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2 + \varepsilon} \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i - \overline{y}_i)^2} \tag{18}$$

where m is the number of samples, y_i and $\hat{y}_i$ denote the measured (true) and predicted values of the i-th sample, respectively; $\overline{y}_i$ is the mean of the measured values; and e is a small constant introduced to avoid division by zero in sMAPE. In general, smaller RMSE/MAE/sMAPE and a larger R^2 indicate better model performance.

3.4.3. SHAP-Based Interpretation

To enhance the transparency and interpretability of the proposed machine-learning models, the SHapley Additive exPlanations (SHAP) method is employed to quantify the contribution of each input feature to the model predictions. SHAP is grounded in cooperative game theory and attributes the prediction of a given sample to individual features by computing Shapley values, thereby providing a consistent and locally accurate explanation. For a trained model, the prediction can be decomposed as:

$$f(\mathbf{x}) = \phi_0 + \sum_{j=1}^{p} \phi_j \tag{19}$$

where $\mathbf{x}$ is the input feature vector with p features, f_0 denotes the base value (i.e., the expected model output over the background dataset), and f_j represents the SHAP value of the j-th feature. A positive ϕ_j indicates that the feature increases the prediction relative to the base value, whereas a negative ϕ_j implies an opposite effect. By aggregating SHAP values across all samples (e.g., using the mean absolute SHAP value), the global importance ranking of features can be obtained, while dependence plots further reveal the nonlinear influence patterns and interaction trends among key variables. This SHAP-based interpretation facilitates identifying the dominant factors governing gushing-induced ground disturbance and supports a mechanistic understanding of the data-driven models.

4. Database and Processing

4.1. Database Generation

Based on the validated case-study presented in Section 2.2, this study develops a numerical database for predicting the response of a single tunnel subjected to water–soil gushing. Figure 9 presents the representative simulation setup, in which the model size is 100 m × 50 m. The soil parameters, boundary conditions, and other fundamental settings are kept the same as those adopted in the case study. Using the validated two-phase coupled MPM model, a series of plane-strain parametric simulations is performed to generate the machine-learning database. The input variables are selected to characterise the major factors controlling gushing-induced ground disturbance. According to previous studies [9,39], six parameters are considered in this study, namely the normalised tunnel depth H/D, gushing location angle θ, gushing intensity, and key soil parameters (E, ϕ, k), as shown in Figure 9. Specifically, five gushing-location scenarios are considered by prescribing the circumferential leakage angle (θ = 0°, 45°, 90°, 135°, and 180°), and three levels of tunnel burial depth (H/D = 1, 2, and 3) are examined to represent different confinement conditions. For the soil parameters, three key properties (E, ϕ, k) are treated as continuous variables and sampled using Latin hypercube sampling (LHS), which ensures an approximately uniform coverage of the multidimensional parameter space and efficiently captures uncertainty when the sample size is limited. It should be noted that E, ϕ, and k were sampled independently to provide broad and uniform coverage of the feasible parameter space for surrogate-model training, rather than to reproduce a site-specific joint distribution. In natural sandy soils, these parameters are generally correlated. For example, denser sand tends to exhibit a higher friction angle and stiffness, while permeability depends on porosity and gradation. Thus, the present database should be interpreted as a broad parametric envelope; for a specific project, the input ranges and their correlations should be constrained using site-investigation data. As discussed in Section 2.2, the cumulative inflowed soil mass, termed soil gushing mass (SGM), is adopted to represent gushing intensity, since it more directly reflects the severity and development stage of the incident than physical time. Accordingly, the resulting database is physically grounded and sufficiently representative for machine-learning model development. As a result, the final database contains 39,810 input–output pairs.

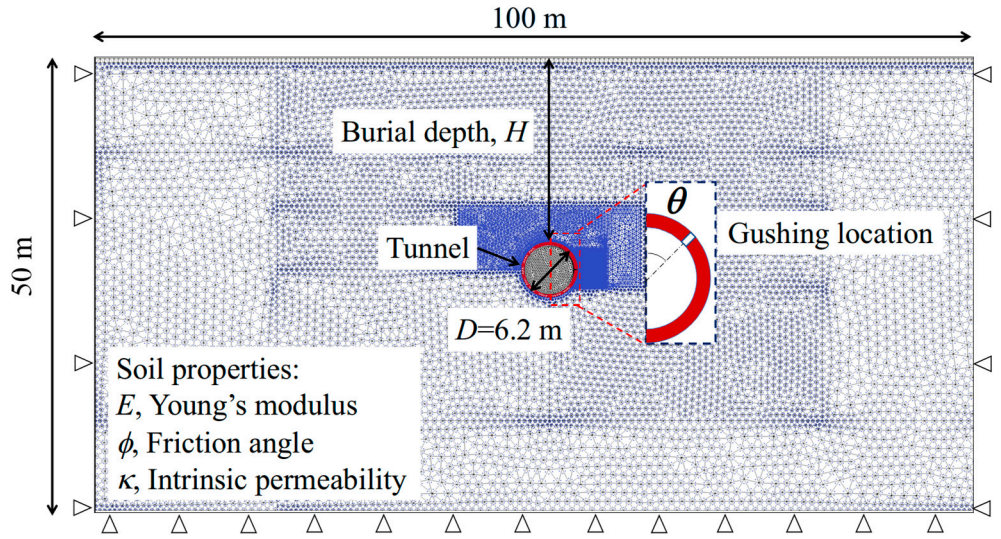

Figure 9. Numerical modelling with typical geometry and variables.

In terms of the output variables, the analysis in Section 2.2 indicates that gushing-induced ground disturbance involves not only surface settlement, but also pronounced subsurface soil deformation. Therefore, a single response quantity is insufficient to characterise the overall disturbance pattern. To provide a more comprehensive description, a representative case from the dataset is selected for illustration. Figure 10 presents the evolution of ground-surface settlement during the gushing process. As shown in Figure 10a, the settlement approaches zero when the distance from the tunnel axis exceeds approximately 20 m, indicating that the adopted model width of 100 m is sufficient to minimise boundary effects. The maximum settlement occurs approximately above the gushing point. With increasing gushing mass, the settlement trough becomes progressively deeper, whereas its lateral extent changes only slightly. Figure 10b further shows that the maximum ground settlement, denoted as S_m, increases continuously with gushing development, reflecting the persistent influence of soil loss on surface deformation.

