Tunnelling for a Better Life – Yan, Celestino, Thewes & Eberhardt (Eds)

Leakage prediction and post-grouting assessment in headrace tunnel of a hydropower project

Tek Bahadur Katuwal*
Norwegian University of Science and Technology (NTNU), Trondheim, Norway
Tribhuvan University, IOE, Pashchimanchal Campus, Pokhara, Nepal

Krishna Kanta Panthi
Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Chhatra Bahadur Basnet
Tribhuvan University, IOE, Pashchimanchal Campus, Pokhara, Nepal

Sailesh Adhikari
Norwegian University of Science and Technology (NTNU), Trondheim, Norway
Tribhuvan University, IOE, Pashchimanchal Campus, Pokhara, Nepal

ABSTRACT: In the Himalayan region, tunnels are often constructed through complex and varying geological formations having rock mass with higher degree of jointing, faulting, folding, and weakness/shear zones. Such rock mass condition significantly increases the rock mass permeability which enables a higher possibility of water leakage into and out of the headrace tunnels built for hydropower projects and is a challenging situation for tunnel stability. Therefore, comprehensive leakage assessment and effective pre- and post-grouting application are essential in hydropower tunnels. In this research, the water leakage was predicted by using three machine learning approaches such as Support Vector Regression (SVR), Decision Tree (DT) regression, and K-Nearest Neighbors (KNN) models. The water leakage/inflow was predicted in one of the hydropower tunnels based on the geological condition of rock mass, rock mass quality, and hydro-geological conditions. The effective post-grouting method was applied to mitigate the potential water leakage and to enhance the rock mass quality and stability of the hydropower tunnel. It was observed that the injection grouting technique helps to make tunnels less permeable, reduces instability conditions, and ensures the long-term safety and structural integrity of the hydropower tunnels.

Keywords: Complex geology, Machine Learning, Leakage, Stability, Post-grouting

1 INTRODUCTION

The tunnel construction in the Himalayan region faces severe tunneling challenges due to the complex geological conditions characterized by a higher degree of jointing, high degree of faulting, folding, weathering, and the presence of weakness/shear zones within the rock mass (Panthi 2006). These complex geological conditions are due to the consequence of active tectonic activities in the region. These conditions significantly increase rock mass permeability and make headrace tunnels of hydropower projects susceptible to water leakage. Thus, groundwater inflow and leakage are the most common and challenging issues in hydropower tunnelling projects in the Himalayan region (Panthi 2006). Water ingress into these tunnels not only affects the stability of the tunnel but also poses operational and safety risks, which could lead to project delay and huge economic loss (Panthi and Nilsen 2010). Therefore, an accurate water inflow/leakage assessment is essential for timely measure and control of water inflow/leakage, and the selection of appropriate grouting techniques to enhance the strength of rock mass and to secure stability of water tunnels.

Many researchers highlight the use of analytical, empirical, and numerical techniques and theoretical frameworks for the prediction of water ingress and grouting capability. However, it is challenging to predict the accurate amount in tunnelling projects (Holmøy 2008 and Stille 2015). To mitigate these limitations, machine learning (ML) techniques could be an appropriate approach to predict the water inflow/leakage and to decide effective grouting

*Corresponding author: tek.b.katuwal@ntnu.no

DOI: 10.1201/9781003495505-404

measures. However, this is challenging due to the limited availability of field data related to rock mass properties, structural geology, groundwater level, project topography, and effective stresses.

This paper attempts to present different machine learning (ML) approaches for comprehensive assessment and prediction of water inflow/leakage, and the application of injection grouting to reduce water leakage and enhance tunnel stability in the headrace tunnel of 54 MW Super Dordi Hydropower Project (SDHEP).

2 PROJECT BACKGROUND

2.1 *Brief about SDHEP*

The Super Dordi Hydropower Project "Kha" (hereafter referred to as SDHEP) is a Run-of-River (ROR) project, which is located at Dordi Rural Municipality, Lamjung District, Gandaki Province of Nepal, which is about 188 km northwest of Kathmandu. Geographically, this project area is located between Longitudes 84°34'15" E and 84°31'00" E, and Latitudes 28° 18'43" N and 28°16'20" N. The project is a high-head scheme with a gross head of 638 m and has a design discharge is 9.9 m^3/sec. The total installed capacity of this project is 54 MW. The main components of this project are low-head diversion dam (18.5 m X 3.3 m), D – shaped gravel trap (36 m X 4 m X 2.3-4 m), double-chamber desander basins (123 m X 11 m X 13 m), D – shaped headrace tunnel (2.8 m Dia and 5.2 km length), surge shaft (6 m Dia. and 49 m height), pressure shafts (2.2 m dia. and 1052 m long including penstock branches to two turbines), underground powerhouse (51.3 m X 15 m X 29.6 m), tailrace tunnel, and access tunnel (PHCPL 2022). In order to address the challenges associated with frequently occurring high tectonic activity and the risk of landslides, and to ensure long-term stability, the major water conveyance structures such as the desander, headrace tunnel, surge shaft, penstock shaft, powerhouse cavern, and tailrace tunnel are strategically constructed underground.

2.2 *Project geology*

Geologically, the SDHEP area is located in the Higher Himalayan range. The major rock types are schist and gneiss as shown in Figure 1. Predominantly, the geological composition of the area consists of banded gneiss with schist partings (Adhikari et al. 2023, Katuwal et al. 2023). These geological formations extensively appear in the Higher Himalayan region, spanning from east to west across the entire Himalayan range.

