

Article

A Lightweight Physics-Guided Dual-Target Neural Network for Cavern-Volume and Maximum-Radius Prediction in Solution-Mined Salt Caverns

Wanli Liu ^{1,2} , Zhigang Shan ^{1,2}, Ding Xia ^{1,2,*} , Tao Chen ³ and Junwei Ma ^{3,4} 

¹ Zhejiang Huadong Geotechnical Investigation & Design Institute Corporation Limited, Hangzhou 310030, China; liu_wl@hdec.com (W.L.); shan_zg@hdec.com (Z.S.)

² PowerChina Huadong Engineering Corporation Limited, Hangzhou 311122, China

³ Badong National Observation and Research Station of Geohazards (BNORSG), China University of Geosciences, Wuhan 430074, China; chent@cug.edu.cn (T.C.); majw@cug.edu.cn (J.M.)

⁴ Hubei Key Laboratory of Operation Safety of High Dam and Large Reservoir, Yichang 443100, China

* Correspondence: xia_d@hdec.com

Abstract

Reliable prediction of final cavern responses is required for rapid screening of solution-mined salt caverns used in large-scale underground energy storage. Final cavern volume and maximum radius jointly control storage performance and geometric safety, but their simultaneous and physically consistent prediction remains challenging. In this study, a lightweight physics-guided neural network (LPGNN) was proposed for dual-target prediction of these two design-critical responses. The 25 recorded stage-wise construction variables were screened to retain 13 varying independent variables and are supplemented by five nonredundant process–geometric features describing leaching-time allocation, mean depth configuration, and staged depth adjustment. A shared residual encoder and two target-specific heads were trained using Huber loss and original-scale relative error for data fitting and a soft constraint related to source-defined cavern volume-coefficient consistency. The framework was evaluated using 1253 simulated five-stage solution-mining cases. On the independent test set, the LPGNN achieved mean absolute percentage error (MAPE) of 1.4779% for cavern volume and 2.8746% for maximum radius, with corresponding R^2 values of 0.9506 and 0.8602, respectively. Correlation, Mantel-test, and feature-response analyses indicate that cavern volume is mainly associated with the overall leaching scale and vertical-control configuration, whereas maximum radius is more sensitive to staged depth adjustment and localized radial development. These results suggest that the proposed LPGNN can serve as an accurate and physically constrained surrogate for preliminary comparison and rapid screening of solution-mined salt-cavern construction schemes.

Keywords: lightweight physics-guided neural network (LPGNN); salt-cavern design; cavern-volume prediction; maximum-radius prediction; soft physical constraint



Academic Editors: Chrysothemis Paraskevopoulou, Petros Koutsovitis and Nikolaos Koukoulas

Received: 6 July 2026

Revised: 20 July 2026

Accepted: 23 July 2026

Published: 30 July 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Large-scale underground energy storage is increasingly required for sustainable low-carbon energy systems. Among available subsurface storage options, solution-mined salt caverns are widely regarded as promising facilities for compressed-air energy storage (CAES), hydrogen storage, and natural-gas reserve regulation because rock salt is characterized by low permeability, favorable creep accommodation, and damage self-healing [1–3].

However, cavern performance is controlled not only by storage demand but also by complex geological conditions, repeated injection–withdrawal cycles, and thermo-mechanical or creep-fatigue loading [4,5]. Therefore, engineering feasibility must be assessed in terms of storage capacity, cavern geometry, and long-term stability [6,7]. During construction design, final cavern volume and maximum radius are two particularly important responses: cavern volume describes the final geometric storage space, whereas maximum radius informs assessments of cavern regularity, wall clearance, inter-cavern spacing, and operational safety [8].

Final cavern volume and maximum radius are largely determined during staged solution mining. Freshwater injection, brine withdrawal, tube positions, blanket-pad depth, flow rate, and stage duration jointly govern brine transport, salt dissolution, insoluble-particle accumulation, and cavern-boundary evolution [9–12]. Previous mechanistic and numerical studies have clarified these coupled processes and shown that reliable morphology prediction requires explicit representation of mass transfer, brine convection, and insoluble-particle sedimentation, particularly in bedded salt formations or formations with high insoluble contents [13–15]. Although high-fidelity simulations remain indispensable for detailed cavern design, their computational cost limits rapid exploration of large construction-parameter spaces [16]. Efficient surrogate models are therefore needed for preliminary design screening, provided that they preserve the principal relationships between staged construction controls and final cavern geometry.

Machine learning models have increasingly been used as surrogate tools for analyses of salt-cavern construction and operation [17]. For construction-stage prediction, Li et al. trained a backpropagation artificial neural network (BPANN) on 1253 simulated cases to predict cavern volume and maximum radius, thereby accelerating construction-parameter screening [8]. Wang et al. later extended this concept from scalar design-index prediction to full-geometry prediction and target-shape-oriented construction design [18]. For operation-stage assessment, Zhao et al. applied deep learning models to forecast long-term displacement and stability responses of CAES salt caverns [19,20]. Collectively, these studies show that data-driven surrogates can reduce the need for repeated forward simulations and improve the efficiency of preliminary assessments.

Despite this progress, several issues remain unresolved. First, cavern volume and maximum radius arise from the same leaching process but describe different aspects of cavern development; their coupling should therefore be represented without assuming identical response mechanisms. Second, the original stage-wise variables do not directly summarize process-level controls such as stage-duration variability, mean vertical-control positions, and staged depth-adjustment amplitudes. This limits the direct interpretability of the link between model inputs and the geotechnical processes governing cavern formation. Third, empirical error minimization alone does not explicitly encourage compatibility between the two predicted responses and the source-defined cavern volume coefficient. These issues are particularly relevant for simulation-derived datasets that contain strongly correlated construction variables [21,22].

Physics-informed and physics-guided learning provides a practical way to address these limitations [23–25]. In these frameworks, physical knowledge can be introduced through feature engineering, auxiliary quantities, or soft constraints in the training objective, rather than through empirical fitting alone [26,27]. For solution-mined salt caverns, such guidance should remain lightweight: the computational efficiency of the surrogate should be preserved while sufficient process–geometric information and engineering consistency are incorporated to discourage implausible joint predictions. A physics-guided surrogate is therefore best positioned as a complement to high-fidelity leaching simulations and geomechanical stability analyses, rather than as a replacement for mechanistic modeling.

In this study, a target-focused lightweight physics-guided neural network (LPGNN) was developed to predict cavern volume and maximum radius simultaneously. The five-stage construction sequence was represented by 13 independent stage-wise variables with nonzero variance and five nonredundant process–geometric descriptors of duration allocation, mean depth configuration, and staged depth adjustment. A residual shared encoder with two target-specific heads was then used to learn common construction–response patterns while retaining target-specific prediction pathways. A soft consistency term based on the source-defined cavern volume coefficient was incorporated into the training objective together with Huber and original-scale relative-error terms. The framework was evaluated through validation-controlled model selection, independent test performance, comparisons with benchmark and ablation models, repeated training across five seeds, and geotechnical feature–response analyses.

2. Data and Methods

2.1. Data Source

This study used the publicly available numerical dataset describing the staged solution mining of energy-storage salt caverns, available at <https://doi.org/10.1016/j.energy.2022.124238>. The dataset comprised 1253 valid cases generated for a homogeneous salt formation with 10% insoluble content, a fixed injection flow rate of 60 m³/h, and a prescribed final cavern height of 80 m. Geological properties were not varied across the cases; consequently, the dataset represents the effects of construction parameters within a single fixed geological and numerical domain. In each case, stage-specific geometric and operational settings were paired with final cavern responses, providing a consistent basis for learning the relationship between construction design and final cavern geometry.

The source files contained 25 stage-wise columns: inner-tube depth, outer-tube depth, blanket-pad depth, dissolution duration, and flow rate for each of five stages. The five flow-rate columns, Stage-1 inner-tube depth, and Stage-5 blanket-pad depth were invariant, and each outer-tube depth was determined by subtracting 5 m from the corresponding blanket-pad depth. These columns were retained as source-simulation metadata but excluded from the learnable input matrix. The remaining independent stage-wise inputs comprised four inner-tube depths, four blanket-pad depths, and five stage durations. The depth values were expressed in a shifted relative vertical coordinate and should not be interpreted as site-specific burial elevations.

The source responses comprised final effective cavern volume (m³), maximum radius (m), the cavern volume coefficient, and the leaching-efficiency coefficient (m³/day). Cavern volume and maximum radius were used as the two prediction targets. The cavern volume coefficient served only as an auxiliary reference in the final consistency loss. The leaching-efficiency coefficient was retained during source-data verification but was not used as an input, target, or loss term in the selected model. Cavern volume is a geometric quantity and does not represent the working inventory of compressed air, hydrogen, or natural gas.

2.2. LPGNN for Cavern Volume and Maximum-Radius Prediction

The overall workflow of the proposed LPGNN for cavern-volume and maximum-radius prediction is illustrated in Figure 1. After fixed and deterministic columns had been removed, the predictor matrix contained 13 independent stage-wise variables and five nonredundant process–geometric descriptors. Pearson correlations and Mantel tests were used to characterize associations among the retained inputs and between each input and the two target responses [28]. The resulting 18 variables were used as inputs to a residual shared-encoder, dual-head LPGNN. The network was trained using Huber loss on the standardized log scale, original-scale relative error, and a consistency term based on the

source-defined cavern volume coefficient. Model behavior was interpreted using SHAP, PDP/ICE curves, and response-surface analysis, whereas consistency with the auxiliary coefficient was evaluated separately [29,30].