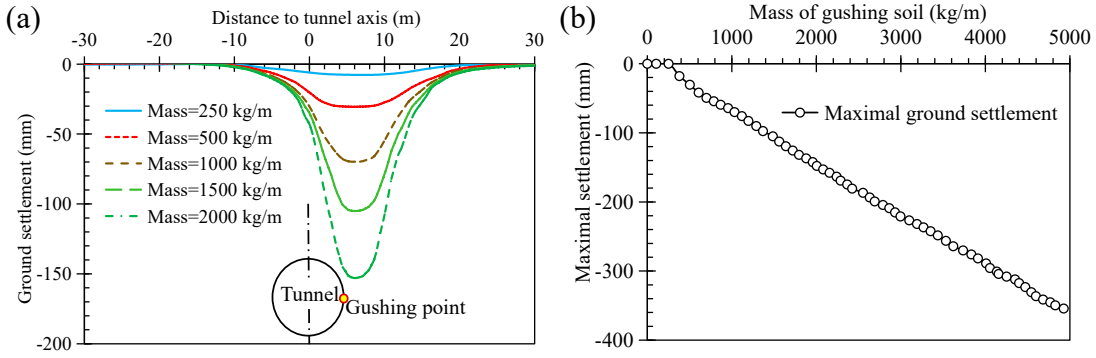

Figure 10. Development of ground settlement during gushing: (**a**) settlement trough; (**b**) maximal ground settlement.

Figure 11 shows the evolution of subsurface soil behaviour. The results indicate that the defined flow zone can reasonably capture the region of large deformation induced by gushing. To quantify its geometric characteristics, two additional descriptors are introduced, namely the flow zone centroid angle θ_f and the flow zone width d_f. These two descriptors provide interpretable indicators of the disturbed region: d_f reflects the lateral extent of large deformation, and hence the potential ground-loss footprint, while θ_f indicates the preferential direction of soil loss and the likely location of void formation

and collapse. Together, S_m, θ_f, and d_f form a set of disturbance-oriented response descriptors that characterise the intensity, extent, and spatial bias of gushing-induced ground disturbance from both surface and subsurface perspectives. These three physically interpretable and complementary variables are therefore selected as the target outputs for machine-learning prediction. As a summary, Table 2 lists the input and output variables used in this study and their mathematical statistics; the listed ranges of E, ϕ, and k define a broad sandy-ground parametric envelope rather than a single sand type with fixed density and gradation.

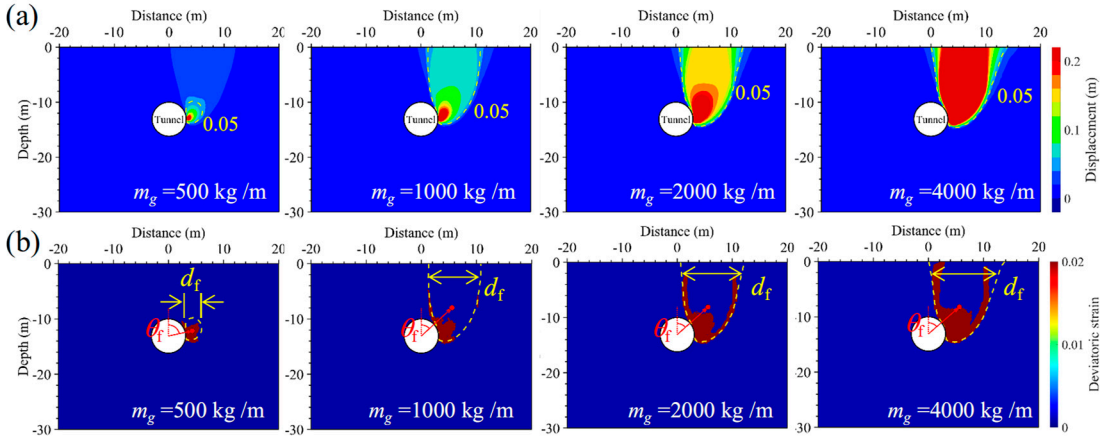

Figure 11. Evolution of soil behaviours caused by different soil gushing mass: (**a**) soil displacement; (**b**) deviatoric strain.

Table 2. Statistic of input and output variables in the database.

Variable	Parameter Type	Data (39,810 Sets)			
		Min	Max	Mean	Std
Gushing location, θ (°)	Input	0	180	68.96	54.56
Tunnel depth ratio, H/D	Input	1	3	1.98	0.82
Elastic modulus, E (kPa)	Input	5375	49,625	28,616.96	12,800.27
Friction angle, ϕ (°)	Input	15.29	49.71	31.41	10.18
Permeability, k ($\times 10^{-4}$ m/s)	Input	1.79	97.29	50.50	27.74
Gushing intensity, SGM ($\times 10^3$ kg)	Input	0.02	9.12	2.15	1.40
Maximum ground settlement, S_m (m)	Output	−0.65	−0.01	−0.16	0.12
Flow zone centroid angle, θ_f (°)	Output	−2.28	190.85	50.42	47.54
Flow zone width, d_f (m)	Output	0.01	29.56	7.95	5.09

4.2. Database Processing

Before developing the machine-learning models from the original database, several pre-processing procedures were carried out to improve data quality and model reliability. First, the dataset was divided at the group level, with 70% of the scenarios used for model development and the remaining 30% reserved for testing so as to prevent information leakage among correlated samples. Second, the input variables in the training set were normalised to [0, 1] to reduce the influence of differences in variable scale and to facilitate model training. For variables with skewed distributions, logarithmic transformation was applied prior to normalisation. The scaled value X_{scaled} was calculated as:

$$X_{\text{scaled}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \tag{20}$$

Figure 12 summarises the overall modelling workflow of ML algorithms. The procedure begins with the MPM-generated database, from which the raw samples are processed through data cleaning, group-based data splitting, and variable transformation/scaling. The processed data are then divided into a training-validation set and an independent testing set. Using the training-validation set, Optuna is adopted to optimise the hyperparameters of MLP, RF, SVR, and XGBoost, with GroupKFold (k = 5) applied to ensure reliable group-consistent validation. Here, a group is defined as one complete MPM simulation scenario, i.e., a single combination of tunnel depth ratio, gushing location, friction angle, Young's modulus, and permeability; all stage-wise samples generated as the *SGM* increases within that scenario share the same group label. The database comprises 900 such scenarios. The 70%/30% split and the subsequent 5-fold GroupKFold validation were both performed over these scenarios, ensuring that all samples from any given scenario fall entirely within either the development set or the independent test set, and that samples from one scenario never appear simultaneously in the training and validation folds. The optimised models are subsequently retrained on the full development set and assessed on the independent test set, while Ridge regression is introduced as a linear benchmark under the same protocol. Finally, SHAP analysis is performed to interpret the trained models in terms of global feature importance and the influence of governing factors on gushing-induced ground disturbance.

