


PRASHANTH RAGAM^{1*}, BONAGIRI DIVYA¹ **MODERN APPROACHES TO AIR OVERPRESSURE PREDICTION: EVALUATING ENSEMBLE MACHINE LEARNING AND CLASSICAL STATISTICAL MODELS**

Mining provides essential raw materials for various sectors but carries significant risks due to hazardous processes. Taking valuable minerals or other geological materials out of the earth is known as mining. Resources like coal and metals, placer, underground, and surface mining are essential, but they also have negative environmental effects, like air pollution from blasting and water pollution. Air noise, frequently caused by industrial operations like mining and construction, can harm wildlife and human health. Transportation, equipment, and blasting activities are examples of sources. To reduce the negative effects of high noise levels on the environment, stress, and hearing loss, noise management and predictive models are crucial by establishing correlations between variables such as charge weight, distance, and geological conditions. Statistical predictor equations calculate blast-induced Air Overpressure (AOp). In India, DGMS regulations ensure mining and blasting operations minimise environmental impacts and keep AOp levels safe for nearby communities. In this study, SVR, RF, GB, BPNN and an ensemble hybrid XGBoost-RF model were developed to predict blast-induced AOp and compared with traditional statistical prediction equations. The performance of these models was evaluated using four metrics: RMSE, MSE, MAE, and R^2 . The results showed high SVR of R^2 (0.9324), MSE (1.7591), MAE (1.3227), and RMSE (1.3263), GB of R^2 (0.9970), MSE (0.8586), MAE (0.5708), RMSE (0.9266), RF of R^2 (0.5267) MSE (17.7087), MAE (3.2990), RMSE (4.2081), BPNN of R^2 (0.9898), MSE (0.0025), MAE (0.0504), RMSE (0.0505), Ensemble hybrid model (XGBoost-RF) of R^2 (0.9991), MSE (0.2176), MAE (0.2828), and RMSE (0.4665), respectively demonstrate the effectiveness of these approaches. The R-squared values for the statistical models were as follows: USBM (0.44), NAASRA (0.3692), OLLOFSO (0.2093), HOLMBERGE (0.4325), MCKENZIE (0.1386), and the Generalised Equation (0.1386). These results highlight the comparatively lower accuracy of traditional statistical approaches in predicting blast-induced AOp compared to machine learning models.

Keywords: Blasting; AOp; Machine learning; R^2 ; RMSE; MSE

¹ VIT-AP UNIVERSITY, SCHOOL OF COMPUTER SCIENCE AND ENGINEERING, INAVOLU, BESIDE AP SECRETARIAT, AMARAVATI, ANDHRA PRADESH, 522237, INDIA

* Corresponding author: prashanth.rajam429@gmail.com



© 2025. The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0). The Journal license is: <https://creativecommons.org/licenses/by/4.0/deed.en>. This license allows others to distribute, remix, modify, and build upon the author's work, even commercially, as long as the original work is attributed to the author.

TABLE 1

Input Parameters Abbreviation

Input parameter	Abbreviation
S	Spacing
B	Burden
F	Rock hardness
PF	Powder Factor
D	Distance.
Q	Charge weight
HD	Hole Depth
ST	Stemming
MCD	Maximum Charge per Delay
HDI	Hole Diameter
STL	Stemming Length
RQD	Rock Quality Designation
DI	Distance of Instrument (DI)
MC	Maximum Charge
N	Number of Drill Holes (N)
CSD	Cylindrical Scaled Distance
ML	Maximum Load
LH	Largest Hole
SH	Smallest Hole
SD	Sub-Drilling
T	Temperature
DV	Detonation Velocity
NBHB	Number of Blast Holes per Blast
MIC	Maximum Instantaneous Charge (kg)
DMS	Distance between the Monitoring Station
BP	Blasting Point (m).
HL	Hole Length
MCWD	Maximum Charge Weight per Delay
DBC	Depth of Burial of Charge
TFR	Total charge Fired in a Round
DM	Distance of Measurement.
NOH	Number of Holes

TABLE 2

Machine Learning Full Forms

ML Algorithms	Abbreviation
1	2
PSO	Particle Swarm Optimisation
FOA	Fruit Fly Optimisation Algorithm
SGA	Simple Genetic Algorithm
ANN	Artificial Neural Network

1	2
RWGW0-ANN	Grey Wolf Optimisation- Artificial Neural Network
Sa-WKRR	Self-adaptive Weighted Kernel Ridge Regression
KRVFL	Random vector functional link
TKRR	Total knee Replacement
RVFL	Random vector functional link
edRVFL	Ensemble deep Random vector functional link
ANN	Artificial neural networks
SVM	Support vector machine
MARS	Multivariate Adaptive Regression Splines.
MVR	Multivariate regression
AI	Artificial Intelligence
RF	Random Forest
SVR	Support Vector Regression
MLP	Multi-Layer Perceptron
LSTM	Long Short-Term Memory
CGO-LSTM	Chaos Game Optimisation -Long Short-Term Memory
MLR	Multiple Linear Regression
GEP	Gene Expression Programming
GRP	General Routing Problem
SRP	Secure Remote Password
XGB-GWO	Extreme gradient boosting-gray wolf optimisation
BI-ENN	Bidirectional Encoder Neural Network.
BPNN	Back Propagation Neural Network
GMDH	Group Method of Data Handling
BA	Bat Algorithm
PSO-ANN	Particle Swarm Optimisation-Artificial Neural Network
CM	Conventional Models
NN	Neural Networks
GPE	Generalised Predictor Equations
DGMS	Directorate General of Mines Safety

1. Introduction

Due to the inherent risks present at each stage of the process, mining stands as one of the most demanding occupations within the industry [1]. Mineral resources serve as the foundation of any industrialised nation, providing essential raw materials for both metallic and non-metallic industries. These resources are extracted utilising a variety of methods, including underground mining and open-cast mining. According to Cheng and Huang (2006), only 20% of the group's explosive vitality is employed to shatter the bulk of the rock, while the remaining is a loss of energy. Due to ground vibration, fly rock, air noise, over break, and other effects. High ground vibration levels caused by blasting can lead to structural damage or failure [2]. The industrial growth of mineral-based industries has been driven by the demand for financial minerals [3]. The process of removing minerals from the surface of the Earth is called mining. The typical mine life cycle comprises four main phases: operations, development and planning, closure, and

exploration. Prospective mining sites are surveyed and mapped during the exploration phase to verify the presence of mineral deposits. The construction of necessary infrastructure, including processing facilities, roads, power lines, and administrative buildings, occurs during the development and planning phase. The actual extraction and processing of the minerals is the main focus of the operations phase [4]. The mining industry is intricate and encompasses various disciplines, operating within constantly changing environments. Artificial lighting and ventilation are necessary, various gases are emitted, dust is produced, and noise from blasting or rock breaking needs to be managed. Additionally, the industry involves navigating through diverse ergonomic positions and operating heavy machinery, vehicles, conveyors, and more. Mining operations encompass numerous apparent and concealed health and safety risks, thereby increasing overall risk. Identifying hazards (the origins of danger or risk factors) within this intricate system is challenging. Uncertainty is exacerbated by insufficient investigation and vague categorisation. Safety investigations acknowledge two forms of uncertainty: inherent variability within the system and gaps in knowledge [5]. Drilling and Blasting (DAB) represent the dominant methods for fragmenting stone. In the mining industry, explosives serve as an essential energy source for fracturing, excavating, and displacing rock masses. Upon detonation, a significant amount of energy is released. However, only around 20% of this energy can be effectively utilised for rock fragmentation, with the remainder dissipating as fly rock, ground tremor, air blast, and back breaks. The impacts of blasting-related air noise include potential damage to nearby ecosystems, such as collapsing trees and disruption to wildlife habitats. Moreover, there's an increased risk of failure in overburden dumps due to the vibrations caused by the blast. The increased air noise level has resulted in increased complaints from residents near mining operations, highlighting the necessity for effective monitoring and control measures to alleviate its impact [2]. Nonetheless, just 20-30% of the explosive energy is used for breakdown. At the same time, the substantial leftover dynamism during the charging process can lead to unwanted consequences such as flying boulders, vibrations, back breaks, over breaks, and loud noises that affect nearby communities [6]. Fig. 1 illustrates Classification and Impact of Blasting Effects in Mining Operations, illustrating blasting processes' various environmental and operational consequences. TABLE 1 and TABLE 2 represent machine learning algorithms and a list of shorthand input parameters, providing comprehensive information for analysis and comparison in the study.

Air Overpressure (AOp) produced by blasting operations originates from the four main sources: the air pressure pulse, gas release pulse, rock pressure pulse, and stemming release pulse. Professionals in various disciplines, such as science, technology, and earth sciences, recognise AOp as a significant environmental danger that affects public safety, health, and the mining industry as a whole. When the frequency of AOp is greater than 20 Hz, it is considered noise, and when it is less than this, it is considered a concussion. Low-frequency AOp can travel long distances and has the potential to intentionally harm structures, including structures, reservoirs, and incline inclinations. Furthermore, AOp can generate greater regularity shakes on doors and windows, dishes, and other objects that are frequently mistaken for noise [7]. Noise pollution is often overlooked in many noisy workplaces. The Environmental Protection Agency (EPA) states that normal human conversation typically reaches about 60 dBA. Extended exposure to an environmental volume of noise of 70 dBA can result in feelings of irritability, nervousness, and difficulty concentrating. Being near sound levels higher than 85 dBA can result in lasting damages to hearing and may indirectly increase the risk of workplace accidents [8]. Noise is defined as sound that lacks a pleasing musical quality or is considered unwanted. The advent of mechanisation has led to the use of large, high-capacity machinery, which has intensified the noise problem

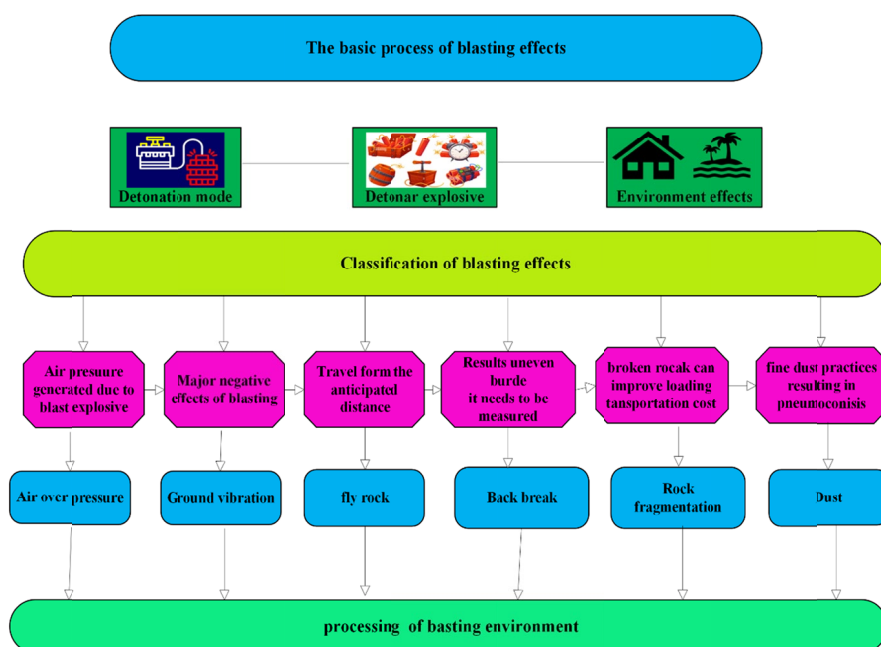


Fig. 1. Classification and Impact of Blasting Effects in Mining Operations

in mines. Extended increased noise exposure can cause both auditory and non-auditory health issues for miners. The mining sector recognises the importance of implementing engineering controls to mitigate noise exposure during underground operations. However, manufacturers are often discouraged from creating quieter machinery due to the relatively limited market for mining equipment. Furthermore, the specialised design requirements imposed by the difficult mining environment hinder the integration of noise control technologies from other industries [9]. During blasting operations, the detonation of explosives generates air blast pressure waves, which are transient and usually last for a brief moment (Wharton et al., 2000). The phrase „blast noise“ can be somewhat misleading, as air blast and noise are separate phenomena. The pressure wave that results from an explosive charge detonating is referred to as an „air blast“ and usually happens at frequencies lower than 20 Hz. This low-frequency air vibration feels like a concussion even though it is not audible. In contrast, noise includes the audible and infrasonic components of the spectrum, ranging from 20 Hz to 20 kHz. When assessing AOp, both the pressure wave (air blast) and the audible/infrasonic components (noise and concussion) are measured in Pascals (Pa) or decibels (dB) across the frequency spectrum. Transient pressure changes, which are typified by discrete pulsating fluctuations both above and below the standard pressure in the air, are reflected in air blast overpressure. These modifications are interpreted as sound waves moving through space. Regardless of whether the highest pressure change is favourable or unfavourable, the accomplished magnitude of the modification is considered to determine the maximum overpressure values. By counting how many pressure fluctuations occur in a second, the frequency of blast overpressure can be determined [10]. Among the greatest hazardous environmental disruptions linked to blasting projects is air blasting. It is challenging to forecast AOp brought on by blasting

with accuracy because several factors affect how airwaves are transmitted. AOp levels are greatly impacted by factors like blast design, wind direction and speed, topography, and ethereal conditions (such as warmth and moisture). Blasting-induced AOp has the potential to cause structural damage and significantly disrupt the surrounding residents' environment [11]. For mines and contractors operating near cities, a plethora of businesses and experts offer pre-blast structural assessments and seismic monitoring services. This work is crucial due to the potential for lawsuits and damage claims arising from nearby blasting activities. Human reactions to vibration and air overpressure are highly subjective, with many individuals being sensitive to even low levels of vibration and blast noise. Residents often report feelings of annoyance when blasting occurs close to populated regions, attributing this to ground vibration and noise from the blasts. Property owners may sometimes suspect that blasting is damaging their buildings. As the name suggests, overpressure describes the brief pressure changes brought on by explosions in air, water, or any other fluid medium. Acoustic waves move through the medium at the speed of sound, causing these pressure changes. When informing the public about the impacts of blasting, it is recommended to use the term „Air Overpressure“ instead of „air blast,“ as the latter may carry negative connotations. „Air Overpressure“ is a more appropriate term for educating the public about blasting effects, as „air blast“ might evoke unfavourable associations [12]. Fig. 2 depicts an Overview of the Article Structure: Methodology and Key Findings for Predicting AOp in Mining.