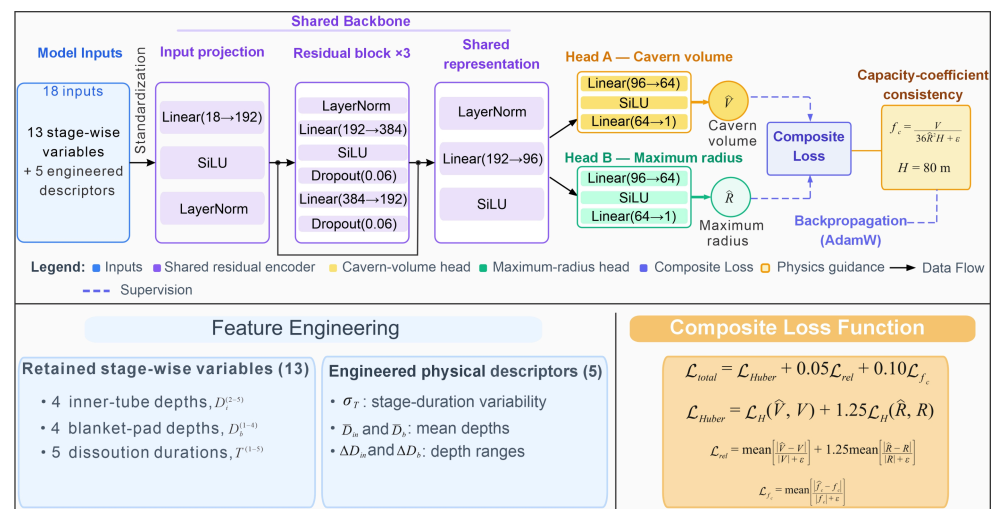


Figure 1. Overall workflow and architecture of the proposed LPGNN.

2.2.1. Feature Engineering

After zero-variance and deterministic-redundancy screening, five nonredundant process–geometric descriptors were derived from the five-stage construction records and retained for modeling (Table 1). These descriptors supplemented 13 independent stage-wise variables with nonzero variance, yielding an 18-variable input matrix that represented both the construction sequence and selected process-level summaries.

The retained descriptors comprised the standard deviation of stage durations, mean inner-tube depth, mean blanket-pad depth, inner-tube depth range, and blanket-pad depth range. Together, they represented inter-stage time allocation, the mean vertical-control configuration, and the amplitude of staged depth adjustment. Total duration and mean duration were excluded because they were nearly invariant and deterministically related. Mean outer-tube depth and outer-tube depth range were excluded because they were deterministic transforms of blanket-pad depth, whereas the tube–blanket gap was excluded because it was constant.

All descriptors were calculated from the recorded five-stage sequence before the excluded columns were removed. The five flow-rate columns were fixed at 60 m³/h; Stage-1 inner-tube depth and Stage-5 blanket-pad depth were fixed boundary conditions; and each outer-tube depth equaled the corresponding blanket-pad depth minus 5 m. Targets and auxiliary coefficients were never included in the predictor matrix. This procedure preserved informative process summaries while preventing invariant or deterministically duplicated quantities from entering the learnable input space.

Pearson correlations and Mantel tests were used as descriptive diagnostics to characterize, rather than select, the 18 retained inputs (Figure 2). The Pearson matrix reflected both the prescribed construction geometry and the derivation of the engineered descriptors. In particular, Stage-1 blanket-pad depth showed a perfect negative correlation with blanket-pad depth range, whereas Stage-5 inner-tube depth was strongly positively correlated with inner-tube depth range. The Mantel tests showed that dissimilarities in cavern volume were associated most strongly with Stage-2 blanket-pad depth, followed by Stage-3 blanket-pad depth and mean blanket-pad depth. Smaller but significant associations were observed for inner-tube depth range and Stage-5 inner-tube depth.

Table 1. Retained stage-wise variables and engineered process–geometric descriptors used as LPGNN inputs, with symbols, physical interpretations, and dataset statistics.

Feature Type	Feature	Symbol/Formula	Physical Interpretation	Mean $\pm$ SD	Range	CV
Retained stage-wise feature	Inner-tube depth (Stages 2–5; 4 variables)	$D_{in}^{(s)}, s = 2, \dots, 5$	Freshwater-injection depth at each varying stage; controls the vertical position of active dissolution.	-2009.03 ± 17.21 m	-2035.00 to -1970.00 m	0.009
	Blanket-pad depth (Stages 1–4; 4 variables)	$D_b^{(s)}, s = 1, \dots, 4$	Blanket-pad position at each varying stage; controls roof protection and the upper leaching boundary.	-1995.48 ± 18.88 m	-2030.00 to -1965.00 m	0.009
	Dissolution duration (Stages 1–5; 5 variables)	$T^{(s)}, s = 1, \dots, 5$	Time allocated to each construction stage; represents inter-stage redistribution of the fixed-duration leaching schedule.	119.98 ± 55.37 day	0.00 to 305.00 day	0.462
	Stage-duration standard deviation	$\sigma_T = \sqrt{\frac{1}{5-1} \sum_{s=1}^5 (T^{(s)} - \bar{T})^2}$	Unevenness of dissolution-time allocation among the five stages.	60.97 ± 10.72 day	35.07 to 106.85 day	0.176
Engineered physical descriptor	Mean inner-tube depth	$\bar{D}_{in} = \frac{1}{5} \sum_{s=1}^5 D_{in}^{(s)}$	Average freshwater-injection position over the complete construction sequence.	-2015.23 ± 3.22 m	-2024.60 to -2006.80 m	0.002
	Mean blanket-pad depth	$\bar{D}_b = \frac{1}{5} \sum_{s=1}^5 D_b^{(s)}$	Average roof-protection and upper-boundary control position over the five stages.	-1988.38 ± 3.98 m	-2004.00 to -1980.00 m	0.002
	Inner-tube depth range	$\Delta D_{in} = \max_s D_{in}^{(s)} - \min_s D_{in}^{(s)}$	Amplitude of staged adjustment of the freshwater-injection position.	54.04 ± 8.29 m	38.00 to 70.00 m	0.153
	Blanket-pad depth range	$\Delta D_b = \max_s D_b^{(s)} - \min_s D_b^{(s)}$	Amplitude of staged adjustment of the roof-protection and upper leaching boundary.	59.70 ± 6.20 m	50.00 to 70.00 m	0.104

Note: The final input matrix contained 13 independent stage-wise variables with nonzero variance and five engineered descriptors. Fixed flow-rate and boundary columns, deterministic outer-tube depths, and redundant aggregate descriptors were excluded. Statistics for stage-wise variables were calculated by pooling the retained stage columns across all cases. The standard deviation of stage duration was calculated using the sample standard-deviation formula across the five stages. SD denotes standard deviation; CV denotes SD divided by the absolute mean.

Dissimilarities in maximum radius showed weaker and more distributed associations. The strongest associations involved Stage-4 blanket-pad depth, Stage-3 blanket-pad depth, and Stage-3 dissolution duration. Inner-tube depth range, duration variability, Stage-5 inner-tube depth, and blanket-pad depth range were also significant. By contrast, neither mean inner-tube depth nor mean blanket-pad depth was significantly associated with dissimilarities in maximum radius.

These descriptive results suggest that cavern volume is associated primarily with mid-stage vertical control, whereas maximum radius shows weaker, distributed associations with later-stage positioning and temporal–depth variability.

2.2.2. Physics-Guided Neural Network Architecture

Before model training, all input features were standardized using means and standard deviations estimated from the training subset. Cavern volume and maximum radius were first transformed using the log1p function and then standardized using the corresponding training-subset statistics. This preprocessing improved numerical stability and reduced the scale imbalance between the two targets [31,32].

The final LPGNN used a residual shared encoder (Figure 1). An input projection mapped the 18 retained variables to a 192-dimensional representation, followed by a SiLU activation and LayerNorm. Each of the three residual blocks applied pre-normalization, a 192-to-384 linear expansion, SiLU activation, dropout, a 384-to-192 linear projection, and a

second dropout operation. The projected output was then added to the block input through an identity skip connection. A final LayerNorm and a 192-to-96 linear projection with SiLU activation formed the shared 96-dimensional latent representation. The dropout probability was 0.06. Each target-specific head mapped the shared representation through a 96-to-64 linear layer and SiLU activation and produced one standardized log-scale output [33]. One head predicted cavern volume, whereas the other predicted maximum radius. The predictions were inverse-standardized and transformed using $\expm1$ before losses and evaluation metrics were calculated on the original physical scale.