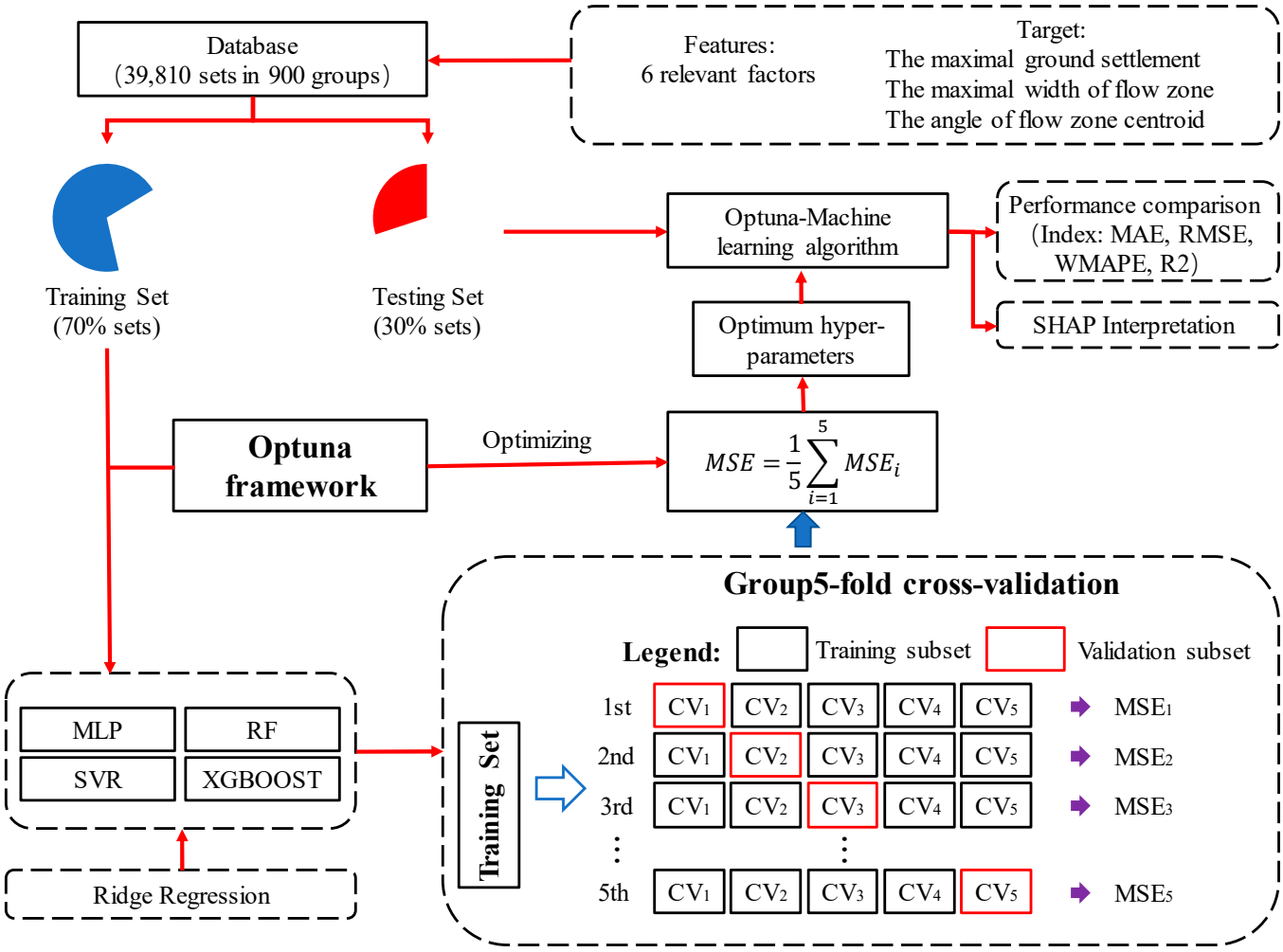

Figure 12. Modelling process of Optuna-ML.

5. Results and Discussion

5.1. Optimal Hyper-Parameters

Using Optuna to automatically search the hyper-parameter space, the optimisation trajectories of the four candidate ML models are visualised in Figure 13, where the fitness T_L is tracked against the trial number. Overall, all models reach a stable fitness plateau well before the pre-defined termination at 50 iterations, indicating that Optuna can efficiently identify near-optimal configurations for this dataset. A clear distinction is observed in the convergence behaviours of the four Optuna-based optimisers. Among them, Optuna-SVR shows the most pronounced early-stage improvement across all three targets, with the objective value dropping sharply within the first few iterations and then rapidly reaching a plateau. This pattern indicates that SVR is highly sensitive to hyper-parameter selection, and that appropriate tuning can quickly shift the model from a poor initial configuration to a near-optimal solution. In contrast, Optuna-RF exhibits a comparatively smooth and flat convergence trajectory, with only limited reductions in the objective value throughout the search process, suggesting that RF is less reliant on fine hyper-parameter adjustment and remains relatively robust within a reasonable parameter range.

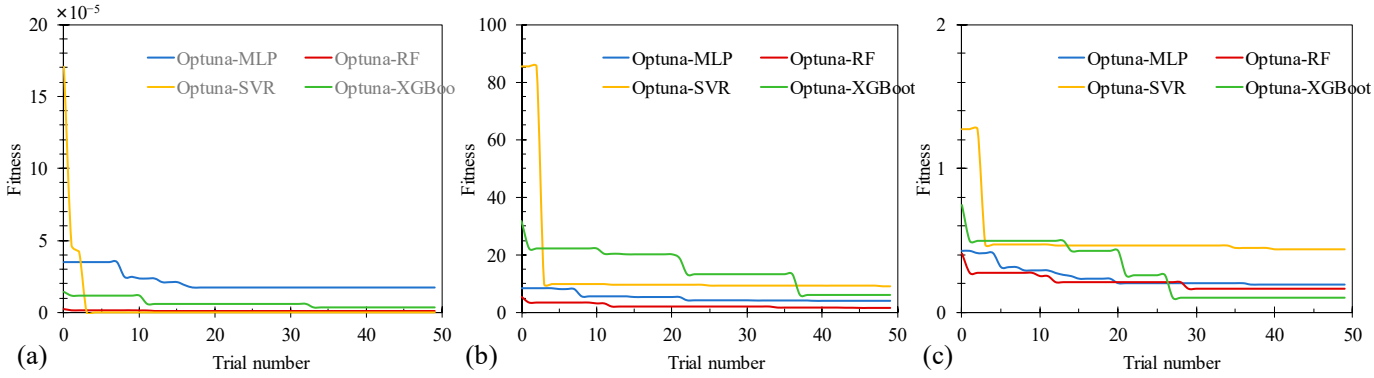

Figure 13. Evolutionary process of each Optuna-ML model for: (**a**) maximal ground settlement; (**b**) flow zone centroid angle; (**c**) flow zone width.

The other two models display intermediate behaviours. Optuna-MLP generally converges in a progressive manner, with the objective value decreasing steadily before stabilising after approximately 15–30 iterations. This trend reflects the continuous search landscape associated with neural-network hyper-parameters and the gradual refinement of network architecture and training settings. By comparison, Optuna-XGBoost shows a more stepwise convergence pattern, characterised by several discrete reductions at specific iterations prior to stabilisation. Such behaviour is consistent with the tree-based boosting mechanism, in which changes in key hyper-parameters may produce non-continuous gains in predictive performance.