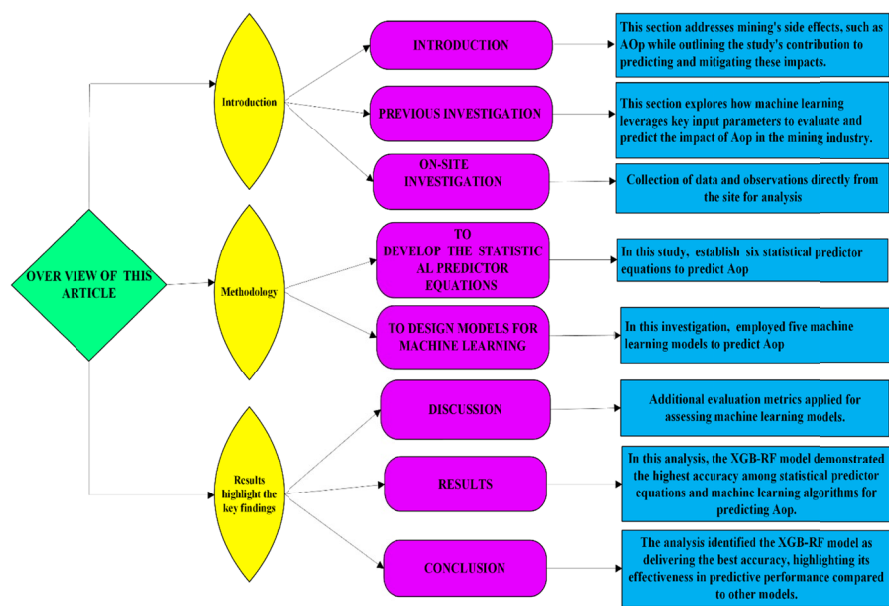


Fig. 2. The Architecture of this Article

2. Previous work(s)

In 2023, researchers Hoang Nguyen, Xuan-Nam Bui, Carsten Drebenstedt, and Yosoon Choi utilised three optimisation algorithms, PSO, FOA, and SGA, to enhance an ANN model for

predicting AOp, Analysing a dataset of 312 samples, they determined that the RWGWO-ANN model achieved the highest prediction precision at 96.2%. This represented a substantial improvement of 16%-20% over the traditional ANN model. The spacing parameter (S) was found to be the most significant changeable, closely followed by the B. Other variables, such as f, PF, Q, and D, were also examined however, their importance was ranked lower than that of B and S in the development of the RWGWO-ANN model for forecasting AOp [12]. In 2023, Ruixuan Zhang, Yuefeng Li, and Yilin Gui developed and validated a model using two exploding records. They compared the performance of their proposed Sa-WKRR model against nine other machine learning models, including KRR, KRVFL, Sa-WKRVFL, transductive KRR (TKRR), edRVF, ANN, RF, SVM, and MARS. Utilising the grid search method, they assessed the optimal performance of these models based on three metrics: root mean squared error (RMSE), mean absolute percentage error (MAPE), and correlation coefficient (R). The results demonstrated that the Sa-WKRR model outperformed all others in both datasets, achieving RMSE values of 0.46/1.98, MAPE values of 0.30%/1.20%, and R values of 0.9991/0.9235 in the first case study, as well as RMSE values of 1.03/3.22, MAPE values of 0.76%/2.74%, and R values of 0.9965/0.9373 in the second case study. These results indicated that the Sa-WKRR model was the most effective and reliable method for predicting blast-induced AOp when contrasted to the other machine-learning models. The key factors included in the model were HD, ST, B, S, MCD, PF, and D [13].

In 2023, Charles Chewu, Tonderai Chik, Desire Runganga, Elia Chipfupi, and Tatenda Nyamagudza developed models utilising ANN and MVR. The most effective ANN model was constructed using the Bayesian Regularisation training method, comprising nine input parameters, 10 hidden neurons, and 3 output parameters. This model achieved a coefficient of determination (R^2) of 0.9983. The optimal ANN model significantly enhanced blast efficiency by reducing ground vibration and airblast levels to below the corporate limits of 20 meters, 9.5 millimetres per second, and 115 decibels. The analysis was based on a database containing 100 blast data sets collected from various locations. The key parameters considered in the model included S, B, HD, HDI, STL, PF, and MCD [14]. In 2023, researchers Shahab Hosseini, Blessing Olamide Taiwo, Yewuhalasht Fissha, Vikram Sakinala, N Sri Chandrahas, Oluwaseun Victor Famobuwa, and Adams Abiodun Akinlabi introduced a novel system for developing AI techniques utilising MARS, RF, SVR, MLP, and LSTM in the Zanzibar region of Iran, employing the CGO algorithm as a metaheuristic approach. They established a hybrid LSTM model, CGO-LSTM, to improve the computational efficiency of these algorithms, highlighting a new form of swarm intelligence. The AOp predictive model was validated against actual recorded values. Soft computing models were created using data from 90 blasting rounds, with the dataset divided into 80% for training and 20% for testing. The model inputs included seven parameters relevant to AOp estimation: Distance of Instrument (DI), B, S, STL, PF, MC, and N. Each of these parameters played a significant role in the AOp estimation process. The accuracy of the models was evaluated using various statistical metrics, including MAE, MAPE, NS, R , RMSE, R^2 , VAF, WI, WMAPE, PI, BI, SI, P, and MRE. The R^2 values for training and testing were as follows: MLP (0.9321, 0.9349), LSTM (0.9861, 0.9851), SVR (0.9391, 0.9473), RF (0.9485, 0.9559), MARS (0.9686, 0.9231), CGO-MLP (0.9529, 0.9637), CGO-LSTM (0.9987, 0.9987), CGO-SVR (0.9453, 0.9425), CGO RF (0.958, 0.9404), and CGO-MARS (0.9808, 0.9611). The results demonstrated that the CGO algorithm significantly enhances the accuracy and performance of AI models [15].

In 2024, Key Fonseca de Lima, Anderson da Cunha Meireles, Nilson Barbieri, and Luan Demarco Fiorentina created a model utilising 12 input parameters through Multiple Linear Regression (MLR) with direct selection. The parameters included Distance, Cylindrical Scaled

Distance, Cubic Scaled Distance, Hole Depth, Maximum Load, Largest Hole, Smallest Hole, B, S, SD, N, and T. With a database comprising 31 observations, the study developed two optimisation methods and two computer models for shock wave prediction: one based on MLR and the other employing GEP. The GEP models, developed using GeneXpro Tools 5.0 software, demonstrated high effectiveness in predicting AOp. The performance metrics of the model revealed an MSE of 7.152, a Root Mean Squared Error (RMSE) of 2.674, and an R^2 value of 0.835, indicating that GEP was the most appropriate approach for AOp prediction [16]. In 2009, C. Kuzu, A. Fisne, and S.G. Ercelebi noted that relying solely on blast design variables such as ST, S, B, and detonation velocity for controlling airblast can be unreliable due to the considerable and unpredictable geological variability. The issues related to GRP and SRP become especially significant when geological deficiencies are present in the rock. To establish a connection between AOp and the key variables D and W, along with other less critical geological and technological factors, a scaling ratio defined as $SD = D/W - 0.33$ and the equation $AOp = H(SD) - \beta$ were employed to reduce AOp values. Therefore, it is crucial to determine the levels of SD and their associated AOp values. To facilitate this, 98 shots were carried out in a quarry to formulate prediction equations based on empirical data collected under varying circumstances [17]. In 2023, Mohammad Mirzei Kalateh Kazemi, Zohreh Nabavi, and Manoj Khandelwal employed several critical input parameters to predict AOp. These parameters included B, NH, HD, S, PF, D, RQD, ST, and MCD. The performance metrics for the predictive models applied to the test data were as follows: XGB-GWO (MAE: 0.69; RMSE: 1.42; R^2 : 0.983), GEP (MAE: 0.63; RMSE: 1.04; R^2 : 0.989), and empirical models (MAE: 5.92; RMSE: 6.68; R^2 : 0.53). Overall, the XGB-GWO hybrid model exhibited the best performance. For the model development, a database consisting of 66 data pairs was compiled, with 20% of the dataset set aside for testing purposes to ensure consistency in the model [18].

In 2020, Victor Amoako Temeng, Yao Yevenyo Ziggah, and Clement Kweku Arthur identified four main factors that influence AOp: stemming length (in meters), the number of blast holes per blast, maximum instantaneous charge (in kilograms), and the distance from the monitoring station to the blasting point (in meters). These factors served as input parameters, while AOp (in decibels) was designated as the output parameter. The study utilised a total of 171 data sets to develop predictive models for AOp, allocating 80% (137 data sets) for training and fitting the model, and the remaining 20% (34 data sets) for testing its accuracy. A novel ANN termed the brain-inspired emotional neural network (BI-ENN), was introduced for predicting AOp. The performance of BI-ENN was benchmarked against two traditional AOp prediction techniques (the generalised predictor and the McKenzie formula), as well as three established AI techniques: BPNN, GMDH, and SVM. The findings revealed that BI-ENN was the most effective approach, achieving the lowest values for RMSE, MAPE, and NRMSE, along with the highest values for R, VAF, and PI (1.0941, 0.8339%, 0.1243%, 0.8249, 68.0512%, and 1.2367, respectively), demonstrating its suitability for monitoring and managing AOp [7]. In 2021, a study by Hesam Dehghani and Farid Ali Mohammad Nia meticulously documented the parameters of 62 blasting operations and measured the resulting AOp for each instance. The recorded parameters included Hole Length, PF, MCD, STL, B, S, NH, and D. To predict the occurrence of AOp in open-pit mining, the researchers explored several methodologies, such as the bat algorithm, PSO-ANN model, and traditional modelling approaches. They created a novel mathematical formula to estimate AOp based on these blasting design parameters, achieving a correlation coefficient of 83% and an RMSE of 2.94. The equation, which is sensitive to location, may serve as a useful reference for other open-pit mining operations [10].

In 2007, researchers C. Sawmliana, P. Pal Roy, R.K. Singh, and T.N. Singh employed AOp data from four distinct mines across India to train a neural network and formulate a generalised prediction equation. The neural network was trained using 70 data sets and validated with an additional 25. Furthermore, both the NN and the generalised prediction equations were evaluated against 15 AOp data sets sourced from two other mines. The results indicated that the neural network achieved an average prediction error of 2.05%, while the generalised equation had a higher error rate of 5.97%. Additionally, the correlation between the actual and predicted AOp values was found to be stronger for the neural network, with a correlation coefficient of 0.931 compared to 0.867 for the generalised equation. This suggests that the neural network delivered more accurate predictions [11]. TABLE 3 summarises previous studies on AOp Prediction.