The upper part of the headrace tunnel (HRT) encountered the slightly to moderately weathered and medium foliated strong gneiss. A relatively short stretch of the tunnel section encountered fresh to slightly weathered gneiss, which is more dominant with quartz content. Likewise, the downstream section of HRT, surge shaft, penstock shaft, and powerhouse cavern encountered the slightly to moderately weathered and medium foliated strong gneiss. Mainly, the bended gneiss with occasional schist parting is encountered along the alignment of the headrace tunnel. The topography of the project area is covered by mainly two soil types, i.e., colluvial and alluvial soil. The colluvial soil consists high content of fine silt and clay which is exposed on the surface of the HRT alignment. Likewise, the surface of powerhouse and headworks are exposed to alluvial soil which consists of predominantly coarse-grained material such as sand, gravel, cobble, and boulder.

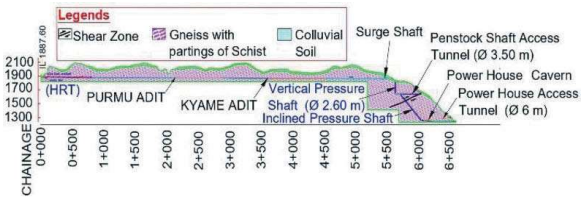

Figure 1. Longitudinal Profile of SDHEP.

3 MACHINE LEARNING TECHNIQUES

Nowadays, machine learning (ML) techniques have been widely applicable to predict water ingress in tunnel projects. The ML approach shows good performance in predicting water inflow/leakage in tunnel projects (Mahmoodzadeh et al. 2023). Therefore, in this paper, three ML regression models such as Support Vector Regression (SVR), Decision Tree (DT) regressor, and K-Nearest Neighbors (KNN) are used to predict the water inflow/leakage at the SDHEP headrace tunnel. These regression models help to predict the dependent variables from several features or independent variables. The most appropriate regression model has been established by comparing the eight statistical indices such as R-squared (R^2), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Relative Root Mean Squared Error (RRMSE), Mean Absolute Percentage Error (MAPE), Mean Relative Error (MRE), and Variance Accounted For (VAF). For this purpose, the research methodology is set up mainly with four components such as dataset preparation, model selection, training and validation of selected model, and result analysis and discussions.

3.1 *Database study*

To predict the water ingress/leakage in the hydropower tunnels, it is necessary to collect the filed data and select the effective parameters. Therefore, in this case study, the number of datasets and types of effective parameters such as Rock Quality Designation (RQD), Joint number (Jn), Joint Roughness (Jr), Joint alteration number (Ja), Q-Classification value (Q), static head (h), overburden (H), shortest distance from the tunnel to topography (d), and specific discharge (q),

average Lugeon Unit (LU) are collected based on actual data received from the project. The summary of the field datasets is presented in Tables 1 and 2, which are used for training and testing parameters. These tables depict the statistical description of selected parameters by defining the minimum, standard deviation, maximum, mean values, and different percentile values.

Table 1. Field datasets for rock mass quality.

	RQD	Jn	Jr	Ja	Q
Count	175	175	175	175	175
Mean	62.17	6.19	1.57	1.55	7.77
Std	9.52	1.17	0.3	0.55	5.69
Min	30	3	1	1	1.1
25%	55	6	1.5	1	3.33
50%	65	6	1.5	2	6
75%	70	6	1.5	2	10.24
Max	80	9	3	3	30

Table 2. Field datasets for topography and water inflow.

	h (m)	d (m)	H (m)	q (lit/min/m)	LU
Count	175	175	175	175	175
Mean	11.3	139.35	173.5	3.52	1.61
Std	7.86	20.53	21.7	6.21	2.82
Min	0.19	63	120	1.28	0.56
25%	3.39	128.5	163	1.95	0.91
50%	14.61	138	170	2.26	1.05
75%	19.96	150	179	2.88	1.3
Max	20.56	191	231	59.51	27.3

3.1.1 *Correlation analysis*

In machine learning, the correlation heatmap is often used to establish and visualize the strength of relationship (multicollinearity) between the multiple feature variables and dependent variables. The accuracy of the predicted model is significantly influenced by the multicollinearity numerical values. In this heat map, Color-coding cells make it easy to see how variables are related with a quick look.

In this research, the correlation heatmap analysis (using the Pearson correlation coefficient) was performed between collected database from the head-race tunnel. This process illustrates the interrelation between independent features and dependent variable. The range of this analysis lies within [-1, 1]. A high value of [+1] means these variables are highly correlated and have a high positive effect and vice versa. In Figure 2, the positive correlation value of 0.62 is seen between the RQD and Q-value, which is remarkable since an increase in the RQD value will give an increased quality of rock mass. The Figure also depicts that the rate of water ingress/leakage positively correlated with the overburden, and joint alteration number. On the other hand, parameters such as RQD, Q-value, and joint roughness are inversely correlated. Moreover, Figure 2 indicates that the correlation between the available datasets seems relatively low. This may be due to the fact that the water ingress and grout consumption are also influenced by variables such as persistence, joint volume, spacing of joints, and infilling material characteristics.

Figure 2. Correlation matrix for database.