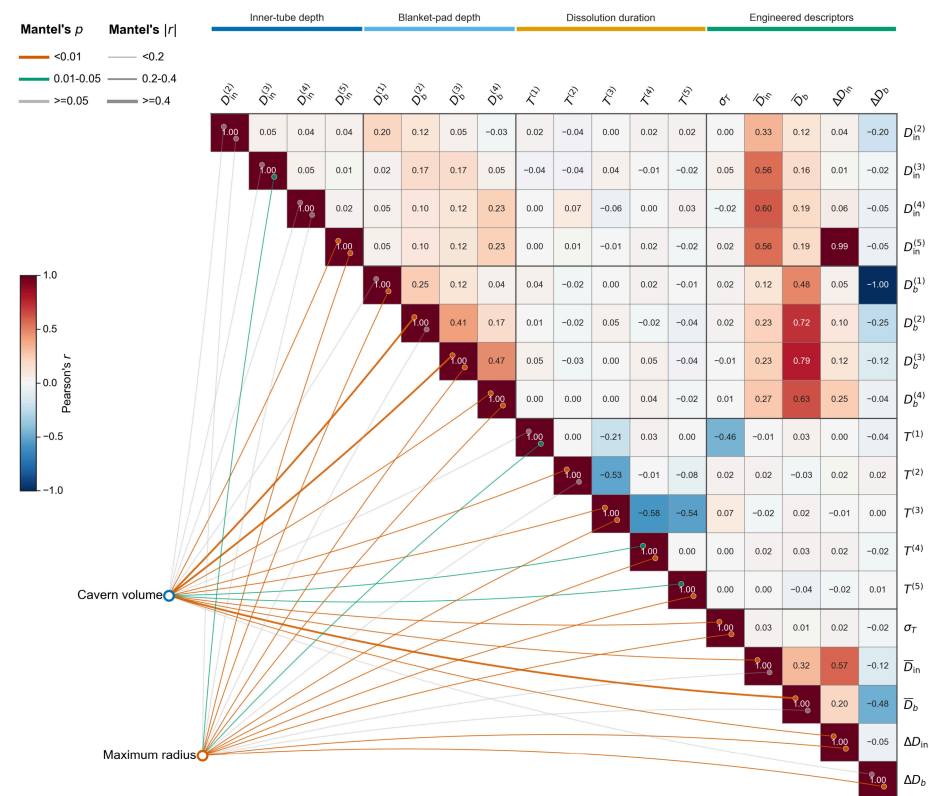


Figure 2. Pearson correlation structure and Mantel-test associations of the 18 LPGNN inputs. The upper triangular matrix reports Pearson correlation coefficients among the 18 retained inputs. Curved links connect cavern volume and maximum radius to each input; link width denotes the absolute Mantel correlation category, and link color denotes the two-sided permutation p -value category.

2.2.3. Physics-Guided Loss Design

The model was trained by minimizing a composite objective that combined robust data fitting with soft physics-guided regularization. The data-fitting component used Huber losses for cavern volume and maximum radius after both targets had been transformed to the standardized log scale. Huber loss reduced the influence of occasional large residuals while retaining sensitivity to typical prediction errors. Because the two responses differed in scale and prediction difficulty, an additional weight was assigned to the maximum-radius branch to balance learning between the two targets [34]:

$$\mathcal{L}_{\text{data}} = \mathcal{L}_{\text{Huber}}(\hat{V}_s, V_s) + w_r \mathcal{L}_{\text{Huber}}(\hat{R}_s, R_s) \quad (1)$$

where $\hat{V}_s$ and $\hat{R}_s$ denote the predicted standardized log-transformed cavern volume and maximum radius, respectively; V_s and R_s denote the corresponding reference values.

Because model selection and performance reporting emphasized percentage error, an additional relative-error term was evaluated after the predictions had been inverse-transformed to the original physical scale:

$$\mathcal{L}_{rel} = \frac{1}{n} \sum_{i=1}^n \left[\frac{|\hat{V}_i - V_i|}{|V_i| + \varepsilon} + w_R \frac{|\hat{R}_i - R_i|}{|R_i| + \varepsilon} \right] \quad (2)$$

The cavern volume coefficient provided the only retained physics-guided coupling between the two output heads. In the source dataset, the final cavern height was fixed at $H = 80$ m, and the inter-cavern safety distance was defined as four times the maximum radius. Under these source-data assumptions, the coefficient was reconstructed as follows:

$$\hat{f}_{c,i} = \frac{\hat{V}_i}{36\hat{R}_i^2 H + \varepsilon} \quad (3)$$

The reconstructed coefficient was compared with the source cavern volume coefficient using the mean relative error. The constant $\varepsilon = 10^{-6}$ prevents division by zero:

$$\mathcal{L}_{f_c} = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{f}_{c,i} - f_{c,i}|}{|f_{c,i}| + \varepsilon} \quad (4)$$

The complete training objective is defined as follows:

$$\mathcal{L}_{total} = \mathcal{L}_{data} + \lambda_{rel} \mathcal{L}_{rel} + \lambda_{f_c} \mathcal{L}_{f_c} \quad (5)$$

2.2.4. Hyperparameter Setting and Optimization Workflow

To ensure reproducible model development and controlled comparisons, all candidate models followed a common data-splitting, preprocessing, and optimization protocol. After feature construction, the 1253 cases were partitioned using joint stratification of the two targets and a fixed split seed. The resulting subsets contained 801 training cases, 201 validation cases, and 251 independent test cases (64:16:20). The test indices were isolated before configuration selection and were not accessed during standardization, early stopping, or epoch selection. After the configuration and training duration had been frozen, the selected model was refitted for 163 epochs on the combined training and validation subsets.

The maximum-radius branch weight, w_r , was determined exclusively from validation performance. Candidate values of 0.75, 1.00, 1.25, 1.50, and 2.00 were evaluated under otherwise identical model and loss settings across three prespecified seeds. The maximum-radius branch weight was set to 1.25, which gave the lowest balanced mean absolute percentage error (MAPE). The key hyperparameters of the final optimized model were as follows: learning rate = 6×10^{-4} , dropout = 0.06, batch size = 64, relative-error weight 0.05, cavern volume-coefficient weight 0.10, Huber $\delta = 1.0$, and AdamW weight decay 5×10^{-5} . The median seed-specific validation-optimal epoch was 163.

The proposed LPGNN was compared with pure data-driven baselines and physics-guided LPGNN variants. Pure data-driven baselines included random forest (RF), support-vector regression SVR, PSO-SVR, CatBoost, XGBoost, and LightGBM [35–40]. The LPGNN-based models included the optimized model, LPGNN-opt; an input-representation reference model using only the 25 original stage-wise variables, LPGNN-raw25; a baseline physical-regularization variant, LPGNN-base; a maximum-radius-emphasized variant, LPGNN-Rmax; and a smaller-batch variant, LPGNN-SB. Additional physics-guided candidates, including PGHE-risk, PGHE-avg, PGS-I, PGMTL, and PGWT, were used to examine whether alternative forms of physical regularization or hybrid prediction could further im-

prove performance. Because the same data split, preprocessing strategy, target definitions, and evaluation metrics were used for all models, the comparison directly reflected the effects of model class, input representation, and physics-guided loss design. All model were conducted under Windows 11 on a workstation equipped with an Intel Core i9-14900HX CPU, 32 GB RAM, and an NVIDIA GeForce RTX 4060 Laptop GPU with 8 GB VRAM, using Python 3.12.5, PyTorch 2.11.0 (CUDA 12.8), scikit-learn 1.8.0, NumPy 2.0.1, and pandas 2.2.2.

2.3. Evaluation Metrics

Model evaluation focused on the predictive accuracy of the two primary engineering responses: cavern volume and maximum radius. All metrics were calculated on the independent test set after the predictions had been inverse-transformed from the standardized log scale to their original physical units.

Let y_i and $\hat{y}_i$ denote the reference and predicted values for sample i , respectively, and n be number of independent-test samples. Four metrics were used: MAPE, mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination (R^2).

MAPE was used as the primary comparison metric because it provides a dimensionless percentage-based error measure, allowing errors in cavern volume and maximum radius to be compared on a common scale. A small ε was included to avoid numerical instability.

$$\text{MAPE} = \frac{100}{n} \sum_i \left| \frac{y_i - \hat{y}_i}{y_i + \varepsilon} \right| \quad (6)$$

MAE quantified the mean absolute prediction error in the original physical units of each target:

$$\text{MAE} = \frac{1}{n} \sum_i |y_i - \hat{y}_i| \quad (7)$$

RMSE was included because squaring the residuals gives greater weight to large deviations, which may be important in engineering assessment:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_i (y_i - \hat{y}_i)^2} \quad (8)$$

The coefficient of determination was used to quantify the proportion of variance in the reference values explained by the predictions:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (9)$$

where $\bar{y}$ denotes the mean reference value.

In addition to target-prediction metrics, agreement with the source-defined cavern-volume coefficient was evaluated on the independent test set. The case-wise coefficient-consistency error was defined as

$$e_{f_c,i} = \frac{|\hat{f}_{c,i} - f_{c,i}|}{|f_{c,i}| + \varepsilon} \times 100\%. \quad (10)$$

Because the relation was implemented as a soft regularizer rather than a hard admissibility condition, the distribution was summarized using the mean, median, 95th percentile, maximum, and the descriptive proportions within 1%, 2%, and 5% tolerances.