It is also evident that the final converged objective values vary across the three output targets, indicating that the relative suitability of the models is output-dependent. For instance, RF achieves the lowest objective value for flow zone centroid angle, whereas XGBoost performs best for flow zone width, while SVR and RF remain particularly competitive for maximal ground settlement. Once the optimisation curve becomes essentially stationary, the corresponding hyper-parameter combination is taken as the optimal setting and is used to retrain the final model. The resulting optimal hyper-parameters for all models are listed in Table 3 for reproducibility and subsequent comparison.

Table 3. Optimal hyper-parameters tuned in each Optuna-ML model.

Algorithm	Hyper-Parameter	Optimum Value		
		S_m	θ_f	d_f
MLP	hidden_layer_numbers	2	2	2
	hidden_layer_sizes	127	125	104
	learning_rate_init	0.00395	0.002377	0.004283
SVR	c_penalty	16.14365	52.34702	17.385922
	epsilon	0.00128	0.012844	0.05142
	g_width_index	2.873511	2.524072	2.669898
RF	n_estimators	1079	735	871
	max_depth	28	27	25
	min_samples_split	3	2	2
XGBoost	n_estimators	2206	2339	2233
	max_depth	7	9	8
	learning_rate	0.010486	0.010621	0.016961
	subsample	0.79059	0.898184	0.845293

It should be noted that the hyperparameters were selected using only the GroupKFold validation MSE on the development set, whereas the independent scenario-level test set was reserved exclusively for the final evaluation and was never used during tuning. To examine the robustness of the optimization, the tuning–retraining–evaluation procedure was repeated under five random seeds and with two validation objectives (MSE and MAE); the results are summarized in Table 4. The model ranking was preserved in all cases. RF was essentially invariant (test-RMSE coefficient of variation $\leq$ 0.9%), whereas MLP and XGBoost showed only small-to-moderate seed-to-seed variation (CV $\leq$ 8.7%) arising from stochastic weight initialization and boosting; switching the objective from MSE to MAE changed the test RMSE by at most 9.7% without altering the ranking. These results confirm that the conclusions are robust to the random seed and the optimization objective.

Table 4. Robustness of hyperparameter optimization across random seeds.

Model	RMSE CV over 5 Seeds, % (S_m/d_f/θ_f)	ΔRMSE, MSE→MAE, % (S_m/d_f/θ_f)
MLP	4.33/8.73/4.60	0.50/9.68/2.83
RF	0.15/0.21/0.94	0.03/0.07/0.19
XGBoost	1.07/1.81/7.58	0.19/3.27/3.88

5.2. Performance of ML Models

Using the optimal hyperparameters summarised in Table 3, the Optuna-ML models are developed for prediction. Model performance is evaluated according to the metrics defined in Section 3.4.2. As illustrated in Figure 14, the predictive accuracy of the candidate models is compared using three complementary metrics, namely MAE, RMSE, and sMAPE. A highly consistent trend is observed across all tasks: the nonlinear models (MLP, RF, and XGBoost) markedly outperform the linear baseline (Ridge), and the model ranking implied by the three metrics is largely coherent. In particular, Ridge yields the largest MAE/RMSE and the highest sMAPE in all three outputs, indicating that the gushing-induced ground disturbance is strongly nonlinear with pronounced feature coupling, and therefore cannot be adequately characterised by a purely linear regression model.

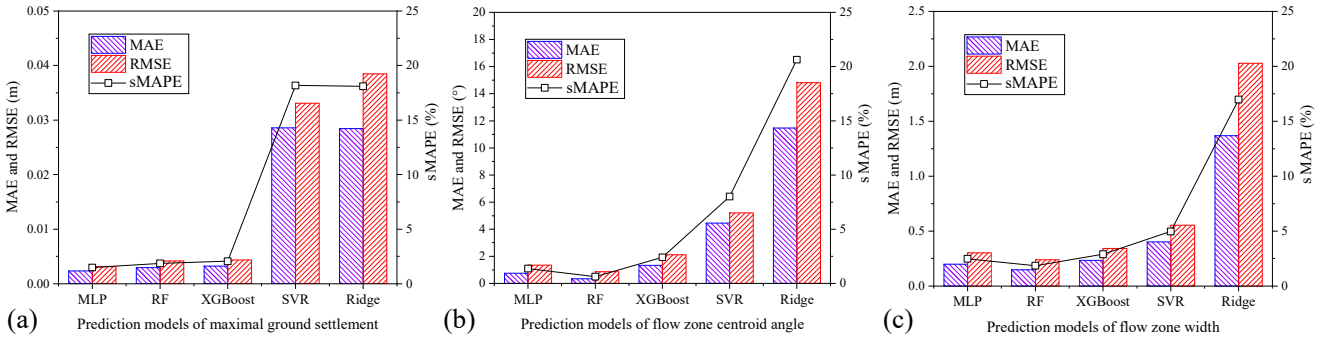

Figure 14. Performance of each model for: (**a**) maximal ground settlement; (**b**) flow zone centroid angle; (**c**) flow zone width.

Across the three prediction targets, a consistent performance hierarchy can be observed. For the maximum ground settlement, S_m, MLP, RF, and XGBoost all maintain relatively low MAE and RMSE, with sMAPE remaining in a narrow range, indicating stable predictive capability for settlement response, whereas SVR and especially Ridge show markedly larger errors and much higher sMAPE, suggesting limited generalisation. For the centroid angle of flow zone θ_f, the three nonlinear models again outperform SVR and Ridge, with relatively low and stable errors, indicating their stronger ability to capture the nonlinear evolution of flow-zone geometry. A similar trend is observed for d_f, where MLP, RF, and XGBoost remain within the lower error range, while SVR and Ridge perform noticeably worse. Overall, MLP, RF, and XGBoost all demonstrate favourable predictive performance across the three outputs, whereas Ridge, as the linear baseline, is consistently inferior, further supporting the use of nonlinear surrogate models for predicting gushing-induced ground disturbance.

Figures 15–17 present a comparison between the predicted and actual values of the five models, with the 45° line representing perfect prediction. The coefficient of determination R^2 is also reported in each model. For S_m, MLP, RF, and XGBoost generate point clouds that closely follow the 1:1 line, indicating agreement over the full response range. SVR also achieves a high R^2, but with slightly greater scatter and mild deviation at the extremes. By contrast, Ridge shows a wider dispersion and a lower R^2, indicating clear underfitting. A similar pattern is observed for θ_f and d_f, where MLP, RF, and XGBoost again maintain strong agreement with the reference line, while SVR exhibits increased dispersion and Ridge performs substantially worse, with segmented prediction patterns that reflect limited ability to capture the nonlinear evolution of flow-zone geometry. Overall, the parity plots consistently indicate that MLP provides accurate and reliable surrogate predictions across all three outputs, whereas SVR is slightly less stable and Ridge is unable to adequately represent the strong nonlinearity and feature coupling involved in gushing-induced ground disturbance.

To exploit the complementary strengths of the models, a validation-error-weighted average and a non-negative linear stacking model of MLP, RF, and XGBoost were constructed on the out-of-fold development-set predictions and evaluated on the independent test set (Table 5). The weighted average did not surpass the best single model, whereas the non-negative stacking ensemble matched or slightly improved on the best single model (MLP) for all three targets, providing the most reliable overall predictor.