3.1.2 *Principal Component Analysis (PCA)*

The PCA is a multivariate statistical approach that is used for the dimensionality reduction of project databases. For this purpose, the principal components have been established to describe the linear combination of features and to define the variation in the selected database. Figure 3 illustrates the screen plot of principal components for the ingress/leakage of water. This Figure clearly explains that 90% of selected database variance is explained by the six major principal components. The variability of each variable in these principal components are presented in Table 3.

Table 3. A Summary of principal components coefficients.

Variables	PC1	PC2	PC3	PC4	PC5	PC6
h	-0.27	-0.15	-0.44	-0.35	0.29	-0.54
H	0.15	0.04	-0.58	0.5	0.41	0.13
RQD	-0.4	0.23	0.31	0.01	0.23	0.45
Jn	0.19	-0.28	0.55	-0.13	0.53	-0.32
Jr	-0.21	0.62	-0.03	-0.2	-0.29	-0.34
Ja	0.16	0.34	0.26	0.63	0	-0.49
Q	-0.39	0.37	0.02	-0.01	0.54	0.07
LU	0.5	0.32	-0.06	-0.28	0.16	0.09
q	0.49	0.33	-0.06	-0.3	0.15	0.13
Eigenvalues	2.89	1.50	1.34	1.11	0.82	0.78
Variability (%)	31.88	16.52	14.78	12.26	9.08	8.57
Cumulative (%)	31.88	48.40	63.18	75.77	84.53	93.10

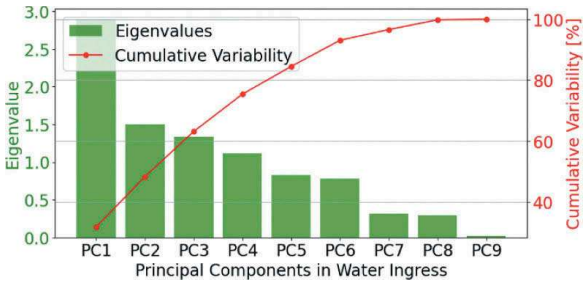

Figure 3. Screen plot for Principal Components.

3.1.3 *Variance Inflation Factor (VIF)*

Dimension of the given data set is primarily reduced by the PCA analysis without quantifying the multicollinearity. To address this limitation, the Variance Inflation Factor (VIF) has been applied. The correlation strength between the independent variables has been established by using the VIF. Moreover, the VIF can provide information about the severity and presence of multicollinearity in the databases, which is illustrated in Table 4. The VIF = 1 indicates no multicollinearity between the selected variables. When the value of the VIF is greater than 10, it is a sign of a high degree of multicollinearity between the variables and it will be problematic in the regression models, therefore, eliminating these features will be easier for regression analysis. However, the selection of independent variables depends upon the selection of the regression model (Cheng et al. 2022). In this research, all parameters are considered for the establishment of a regression model.

Table 4. A summary of Variance Inflation Factor (VIF).

For q	H	h	RQD	Jn	Jr	Ja	Q	LU
VIF	45.9	4.3	65.3	25.1	31.5	11.8	4.8	4.9

3.2 *Data distribution*

The selected datasets have been organized into distinct plots to enable quick visualization and enhance understanding. To accomplish this, box plots were initially generated for the chosen datasets. It was observed that the Jn, Jr, and Ja datasets displayed limited data variability and were not well-distributed. To address this limitation, the plot performance has been enhanced by incorporating both box and histogram plots for the selected datasets as shown in Figure 4. Histograms display how data is distributed in terms of shape and frequency. In contrast, box plots are useful for a rapid assessment of central tendencies, data spread, and the presence of outliers.

In Figure 4, except the box plot data of h, and RQD, the data are out of the middle of the box. Therefore, to address this issue and combine both of these distribution representations into a single plot,

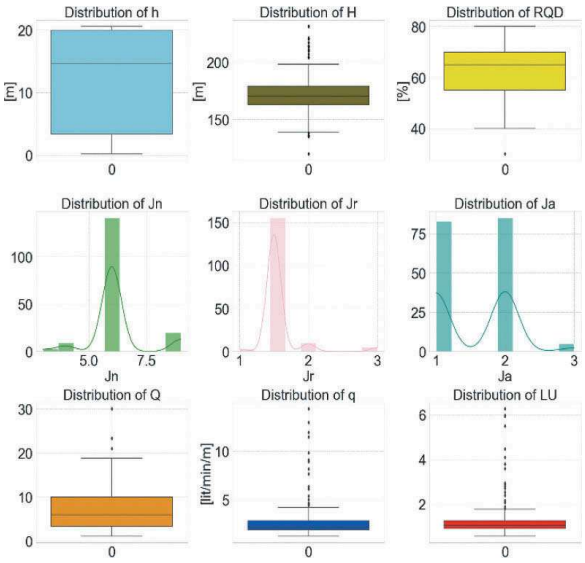

Figure 4. Multivariate data visualization of selected parameters.

the violin plot is found to be a more effective choice for visualizing of data that has multiple peaks or a skewed distribution. Likewise, a violin plot is the configuration of the data set that enhances clarity and understanding. Moreover, the violin plots illustrate the summary of statistics and probability density function (or density) of each feature and dependent variables. Figure 5 illustrates the distribution of the statical quartile summary of filed data set. In these Figures, a wider region of density plot indicates the more frequent occurrence and vice versa.