TABLE 3

Overview of earlier studies accomplished in the area of AOp prediction

S. No	Author	Year	ML Models	Input Parameters	Dataset	R^2
1	Hoang Nguyen et al.	2023	PSO,FOA,SGA, ANN,RWGWO-ANN	S, B, F, PF, Q, D	312	—
2	Ruixuan Zhang et al.	2023	Sa-WKRR(KRVFL), TKRR,RVFL, edRVFL ANN,RF,SVM, MARS	HD, ST, B, S, MCD, PF and D	76	—
3	Charles Chewu et al.	2023	ANN, MVR	S, B, HD, HDI, STL, PF, MCD, and RQD	1000	0.9983
4	Shahab Hosseini et al.	2023	AI, MARS, RF, SVR, MLP, and LSTM, CGO-LSTM	DI, B, S, STL, PF, MC, and N	90	—
5	Key Fonseca de Lima et al.	2024	MLR, GEP	D, CSD, HD, ML, LH, SH, B, S, SD, NOH, and T	31	0.835
6	C. Kuzu et al.	2009	GRP and SRP	ST, S, B, and DV	98	—
7	Mohammad Mirzei Kalateh Kazemi et al.	2023	XGB-GWO, empirical, GEP	B, NOH, HD, S, PF, D, RQD, ST, and MCD	66	XGB-GWO, 0.983 GEP 0.989 empirical 0.53
8	Victor Amoako Temeng et al.	2020	BI-ENN, BPNN, GMDH, SVM	STL, NBHB, MIC, DMS, BP	171	—
9	Hesam Dehghani et al.	2021	Bat algorithm, PSO-ANN model, conventional models	HL, PF, MCD, STL, B, S, NOH and D	62	—
10	C. Sawmliana et al.	2007	NN& generalised predictor equations	MCWD, DBC, TFR, and DM	70	—

3. Dataset collection

The first step of the proposed system involves gathering data from the public resource GitHub website to estimate AOp resulting from blasting operations. The dataset consists of 994 samples, each containing 13 key parameters relevant to the study, including Hole Count, Depth,

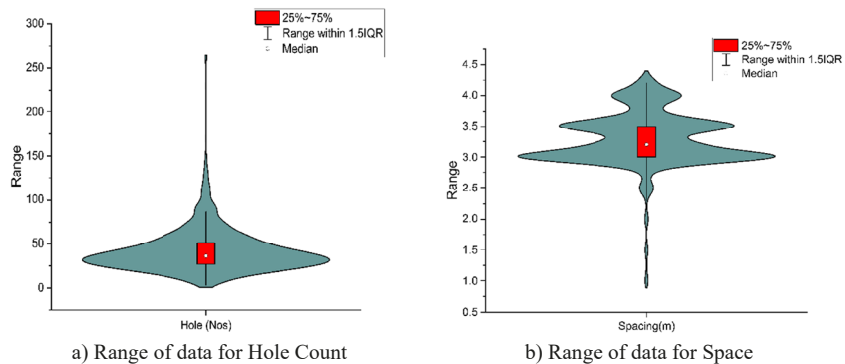
Spacing, Burden, Stemming, Decking, Total Drill Length, Explosive Amount, Volume, Powder Factor, Average Charge per Hole (Av. CPH), Maximum Charge per Delay (MCPD), and Seismic Distance. These input parameters are paired with corresponding output measurements, such as AOp intensity [33]. These input parameters are corresponding output measurements, specifically the AOp intensity, which measures the peak overpressure caused by the blast. Combining these diverse parameters helps capture the complex relationship between the blasting parameters and the resulting Air Overpressure, enabling more accurate prediction models. TABLE 4. Describes a Summary of the Collected Data. The dataset is further enhanced using specific measurement tools and historical data, which ensure the reliability and precision of the output values. This makes it suitable for advanced machine learning and statistical modelling techniques. Fig. 4(a) to (n) illustrates this. The violin plots, illustrate the AOp distribution in dBL, highlighting its range, variability, and key statistical features.

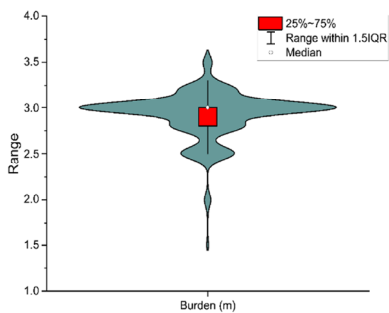
TABLE 4

An overview of the obtained data

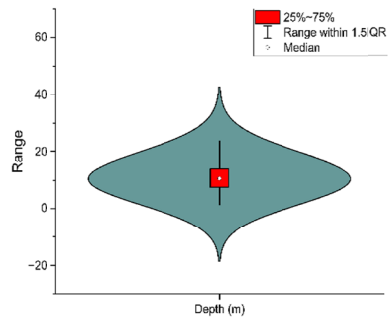
Parameters	Units	Minimum	Maximum	Mean	SD
Hole	Nos	3	280	42.09	26.61
Depth	m	1.39	70	11.17	5.24
Spacing	m	1	4.3	3.29	0.44
Burden	m	0.6	4	2.92	0.30
Stemming	m	1.7	5	2.95	0.22
Decking	m	2.8	60.4	13.28	6.20
Total Drill	RMT	24	714714.78	1147.25	22692.79
Explosive	kg	30	20750	2403.64	2210.97
Volume	m ³	0.27	61222	4452.28	4990.08
Powder Factor	Kg/m ³	0.057	8.33	0.51	0.30
Av. CPH	—	1.39	205	52.33	35.43
MCPD	Kg/D	2.78	540	95.02	71.45
Seis. Dist.	M	0.508	740	71.45	123.70

This dataset is highly valuable for developing predictive models using machine learning algorithms to enhance the understanding of AOp and mitigate its effects. Fig. 4. Demonstrates

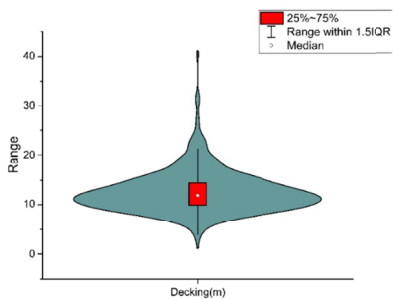




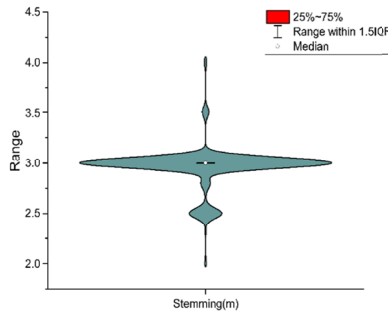
c) Range of data for Burden



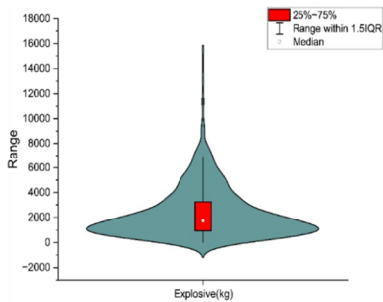
d) Range of data for Depth



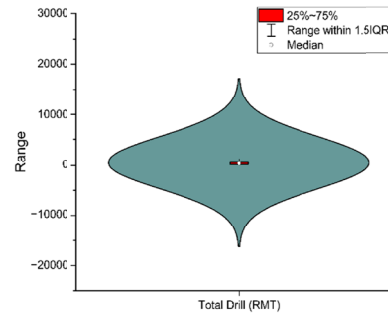
e) Range of data for Decking



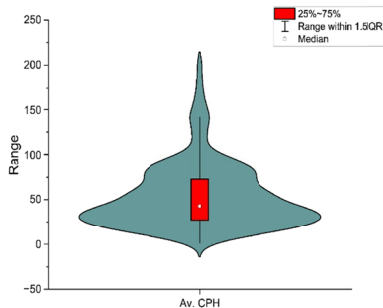
f) Range of data for Stemming



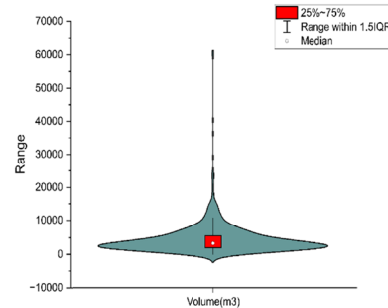
g) Range of data for Explosive



h) Range of data for Total drill



i) Range of data for Average Charge per Hole



j) Range of data for Volume

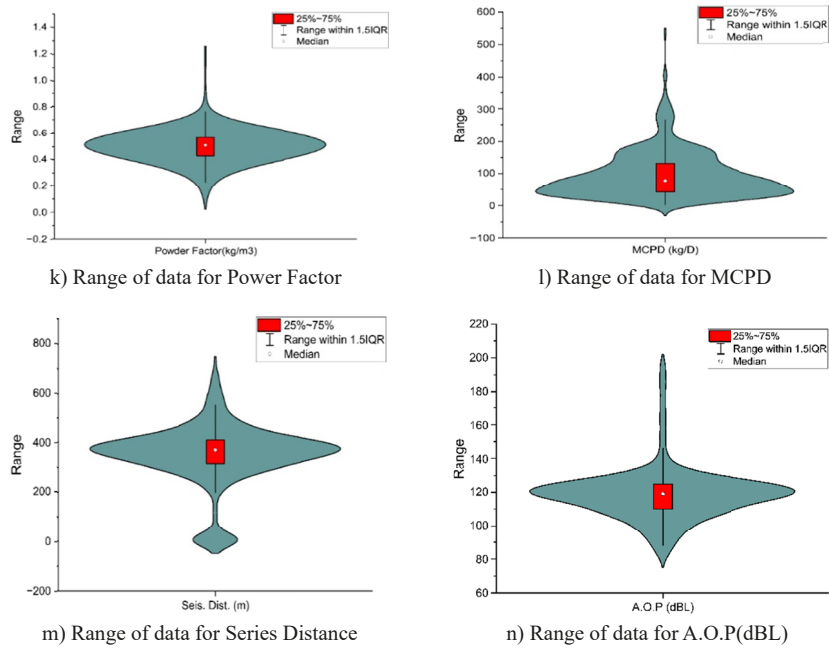


Fig. 3. Violin plots are used to display the range of the measured values for each parameter

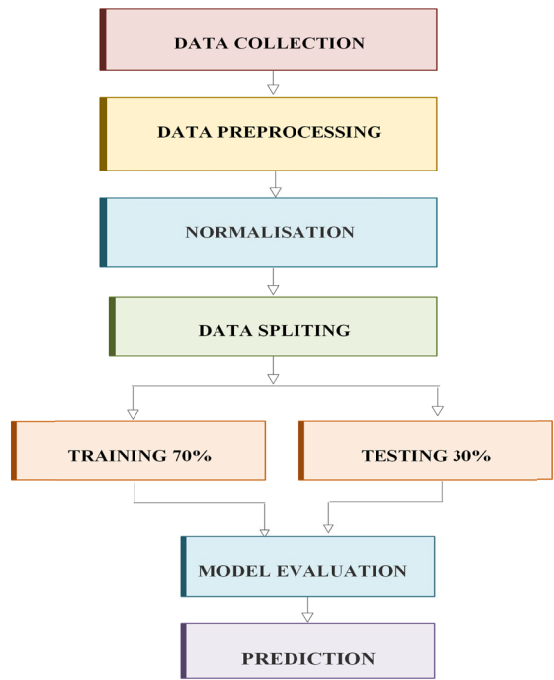


Fig. 4. Methodology of the current study

the methodology used in this study. It can be used for training and testing models to improve safety and efficiency in blasting operations.

4. Directorate General of Mines Safety (DGMS) rules for AOp

Numerous efforts have been made to regulate AOp levels because of their potential to cause harm. AOp, measured in decibels (dB) or pascals (Pa), denotes the intensity of sound, with 20 Hz being the threshold for human auditory perception. TABLE 5 outlines the rules for AOp values and limits according to DGMS regulations. In addition, there is a risk of an ear concussion. While general window breakage happens at 171 dB and sporadic window breakage happens at 151 dB, structural damage can happen at an AOp of 180 dB. As per the suggestions made by Siskind et al. and referenced by the United States Bureau of Mines (USBM), a maximum AOp limit of 134 dB has been advised. Typical AOp is associated with various types of damage resulting from its effects. When a blast occurs, whether from controlled demolition, mining activities, or military operations, it generates a significant amount of air pressure known as AOp. This air overpressure can result in various types of damage, which are crucial for implementing safety measures, designing blast-resistant structures, and minimising blasting activities that can significantly affect nearby areas and communities. It is essential to implement thorough planning and coordination to conduct these operations safely and efficiently. This approach should include strategies to minimize potential risks and damages [19,20].