3. Results and Discussion

3.1. Training Convergence of LPGNN-Opt

Figure 3 summarizes the validation-only model-selection results for the configuration later frozen as LPGNN-opt across three prespecified seeds. The mean smoothed training objective decreased from 0.346 at epoch 1 to 0.035 at epoch 50 and 0.011 at epoch 163. Over the same interval, validation MAPE decreased from 3.745% to 2.221% for cavern volume and from 5.086% to 3.943% for maximum radius. Subsequent changes were small: at epoch 200, the corresponding validation MAPEs were 2.216% and 3.957%. The median of the three seed-specific optimal epochs was therefore fixed at 163 before the final refit.

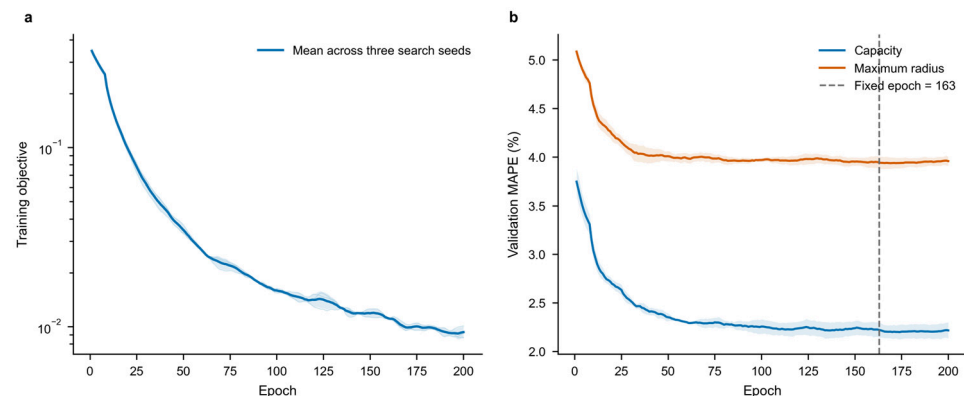


Figure 3. Validation-guided training behavior of LPGNN-opt. (a) Training objective smoothed with a 15-epoch moving average and then averaged across the three prespecified model-selection seeds; the shaded band denotes mean $\pm$ SD. (b) Corresponding validation MAPEs for cavern volume and maximum radius. The vertical dashed line marks the fixed training duration of 163 epochs, defined as the median validation-optimal epoch before the independent test set was accessed.

3.2. Comparative Performance Against Data-Driven and Physics-Guided Models

Table 2 summarizes the predictive performance of LPGNN-opt and the comparison models. The published BPANN model of Li et al. [8] was first used as an external benchmark. Compared with BPANN, LPGNN-opt reduced the cavern volume MAPE from 1.838% to 1.4779% and the maximum-radius MAPE from 3.144% to 2.8746%. The corresponding relative reductions were 19.6% and 8.6%, respectively. Although the numerical margins were modest, improvements were obtained for both responses simultaneously. This is important for salt-cavern design because cavern volume and maximum radius represent different engineering requirements; an improvement in storage-volume prediction should not be achieved at the expense of geometric-control prediction. This balanced improvement can be attributed to the target-focused formulation of LPGNN-opt. In this formulation, cavern volume and maximum radius were retained as the two primary prediction targets, whereas the source-defined cavern volume coefficient was retained only as a soft consistency link between the two output heads. As a result, the model was guided toward physically reasonable joint predictions while the main optimization priority remained focused on the two design-critical responses. The LPGNN-opt contained 480,738 trainable parameters, corresponding to 1.834 MiB of raw float32 weights. For comparison, the published 15-512-512-512-2 BPANN contains 534,530 trainable parameters under a standard dense implementation with biases; one LPGNN member therefore contains 10.1% fewer trainable parameters.

Table 2. Benchmark comparison between pure data-driven surrogate models and physics-guided candidate models.

Group	Model	Cavern Volume MAPE (%)	Maximum-Radius MAPE (%)	Model Description
Pure data-driven	CatBoost	2.3126	4.3250	Gradient-boosted decision-tree baseline; the strongest pure data-driven model in this comparison.
	XGBoost	2.3524	4.4438	Extreme gradient-boosting tree ensemble used as a pure data-driven tabular baseline.
	LightGBM	2.3771	4.4751	Leaf-wise gradient-boosting tree ensemble used as a pure data-driven tabular baseline.
	RF	2.9132	5.0047	Random forest ensemble baseline using bagged decision trees.
	PSO-SVR	2.5936	4.2986	Support-vector regression optimized using particle swarm optimization.
Physics-guided	LPGNN-opt	1.4779	2.8746	Final shared-encoder dual-head LPGNN selected as the optimized model.
	LPGNN-raw25	1.9675	3.3048	Strict ablation variant using the same split, dual-head architecture, physics-guided loss terms and training controls as the LPGNN but using only the 25 original stage-wise construction variables. It is included to isolate the contribution of engineered physical descriptors.
	LPGNN-base	1.7450	3.0910	Dual-target LPGNN variant with baseline-like physical regularization, used to evaluate the effect of the final lighter-physics weighting strategy.
	LPGNN- $R_{\max}$	1.8881	3.1865	Dual-target LPGNN variant that increases emphasis on the maximum-radius branch to test target-balancing sensitivity.
	LPGNN-SB	1.8337	3.2618	Dual-target LPGNN variant trained with a smaller mini-batch size to examine optimization sensitivity to batch-level sampling.
	Physics-guided hybrid ensemble (PGHE-risk)	2.2758	3.7348	Post hoc physics-guided ensemble that combines candidate predictions using risk-aware weights derived from physical-consistency behavior.
	Physics-guided hybrid ensemble (PGHE-avg)	2.2807	3.7325	Simple average ensemble of selected physics-guided candidates, used to test whether ensemble averaging improves robustness.
	Physics-guided strategy I (PGS-I)	2.5629	4.1196	Physics-guided weighting strategy that strengthens radius-related influence through inverse-maximum-radius-squared weighting.
	Physics-guided multi-task learning network (PGMTL)	4.6543	5.9965	Earlier physics-guided multi-output network that treats multiple related physical quantities as coequal learning targets.
	Physics-guided weighted-training baseline (PGWT)	2.6445	4.3842	Weighted-training baseline that introduces physical preference through the training objective without the final target-focused LPGNN formulation.

The comparison with pure data-driven models further indicates that the performance difference did not arise from nonlinear regression alone. Among the newly trained data-driven baselines, CatBoost achieved the best overall performance, with MAPEs of 2.3126%

for cavern volume and 4.3250% for maximum radius. These errors were reduced by LPGNN-opt to 1.4779% and 2.8746%, corresponding to relative reductions of approximately 36.1% and 33.5%, respectively.

The LPGNN-based comparison models were used to examine the contribution of the input representation and loss design. When only the 25 original stage-wise variables were used, the resulting LPGNN-raw25 model produced higher MAPEs of 1.9675% for cavern volume and 3.3048% for maximum radius. This result shows that the process-geometric features generated through feature engineering provided useful information beyond the raw construction variables. LPGNN-base achieved performance close to that of LPGNN-opt but remained slightly less accurate, indicating that the final weighting of the physical consistency terms was more suitable for this dual-target task. In contrast, greater emphasis on the maximum-radius branch in LPGNN- $R_{\max}$ and the smaller-batch setting in LPGNN-SB did not improve the final prediction accuracy.

Other physics-guided candidates were also less effective than LPGNN-opt. PGHE-risk and PGHE-avg yielded cavern volume MAPEs of approximately 2.28% and maximum-radius MAPEs of approximately 3.73%. The PGS-I, PGMTL, and PGWT models produced still larger errors. These results indicate that more complex weighting strategies, additional output structures, or post hoc combinations do not necessarily lead to better prediction. The effectiveness of physical guidance depends on how engineering relationships are incorporated into the model.

Overall, the best joint accuracy was achieved by LPGNN-opt. The comparison results show that the performance improvement was mainly associated with two design choices: the use of process-geometric feature engineering to represent staged cavern construction, and the use of moderately weighted consistency constraints to regularize the coupled cavern volume-radius response. Therefore, LPGNN-opt provided a more effective balance between predictive accuracy and engineering consistency than either pure data-driven models or more elaborate physics-guided alternatives. The following section further examines the observation-prediction agreement of LPGNN-opt for the two target responses.

3.3. Prediction Accuracy and Target-Specific Error Structure

Reference and predicted responses are compared in Figure 4. For cavern volume, LPGNN-opt achieved $R^2 = 0.9506$, $RMSE = 1674.21 \text{ m}^3$, $MAE = 1080.26 \text{ m}^3$, and $MAPE = 1.4779\%$. The bootstrap 95% interval for MAPE was 1.2553% to 1.7347%. For maximum radius, the independent-test ensemble achieved $R^2 = 0.8602$, $RMSE = 1.2103 \text{ m}$, $MAE = 0.9015 \text{ m}$, and $MAPE = 2.8746\%$; the bootstrap 95% MAPE interval was 2.5703% to 3.1953%.