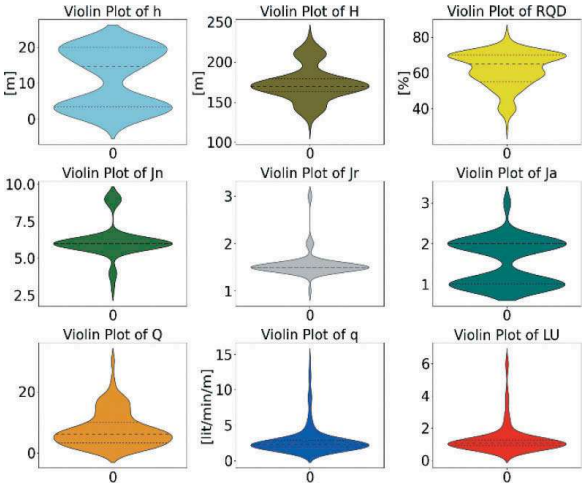

Figure 5. Datasets distribution in violin plot.

3.3 *Data normalization*

Data normalization constitutes a fundamental process within the domain of machine learning (ML). This procedure is essential due to the inherent variability in the dimensions and units of input data,

which can fluctuate with changing input data. The act of normalization serves to rescale all input features and variables to a uniform scale. Table 1 and 2 depicts that the range of specific discharge varies from 1.28 lit/min/m to 59.51 lit/min/m, whereas the joint alteration number (Ja) varies from 1 to 3 and the height of overburden varies from 120 m to 131 m. This indicates that the magnitude of input variables are in different scales. Therefore, data normalization should be conducted for better correlation and prediction of machine learning models.

3.4 *Statistical analysis of selected model*

Statistical evaluation of selected model plays a crucial role in understanding and predicting the efficient model for headrace tunnels of hydropower projects. Therefore, various statistical indices such as R-squared (R^2), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Relative Root Mean Squared Error (RRMSE), Mean Absolute Percentage Error (MAPE), Mean Relative Error (MRE), and Variance Accounted For (VAF) are calculated by using Equation 1 to Equation 8, respectively. The linear correlation between the predicted and actual values is established by using R^2.

$$R^2 = 1 - \frac{\text{sum of squared regression (SSR)}}{\text{sum of squared total (SST)}} \quad (1)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i^a - y_i^p| \quad (2)$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} \left(y_i^a - y_i^p\right)^2 \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(y_i^a - y_i^p\right)^2} \quad (4)$$

$$\text{RRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(\frac{y_i^a - y_i^p}{y_i^a}\right)^2} \quad (5)$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i^a - y_i^p}{y_i^a}\right| * 100\% \quad (6)$$

$$\text{MRE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i^a - y_i^p}{y_i^a}\right| \quad (7)$$

$$\text{VAF} = 1 - \left|\frac{va\ \left(y_i^a - y_i^p\right)}{var\ \left(y_i^a\right)}\right| * 100\% \quad (8)$$

Where, actual and predicted values of variables are represented by y_i^a and y_i^p respectively, and n in the equations is the total number of datasets that are used in selected machine learning models.

4 WATER INFLOW PREDICTION MODEL AND RESULTS

In this study, Support Vector Regression (SVR), Decision Tress (DT), and K-Nearest Neighbors (KNN) models are applied for the assessment and prediction of water ingress in the headrace tunnel of SDHEP in Anaconda version 3.6 with computation in Python.

4.1 *Support Vector Regression (SVR)*

Support Vector Regression (SVR) is a machine-learning technique specifically adapted for predicting water ingress in the headrace tunnel. It connects the capabilities of support vector machines to model and predict water ingress into hydropower tunnels, which supports rock engineers to make well-informed choices for effective management and mitigation of water ingress challenges. The Figure 6 illustrates the water inflow prediction by using this SVR model.

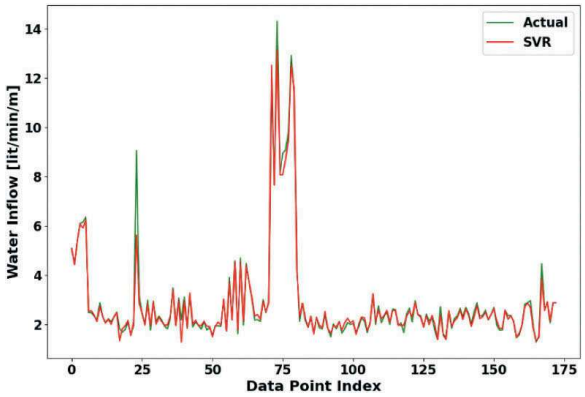

Figure 6. SVR model for water inflow prediction.

As Figure 6 indicates, there is a good correlation between the model and different dependent variables. Table 5 provides information about the statistical indices R^2, MAE, MSE, RMSE, RRMSE, MAPE, MRE, and VAF which are evaluated as 0.99, 0.09, 0.02, 0.12, 0.05, 3.47, 0.48, and 99.6, respectively, where the SVR model indicates high degree of accuracy. Therefore, this ML model has a good capacity for the prediction of water leakage/inflow in hydropower tunnels.

Table 5. A summary of statistical indices of SVR model.