Table 5. Test-set performance of single models and ensemble strategies.

Target	Model	RMSE	sMAPE (%)
S_m	MLP/RF/XGBoost	0.00677/0.00921/0.00970	5.06/6.14/9.55
	Weighted avg/Stacking	0.00689/0.00663	5.04/4.70
θ_f	MLP/RF/XGBoost	2.171/4.363/4.589	5.15/9.30/15.27
	Weighted avg/Stacking	2.623/2.145	7.21/5.21
d_f	MLP/RF/XGBoost	0.416/0.665/0.646	3.17/4.85/5.36
	Weighted avg/Stacking	0.450/0.419	3.49/3.39

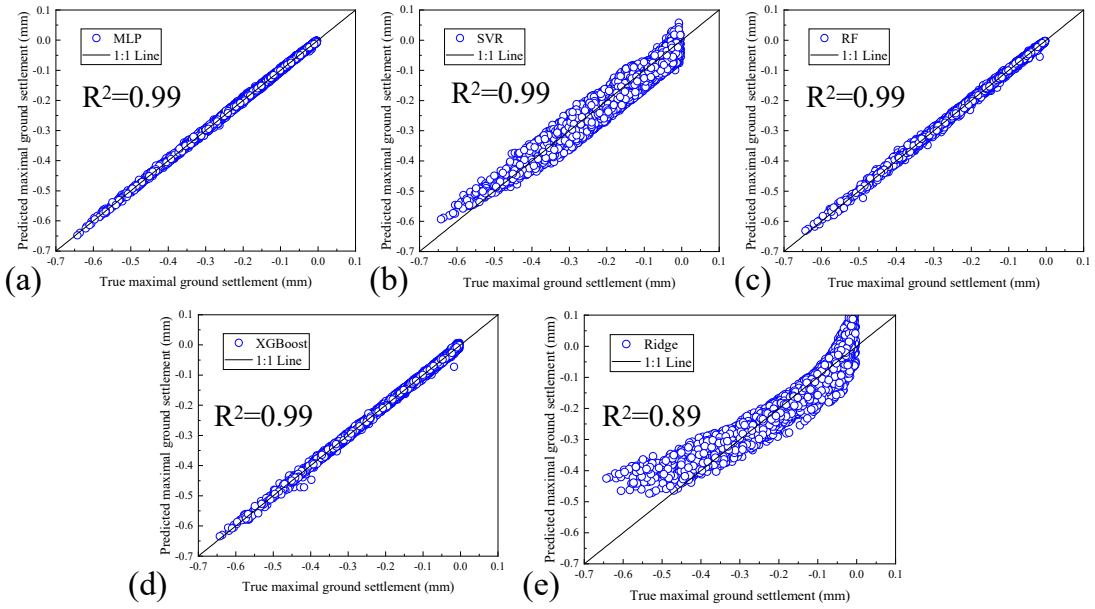

Figure 15. Analysis of predicted maximal ground settlement with true values: (**a**) MLP; (**b**) SVR; (**c**) RF; (**d**) XGBoost; (**e**) Ridge.

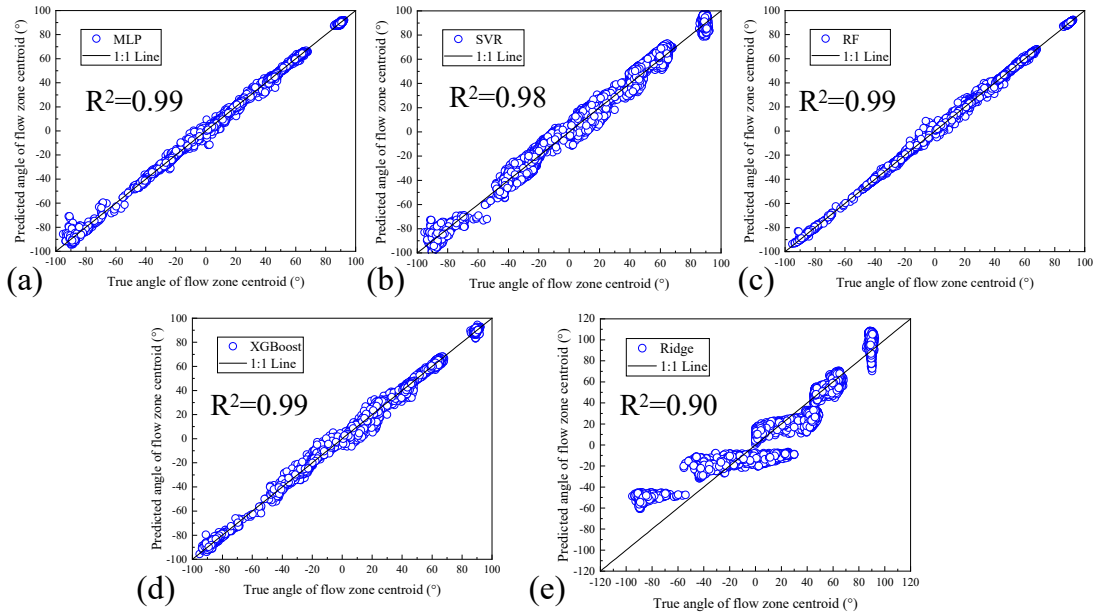

Figure 16. Analysis of predicted flow zone centroid angle with true values: (**a**) MLP; (**b**) SVR; (**c**) RF; (**d**) XGBoost; (**e**) Ridge.