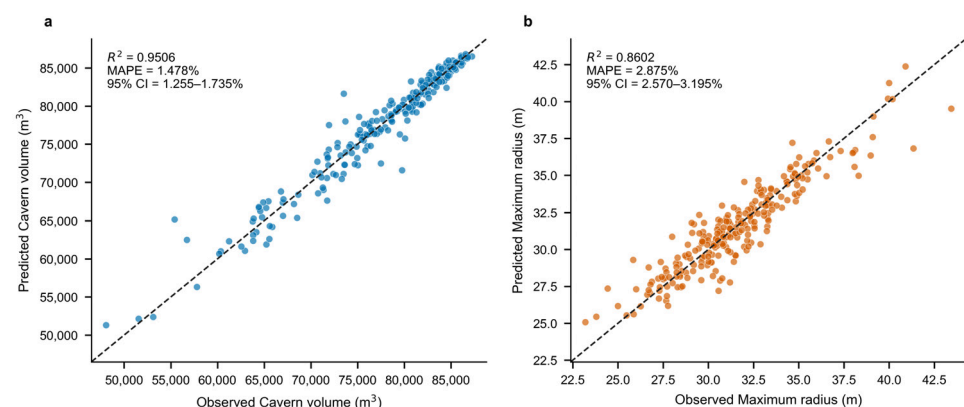


Figure 4. Independent-test reference-prediction agreement for LPGNN-opt for (a) cavern volume and (b) maximum radius. Dashed lines denote perfect agreement. Insets report R^2 , MAPE, and the 95% bootstrap interval of MAPE obtained by resampling the independent-test cases.

The target-specific difference can be interpreted in light of the physical character of the two responses. Cavern volume is an integrated volumetric response associated with the overall leaching scale, mean vertical-control configuration, and stage-duration allocation. Maximum radius, by contrast, is a localized radial extreme that is more sensitive to staged geometric adjustment and local boundary evolution. Its lower R^2 should therefore be interpreted conservatively: maximum radius may depend on higher-order stage interactions and spatial boundary evolution that are not fully represented by the fixed-length tabular inputs. The performance difference reflects both the distinct physical character of the targets and limitations of the current input representation.

3.4. Geotechnical Significance of Process–Geometric Descriptors and Prediction Reliability

The process–geometric descriptors were designed to express the original five-stage construction records in coordinates with clearer geotechnical meaning. The 13 independent stage-wise variables were retained, and five descriptors were added to represent stage-duration variability, mean inner-tube and blanket-pad positions, and the ranges of staged depth adjustment. This representation made the temporal allocation and vertical organization of the leaching sequence more explicit than the stage-wise variables alone, thereby supporting geotechnical interpretation of the learned construction–response relationships.

The engineering value of the feature-engineered variables was examined by comparing LPGNN-opt with LPGNN-raw25 under the same modeling conditions. The same data partition, network architecture, physics-guided loss formulation, and training settings were used in both models. The LPGNN-raw25 model was trained using only the 25 original stage-wise construction variables. In this way, the effect of feature representation on the prediction of cavern volume and maximum radius was evaluated under a controlled setting. Improved prediction accuracy was obtained for both engineering responses after the process–geometric features were included. For cavern volume prediction, the MAPE was reduced from 1.9675% in LPGNN-raw25 to 1.4779% in LPGNN-opt. For maximum-radius prediction, the MAPE was reduced from 3.3048% to 2.8746%. The simultaneous improvement of the two targets indicates that the constructed features provided complementary construction-process information that could not be fully represented by the original stage-wise variables alone.

The larger improvement in cavern volume prediction is consistent with the volumetric nature of cavern volume. Cavern volume is mainly controlled by the cumulative dissolution scale and the effective vertical range of leaching. These controls were represented by total leaching duration, mean depth configuration, and tube–blanket spacing, which describe the overall cavern-forming conditions more directly than individual stage values. By contrast, maximum radius is governed more strongly by localized boundary expansion and by the redistribution of the active dissolution zone during different construction stages. Therefore, depth-adjustment ranges and duration variability were more relevant to maximum-radius prediction because they describe changes in the position and persistence of the leaching interface. The smaller but consistent improvement in maximum-radius prediction suggests that part of the localized radial-growth behavior was captured by feature engineering, although this response remained more difficult to represent using stage-aggregated construction information.

Robustness to random initialization was assessed through five independent training runs in which the data partition, architecture, objective, and hyperparameters were held fixed. For cavern volume, the mean $\pm$ SD values were $1.573 \pm 0.034\%$ for MAPE and 0.9449 ± 0.0030 for R^2 . For maximum radius, the corresponding values were $2.988 \pm 0.075\%$ and 0.8500 ± 0.0084 . Across the five runs, cavern-volume MAPE ranged from 1.524% to 1.615%, whereas maximum-radius MAPE ranged from 2.919% to 3.102% (Figure 5). This

limited spread indicates low sensitivity to model initialization under the adopted training protocol; it does not constitute a site-level uncertainty estimate.

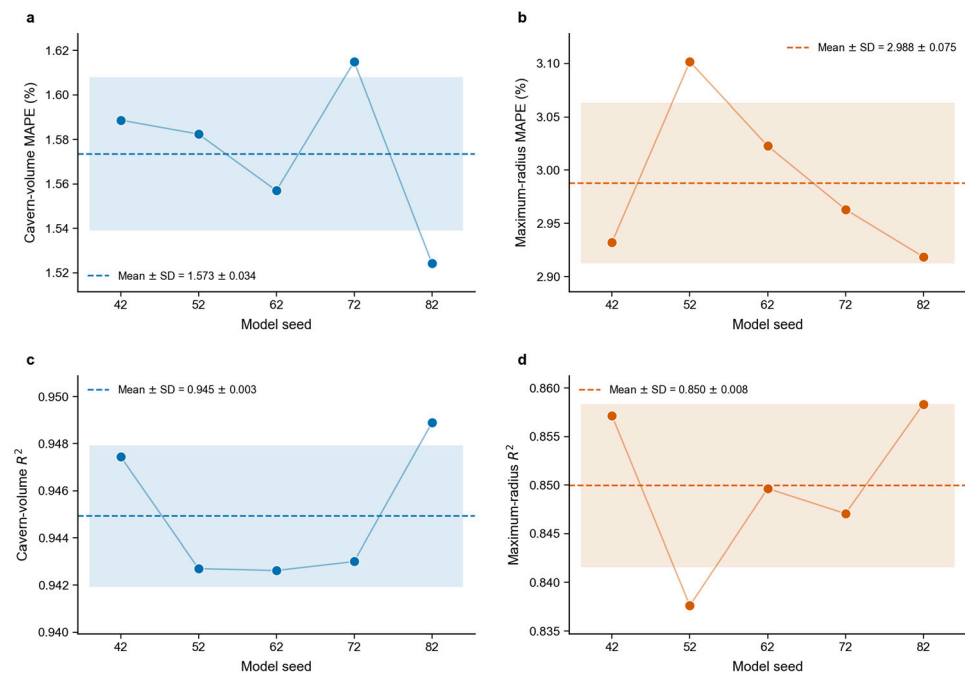


Figure 5. Prediction reliability of LPGNN-opt across five independent training runs. Panels show (a) cavern-volume MAPE, (b) maximum-radius MAPE, (c) cavern-volume R^2 , and (d) maximum-radius R^2 . Dashed lines denote the mean values, and shaded bands represent the mean $\pm$ standard deviation.

Overall, the results demonstrate that process–geometric feature engineering improved both the predictive accuracy and the geotechnical interpretability of LPGNN. By reorganizing the original construction records into features related to leaching duration, vertical-control configuration, staged adjustment, and tube–blanket spacing, a closer relationship was established between construction design and cavern response. The repeatable performance obtained across independent training runs further supports the use of the proposed framework for preliminary design assessment and rapid screening of solution-mined salt-cavern construction schemes.

3.5. Geotechnical Controls on Cavern Development and Physical Consistency of Predictions

The learned construction–response relationships were interpreted using PDP/ICE curves, permutation importance, and SHAP analysis (Figures 6–8). Figure 6 shows average and sample-specific response trends, Figure 7 quantifies permutation importance, and Figure 8 presents SHAP values that describe the magnitude and direction of each input’s contribution to individual predictions [41]. Because the construction variables are coupled through the staged leaching design, the three analyses were interpreted jointly as model-based associations rather than independent causal effects. Their combined evidence was used to identify the principal geotechnical controls represented by the model.

Figures 6–8 identified two principal groups of construction controls. The first comprised the vertical-control conditions around Stages 2 and 3, particularly the Stage-2 and Stage-3 blanket-pad depths. The second comprised staged geometric adjustments, including inner-tube depth range and selected stage-duration and tube-position variables. Both groups were associated with the final cavern responses, but their roles differed between cavern volume and maximum radius. Cavern volume represents the integrated volumetric outcome of the leaching process, whereas maximum radius represents a localized radial extreme.