R2	MAE	MSE	RMSE	RRMSE	MAPE	MRE	VAF
0.99	0.09	0.02	0.12	0.05	3.47	0.48	99.6

4.2 *Decision Tree (DT)*

A Decision Tree (DT) is the most widely used predictive model for regression analysis in the context

of hydropower tunnels to forecast and assess the probability of water ingress. It's a visual representation of a decision-making process that considers various factors and their interactions to determine the probability of water inflow/leakage in/from the tunnels. This tool helps rock/tunnel engineers make decisions to prevent and manage water ingress effectively, ensuring the safety and efficiency of hydropower operations.

The Figure 7 illustrates the comparison between the actual field data set and the water inflow prediction by using this DT model. Likewise, Table 6 establishes the statistical indices R^2, MAE, MSE, RMSE, RRMSE, MAPE, MRE, and VAF which are evaluated as 0.97, 0.14, 0.08, 0.29, 0.10, 4.67, 0.98, and 97.6, respectively, which indicate that the DT models have a good capacity for the prediction of water inflow. The statical indices indicated that this model presents a good correlation with the features and dependent variables.

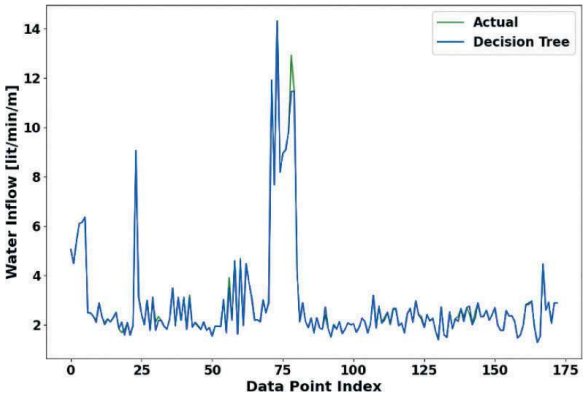

Figure 7. DT model for water inflow prediction.

Table 6. A summary of statistical indices of SVR model.

R2	MAE	MSE	RMSE	RRMSE	MAPE	MRE	VAF
0.97	0.14	0.08	0.29	0.10	4.67	0.98	97.6

4.3 *K-Nearest Neighbors (KNN)*

A K-Nearest Neighbors (KNN) is the most widely used predictive model for regression. Figure 8 illustrates the comparison between the actual field data set and the KNN prediction model for water inflow in the case tunnel project. Table 7 demonstrates the statistical indices R^2, MAE, MSE, RMSE, RRMSE, MAPE, MRE, and VAF which are evaluated as 0.89, 0.45, 0.38, 0.61, 0.22, 17, -1.22, and 88.9, respectively. The statical indices indicated that this model presents a good correlation with the features and dependent variables. All outcomes indicate that the KNN models have a good capacity and are acceptable for the prediction of water inflow in hydropower tunnels.

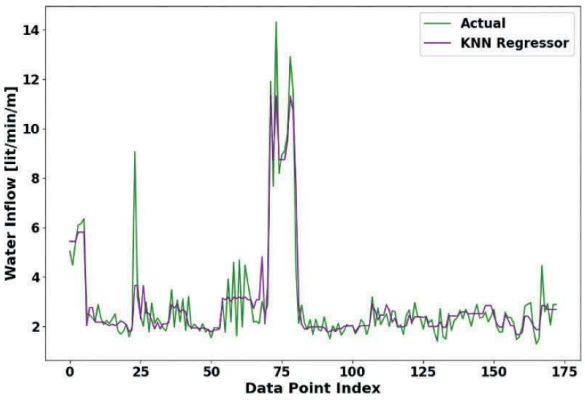

Figure 8. KNN model for water inflow prediction.

Table 7. A summary of statistical indices of SVR model.

R2	MAE	MSE	RMSE	RRMSE	MAPE	MRE	VAF
0.89	0.45	0.38	0.61	0.22	17	-1.22	88.9

5 GROUT CONSUMPTION

Post-grouting technique is very often used to control water leakage from headrace tunnels of hydropower projects in the Himalayan region. The measurement of the Lugeon value, which assesses rock permeability, plays a pivotal role to evaluate the effectiveness of post-grouting. At SDHEP, the post-grouting technique was used to control potential water leakage from the headrace tunnel. The post-grouting technique involves the injection of cement grout into the surrounding rock mass to achieve substantial sealing of joints and discontinuities and reduce permeability. A typical drill hole pattern adopted while post-grouting is illustrated in Figure 9. This Figure demonstrates the drill hole length, and drilling pattern in both the hill and valley sides of the headrace tunnel are shown. Likewise, a detailed description of the parameters associated to post-grouting is illustrated in Table 8.

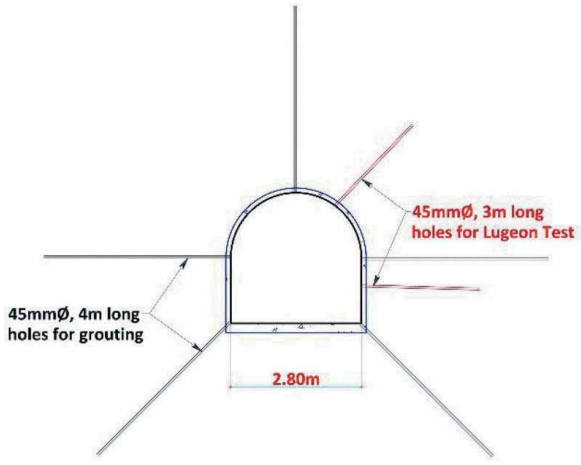

Figure 9. Detail illustration of drill hole arrangement.