TABLE 5

Types of damages due to blasting in AOp

S. No	Type of Structural Damage	AOp Value in dB(L)
1.	General window breakage	171
2.	Occasional window breakage	151
3.	Long-term history of application as a safe project specification	140
4.	Bureau of Mines recommendation following a study of large-scale surface mine blasting	134
5.	Plaster Cracks	180
6.	Loose window sash rattles	176
7.	Failure of Badly Installed Window Panes	140-145
8.	Failure of Correctly Installed Window Panes Over	168
9.	All Window Panes Fail	176

4.1. The typical effects of pressure on building structures have significant values

AOp effects on building structures, such as blast-induced Air Overpressure, can lead to varying levels of damage depending on intensity. TABLE 6 highlights the significant impact of pressure on building structures. High-pressure waves can lead to the shattering of windows, cause structural vibrations, or even result in cracks in walls. Continuous exposure to moderate pressure can weaken materials over time, while extreme cases might compromise structural integrity. Understanding these effects is vital for designing resilient buildings [21].

TABLE 6

The effects of pressure on building structures can be substantial

S. No	(dB-L)	Effect
1.	120	Significantly compromised traditional structure
2.	110	Common instances of window breakage
3.	100	Safety regulations for personnel, GB6722, China
4.	90	Instances of window breakage
5.	80	Acceptable limit to mitigate damage to glass and plaster, USBM
6.	70	Detachment of loose plaster flakes, USBM RI 8485, and regulatory standards
7.	60	Vibration of windows and associated feelings of irritation

5. Statistical empirical predictor to estimate the AOp

The measurement of air noise equations using a statistical empirical predictor involves employing advanced statistical models to analyse and predict the intricate patterns of airborne sound, providing a robust framework for assessing and understanding the acoustic characteristics in various environments [22]. TABLE 7 lists the adopted statistical empirical predictor to measure the air noise. Most authors exclusively employed two input parameters: burden and distance, as identified in a previous study. Because the error and accuracy are not perfect, using the statistical empirical equation indicates that it is possible to employ machine learning techniques. In this context, (P) represents the overpressure or air blast, and (Q) stands for the mass of the explosive charge.

TABLE 7

Measurement of the air noise equations using statistical empirical predictor equations

Prediction Model	Equations
1	2
USBM	$P = \beta_1 * \left(\frac{D}{Q^{0.33}} \right)^{\beta^2}$
NAASRA	$P = 140^3 \sqrt{\frac{Q}{200}} \frac{1}{d}$
Olofso; Persson et al.	$P = 0.7 * \frac{Q^{\frac{1}{3}}}{D} \text{ (mbar)}$
Holmberg-Persson	$P = K * 0.7 * \frac{Q^{\frac{1}{3}}}{D} \text{ (mbar)}$
Mckenzie	$P = 165 - 24 \log \left(\frac{D}{Q^{\frac{1}{3}}} \right) \text{ (dB)}$

1	2
Generalised equation	$AOP = K * \left(\frac{D}{\frac{1}{Q_{\max}^3}} \right)^{-b}$

NOTE: In this context, (*P*) represents the overpressure or air blast, (*Q*) stands for the mass of the explosive charge (kg), (*D*) is the distance (*m*) of the charge from the monitoring point, and (*H*) and (*β*) are the site-specific factors. The effects of blasting extend beyond immediate physical impacts, encompassing environmental consequences, structural implications, and potential socio-economic ramifications, necessitating comprehensive studies and strategic mitigation approaches to address and minimise this multifaceted repercussion.

Most authors exclusively employed two input parameters: burden and distance, as identified in a previous study. Because the error and the accuracy are not perfect, using the statistical empirical equation indicates that machine learning techniques can be employed.

6. Traditional models for predicting Air Overpressure (AOp)

Conventional noise prediction equations, developed by academicians, researchers, and field engineers, estimate sound levels based on empirical and theoretical models. These equations often use parameters such as maximum charge per delay, distance, and environmental conditions to predict noise propagation. While widely applied, their accuracy can be limited to assumptions that simplify complex real-world scenarios. They are foundational in ecological impact assessments, offering baseline predictions for industrial and construction activities. Modern advancements, such as machine learning, aim to improve accuracy by incorporating non-linear and site-specific variables.

6.1. U.S. Bureau of Mines (USBM)

The United States Bureau of Mines (USBM) introduced a widely recognised predictive model to estimate AOp generated during blasting operations.

The equation is expressed as:

USBM

$$P = \beta_1 * \left(\frac{D}{Q^{0.33}} \right)^{\beta_2}$$

(1)

To estimate the AOp using the USBM predictive equation, the following methodology is employed:

- Input:
- D* – Input variable representing a parameter (distance)

Q – Input variable representing another parameter

*β*1 – Constant multiplier

*β*2 – Exponent applied to the formula
- Output:
- P* – Result calculated using the given formula

Steps:

1. Begin by reading the input values:
 - Read D (e.g., distance or related parameter).
 - Read Q (e.g., flow rate or similar measure).
 - Define constants β_1 and β^2 .
 2. Calculate the intermediate value of $(Q^{0.33})$:
 - Raise (Q) to the power of 0.33 (cube root of Q).
 3. Compute the ratio $(D/Q^{0.33})$:
 - Divide (D) by the computed value of $(Q^{0.33})$.
 4. Raise the result of the division to the power (β^2) :
 - Take the ratio from step 3 and raise it to the power of (β^2) .
 5. Multiply the result by (β_1) to calculate (P) :
 - Multiply (β_1) by the result obtained in step 4.
 6. Output the calculated value of (P) .
-

The equation (1) $P = \beta_1 * \left(\frac{D}{Q^{0.33}} \right)^{\beta^2}$ Determine the β_1 and β^2 values, then substitute the equation.

$$P = 0.009 * \left(\frac{D}{Q^{0.33}} \right)^{2.2664}$$

$$P = 0.2162$$

By implementing the outlined process, an R^2 value of 0.44 was obtained for predicting AOp. This indicates moderate predictive performance for the USBM.

6.2. NAASRA

NAASRA developed a predictor equation to evaluate blast-induced air noise, expressed as:

$$P = 140 \sqrt[3]{\frac{Q}{200}} \frac{1}{d} \text{ (Kpa)} \quad (2)$$

The following methodology is adopted to determine AOp using this equation:

Input:

- $Q1$ – Input quantity (e.g., related to pressure or load)
- d – Input parameter (e.g., distance or another physical property)

Constants:

- Constant_1 = 140 # Numerator in the formula (in KPa)
- Constant_2 = 200 # Denominator in the formula (in KPa)

Output:

- P – Resultant pressure in kilopascals (KPa)

Steps:

1. Start by defining the constants:
 - Set Constant_1 to 140 (numerator of the equation).
 - Set Constant_2 to 200 (denominator of the equation).
 2. Read the input values:
 - Accept the value of Q_1 .
 - Accept the value of d .
 3. Calculate the ratio of constants:
 - Compute the ratio Constant_1 / Constant_2.
 4. Multiply the ratio by the input value Q_1 :
 - Let intermediate_result = (Constant_1 / Constant_2) \times Q_1 .
 5. Adjust for the effect of d (if d is a multiplier or influence):
 - Incorporate d into the calculation (e.g., multiplying intermediate result by d).
 6. Store the final result as P :
 - Assign P = intermediate_result.
 7. Output the calculated value of P with the appropriate units (KPa).
- End.
-

Using the implemented approach, the prediction of AOp achieved an R^2 value of 0.3692, indicating a moderate level of accuracy for the NAASRA method.

6.3. OLLOFSO

The OLLOFSO derived the following equation to estimate the AOp. The equation is:

$$P = 0.7 * \frac{Q^{\frac{1}{3}}}{D} \text{ (mbar)} \quad (3)$$

To analyses and predict the AOp, apply the following methodology:

Input:

- Q – Explosive charge quantity or energy release (e.g., in kg or equivalent units)
- D – Distance from the explosion source (e.g., in meters)

Constants:

Coefficient = 0.7 # Scaling factor for pressure estimation

Output:

- P – Estimated peak overpressure (e.g., in millibars)

Steps:

1. Define the scaling constant:
 - Set Coefficient to 0.7.
2. Accept input values:
 - Read Q , explosive charge quantity.
 - Read D , distance from the explosion source.

3. Apply the cube root transformation to Q :
 - Compute $Q^{1/3}$ (cube root of Q).
 4. Calculate the peak overpressure:
 - Use the formula $P = \text{Coefficient} \times (Q^{1/3}/D)$.
 - Divide the cube root of Q by D to account for the distance effect.
 - Multiply the result by the Coefficient to get the peak overpressure.
 5. Store the result in P .
 6. Output P with the appropriate units (e.g., mbar).
-

The applied method resulted in an R^2 value of 0.2093 for predicting AOp, reflecting a relatively low predictive performance for the OLLOFSO approach.

6.4. Holmberg-Persson

Holmberg-Persson adopted the predictor equation to estimate the AOp levels induced by blasting operations. The stated equation is as follows:

$$P = K * 0.7 * \frac{Q^{\frac{1}{3}}}{D} \quad (4)$$

The following procedure was taken during the evaluation of AOp:

Input:

- Q – Explosive charge quantity or energy release (e.g., in kg or equivalent units)
- D – Distance from the explosion source (e.g., in meters)
- K – Correction factor (unitless)

Constants:

Coefficient = 0.7 # Scaling factor for pressure estimation

Output:

- P – Estimated peak overpressure (e.g., in millibars or kilopascals)

Steps:

1. Define the scaling constant:
 - Set the Coefficient to 0.7.
 2. Accept input values:
 - Read Q , the explosive charge quantity.
 - Read D , the distance from the explosion source.
 - Read K , the correction factor.
 3. Apply the cube root transformation to Q :
 - Compute $Q^{1/3}$ (cube root of Q).
 4. Calculate the peak overpressure:
 - Use the formula $P = K * \text{Coefficient} * (Q^{1/3}/D)$.
 - Divide the cube root of Q by D to account for the distance effect.
 - Multiply the result by the Coefficient and K to get the peak overpressure.
 5. Store the result in P .
 6. Output P with the appropriate units (e.g., mbar or kPa).
-

The prediction of AOp using this approach produced an R^2 value of 0.4325, indicating moderate predictive capability for the Holmberg-Persson method.

6.5. Mckenzie

Mckenzie developed the predictor equation to estimate the AOp and stated it as:

$$P = 165 - 24 \log \frac{D}{Q^{\frac{1}{3}}} \text{ (dB)} \quad (5)$$

The procedure has been followed to calculate the AOp:

Input:

D – Distance from the explosion source (e.g., in meters)

Constants:

BaseValue = 165 # Constant base pressure value (in dB)

Coefficient = 24 # Scaling factor for logarithmic component

Output:

P – Estimated air overpressure (in decibels, dB)

Steps:

1. Define the constants:
 - Set BaseValue to 165.
 - Set the Coefficient to 24.
 2. Accept input value:
 - Read D , the distance from the explosion source.
 3. Compute the logarithmic term:
 - Calculate $\log(D)$, where D is the distance.
 4. Calculate the air overpressure:
 - Use the formula $P = \text{BaseValue} - \text{Coefficient} * \log(D)$.
 5. Store the result in P .
 6. Output P with the appropriate units (e.g., dB).
-

Using the McKenzie method, the prediction of AOp resulted in an R^2 value of 0.1386, indicating a low predictive performance.

6.6. Generalised equation

One of the popular predictors to estimate the AOp due to blasting operation is the generalised equation, which is represented as:

$$AOP = K * \left(\frac{D}{Q_{\max}^{\frac{1}{3}}} \right)^{-b} \quad (6)$$

The AOp levels are determined and analysed using the following methodology:

Input:

- D – Distance from the explosion source (e.g., in meters)
- Q_{\max} – Maximum charge weight or energy release (e.g., in kg or equivalent units)
- K – Empirical constant derived from regression or field observations
- b – Exponent determined through regression or field observations

Output:

- AOP – Estimated Air Overpressure (e.g., in millibars or kilopascals)

Steps:

1. Define the empirical constants:
 - Set K to the determined constant value.
 - Set b to the determined exponent value.
2. Accept input values:
 - Read D , the distance from the explosion source.
 - Read Q_{\max} , the maximum charge weight or energy release.
3. Calculate the charge effect:
 - Compute $Q_component = Q_{\max}^{1/3}$
($Q_component = Q_{\max}^{1/3}$)
4. Calculate the distance effect:
 - Compute $D_component = D^{-b}$
($D_component = D^{-b}$)
5. Compute the Air Overpressure (AOp):
 - Use the formula $AOP = K * (Q_component) / D_component$
($AOP = K * Q_component / D_component$)
6. Store the result:
 - Assign the calculated value to AOp.
7. Output the Air Overpressure:
 - Display or return AOp with the appropriate units (e.g., mbar or kPa).

The equation (6) $AOP = K * \left(\frac{D}{Q_{\max}^3} \right)^{-b}$ Determine the k and $-b$ values, then replace the equation.

$$AOP = 0.009 * \left(\frac{D}{Q_{\max}^3} \right)^{2.2609}$$

$$AOp = 0.2166$$

The prediction of AOp using the Generalised equation yielded an R^2 value of 0.1386, reflecting a low predictive capability.