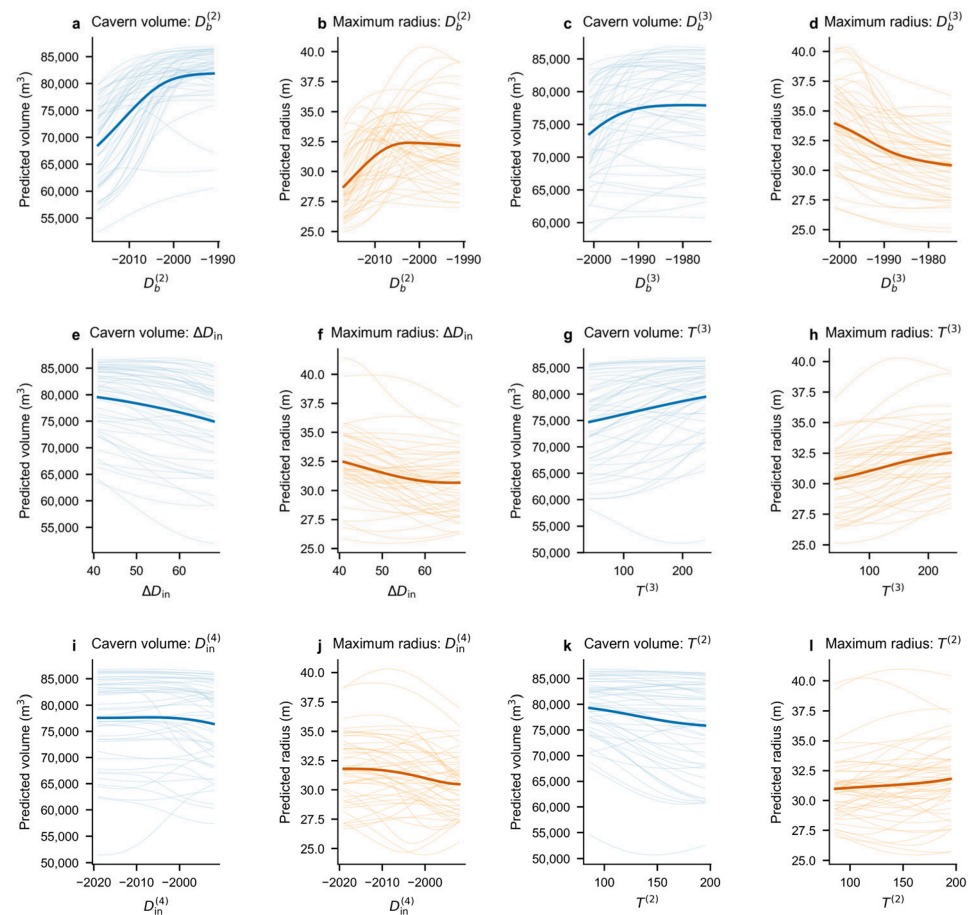


Figure 6. PDP and ICE responses of selected construction variables and process–geometric features for the cavern volume and maximum-radius branches of LPGNN-opt. Thin curves represent 50 reproducibly sampled independent-test cases, and thick curves show the corresponding partial dependence. Each variable was varied over the 5th–95th percentile of the training–validation distribution while the remaining model inputs were held fixed. The curves are conditional model responses rather than independent causal effects.

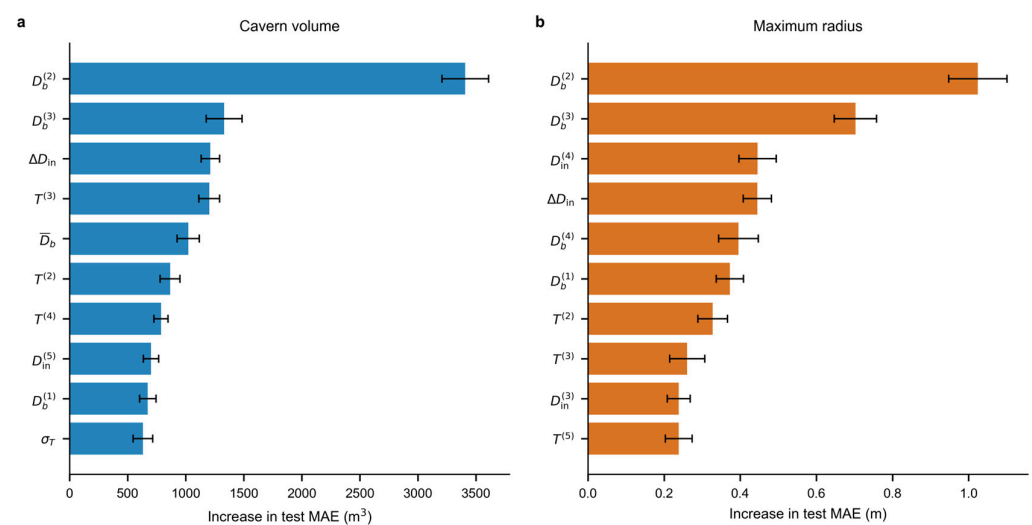


Figure 7. Permutation importance of selected construction variables and process–geometric features for cavern volume and maximum-radius prediction. Bars report the mean increase in MAE after independently permuting each input for (a) cavern volume and (b) maximum radius; error bars denote SD across 30 permutations. Larger values indicate a greater loss of predictive accuracy when the corresponding input information is disrupted.

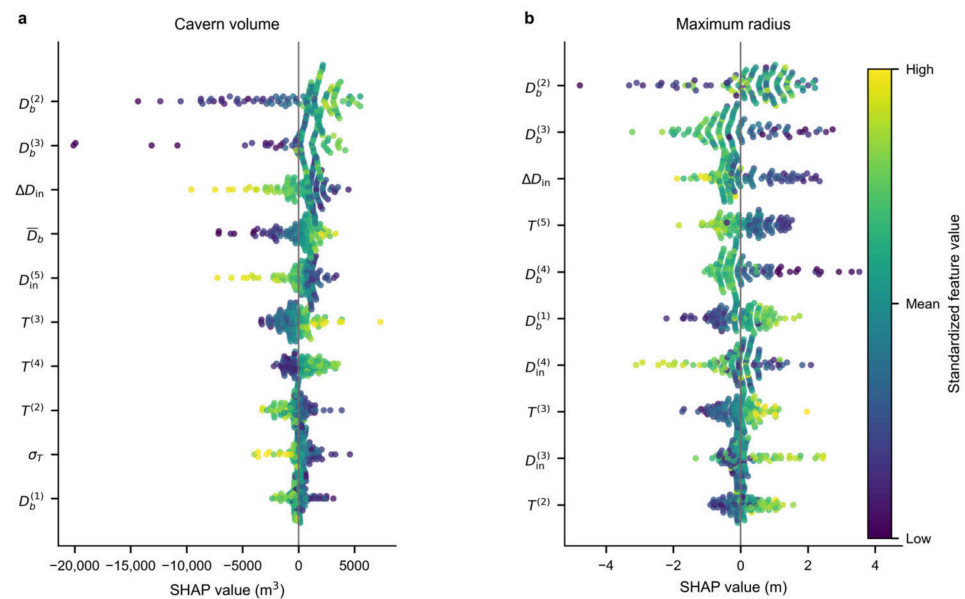


Figure 8. SHAP summary plots for (a) cavern-volume and (b) maximum-radius prediction. Expected-gradient SHAP values were computed for 120 reproducibly sampled independent-test cases and averaged across the five LPGNN-opt members. Horizontal position gives the contribution in the original target unit, and color denotes the standardized feature value. Features are ordered separately for the two targets by mean absolute SHAP magnitude.

For cavern-volume prediction, Stage-2 blanket-pad depth was the most influential input: permuting this variable increased test MAE by $3408 \pm 201 \text{ m}^3$. It was followed by Stage-3 blanket-pad depth, inner-tube depth range, and Stage-3 dissolution duration. These inputs describe the elevation of the active leaching interval and the allocation of dissolution time during the transition from initial cavern formation to subsequent enlargement. Their importance is consistent with the role of vertical control in defining the effective dissolution space and final cavern volume. Because the source simulations prescribed each outer-tube depth at 5 m below the corresponding blanket pad, outer-tube depths were excluded, and the retained blanket-pad depths represent the independent vertical-control variables. Because stage durations were coupled through the prescribed five-stage schedule, the Stage-3 duration effect should be interpreted as inter-stage time allocation rather than simply as a longer total leaching time.

For maximum-radius prediction, the largest permutation effects were associated with inputs describing the migration and local persistence of the active dissolution zone. Stage-2 and Stage-3 blanket-pad depths ranked first and second, with MAE increases of $1.024 \pm 0.077 \text{ m}$ and $0.703 \pm 0.055 \text{ m}$, respectively, followed by Stage-4 inner-tube depth, inner-tube depth range, and Stage-4 blanket-pad depth. This ranking is consistent with maximum radius being governed by localized radial enlargement and by the stage at which dissolution is concentrated near a particular elevation. Compared with cavern volume, maximum radius was therefore more sensitive to stage-to-stage changes in injection position, blanket-pad control, and duration allocation.

Permutation importance and SHAP identified the same three leading inputs for both targets: Stage-2 blanket-pad depth, Stage-3 blanket-pad depth, and inner-tube depth range. The PDP/ICE curves showed smooth average trends but appreciable dispersion among individual trajectories, indicating that the response to one input depended on the broader construction sequence. These patterns are consistent with, but do not independently validate, the use of a shared encoder with separate output heads. The ICE dispersion is not a predictive probability distribution; it reflects conditional heterogeneity in deter-

ministic model responses. Likewise, the five-seed spread measures sensitivity to random initialization, whereas the bootstrap confidence interval quantifies sampling variation in test-set MAPE. Neither is a calibrated prediction interval for an individual cavern or a new geological site.