At SDHEP the post-grouting not only enhanced the structural integrity of the rock mass but also established a water-resistant barrier, thereby minimizing the risk of water leakage from the tunnel. Consequently, the combination of post-grouting and the reduction in Lugeon value ensured the long-term structural integrity and operational reliability of the water tunnel.

Table 8. Parameter for Injection Grouting.

Parameters	Values
Drill hole length (m)	4
Drill hole diameter (mm)	45-48
Packer diameter (mm)	42
Placement of Packer inside (m)	2
Minimum Grouting Pressure (bar)	4-6
Grout pressure at seepage area (bar)	15
Bentonite (%)	5
Plasticizer (%)	1.5-3
Retarder (if mixed outside) (%)	1.5
W/C ratio	1:1

5.1 *Grout consumption assessment*

The effective post-grouting work is very challenging since it is extremely difficult to achieve targeted pressure without losing unnecessary extra grout material. Therefore, the effective post-grouting work is a challenging task since the grout take is significantly affected by the rock mass properties, topographical condition, structural geology, joint infilling material, groundwater level and effective rock stress conditions. The summary of the field datasets is presented in Table 9, which illustrates the rock mass class (RMC) count and statistical variation of selected parameters that defines minimum, maximum, mean values, standard deviation, and different percentile values.

The pairwise data relationships between different parameters used in grout consumption assessment are illustrated in Figure 10. The Figure shows pairwise plot or scatterplot matrix. This plot helps to visualize and explore different independent and dependent variables. The plot can describe how different variables are relative to each other based on different types of rock mass classes. Moreover, the plot provides a quick overview of the relationship between grout consumption (GC), joint aperture (e), hydraulic

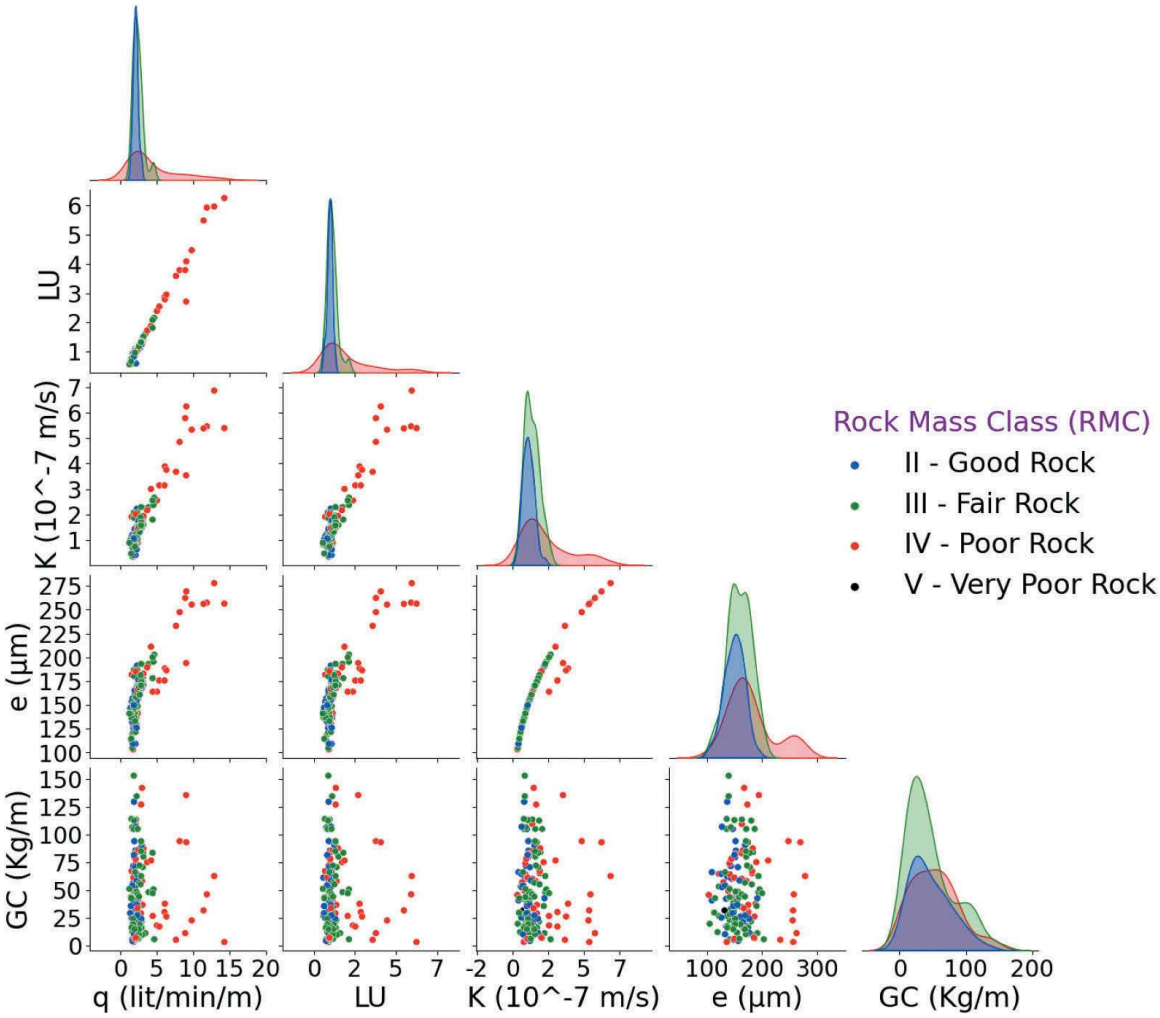

Figure 10. Grout consumption data relationship.