7. Machine Learning Models for Predicting Blast-Induced Air Overpressure (AOp)

Machine learning techniques, including Support Vector Regression (SVR), Gradient Boosting (GB), Random Forest (RF), Backpropagation, and XGBoost-RF, are applied to predict blast-induced AOp. These models aim to enhance the accuracy of AOp predictions by analysing complex relationships within the data. Each algorithm contributes unique strengths, such as handling non-linear relationships or reducing overfitting. The integration of these methods enables the achievement of robust and reliable predictions for blast events. This approach offers significant improvements over traditional predictive models in the field of blasting.

7.1. Best Hyperparameters in ML for Predicting AOp

In machine learning models for predicting blast-induced AOp, selecting the best hyperparameters is crucial for achieving high accuracy and minimising error. Optimal hyperparameters (such as learning rate, maximum depth, number of estimators, etc.) can significantly improve model performance by fine-tuning the model to capture complex patterns in AOp data more effectively. The tables below show Support Vector Regression, Gradient Boosting, Random Forest, BBPNN, and XGBoost Random Forest models with optimal hyperparameters for best performance.

7.1.1. Support Vector Regression (SVR)

SVM is a supervised machine learning method based on statistical learning theory. It excels in both classification and regression tasks [23]. It is regarded as more effective than other ANNs, including perceptron networks [24]. SVM was first introduced by Vapnik and is divided into two main categories: Support Vector Classification (SVC) and Support Vector Regression (SVR). SVM transforms data into a high-dimensional feature space, creating predictive models based on a select group of support vectors. It can effectively generalise complex gray-level patterns with just a limited number of support vectors, which makes it beneficial for image compression. SVR, a regression adaptation of SVM, was proposed by Vapnik, Steven Golowich, and Alex Smola in 1997, and builds models that focus on a subset of training data, ignoring data points within a margin of error (ϵ) [25]. The methodology illustrated below has been implemented in the SVR approach.

Algorithm for SVR Model Evaluation

Input:

- Dataset: $(a_i, b_i), i = 1, 2, \dots, n$
- Number of replicates: $n_{repl} = 30$

SVR Parameters:

- Kernel: Radial Basis Function ('rbf')
- Regularisation Parameter ($C = 100$)
- Kernel Coefficient ($\gamma = 0.1$)
- Epsilon ($\epsilon = 0.1$)

Output:

- Aggregated Performance Metrics: MAE, MSE, RMSE, R^2 .
 1. $t \leftarrow 0$ replicate loops

2. Initialise empty lists for performance metrics: MAE and MSE.
3. **For** $t = 1$ to n_{repl}
4. Shuffle the dataset randomly
5. Train an SVR model using $(a_i^{train}, b_i^{train})$ with parameters
6. Kernel: 'rbf' $C = 100, \gamma = 0.1, \varepsilon = 0$.
7. Use the trained model to predict b_i for using (a_i^{test}, b_i^{test})
8. Compute residuals for b_i^{test}
9. $r_i = b_i^{test} - b_i$
10. Calculate performance metrics for this replicate: MAE, MSE, RMSE, R^2 .
11. Append MAE, MSE, RMSE, R^2
12. **End For**
13. Compute aggregated performance metrics
14. Return MAE, MSE, RMSE, R^2

Following the process below, identify the best hyperparameters. Create a Support Vector Regression (SVR) model with optimised hyperparameters for the best performance. The selected parameters enhance accuracy and model fit. TABLE 8 exhibits an SVR with optimal hyperparameters for the best performance.

TABLE 8

Support Vector Regression (SVR) with optimal hyperparameters for best performance

Key Parameters	Description
Kernel = 'rbf'	RBF (Radial Basis Function) kernel: This kernel maps the input data into a higher-dimensional space where it becomes easier to separate the data points.
$C = 100$	Regularisation parameter (C): This controls the trade-off between achieving a low error on the training data and having a smooth (simplified) decision boundary.
gamma = 0.1	Gamma parameter for the RBF kernel: This defines how far the influence of a single training example reaches.
epsilon = 0.1	Epsilon in the epsilon: This defines a margin of tolerance where no penalty is given for errors.
random_state = 42	Random seed: This ensures reproducibility by controlling the random processes in the SVR algorithm (e.g., during data shuffling or other stochastic processes)

7.1.2. Gradient Boosting

Gradient boosting is a machine learning method used for both regression and classification tasks that creates a predictive model by combining an ensemble of weaker models. Simply put, it builds a stronger model by integrating multiple weaker predictors. This algorithm is highly effective for developing predictive models and has diverse applications across multiple fields, including data management systems [26]. All aspects of the Harmony Gradient Boosting Machine were discussed. In Gradient Boosting Machines (GBMs), the learning process is sequential, with each new model refining the predictions of the response variables. Boosting algorithms, in general, are straightforward to implement and allow flexibility in experimenting with different

model designs. Additionally, GBMs have demonstrated success in real-world applications as well as in various data mining and machine-learning tasks [27]. LightGBM, created by Microsoft, is an advanced gradient-boosting framework that utilises a tree-based learning algorithm. It is designed for quicker training times, lower memory usage, and enhanced accuracy. LightGBM is compatible with GPU acceleration and parallel processing, and it includes a feature for testing and retaining the best set of parameters. In contrast to traditional gradient boosting techniques, which generally start by fitting a weak learner to input-output pairs and then iteratively reducing the errors through weak models like decision trees, LightGBM employs a different tree-building strategy. It adopts a leaf-wise (best-first) method instead of a level-wise (depth-first) approach, distinguishing it from other boosting algorithms [28]. The methodology depicted has been implemented in the GB approach.

Algorithm for Gradient Boosting Model Performance Evaluation

Input:

- Dataset: $(a_i, b_i), i = 1, 2, \dots, n$
- Number of replicates: $n_{repl} = 30$
- Number of boosting iterations: $n_{estimators} = 100$

Output:

- Aggregated Performance Metrics: MAE, MSE, RMSE, R^2
 - 1) $t \leftarrow 0$ replicate loops
 - 2) Initialise empty lists for performance metrics: MAE, MSE, RMSE, R^2
 - 3) **For** $t = 1$ to n_{repl}
 - 4) Shuffle the dataset randomly.
 - 5) Train a Gradient Boosting (GB) model using $(a_i^{train}, b_i^{train})$ with parameters
 - 6) Number of estimators: $n_{estimators} = 100$
 - 7) Use the trained GB model to predict b_i for using (a_i^{test}, b_i^{test})
 - 8) Evaluate the performance metrics for the test data
 - 9) Append MAE, MSE, RMSE, R^2
 - 10) **End For**
 - 11) Compute aggregated performance metrics
 - 12) Return MAE, MSE, RMSE, R^2
-

By following the above process, identify the best hyperparameters. Gradient Boosting model with optimised hyperparameters for optimal performance, enhancing accuracy and model fit. TABLE 9 describes GB with optimal hyperparameters for best performance.

TABLE 9

GB with optimal hyperparameters for best performance

Parameters	Description
$n_{estimators} = 30$	Number of boosting iterations ($n_{estimators}$): This specifies the number of trees (or weak learners) built sequentially in the Gradient Boosting model.
$random_state = 10$	Random seed: This controls the randomness involved in various algorithm steps (e.g., subsampling the training data, splitting data into training/testing sets).

7.1.3. Random Forest (RF)

Over the past ten years, the Random Forest (RF) algorithm has emerged as a popular tool for data analysis in bioinformatics, which encompasses the application of computer science and information technology to biology and medicine. RF is utilised to address two primary types of challenges: creating prediction models for supervised learning tasks and assessing and ranking variables based on their predictive capabilities. This ranking is accomplished through variable importance measures (VIMs), which the RF algorithm automatically computes for each predictor. Importantly, RF VIMs are recognised for their effectiveness in identifying predictors that interact, meaning they can predict outcomes only when combined with other predictors [29]. Random Forest is an ensemble learning method categorised under homogeneous base learners in the realm of positive classifiers. As its name demonstrates, every foundational pupil consists of decision trees, resulting in a simpler structure compared to similar techniques. The Random Forest model provides two primary advantages: one related to computation and the other to statistical performance. From a computational perspective, Random Forest effectively manages both regression and classification tasks, with training techniques for predicting that are notably rapid, making it one of the faster traditional classifiers. Additionally, it excels in addressing high-dimensional problems directly [30]. The Random Forest comprises numerous decision trees that grow to their full extent without the need for pruning. Increasing the number of trees generally enhances prediction accuracy while reducing the risk of overfitting. The algorithm provides an overall estimation and includes the advantage of automatic feature selection and other benefits [31]. The methodology represented below has been implemented in the RF approach.

Algorithm for Random Forest Performance Evaluation

Input:

1. Dataset: $(a_i, b_i), i = 1, 2, \dots, n$
2. Number of replicates: $n_{\text{repl}} = 30$
3. Number of runs per replicate: $n_{\text{runs}} = 100$
4. Number of trees in Random Forest: $n_{\text{trees}} = 100$

Output:

Aggregated Performance Metrics: Test set evaluation metrics (MAE, MSE, RMSE, R^2)

Jungle: Collection of all decision trees trained during the process.

Super-Ensemble: Collection of all Random Forest models.

1. $t \leftarrow 0$ replicate loops
2. Initialise empty lists for performance metrics: MAE, MSE, RMSE, R^2 .
3. Initialise Jungle $\leftarrow \{\}$ (empty set of decision trees).
4. Initialise Super-Ensemble $\leftarrow \{\}$ (empty set of Random Forest models).
5. **For** $t = 1$ to n_{repl} :
6. Shuffle the dataset randomly.
7. **For** $r = 1$ to n_{runs} :
8. Train a Random Forest (RF) model using $(a_i^{\text{train}}, b_i^{\text{train}})$ with Number of trees: $n_{\text{trees}} = 100$
9. Add the trained Random Forest model to the Super-Ensemble

10. Extract individual decision trees from the Random Forest model and add them to Jungle.
11. Use the trained RF model to predict b_i for using (a_i^{test}, b_i^{test})
12. Evaluate the performance metrics for the test data
13. Append MAE, MSE, RMSE, R^2
14. **End For**
15. Aggregate metrics for replicate t :
16. Append MAE, MSE, RMSE, R^2
17. **End For**
18. Compute final aggregated performance metrics
19. Overall prediction values vs actual values
20. Return MAE, MSE, RMSE, R^2 , Jungle, Super-Ensemble

By following the above process, identify the best hyperparameters. Random forest model with optimised hyperparameters for optimal performance, enhancing accuracy and model fit. Below is TABLE 10. Random Forest with optimal hyperparameters for best performance.

TABLE 10

RF with optimal hyperparameters for best performance

Parameters	Description
$test_size = 0.3$	When $test_size = 0.3$, 30% of the data is allocated for testing, while the remaining 70% is used for training the model.
$random_state = 51$	The $random_state$ parameter ensures that the data-splitting process is reproducible.

7.1.4. Back Propagation Neural Network (BPNN)

Back propagation neural network (BPNN) is a fundamental algorithm used to train artificial neural networks, particularly in supervised learning. It works by minimising the error between predicted and actual outputs through two key phases: forward propagation and backwards propagation. In forward propagation, input data moves through the network to generate a prediction. The error, calculated using a loss function, is then propagated backwards through the network in the backwards phase. Gradients of the error concerning each weight are calculated, and the weights are updated using gradient descent to reduce the error. This iterative process continues until the model converges to an optimal solution. Backpropagation is essential for training deep learning models efficiently, but it can face challenges like vanishing gradients in deep networks. Fig. 5 portrays how Backpropagation neural networks enhance the perceptron by enabling multi-layer networks to learn through error correction using gradient descent.