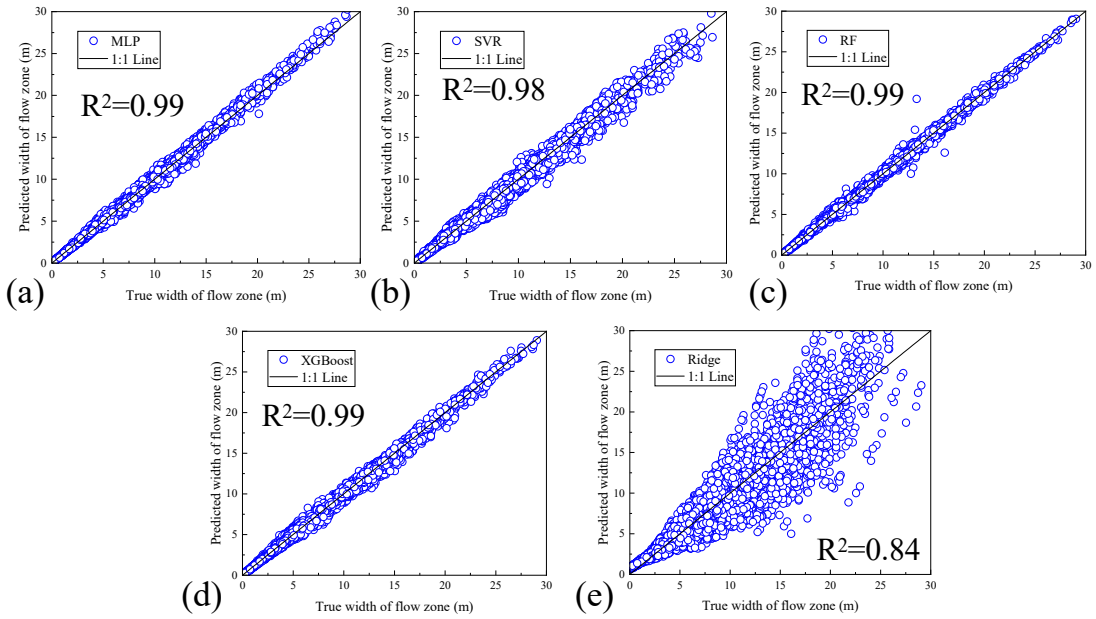

Figure 17. Analysis of predicted flow zone width with true values: (**a**) MLP; (**b**) SVR; (**c**) RF; (**d**) XGBoost; (**e**) Ridge.

5.3. Distribution of Prediction Errors

To further evaluate the reliability of different predictors beyond the aggregated metrics, the prediction residuals are post-processed in terms of the absolute percentage error (APE), defined as

$$\mathrm{APE} = \frac{\left| y_{\mathrm{true}} - y_{\mathrm{pred}} \right|}{\left| y_{\mathrm{true}} \right|} \times 100\% \tag{21}$$

Following common engineering tolerance criteria, the APE is classified into five intervals, i.e., 0–5%, 5–10%, 10–15%, 15–20%, and >20%. The corresponding frequencies for all models are summarised in Figure 18 for the three outputs.

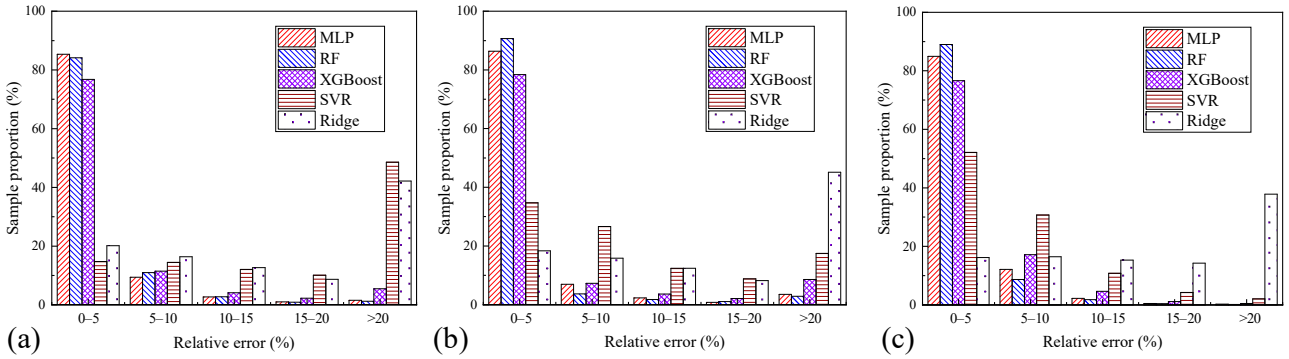

Figure 18. Comparison of the relative error between each model for: (**a**) maximal ground settlement; (**b**) flow zone centroid angle; (**c**) flow zone width.

A clear and consistent trend is observed across all three targets. The nonlinear models, particularly RF and MLP, concentrate most samples in the low-error ranges (APE < 10%), with a dominant proportion falling within 0–5%, indicating both high predictive accuracy and strong robustness. XGBoost shows a similar but slightly less favourable distribution. In contrast, Ridge exhibits the weakest performance, characterised by a reduced proportion of low-error samples and the highest frequency of APE > 20%, reflecting its limited ability to capture the nonlinear gushing-induced disturbance. SVR generally lies between these two groups but still retains a noticeable medium-to-high error tail.

From an engineering prediction perspective, the practical value of a surrogate model depends not only on its average accuracy, but also on its ability to suppress large-error cases. In this respect, MLP and RF demonstrate the most favourable overall behaviour, combining a high proportion of low-APE predictions with a low occurrence of unacceptable errors. These distributional characteristics are consistent with the MAE, RMSE, and sMAPE results and further confirm the advantage of nonlinear machine-learning models for predicting gushing-induced ground disturbance.

5.4. Global Feature Importance by SHAP

SHAP values are employed to interpret the trained models and quantify the relative contribution of each input variable to the three target responses. Figure 19 shows the global feature importance measured by the mean absolute SHAP value, mean(|SHAP|). A broadly consistent ranking is observed across the five models, indicating that the identified controlling factors are robust and physically meaningful.

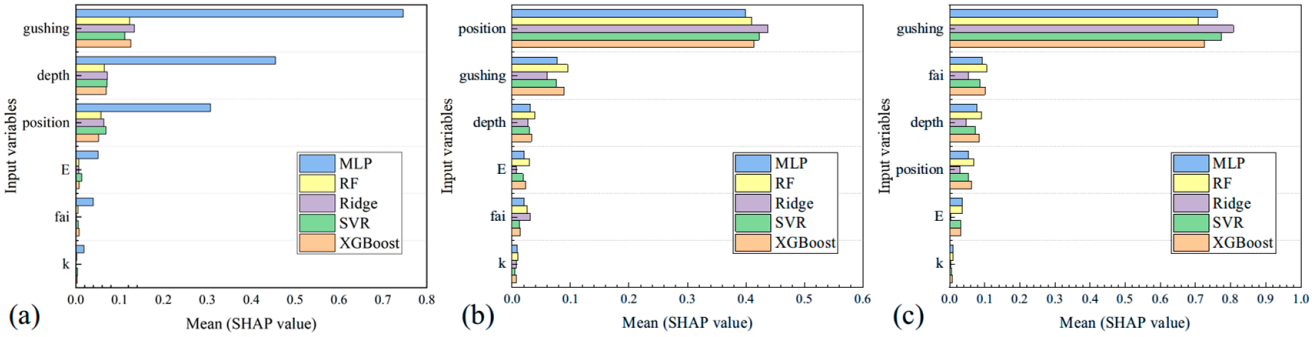

Figure 19. Comparison of global SHAP importance between each model for: (**a**) maximal ground settlement; (**b**) flow zone centroid angle; (**c**) flow zone width.

For S_m, the gushing mass *SGM* is the dominant predictor, followed by tunnel depth H/D, showing that settlement magnitude is governed mainly by gushing severity and overburden confinement. Soil modulus E has only a secondary effect, whereas k and f contribute relatively little within the investigated ranges. For d_f, *SGM* and the gushing location are the two most influential variables, indicating that the scale of the disturbed zone is jointly controlled by gushing intensity and leakage geometry, with smaller contributions from H/D and soil properties. For q_f, the gushing location is the dominant factor, while *SGM* and H/D provide noticeable secondary effects through their influence on flow-zone morphology. Overall, the SHAP results are consistent with the hydro-mechanical behaviour of gushing-induced ground disturbance, suggesting that the models have captured physically meaningful relationships.