Target-specific interactions between the principal construction controls were further examined using two-dimensional response surfaces (Figure 9). The model-predicted surfaces showed that shallower Stage-2 blanket-pad settings were generally associated with increases in both cavern volume and maximum radius. A shallower Stage-3 blanket-pad setting further increased predicted cavern volume but tended to reduce predicted maximum radius, particularly when the Stage-2 blanket pad was already relatively shallow. This contrast suggests that the relative positions of adjacent blanket controls affect how the model distributes dissolution between integrated cavern enlargement and localized radial development. When total dissolution time was held constant, allocating more time to Stage 3 produced a nonlinear increase in both predicted responses, whereas a larger inner-tube depth range generally reduced them. These conditional patterns illustrate why settings favorable for cavern volume need not produce an equivalent change in maximum radius and are consistent with joint prediction through separate output heads. Gray regions outside the support of the combined training and validation data should not be used for engineering interpretation, and the displayed surfaces should be read as model-based associations rather than causal effects.

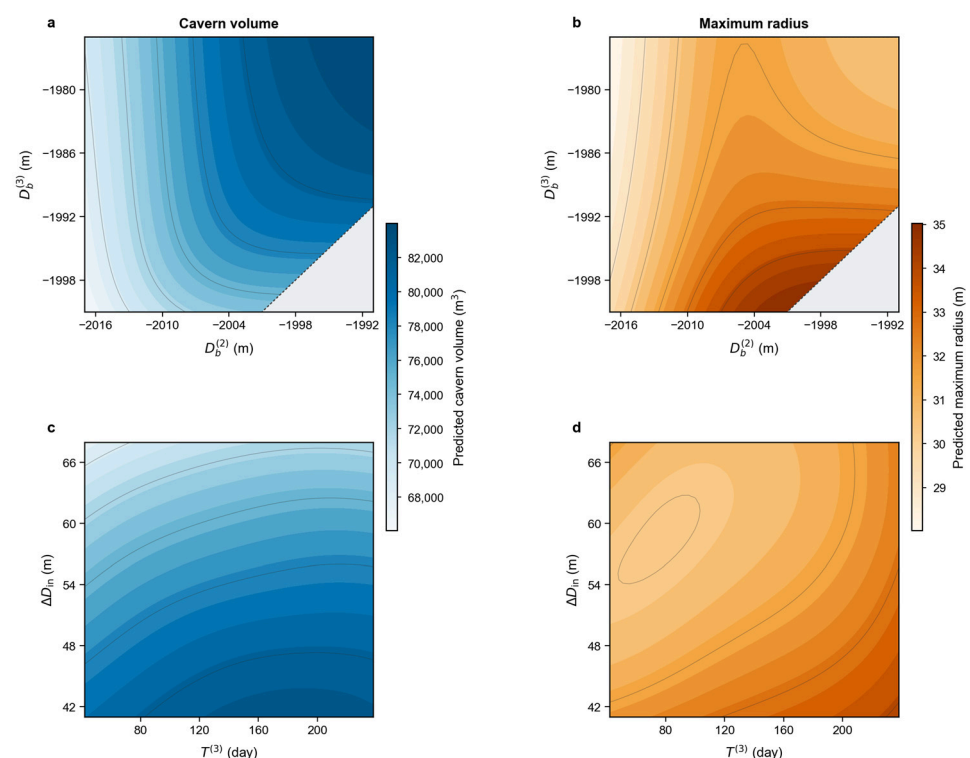


Figure 9. Two-dimensional response surfaces of the LPGNN-opt showing the coupled effects of vertical control and duration–depth adjustment on cavern volume and maximum radius. (a) Predicted cavern volume as a function of Stage-2 and Stage-3 blanket-pad depths. (b) Predicted maximum radius for the same vertical-control variables. (c) Predicted cavern volume as a function of Stage-3 dissolution duration. (d) Predicted maximum radius for the same duration–depth-adjustment variables.

We also trained a strict ablation model, LPGNN-no-fc, using the same inputs, residual dual-head architecture, data split, and original-scale relative-error term, with only λ_{fc} changed from 0.10 to 0. As shown in Table 3, within the 1% coefficient-consistency tolerance band, the LPGNN-opt included 35 of the 251 independent test cases (13.94%), whereas

LPGNN-no-fc included 20 cases (7.97%). This corresponds to 15 additional cases and an absolute increase of 5.98 percentage points for LPGNN-opt, indicating better coverage within the strictest descriptive tolerance band. These mixed differences indicate that removing $\mathcal{L}_{fc}$ cause a pronounced loss of accuracy or consistency in this initialization.

Table 3. Descriptive comparison between the LPGNN-opt and the LPGNN- no-fc ablation.

Model	Cavern Volume MAPE (%)	Maximum-Radius MAPE (%)	$e_{(f_c,i)}$ (%)	Within 1% (%)	Within 2% (%)	Within 5% (%)
LPGNN-opt	1.4779	2.8746	6.025	13.94 (35/251)	23.11 (58/251)	54.98 (138/251)
LPGNN-no-fc	1.5520	3.0216	6.3667	7.97 (20/251)	20.72 (52/251)	52.59 (132/251)

3.6. Methodological Transferability Across Geological Settings and Stage Configurations

The methodological contribution of the LPGNN lies in the formulation of the prediction problem. Stage-wise construction variables are supplemented with physically interpretable summaries, cavern volume and maximum radius are learned jointly through separate output heads, and consistency of the source-defined cavern volume coefficient is used to discourage incompatible predictions. These choices are not tied to a single fitted network. For another dataset, the stage representation and consistency terms could be reconstructed around the available variables while retaining the coupled-target architecture and validation-based training procedure. What is transferable is therefore the modeling strategy, not the fitted weights. The reported accuracy remains specific to the domain represented by the source simulations.

The present dataset does not evaluate transfer across geological settings. All 1253 cases share a homogeneous salt formation with 10% insoluble content, a final cavern height of 80 m, a constant injection rate of 60 m³/h, and five direct-circulation stages. Geological factors, including salt purity, bed thickness, non-salt interbeds, faults and fractures, in situ stress, temperature, permeability, and creep or damage properties, were not varied across the cases; their effects therefore cannot be identified from these data. Accordingly, the fitted LPGNN is limited to predicting construction-stage cavern geometry within this fixed geological and operational domain. Application to a materially different site would require a representative site-specific dataset and model retraining. The model does not assess cavern stability, leakage, well integrity, or allowable operating pressure; these outcomes require site-specific geological characterization and geomechanical analysis.

The fixed five-stage encoding imposes a related limitation. Each case is represented by the same five stage slots within the 18-variable input vector. Repeating identical settings in adjacent slots can encode a scheme with fewer operational adjustments because no effective change occurs at the artificial boundary. This device does not, however, make the network a variable-length model. A design with more than five stages or a substantially different operating sequence lies outside the present input domain. Although operator preferences may determine the number and timing of stages in practice, the current LPGNN can compare only schemes that can be represented consistently within the five-slot format. Changing the number of stages at inference would therefore constitute unsupported extrapolation.

A broader model would require data in which both geological conditions and stage sequences vary. One option would be to represent each stage as an ordered record of tube positions, duration, and flow rate and to process the sequence with a masked recurrent or attention-based encoder. The resulting sequence representation could feed the existing volume and radius heads, while global process descriptors and physical-consistency terms could be calculated from the complete sequence. Geological descriptors could enter

through a separate branch and condition the sequence representation. Equipment limits, admissible stage counts, and scheduling preferences could then be handled explicitly as conditional inputs or optimization constraints rather than being embedded implicitly in a fixed input format.

4. Conclusions

A lightweight physics-guided framework was developed for the preliminary assessment of solution-mined salt-cavern construction schemes. Final cavern volume and maximum radius were selected as complementary design responses. In the proposed formulation, the source-defined cavern volume coefficient was retained as a soft consistency condition rather than an additional prediction target. This preserved the practical priority of the two main design responses while discouraging incompatible joint predictions.

The staged construction process was represented by 13 independent stage-wise variables and five nonredundant process–geometric descriptors of duration variability, mean depth configuration, and staged depth adjustment. Compared with the model using original stage-wise variables, the inclusion of process–geometric features changed the cavern-volume and maximum-radius MAPEs from 1.9675% and 3.3048% to 1.4779% and 2.8746%, respectively. This result indicates that the temporal and spatial organization of the leaching process provides useful process-oriented information for interpreting final cavern geometry.

Accurate and repeatable predictions were obtained by the optimized LPGNN. On the independent test subset, MAPEs of 1.4779% for cavern volume and 2.8746% for maximum radius were achieved, with corresponding R^2 values of 0.9506 and 0.8602. Better joint accuracy was obtained than that of the pure data-driven models, and other physics-guided candidates. Repeated training under different initializations also produced limited variation, suggesting that the prediction performance was stable under the adopted modeling protocol.

Different geotechnical controls were identified for the two target responses. Cavern volume was mainly associated with the overall leaching scale and vertical-control configuration, especially the Stage-2 and Stage-3 blanket-pad positions and the redistribution of dissolution time among stages. Maximum radius was more sensitive to the migration and local persistence of the active dissolution zone, as reflected by the inner-tube depth range and selected later-stage tube positions and durations. These differences suggest that cavern volume and localized radial enlargement should be assessed jointly rather than optimized separately during salt-cavern construction design.

Overall, the LPGNN provides a physics-guided and computationally efficient surrogate for comparing construction schemes within the fixed geological and operational domain represented by the source simulations. It is intended for preliminary screening before detailed leaching simulation and site-specific geomechanical stability analysis.