Algorithm: Backpropagation Neural Network Performance Evaluation

Input:

1. Dataset: $(a_i, b_i), i = 1, 2, \dots, n$
2. Number of replicates: $n_{repl} = 30$
3. Epochs: $epochs = 1000$
4. Learning rate: $lr = 0.1$

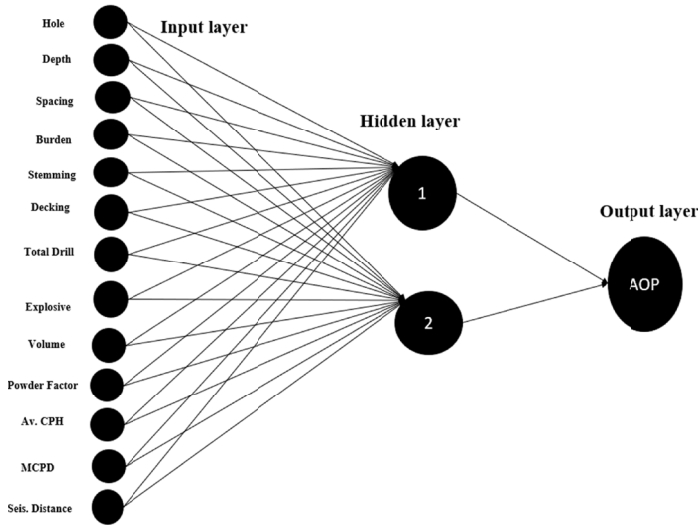


Fig. 5. Neural Network 13-2-1 Architecture for Predicting Air Overpressure

Output:

Aggregated Performance Metrics: Test set evaluation metrics (MAE, MSE, RMSE, R^2)

1. $t \leftarrow 0$ replicate loops
2. Initialise empty lists for performance metrics: MAE, MSE, RMSE, R^2
3. **For** $t = 1$ n_{repl} :
4. Shuffle the dataset randomly.
5. Initialise neural network weights randomly with seed 42 for reproducibility.
6. **For** $e = 1$ to epochs
7. Perform a forward pass to compute predictions b_i for a_i^{train} .
8. Compute the loss (e.g., Mean Squared Error).
9. Backpropagate the error to calculate gradients for each weight.
10. Update weights using the learning rate lr :
11. **If** $e \% 1000$ Log the training loss
12. **End For**
13. Use the trained neural network to predict b_i for (a_i^{test}, b_i^{test})
14. Evaluate the performance metrics for the test data: actual /predicted
15. Compute based on the specific loss function (e.g., MSE).
16. Append MAE, MSE, RMSE, R^2
17. **End For**
18. Compute final aggregated performance metrics: actual /predicted
19. Return MAE, MSE, RMSE, R^2 .

By following the above process, identify the best hyperparameters. Backpropagation model with optimised hyperparameters for optimal performance, enhancing accuracy and model fit. Describe TABLE 11. Backpropagation Neural Network with optimal hyperparameters for best performance.

TABLE 11

Backpropagation Neural Network Model with optimal hyperparameters for best performance

Parameters	Description
np. random.seed (42)	Ensures that the initial conditions (e.g., weight initialisation) are the same for every run, enabling reproducibility in training.
$lr = 0.1$	Sets the step size for updating the model's weights during backpropagation. The gradients calculated during backpropagation will be multiplied by 0.1 to adjust the weights.
$epoch \% 1000 = 0$	Checks whether the current epoch is divisible by 1000. This is useful for periodically logging or saving progress during the backpropagation process.

7.1.5. XG Boost-Random Forest

Gradient-boosted decision trees and other gradient-boosted models are frequently trained using XGBoost. XGBRF, also known as the Random Forest implementation in Extreme Gradient Boosting, enhances prediction accuracy and delivers improved results. While XGBoost is categorised under the boosting algorithm, RF (Random Forest) falls under the bagging algorithm. In this hybrid approach, Random Forests are used as the base estimator. The model structure and inference process are similar for both Random Forests and Gradient-Boosted Decision Trees; however, the training methodology differs between the two (Ragam et al., 2022).

Algorithm: XGBoost-RF Model Evaluation

Function:

Evaluate the performance of an XGBoost-RF model using multiple replicate experiments with specific hyperparameters.

Inputs:

- Dataset: $(a_i, b_i), i = 1, 2, \dots, n$
- Replicates: $n_{repl} = 30$
- $n_estimators$: Number of trees in the ensemble $n_{estimators} = 100$
- max_depth : Maximum depth of each tree ($max_depth = 6$)
- $learning_rate$: Step size for weight updates ($lr = 0.1$)
- $subsample$: Ratio of training samples used per iteration ($subsample = 0.8$)
- $colsample_bynode$: Ratio of features sampled per node ($colsample_bynode = 0.5$)

Outputs:

- Aggregated performance metrics over replicates:
 R^2 : Coefficient of Determination.
 RMSE: Root Mean Squared Error.

Steps:

1. Initialisation

Set the hyperparameters:

- $N_{estimators} = 100$
- $max_depth = 6$
- $lr = 0.1$
- $subsample = 0.8$
- $colsample_bynode = 0.5$

- 1.2. Initialize empty lists: R^2 , RMSE to store results.
- 2. Replicate Experiment Loop
 - For $t = 1$ to n_{repl} :
 - Shuffle the dataset randomly.
- 3. Model Training
 - Train the XGBoost-RF model $(a_i^{train}, b_i^{train})$ using the hyperparameters:
 - $n_{estimators} = 100$
 - $max_depth = 6$
 - $lr = 0.1$,
 - $subsample = 0.8$,
 - $colsample_bynode = 0.5$
- 4. Metric Collection
 - Use the trained model to predict b_i^{test} for a_i^{test}
 - Compute performance metrics for the replicate: RMSE, MSE, MAE, R^2
 - Append RMSE, MSE, MAE, R^2
- 5. Aggregating Results
 - Compute final aggregated performance metrics over all replicates:

Output Results

- Return RMSE, MSE, MAE, R^2 .

By following the above process, identify the best hyperparameters. XGBoost-Random Forest model with optimised hyperparameters for optimal performance, enhancing accuracy and model fit. Portray TABLE 12. XGBoost-Random Forest with optimal hyperparameters for best performance.

TABLE 12

XGBOOST-RF model optimal hyperparameters for best performance

XGBOOST parameters	Description
$n_{estimators} = 100$	Number of trees
$max_depth = 6$	Maximum depth of trees
$learning_rate = 0.1$	Learning rate
$subsample = 0.8$	A random sampling of training data
$colsample_bynode = 0.5$	Random feature sampling for each tree
$tree_method = 'auto'$	Auto method for tree construction
$booster = 'gbtree'$	Use tree-based models
$random_state = 42$	Random seed: This controls the randomness involved in various algorithm steps (e.g., subsampling the training data, splitting data into training/testing sets).

Random Forest parameters	Description
$n_{estimators} = 100$	Number of trees
$max_depth = 6$	Maximum depth of trees
$random_state = 42$	Random seed: This controls the randomness involved in various algorithm steps (e.g., subsampling the training data, splitting data into training/testing sets).

8. Model performance

The effectiveness of various conventional prediction models was evaluated using standard statistical metrics, with coefficient of determination (R^2) being the most commonly used measure. Likewise, the performance of machine learning models was evaluated through various statistical metrics, including the coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), and root mean square error (RMSE).

8.1. Coefficient of determination (R^2)

The equation for the R -squared (R^2) statistic, which represents the proportion of variance in the dependent variable that is predictable from the independent variables, is given by:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (7)$$

where:

y_i are the observed values,

\hat{y}_i are the predicted values,

\bar{y} is the mean of the observed data,

$\sum (y_i - \hat{y}_i)^2$ is the sum of squared residuals (also known as the sum of squared errors or SSE),

$\sum (y_i - \bar{y})^2$ is the total sum of squares (TSS), which measures the total variance in the data.

An R^2 value closer to 1 indicates that the model explains a large portion of the variance, while a value close to 0 indicates that the model does not explain much of the variance.

8.2. Mean Squared Error (MSE)

The Mean Squared Error (MSE) is a metric used to evaluate the accuracy of a regression model. It measures the average of the squares of the errors between the predicted values \hat{y}_i and the actual values y_i .

The formula for MSE is:

$$MSE = \frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2 \quad (8)$$

where:

n is the number of data points.

y_i represents the actual values.

\hat{y}_i represents the predicted values.

MSE conveys how far the predictions are from the actual values. Lower MSE values indicate better model performance, as the errors are less significant. However, since the errors are squared, larger errors are penalised more heavily.

8.3. Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) is a measure of prediction accuracy in regression models. It calculates the average of the absolute differences between predicted values \hat{y}_i and actual values y_i .

The formula for MAE is:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

where:

n is the number of data points.

y_i represents the actual values.

\hat{y}_i represents the predicted values.

Unlike Mean Squared Error (MSE), MAE does not square the errors, meaning it treats all errors equally, without penalising larger errors more than smaller ones. Therefore, MAE provides a more intuitive sense of the average prediction error. Lower MAE values indicate better model performance.

8.4. Root Mean Square Error (RMSE)

The Root Mean Square Error (RMSE) is a commonly used metric for measuring the accuracy of a regression model. It represents the square root of the average of the squared differences between the predicted values: \hat{y}_i and the actual values y_i .

The formula for RMSE is:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$

where:

n is the number of data points.

y_i represents the actual value

\hat{y}_i represents the predicted values.

RMSE is sensitive to large errors because it squares the differences, similar to Mean Squared Error (MSE). Taking the square root brings the error back to the original units of the dependent variable, making it easier to interpret and understand. Lower RMSE values indicate better model performance and are preferred when large errors need heavier penalties.

9. Results and discussion

Statistical equations and machine learning algorithms play a pivotal role in analysing and interpreting data to derive meaningful insights. Results from statistical models, such as R^2 and RMSE quantify model accuracy. Machine learning algorithms like SVR, GB, RF, Backpropagation Neural Networks, and XGBoost -RF demonstrate their efficacy by achieving high predictive evaluation metrics.

9.1. Statistical Equations Results

To predict the AOp, statistical equations usually employ empirical or regression-based models that connect the above variables (charge weight, distance) to the resulting overpressure. These equations are derived from historical blasting data and field measurements, allowing predictions in similar blasting scenarios. Fig. 6 Depicts predicted AOp values using statistical equations, showing the fit with an R^2 value representing predictive accuracy (a) USBM R^2 value, (b) NAASRA R^2 value, (c) OLLOFSO R^2 value, (d) HOLMBERGE R^2 value, (e) Mckenzie R^2 value, (f) GENERALISED EQUATION R^2 value.

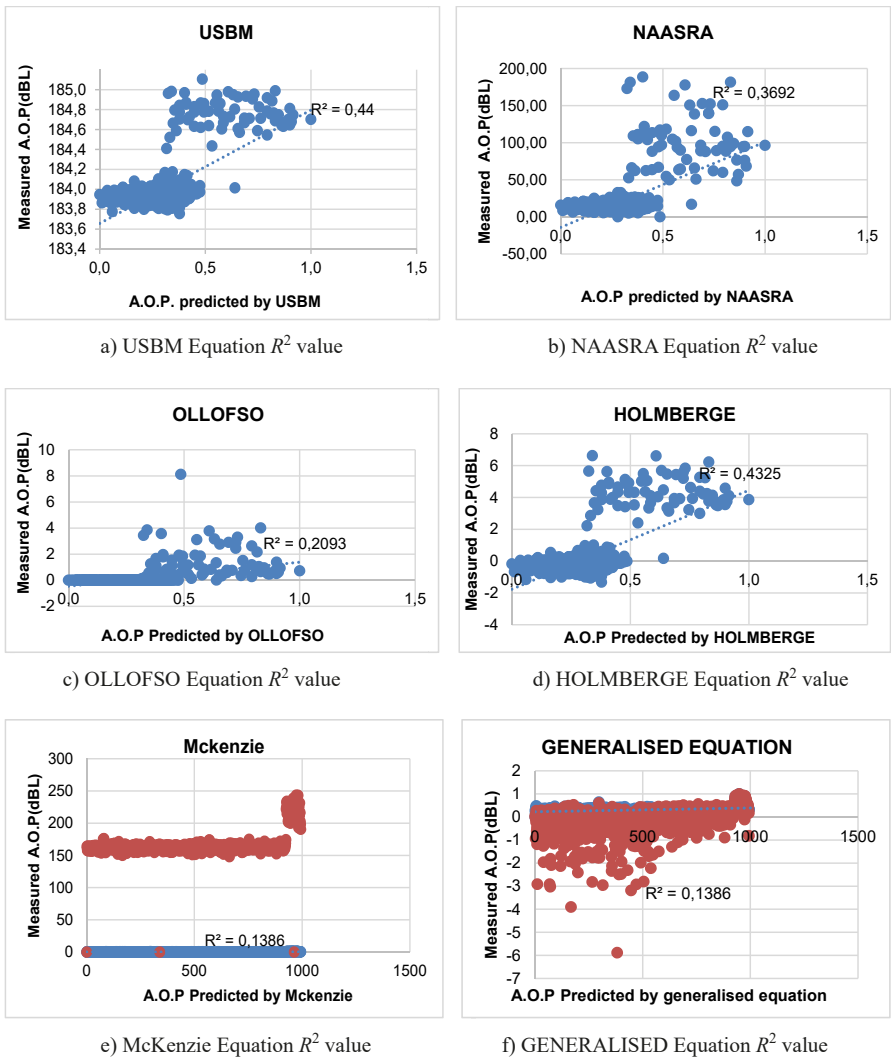


Fig. 6. Predicted Air Overpressure (AOp) values using statistical equations, with R^2 indicating predictive accuracy

9.2. The performance evaluation of the developed statistical models and Machine learning models

The performance evaluation compares the predictive accuracy of both statistical and machine learning models in estimating AOp, including R^2 values, MAD – Mean Absolute Deviation, RMSE – Root Mean Square Error, and NRMSE – Normalised Root Mean Square Error. Fig 7. Demonstrates Statistical Prediction Equations and Corresponding R^2 Values. It assesses how well each model predicts observed AOp values, highlighting the strengths and limitations of each approach. TABLE 13 compares the predictive performance of developed statistical and machine learning models for AOp based on evaluation metrics like R^2 and error rates.