Table 9. Field Datasets for Grout Consumption.

	q (lit/min/m)	LU	K (10^-7 m/s)	e (μm)	GC (Kg/m)
Count	173				
RMC	II = 44, III = 78, IV = 50, V = 1				
Mean	2.88	1.32	1.57	162.04	47.91
Std	2.09	0.95	1.11	30.43	33.68
Min	1.28	0.56	0.35	103.25	3
25%	1.94	0.91	0.94	143.45	22.5
50%	2.24	1.05	1.28	158.78	41.5
75%	2.86	1.26	1.71	174.39	66
Max	14.31	6.26	6.86	277.92	153

permeability (K), Lugeon unit (LU), and water ingress (q) for rock mass class (RMC) II, III, IV, and V. For example, in the context of rock mass classified as Type IV, a high degree of fracture aperture can significantly increase the hydraulic permeability of the rock mass. As the Lugeon unit, a measure of permeability increases, it signifies a higher propensity for water ingress from the hydropower tunnel. Consequently, to mitigate this condition and ensure the stability of the tunnel, large quantities of grout material may be necessitated to seal the fractures and reduce water leakage. Therefore, the matrix is a valuable tool for gaining insights and is particularly useful when dealing with multivariate data, as it offers a quick and comprehensive overview of the relationships between variables, helping in the initial stages of data exploration.

The relationship between water leakage, Lugeon unit, and grout consumption is shown in Figure 11. Likewise, the relationship between hydraulic permeability, hydraulic aperture, and grout consumption is illustrated in Figure 12.

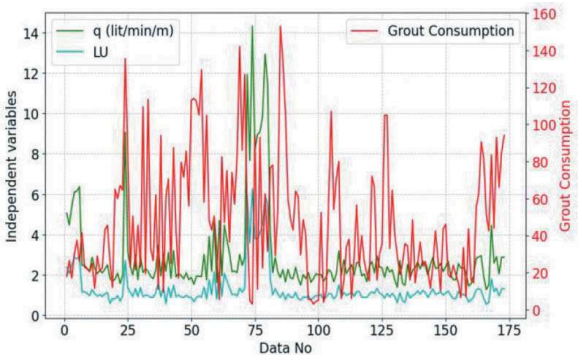

Figure 11. Relationship between water inflow, Lugeon unit, and grout consumption.

6 RESULT ANALYSIS AND DISCUSSION

The water leakage and grout consumption assessments were made on the headrace tunnel of SDHEP. The assessment was conducted using field datasets

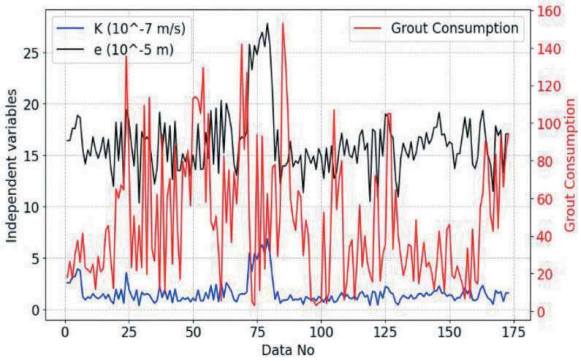

Figure 12. Relationship between hydraulic permeability, hydraulic aperture, and grout consumption.

related to the rock mass quality, hydro-geological conditions, Lugeon unit, and grout consumption.

6.1 *Result comparison of water inflow prediction*

Various machine learning models were employed to predict water leakage from the headrace tunnel. The prediction was then compared to actual datasets. All models demonstrated acceptable and good predictive performance individually. However, it is essential to compare and rank these models to enhance the reliability of selected models. To achieve this, statistical indices were used to evaluate and categorize the performance of each model as good, better, or best, with respective weightings of 1, 2, and 3. The overall ranking of each model was determined by summing assigned weights. Table 10 illustrates the comparative ranking of selected models. The table shows that the SVR, DT, and KNN models perform best, better, and good performance, respectively to predict the ingress of water.

Table 10. Comparison of Water inflow predicted models.

Parameter/Model		SVR	DT	KNN
R^2	Value	0.99	0.97	0.89
	Weightage	3	2	1
MAE	Value	0.09	0.14	0.45
	Weightage	3	2	1
MSE	Value	0.02	0.08	0.38
	Weightage	3	2	1
RMSE	Value	0.12	0.29	0.61
	Weightage	3	2	1
RRMSE	Value	0.05	0.10	0.22
	Weightage	3	2	1
MAPE	Value	3.47	4.67	17
	Weightage	3	2	1
MRE	Value	0.48	0.98	-1.22
	Weightage	3	2	1
VAF	Value	99.6	97.6	88.9
	Weightage	3	2	1
Remarks	Total Weitage	24	16	8
	Rank	1	2	3

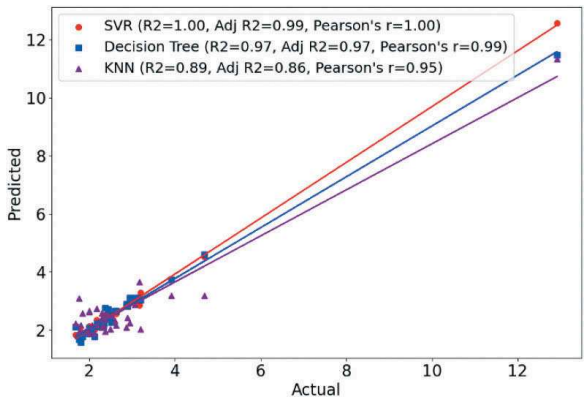

Figure 13. Actual and prediction results in comparisons.