Author Contributions: Conceptualization, W.L., Z.S. and D.X.; methodology, W.L., Z.S., D.X., T.C. and J.M.; software, Z.S. and D.X.; validation, T.C. and J.M.; formal analysis, W.L., Z.S., D.X., T.C. and J.M.; supervision, Z.S. and D.X.; writing—original draft preparation, W.L., Z.S., D.X., T.C. and J.M.; writing—review and editing, Z.S. and D.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of Zhejiang Huadong Geotechnical Investigation & Design Institute Co., Ltd. (ZKY2025-HDYT-02-14), the National Natural Science Foundation of China (Grant Nos. 42577167 and 42177147), and the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) (CUG2642022006).

Data Availability Statement: The source simulation dataset is openly available in <https://doi.org/10.1016/j.energy.2022.124238>. The processed feature table, split indices, trained-model outputs, and code used in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: Authors Wanli Liu, Zhigang Shan and Ding Xia were employed by Powerchina Huadong Engineering Corporation (China). 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 conflict of interest. The authors declare that this study received funding from Zhejiang Huadong Geotechnical Investigation & Design Institute Co., Ltd. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

References

1. Jiang, W.; Huang, L.; Zhang, S.; Ma, H. Advances in research on geological evaluation of compressed air energy storage in underground gas storage facilities. *Coal Geol. Explor.* **2025**, *53*, 62–75. [CrossRef]
2. Zhang, G.; Wang, X.; Xiang, Y.; Pan, J.; Xiong, F.; Hua, D.; Tang, Z. Compressed air energy storage in hard rock caverns: Airtight performance, thermomechanical behavior and stability. *Chin. J. Rock Mech. Eng.* **2024**, *43*, 2601–2626. [CrossRef]
3. Ozarslan, A. Large-scale hydrogen energy storage in salt caverns. *Int. J. Hydrogen Energy* **2012**, *37*, 14265–14277. [CrossRef]
4. Li, H.; Ma, H.; Yang, C.; Wang, X.; Daemen, J.J.K. Creep-fatigue model of rock salt based on a fractal-order derivative considering thermo-mechanical damage. *Comput. Geotech.* **2024**, *174*, 106605. [CrossRef]
5. Yin, D.; Yu, Z.; Chen, S.; Zhang, J.; Qu, X.; Xue, S. Development and experimental study of a rock testing system under cyclic pressurization-storage-depressurization-storage gas disturbance. *J. China Coal Soc.* **2026**, 1–19. [CrossRef]
6. Zhao, K.; Yang, C.; Ma, H.; Daemen, J.J.K. A creep-fatigue model of rock salt and its application to the deformation analysis of CAES salt caverns. *Comput. Geotech.* **2023**, *156*, 105311. [CrossRef]
7. Li, Z.; Wang, Y.; Fan, J.; Zou, Y.; Chen, J.; Jiang, D.; Nelias, D. Deformation of surrounding rocks in salt caverns in compressed air energy storage plants: Insights from laboratory experiments and numerical simulations. *Tunn. Undergr. Space Technol.* **2026**, *170*, 107335. [CrossRef]
8. Li, J.; Wang, Z.; Zhang, S.; Shi, X.; Xu, W.; Zhuang, D.; Liu, J.; Li, Q.; Chen, Y. Machine-learning-based capacity prediction and construction parameter optimization for energy storage salt caverns. *Energy* **2022**, *254*, 124238. [CrossRef]
9. Dai, X. Optimization and Application of Cavern Construction Process Parameters for the Jintan Salt-Cavern Gas Storage. Master's Thesis, China University of Petroleum (East China), Qingdao, China, 2016.
10. Wang, J. Study on Cavern-Formation Laws and Parameter Optimization for Dual-Vertical-Well Solution Mining in Salt-Cavern Gas Storage. Master's Thesis, China University of Petroleum-Beijing, Beijing, China, 2025.
11. Wang, J. Study on the Coupled Model and Evolution Laws of Horizontal Solution-Mined Caverns for Salt-Cavern Gas Storage. Ph.D. Thesis, China University of Petroleum-Beijing, Beijing, China, 2025.
12. Li, J.; Xu, W.; Zheng, J.; Liu, W.; Shi, X.; Yang, C. Modeling the mining of energy storage salt caverns using a structural dynamic mesh. *Energy* **2020**, *193*, 116730. [CrossRef]
13. Li, J.; Shi, X.; Yang, C.; Li, Y.; Wang, T.; Ma, H. Mathematical model of salt cavern leaching for gas storage in high-insoluble salt formations. *Sci. Rep.* **2018**, *8*, 372. [CrossRef] [PubMed]
14. Li, J.; Liu, J.; Li, Q.; Xu, W.; Zhuang, D. Advances in mechanism and modeling method for solution mining of salt cavern storage. *Hazard. Control Tunn. Undergr. Eng.* **2020**, *2*, 1–8. [CrossRef]
15. Yi, W.; Li, Q.; Zhao, X.; Liu, W.; Du, J. Feasibility assessment of solution mining and gas storage in salt caverns: A case study of the Sanshui salt mine. *Front. Energy Res.* **2023**, *11*, 1301765. [CrossRef]
16. Cao, Z. Research on Intelligent Construction Technology for Underground Gas Storage Caverns in Compressed Air Energy Storage Power Stations. Master's Thesis, North China Electric Power University, Beijing, China, 2025.
17. Wang, Y. Study on the Stability of Salt-Cavern Gas Storage Based on a Database System. Ph.D. Thesis, Southwest Petroleum University, Chengdu, China, 2019.
18. Wang, Z.; Chen, J.; Wang, G.; Li, J.; Li, S.; Azhar, M.U.; Ma, S.; Xu, W.; Zhuang, D.; Zhan, L.; et al. Geometry prediction and design for energy storage salt caverns using artificial neural network. *Energy* **2024**, *308*, 132820. [CrossRef]
19. Zhao, K.; Yu, S.; Wong, L.N.Y. A long short-term memory-based model in forecasting the displacement of compressed air energy storage salt caverns. *J. Energy Storage* **2025**, *133*, 117962. [CrossRef]
20. Zhao, K.; Yu, S.; Wong, L.N.Y. Long-term stability forecasting for energy storage salt caverns using deep learning-based model. *Energy* **2025**, *319*, 134854. [CrossRef]
21. Chen, S. Study on Structural Optimization Design and Long-Term Stability of Underground Gas Storage Caverns. Master's Thesis, Shijiazhuang Tiedao University, Shijiazhuang, China, 2025.

22. Ge, X.; Huang, J.; Zhao, T.; Tao, G.; Ma, H.; Wang, W. A review of the application of artificial intelligence in underground engineering for compressed air energy storage. *Rock Soil Mech.* **2026**, *47*, 413–425. [[CrossRef](#)]
23. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [[CrossRef](#)]
24. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [[CrossRef](#)]
25. Willard, J.; Jia, X.; Xu, S.; Steinbach, M.; Kumar, V. Integrating scientific knowledge with machine learning for engineering and environmental systems. *ACM Comput. Surv.* **2023**, *55*, 1–37. [[CrossRef](#)]
26. Zhang, S.; Wang, L.; Zhu, S.-P.; Deng, X.; Fu, S.; Luo, C.; Dong, Y.; Yan, D. Physics-informed neural network for creep-fatigue life prediction of Inconel 617 and interpretation of influencing factors. *Mater. Des.* **2024**, *245*, 113267. [[CrossRef](#)]
27. Yang, H.; Hu, G.; Liu, L.; Li, Y.; Deng, Y.; Wu, J. A physics-informed and SHAP-enhanced modeling framework for predicting strength of cement-stabilized soil. *Case Stud. Constr. Mater.* **2025**, *23*, e05280. [[CrossRef](#)]
28. Mantel, N. The detection of disease clustering and a generalized regression approach. *Cancer Res.* **1967**, *27*, 209–220. [[PubMed](#)]
29. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2017.
30. Goldstein, A.; Kapelner, A.; Bleich, J.; Pitkin, E. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *J. Comput. Graph. Stat.* **2015**, *24*, 44–65. [[CrossRef](#)]
31. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015*; PMLR: Lille, France, 2015; pp. 448–456.
32. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
33. Caruana, R. Multitask learning. *Mach. Learn.* **1997**, *28*, 41–75. [[CrossRef](#)]
34. Huber, P.J. Robust estimation of a location parameter. *Ann. Math. Stat.* **1964**, *35*, 73–101. [[CrossRef](#)]
35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
36. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
37. Kennedy, J.; Eberhart, R. Particle swarm optimization. In *Proceedings of ICNN'95—International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995*; IEEE: Perth, WA, Australia, 1995; pp. 1942–1948. [[CrossRef](#)]
38. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2018.
39. 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, San Francisco, CA, USA, 13–17 August 2016*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794. [[CrossRef](#)]
40. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2017.
41. Lei, D.; Ma, J.; Zhang, Z.; Jiang, Z.; Zhang, G.; Wei, Q.; Yi, W.; Chen, T. Physical and data hybrid-driven modeling for the evolution-state prediction of reservoir colluvial landslides with step-like movement patterns: A case study in the Xiluodu Reservoir Area, China. *Bull. Eng. Geol. Environ.* **2026**, *85*, 499. [[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.