TABLE 13

Evaluation of the Performance of Developed Statistical Models

Statistical Models	R^2	MAD	RMSE	NRMSE
USBM	0.44	1.997855	1.412697	1.188569
NAASRA	0.3692	1.017779	1.227521	1.107935
OLLOFSO	0.2093	2.417985	1.735132	1.317244
HOLMBERGE	0.4325	1.142856	1.010152	1.005063
MCKENZIE	0.1386	1.86096	1.319565	1.148723
GENERALISED EQUATION	0.1386	1.184724	1.016918	1.008424

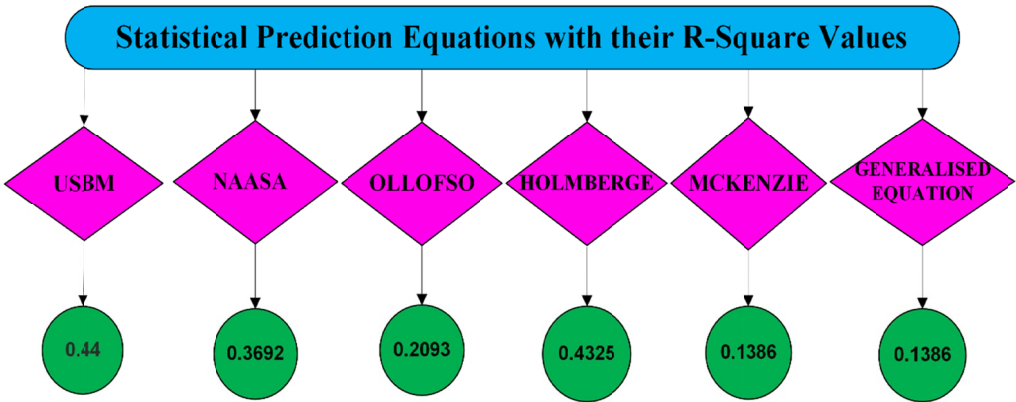


Fig. 7. Statistical prediction Equations with R^2 values

9.3. Machine Learning Algorithm Results

Machine learning algorithm results typically include metrics like R^2 , MSE, MAE, and RMSE to evaluate model performance (TABLE 14). Assessment of the Performance of the Developed Machine Learning Models. These metrics help compare algorithms such as SVR, Gradient Boosting, Random Forest, Backpropagation, and XG Boost-Random Forest to determine the best-performing model. Fig. 13 demonstrates Machine learning models along with their R^2 values.

TABLE 14

Evaluation of the Performance of Developed Machine Learning

ML Algorithms	R^2	MSE	MAE	RMSE
SVR	0.9324	1.7591	1.3227	1.3263
Random forest	0.5267	17.7087	3.2990	4.2081
Gradient Boosting	0.9970	0.8586	0.5708	0.9266
Backpropagation	0.9898	0.0025	0.0504	0.0505
XGBoost-RF	0.9991	0.2176	0.2828	0.4665

9.3.1. Support Vector Regression

The analysis of results is one of the most critical components of this study. The dataset was divided into 70% for training and 30% for testing to ensure a balanced evaluation of model performance. For instance, the Support Vector Regression (SVR) model demonstrated strong predictive capability with the following evaluation metrics: R^2 Score: 0.9324, Mean Squared Error (MSE): 1.7591, Mean Absolute Error (MAE): 1.3227, and Root Mean Squared Error (RMSE): 1.3263. These results indicate that the SVR model is highly effective in capturing the underlying patterns of the dataset. A high R^2 value of 0.9324 suggests that the model explains approximately 93% of the variance in the output variable, which reflects strong prediction accuracy. Additionally, the relatively low values of MSE, MAE, and RMSE show that the errors between predicted and actual values are minimal, further validating the model's reliability. Fig. 8 Represent (a) Performance Evaluation of the Prediction Model: Scatter Plot of Actual vs Predicted Values (b) Range of values in the SVR mode

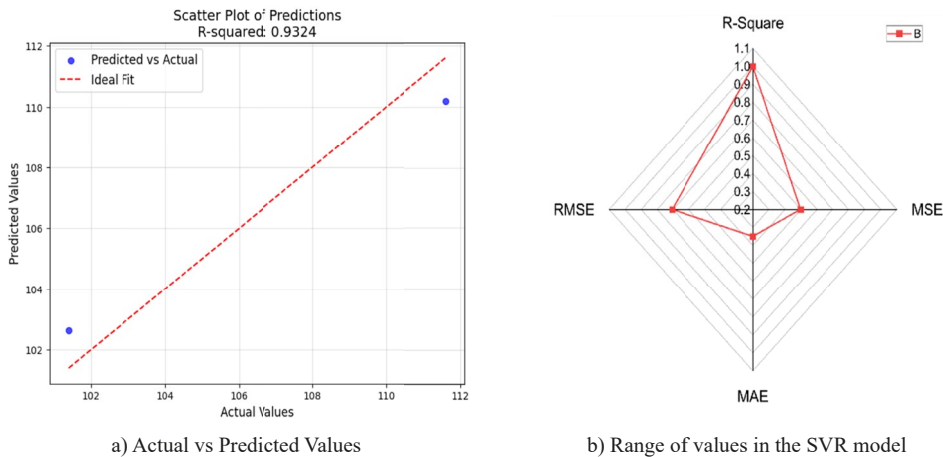


Fig. 8. Comparison of Actual and Predicted Values and Value Range in the SVR Model

9.3.2. Gradient Boosting

The analysis of results is one of the most critical components of this study. The dataset was divided into 70% for training and 30% for testing to ensure a balanced evaluation of model

performance. Among the models evaluated, the Gradient Boosting model delivered outstanding performance. The key performance indicators for this model are as follows: R^2 Score: 0.9970, Mean Squared Error (MSE): 0.8586, Mean Absolute Error (MAE): 0.5708, Root Mean Squared Error (RMSE): 0.9266. An R^2 score of 0.9970 signifies that the model accounts for nearly 100% of the variance in the target variable, indicating an exceptionally high level of predictive accuracy. The low error metrics, particularly the MSE, MAE, and RMSE, further confirm that the predicted values are very close to the actual values. Fig. 9 illustrates (a) GB actual vs predicted values and (b) GB value and ranges.

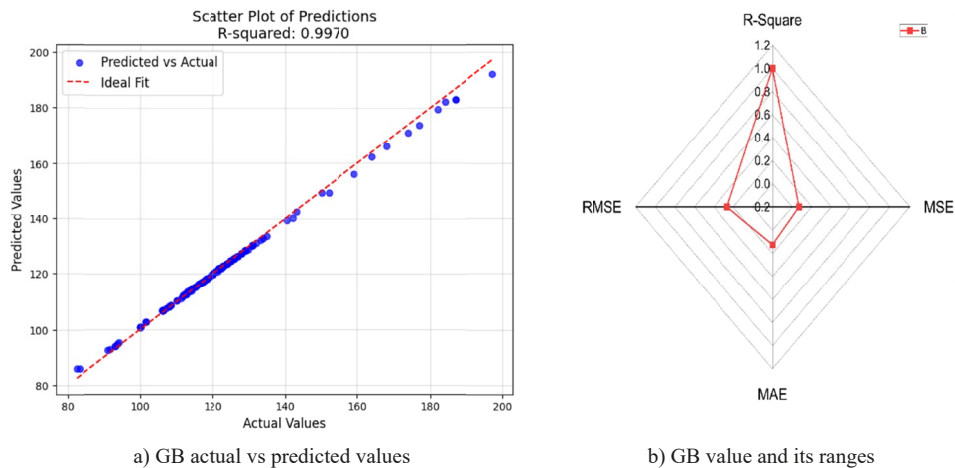


Fig. 9. GB Actual vs Predicted Values and Their Range

9.3.3. Random Forest

A thorough analysis of the results forms a crucial part of this study. To ensure a balanced and reliable evaluation of model performance, the dataset was split into 70% for training and 30% for testing. The Random Forest model, in contrast to other algorithms in the study, showed moderate predictive performance. The evaluation metrics are as follows: R^2 of 0.5267, with a Mean Squared Error (MSE) of 17.7087, a Mean Absolute Error (MAE) of 3.2990, and a Root Mean Squared Error (RMSE) of 4.2081. An R^2 value of 0.5267 indicates that the model could explain only about 52% of the variance in the output variable, reflecting a limited ability to capture the complexity of the dataset. The relatively high values of MSE and RMSE suggest a greater deviation between predicted and actual values, while the MAE confirms that the model struggled to produce precise predictions. As depicted in Fig. 10, random forests produced results within the specified value range.

9.3.4. Backpropagation neural networks

A comprehensive analysis of the results constitutes a key aspect of this research. To facilitate a robust assessment of model performance, the dataset was partitioned into 70% for training and

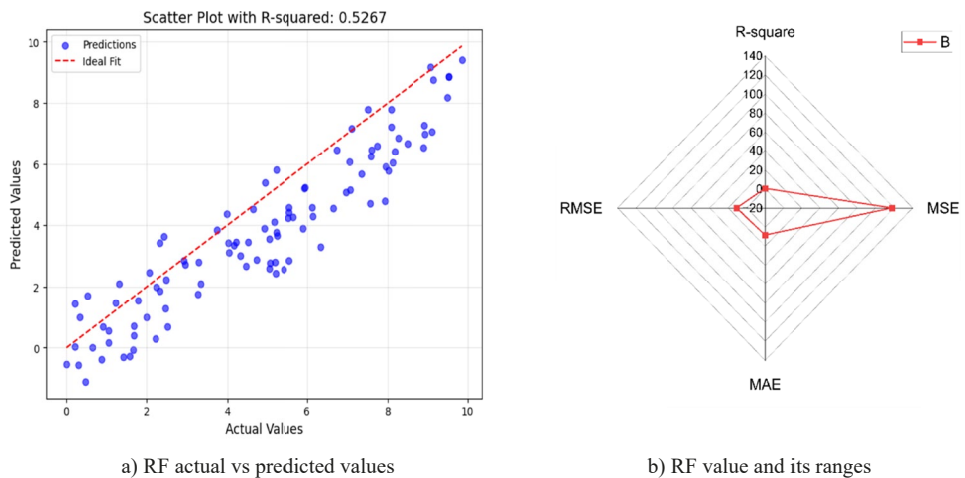


Fig. 10. Random Forest achieved results in the value range

30% for testing. The Backpropagation Neural Networks (BPNN) model demonstrated highly accurate predictive performance. The evaluation metrics are as follows: R^2 of 0.9898, with a Mean Squared Error (MSE) of 0.0025, a Mean Absolute Error (MAE) of 0.0504, and a Root Mean Squared Error (RMSE) of 0.0505. An R^2 value of 0.9898 indicates that the model successfully explains nearly 99% of the variation in the target variable, which reflects a very high level of prediction accuracy. The extremely low values for MSE, MAE, and RMSE further validate the model's precision, showing minimal differences between predicted and actual values. Fig. 11 demonstrates that the Backpropagation neural network model achieved results within the specified value range.

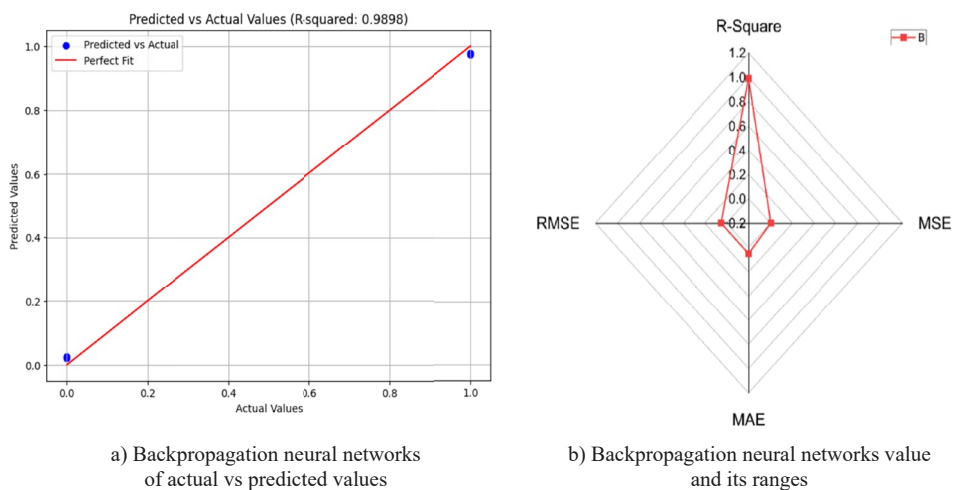


Fig. 11. The Backpropagation model results fall within the specified range

9.3.4. XGBoost- Random Forest

This study includes an important analysis of the results. The dataset was split into 70% for training and 30% for testing to check how well the model performs. The XGBoost-Random Forest hybrid model demonstrated exceptional predictive accuracy, outperforming the individual models evaluated in this study. Its performance metrics are as follows: R^2 of 0.9991, with a Mean Squared Error (MSE) of 0.2176, a Mean Absolute Error (MAE) of 0.2828, and a Root Mean Squared Error (RMSE) of 0.4665. An R^2 value of 0.9991 indicates that the model explains more than 99.9% of the variance in the target variable, showcasing near-perfect predictive power. The remarkably low error metrics, MSE, MAE, and RMSE, suggest minimal deviation between the predicted and actual values, affirming the model's high precision. Fig. 12 exhibits the predicted versus actual value, b) XGBoost-RF value and its ranges.