The performance of the selected models is compared using statistical tools such as R-squared, Adjusted R-squared, and Pearson's r value. The actual test database and predicted results of the selected ML models are presented in Figure 13. In the Figure, SVR shows the strongest correlation between the predicted and actual databases, followed by the DT and KNN regression models. Moreover, Pearson's r measures how closely the data points in the scatterplot cluster around a straight line. The authors found that the SVR model quantifies a strong linear relationship between variables, followed by the DT and KNN models. However, all selected ML models establish a strong relationship between actual and predicted variables.

6.2 *Result comparison of grout consumption*

In the above section, the relationship between different independent variables such as water inflow, Lugeon unit, hydraulic permeability, and hydraulic aperture are plotted with grout consumption volume. These curves show fairly good correlations between these parameters in most of the chainage and rock mass conditions. In some parts, the results show unusual tendencies. However, after injection grouting the ingress of water is fully controlled and the headrace tunnel is well-functioning. Hence, the post-injection grouting was successful and has significantly reduced the water leakage potential and rock mass permeability and improved the rock mass strength, and enhanced tunnel stability.

7 CONCLUSION

A comprehensive overview of water ingress/leakage in the headrace tunnel of SDHEP was conducted based on the field data consisting of geology, hydrogeology, and injection grout parameters. After that, thorough data interpretation, statistical correlation, and regression analysis were performed using machine learning algorithms to predict water leakage from the headrace tunnel. It has been observed that machine learning techniques may be successfully used to assess the water ingress potential of the headrace tunnels of hydropower projects. Different machine learning regression models, such as Support Vector Regression (SVR), Decision Tree (DT), and K-Nearest Neighbors (KNN), have demonstrated the ability to predict water inflow/leakage in the water tunnel. As the field continues to evolve, the integration of machine learning in the management of water tunnels may prove to be a helpful tool for assessing tunnel stability challenges.

ACKNOWLEDGMENTS

This research was supported by NORHED II Project 70141 6; Capacity Enhancement in Rock and Tunnel Engineering at the Pashchimanchal Campus (WRC), Institute of Engineering (IoE), Tribhuvan University (TU), Nepal. The authors like to acknowledge NORAD, Norway for funding the project. In addition, the authors are thankful to the management of Peoples Hydropower Company (P) Ltd for providing access to map and collect necessary information and data related to SDHEP.

REFERENCES

Adhikari, S., Panthi, K. K. & Basnet, C. B. 2023. Stability issues associated with the construction of underground caverns of Super Dordi Hydropower Project, Nepal. 15th ISRM Congress 2023 & 72nd Geomechanics Colloquium, Schubert & Kluckner (eds.) © ÖGGAt: Salzburg Congress, Austria, pp. 1199–1204.

Cheng J., Sun J., Yao K., Xu M., & Cas Y. 2022. A variable selection method based on mutual information and variance inflation factor. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 268, 120652.

Holmøy, K. H. 2008. Significance of Geological Parameters for Predicting Water Inflow in Hard Rock Tunnels. Ph.D. Theses at NTNU 2008:291. Norwegian University of Science and Technology, Trondheim, Norway. ISBN 978-82-471-1284-7.

Katuwal, T. B., Panthi, K. K. & Basnet, C. B. 2023. Challenges associated with the construction of vertical and inclined shafts in the Himalayan Region. 15th ISRM Congress 2023 & 72nd Geomechanics Colloquium, Schubert & Kluckner (eds.) © ÖGGAt: Salzburg Congress, Austria, pp. 1199–1204.

Mahmoodzadeh A., Ghafourian H., Mohammed A. H., Rezaei N., Ibrahim H.H. & Rashidi S. 2023 Predicting tunnel water inflow using a machine learning-based solution to improve tunnel construction safety. Transportation Geotechnics, 40, 100978.

Panthi, K. K. 2006. Analysis of Engineering Geological Uncertainties Related to Tunnelling in HimalayanRock Mass Conditions. Ph.D. Theses at NTNU 2006:41. Norwegian University of Science and Technology. ISBN 82-471-7825-7.

Panthi, K.K., Nilsen, B. 2010. Uncertainty Analysis for Assessing Leakage Through Water Tunnels: A Case from Nepal Himalaya. Rock Mech Rock Eng 43, 629–639.

PHCPL, 2022. Progress Report of Super Dordi Hydropower Project. Peoples Hydropower Company (P) Ltd., pp.1–55.

Stille, H. 2015. Rock Grouting - Theories and Applications. BeFo – Rock Engineering Research Foundation. ISBN 978-91-637-7638-0.