6. Conclusions

This study developed an interpretable machine-learning surrogate framework for predicting gushing-induced ground disturbance around shield tunnels by integrating a validated two-phase MPM database with data-driven modelling. Within the scope of this work, the following conclusions can be drawn:

(1) A physics-based numerical database containing 39,810 samples was established using the validated two-phase MPM framework. By selecting the maximum ground settlement, flow zone centroid angle, and flow zone width as output variables, the database characterises gushing-induced disturbance from both surface and subsurface perspectives, thereby providing a physically meaningful basis for surrogate modelling.

(2) Among the five candidate algorithms, the nonlinear models consistently outperform the linear Ridge baseline for all three targets, confirming that the mapping between the governing factors and gushing-induced disturbance is strongly nonlinear. In particular, MLP, RF, and XGBoost achieve the best overall predictive accuracy, whereas Ridge shows clear underfitting and SVR exhibits relatively weaker robustness.

(3) In terms of comprehensive performance, MLP and RF are the most reliable models. They not only achieve low MAE, RMSE, and sMAPE values, but they also produce the most favourable error distributions, with a high proportion of predictions concentrated in the low-error intervals and a low frequency of large-error cases. This indicates that these two models are more suitable for practical engineering prediction where robustness is as important as average accuracy.

(4) SHAP-based interpretation demonstrates that the learned input–output relationships are physically consistent. For maximum ground settlement, the soil gushing mass SGM is the dominant controlling factor, followed by tunnel depth ratio H/D, indicating the primary roles of gushing severity and overburden confinement. For flow zone width, SGM and gushing location jointly govern the scale of the disturbed zone. For flow zone centroid angle, leakage location is the dominant factor, while SGM and H/D provide measurable secondary effects through their influence on flow-zone morphology.

From an engineering standpoint, the proposed surrogate framework provides an efficient alternative to repeated high-cost numerical simulations and can support several practical tasks. First, it enables pre-event scenario evaluation: by sweeping the governing variables (burial depth ratio, leakage location, and soil parameters), engineers can identify tunnel sections and ground conditions that are most susceptible to severe gushing-induced disturbance before any incident occurs. Second, once leakage information becomes available during an incident, the model allows rapid assessment of gushing severity, translating an estimated soil gushing mass into the expected maximum settlement and flow-zone geometry within seconds rather than hours of simulation. Third, the predicted flow-zone width and centroid angle provide a preliminary identification of the subsurface disturbance zone—its lateral extent and preferential propagation direction—which can guide targeted monitoring, prioritisation of inspection, and emergency-response planning. These capabilities make the framework suitable for risk-informed decision-making and for integration into smart-city underground infrastructure management systems.

It should nevertheless be emphasized that the present framework must be interpreted within its modelling scope. First, the database is generated from two-dimensional plane-strain MPM simulations, so three-dimensional effects such as longitudinal gushing development and spatial arching are not captured. Second, the ground is idealized as homogeneous sandy soil with simplified boundary conditions and a prescribed leakage-channel location and width, rather than layered or composite strata with an explicitly modelled waterproof-system failure. Third, the surrogate models are trained and validated against MPM results, and direct calibration with field monitoring data is still limited. Consequently, the trained models should be applied primarily within the investigated parameter ranges and to similar geological and geometric settings. Future work will incorporate field monitoring data for direct calibration, extend the database to three-dimensional and heterogeneous-stratum conditions and different tunnel geometries through transfer learning and adaptive retraining, introduce variables describing the waterproof-system state, and add uncertainty quantification, so as to broaden the applicable scope of the framework for practical tunnel engineering. Z.G. C.

https://doi.org/10.3390/smartcities9060100

Author Contributions: Conceptualization, X.-C.X. and Y.-X.Z.; methodology, X.-C.X. and Z.-G.C.; formal analysis, X.-C.X. and Z.-G.C.; data curation, X.-C.X.; writing—original draft preparation, X.-C.X.; writing—review and editing, Y.-X.Z.; visualization, X.-C.X. and Z.-G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by National Natural Science Foundation of China (Grant Nos. 52408436, 52238010).

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Chen, S.-R.; Zhang, D.-M.; Xie, X.-C.; Chen, Z.-G.; Chen, H.-H.; Li, H.-Y. Suffusion in shield tunnel surrounding soils under train vibration using an improved DEM-PNM coupling method. *Tunn. Undergr. Space Technol.* **2026**, *168*, 107130. [CrossRef]
2. Xie, X.-C.; Zhang, D.-M.; Zhou, M.-L. Evolution of mechanical properties of shield tunnels induced by water-soil gushing. *Undergr. Space* **2026**, *26*, 82–105. [CrossRef]
3. Gong, C.; Cheng, M.; Ge, Y.; Song, J.; Zhou, Z. Leakage mechanisms of an operational underwater shield tunnel and countermeasures: A case study. *Tunn. Undergr. Space Technol.* **2024**, *152*, 105892. [CrossRef]
4. Lin, Y.; Sun, Q.; Liu, X. Analysis of the mechanism of ground collapse induced by water and sand inflow during tunnel construction: A case study of xi'an Metro Line 8. *Transp. Geotech.* **2025**, *55*, 101664. [CrossRef]
5. Zhang, D.-M.; Xie, X.-C.; Zhou, M.-L.; Huang, Z.-K. An incident of water and soil gushing in a metro tunnel due to high water pressure in sandy silt. *Eng. Fail. Anal.* **2021**, *121*, 105196. [CrossRef]
6. Xie, X.-C.; Zhang, D.-M.; Huang, Z.-K.; Chen, Z.-G.; Zhang, X.-L. Fragility assessment of shield tunnels under water and soil gushing hazards. *Acta Geotech.* **2026**, *6*. [CrossRef]
7. Xie, J.; Molins, C.; Huang, X. A hydro-mechanical coupled method for assessing the influence of localised leakage at gasketed joints on the long-term behaviour of shield tunnels. *Tunn. Undergr. Space Technol.* **2025**, *159*, 106492. [CrossRef]
8. Zhang, D.; Zhang, X.; Du, W. DEM-FEM based numerical analysis on mechanical responses of sandy soil and pipeline to seepage erosion. *Eng. Geol.* **2022**, *310*, 106868. [CrossRef]
9. Zheng, G.; Qiu, H.; Zhang, T.; Cheng, H.; Diao, Y.; Wang, K. Modelling ground and tunnel response to water-soil gushing in stratified soil. *Tunn. Undergr. Space Technol.* **2025**, *161*, 106583. [CrossRef]
10. Karoui, T.; Jeong, S.-Y.; Jeong, Y.-H.; Kim, D.-S. Experimental Study of Ground Subsidence Mechanism Caused by Sewer Pipe Cracks. *Appl. Sci.* **2018**, *8*, 679. [CrossRef]
11. Liang, C.; Wang, Y.; Wu, C.; Zou, H. Experimental investigation on gush disaster in shield tunnel driven through upper-silt and lower-sand strata. *Tunn. Undergr. Space Technol.* **2025**, *158*, 106393. [CrossRef]
12. Zheng, G.; Qiu, H.; Zhang, T.; Sun, J.; Cheng, H.; Diao, Y. Coupled Eulerian-Lagrangian simulation of progressive failure in shield tunnels induced by developing contact loss. *Eng. Fail. Anal.* **2024**, *155*, 107748. [CrossRef]
13. Zhang, X.-L.; Zhang, D.-M.; Huang, Z.-K.; Xie, X.-C.; Yang, J. CFD-DEM investigation into multi-mode evolutionary mechanisms of underground seepage erosion. *Transp. Geotech.* **2026**, *56*, 101838. [CrossRef]
14. Li, Y.; Liu, H.; Yang, L.; Liu, Y. An optimized DEM-SPH model for surge waves induced by riverside landslides. *Int. J. Numer. Anal. Methods Geomech.* **2024**, *48*, 270–286. [CrossRef]
15. Chen, X.; Hu, Y.; Zhang, L.; Liu, Y. 3D large-deformation modelling on face instability and sinkhole formation during tunnelling through non-uniform soils. *Tunn. Undergr. Space Technol.* **2023**, *134*, 105011. [CrossRef]
16. Zheng, Y.; Li, J.; Zhu, T.; Li, J. Experimental and MPM modelling of widened levee failure under the combined effect of heavy rainfall and high riverine water levels. *Comput. Geotech.* **2025**, *184*, 107259. [CrossRef]
17. Zheng, X.; Lyu, J.; Wang, S.; Yang, J.; Yang, F.; Osman, A. Three-dimensional Hybrid Finite Element-Material Point Method for analyzing reinforcement and stability of tunnel face with fiberglass anchor bolts. *Tunn. Undergr. Space Technol.* **2026**, *168*, 107102. [CrossRef]
18. Xie, T.; Zhu, H.; Dong, Y.; Zhou, M.; Wang, B.; Zhang, W.; Zhao, J. Geo-interface modeling with material point method: A review. *J. Rock Mech. Geotech. Eng.* **2025**, *17*, 3950–3972. [CrossRef]
19. Zhang, D.; Shen, Y.; Huang, Z.; Xie, X. Auto machine learning-based modelling and prediction of excavation-induced tunnel displacement. *J. Rock Mech. Geotech. Eng.* **2022**, *14*, 1100–1114. [CrossRef]
20. Zhang, X.; He, D.; Wang, J.; Wang, S.; Gu, M. Machine learning for predicting maximum displacement in soil-pile-superstructure systems in laterally spreading ground. *Eng. Appl. Artif. Intell.* **2025**, *139*, 109701. [CrossRef]

21. Zhang, C.; Zhang, B.; Zhang, D.; Huang, Z.; Huang, H. Machine learning-based resilience assessment of tunnel under explosions: Integrating post-explosion loss and recovery models. *Comput. Geotech.* **2025**, *179*, 107008. [CrossRef]
22. Sheil, B.; Anagnostopoulos, C.; Buckley, R.; Ciantia, M.O.; Febrianto, E.; Fu, J.; Gao, Z.; Geng, X.; Gong, B.; Hanley, K.; et al. Artificial intelligence transformations in geotechnics: Progress, challenges and future enablers. *Comput. Geotech.* **2026**, *189*, 107604. [CrossRef]
23. Zhou, S.; Cheng, X.; Li, Q.; Zheng, G.; Guo, J.; Ma, Y.; Shi, J.; Shi, Z. Failure analysis of water gushing in excavation and application of rapid dewatering and recharge emergency measures. *Eng. Fail. Anal.* **2024**, *159*, 108074. [CrossRef]
24. Soga, K.; Alonso, E.; Yerro, A.; Kumar, K.; Bandara, S. Trends in large-deformation analysis of landslide mass movements with particular emphasis on the material point method. *Géotechnique* **2016**, *66*, 248–273. [CrossRef]
25. Ceccato, F.; Beuth, L.; Vermeer, P.A.; Simonini, P. Two-phase material point method applied to the study of cone penetration. *Comput. Geotech.* **2016**, *80*, 440–452. [CrossRef]
26. Zhang, H.W.; Wang, K.P.; Chen, Z. Material point method for dynamic analysis of saturated porous media under external contact/impact of solid bodies. *Comput. Methods Appl. Mech. Eng.* **2009**, *198*, 1456–1472. [CrossRef]
27. Jassim, I.; Stolle, D.; Vermeer, P. Two-phase dynamic analysis by material point method. *Int. J. Numer. Anal. Methods Geomech.* **2013**, *37*, 2502–2522. [CrossRef]
28. Bardenhagen, S.G.; Guilkey, J.E.; Roessig, K.M.; Brackbill, J.U.; Witzel, W.M.; Foster, J.C. An improved contact algorithm for the material point method and application to stress propagation in granular material. *Comput. Model. Eng. Sci.* **2001**, *2*, 509–522.
29. Anura3D MPM Research Community. Anura3D Version 2023 Source Code. 2023. Available online: https://www.anura3d.com (accessed on 10 June 2026).
30. Zhou, Z.; Tan, Z.; Liu, Q.; Zhao, J.; Dong, Z. Experimental Investigation on Mechanical Characteristics of Waterproof System for Near-Sea Tunnel: A Case Study of the Gongbei Tunnel. *Symmetry* **2020**, *12*, 1524. [CrossRef]
31. *GB 50911-2013*; Code for Monitoring Measurement of Urban Rail Transit Engineering. China Building Industry Press: Beijing, China, 2014.
32. Ho, T.K. The random subspace method for constructing decision forests. *IEEE Trans. Pattern. Anal. Mach. Intell.* **1998**, *20*, 832–844. [CrossRef]
33. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM: San Francisco, CA, USA, 2016; pp. 785–794.
34. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
35. Cortes, C.; Vapnik, V.; Saitta, L. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
36. Miller, A.; Panneerselvam, J.; Liu, L. A review of regression and classification techniques for analysis of common and rare variants and gene-environmental factors. *Neurocomputing* **2022**, *489*, 466–485. [CrossRef]
37. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-generation Hyperparameter Optimization Framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; ACM: Anchorage, AK, USA, 2019; pp. 2623–2631.
38. Zhang, P.; Yin, Z.-Y.; Jin, Y.-F.; Chan, T.H.; Gao, F.-P. Intelligent modelling of clay compressibility using hybrid meta-heuristic and machine learning algorithms. *Geosci. Front.* **2021**, *12*, 441–452. [CrossRef]
39. Xie, X.C.; Ceccato, F.; Zhou, M.L.; Zhang, D.M. Hydro-mechanical behaviour of soils during water-soil gushing in shield tunnels using MPM. *Comput. Geotech.* **2022**, *145*, 104688. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.