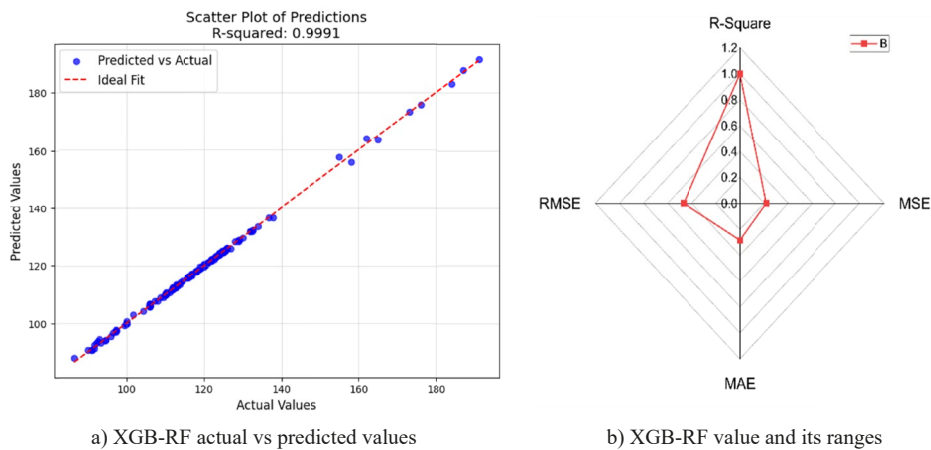


Fig. 12. XGB-RF Actual vs Predicted Values and Their Ranges

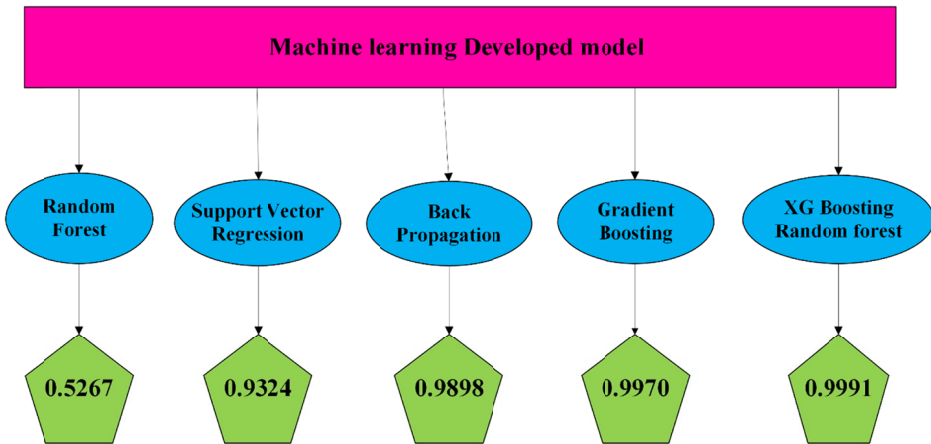


Fig. 13. Machine learning models along with their R^2 values

10. Future work

1. **Dataset Expansion:** Future work will focus on expanding the dataset by incorporating new input fields that influence air noise in mining operations.
2. **Inclusion of Additional Variables:** The model will be enhanced by adding more detailed data, including variables like air conditions, rock characteristics, and precise blast timing.
3. **Improving Model Accuracy:** To enhance model accuracy, we are collecting more comprehensive data to improve the design's predictive precision.
4. **Integration of Real-time Monitoring:** Future efforts will explore integrating real-time monitoring systems to enable more dynamic and timely data collection.
5. **Development of Deep Learning Hybrid Models:** Investigating the implementation of deep learning hybrid models will enable the processing of more granular data, significantly enhancing the model's forecasting capabilities and improving predictive accuracy.
6. **Safety and Environmental Impact:** These developments will lead to a more resilient and flexible model, enhancing safety in mining operations while reducing environmental impacts.

11. Conclusions

- Air noise prediction in mining is crucial for minimising environmental and health impacts caused by blast-induced AOp. Accurate prediction models help mine operators comply with regulatory limits, reduce noise pollution, and improve community relations by proactively managing noise levels during blasting operations.
- Collecting data from mining operations for air noise prediction is challenging due to the dynamic and hazardous environment, requiring specialised equipment and precise calibration to capture accurate measurements under varying field conditions. The collected data from these parameters, such as Holes, Depth, Spacing, Burden, Stemming, Decking, Total Drill, Explosive, Volume, Powder Factor, and Av. CPH, MCPD, and Seismic Distance.
- Statistical prediction equations, like those developed by USBM, NAASRA, Olofsson, Holmberg, and McKenzie, are Generalized Equations. They primarily rely on two key input parameters: burden and distance, as highlighted in prior studies. However, these equations often have limitations in accuracy and may produce errors. To address these issues, machine learning techniques offer a promising alternative, allowing for more precise predictions by incorporating additional variables and complex patterns beyond traditional empirical equations.
- To overcome the limitations of traditional statistical prediction equations, machine learning algorithms like Support Vector Regression (SVR), Random Forest (RF), Gradient Boosting (GB), Backpropagation Neural Networks, and ensemble hybrid models such as XGBoost-RF can be developed for more accurate predictions. These algorithms can handle complex, nonlinear relationships and incorporate a wide range of variables, making them well-suited for predicting Air Overpressure in diverse geological settings.
- While algorithms like SVR, RF, GB, and Backpropagation Neural Networks achieved moderate accuracy in predicting Air Overpressure, an ensemble model combining XGBoost and Random Forest produced significantly improved accuracy, making it the most effective approach for this prediction task.

- The models' performance was evaluated on the test data using four metrics: RMSE, MAE, MSE, and R^2 . To develop a more comprehensive prediction model in the future, considering various blast-induced effects.

References

- [1] P. Ragam et al., Estimation of slope stability using ensemble-based hybrid machine learning approaches. *Front. Mater.* **11**, (2024). DOI: <https://doi.org/10.3389/fmats.2024.1330609>
- [2] C. Mueller, *Mining goes digital*. CRC Press / Taylor & Francis Group, 2019.
- [3] K. Manoj, M. Monjezi, Prediction of flyrock in open pit blasting operation using machine learning method. *Int. J. Min. Sci. Technol.* **23**, 3, 313-316 (2013). DOI: <https://doi.org/10.1016/j.ijmst.2013.05.005>
- [4] V. Munagala, S. Thudumu, I. Logothetis, S. Bhandari, R. Vasa, K. Mouzakis, A comprehensive survey on machine learning applications for drilling and blasting in surface mining. *Mach. Learn. with Appl.* **15**, 100517, (2024). DOI: <https://doi.org/10.1016/j.mlwa.2023.100517>
- [5] S.A. Baghaei Naeini, A. Badri, Identification and categorization of hazards in the mining industry: A systematic review of the literature. *Akademai Kiado ZRt. Jan.* **22**, (2024). DOI: <https://doi.org/10.1556/1848.2023.00621>
- [6] P. Ragam, D.S. Nimaje, Monitoring of blast-induced ground vibration using WSN and prediction with an ANN approach of ACC dungri limestone mine, India. *J. Vibroengineering* **20**, 2, 1051-1062 (2018). DOI: <https://doi.org/10.21595/jve.2017.18647>
- [7] V.A. Temeng, Y.Y. Ziggah, C.K. Arthur, A novel artificial intelligent model for predicting air overpressure using brain inspired emotional neural network. *Int. J. Min. Sci. Technol.* **30**, 5, 683-689 (2020). DOI: <https://doi.org/10.1016/j.ijmst.2020.05.020>
- [8] P.J. Wen, C. Huang, Noise prediction using machine learning with measurements analysis. *Appl. Sci.* **10**, 18 (2020). DOI: <https://doi.org/10.3390/AP10186619>
- [9] D. Prasad Tripathy, S. Kumar Nanda, *Noise Identification, Modeling and Control In Mining Industry*. (2024).
- [10] H. Dehghani, F. Ali Mohammad Nia, Estimation of air overpressure using bat algorithm. *Min. Sci.* **28**, 7, 7-92, (2020). DOI: <https://doi.org/10.37190/msc212806>
- [11] C. Sawmliana, P. Pal Roy, R.K. Singh, T.N. Singh, Blast induced air overpressure and its prediction using artificial neural network. *Trans. Institutions Min. Metall. Sect. A Min. Technol.* **116**, 2, 41-48 (2007). DOI: <https://doi.org/10.1179/174328607X191065>
- [12] H. Nguyen, X.N. Bui, C. Drebenstedt, Y. Choi, Enhanced Prediction Model for Blast-Induced Air Over-Pressure in Open-Pit Mines Using Data Enrichment and Random Walk-Based Grey Wolf Optimization–Two-Layer ANN Model. *Nat. Resour. Res.* **33**, 2, 943-972 (2024). DOI: <https://doi.org/10.1007/s11053-023-10299-w>
- [13] R. Zhang, Y. Li, Y. Gui, Prediction of rock blasting induced air overpressure using a self-adaptive weighted kernel ridge regression. *Appl. Soft Comput.* **148**, (2023). DOI: <https://doi.org/10.1016/j.asoc.2023.110851>
- [14] C. Chewu, T. Chikwere, D. Runganga, Elia Chipfupi, T. Nyamagudza, Prediction of Flyrocks, Airblasts and Ground Vibrations Using Neural Computing and Applications at ZCDC Mine. *Malaysian J. Sci. Adv. Technol.*, 212-221, (2023). DOI: <https://doi.org/10.56532/mjsat.v3i3.135>
- [15] S. Hosseini et al., Mine Induced Airblast prediction: An Application of Chaos Game Optimization based soft computing approaches. May 30, (2023). DOI: <https://doi.org/10.21203/rs.3.rs-2992457/v1>
- [16] K.F. Lima, A. da C. Meireles, N. Barbieri, L.D. Fiorentin, Ground vibration and air overpressure prediction applied to a blasting operation in a Gneiss quarry in southern Brazil. Feb. 05, (2024). DOI: <https://doi.org/10.21203/rs.3.rs-3914158/v1>
- [17] C. Kuzu, A. Fiske, S.G. Ercelebi, Operational and geological parameters in the assessing blast induced airblast-overpressure in quarries. *Appl. Acoust.* **70**, 3, 404-411 (2009). DOI: <https://doi.org/10.1016/j.apacoust.2008.06.004>
- [18] M.M.K. Kazemi, Z. Nabavi, M. Khandelwal, Prediction of blast-induced air overpressure using a hybrid machine learning model and gene expression programming (GEP): A case study from an iron ore mine. *AIMS Geosci.* **9**, 23, 57-381 (2023). DOI: <https://doi.org/10.3934/geosci.2023019>

- [19] S. Jayanthu, C. Naveen, G.V. Rao, B.R.V. Susheel Kumar, Ground vibrations in opencast mine blast on structures vis-à-vis a local environmental effect and its mitigation through mining technology.
- [20] P.P. Roy, C. Sawmliana, R.K. Singh, Strategic planning to reduce ground vibration, air overpressure and flyrock in a mine at a sensitive area. *Curr. Sci.* **123**, 8, 995-1004 (2022). DOI: <https://doi.org/10.18520/cs/v123/i8/995-1004>
- [21] X. Zhou et al., Propagation characteristics and prediction of airblast overpressure outside tunnel: a case study. *Sci. Rep.* **12**, 1 (2022). DOI: <https://doi.org/10.1038/s41598-022-24917-9>
- [22] N.K. Dumakor-Dupey, S. Arya, A. Jha, Advances in blast-induced impact prediction – a review of machine learning applications. *Minerals* **11**, 6 (2021). DOI: <https://doi.org/10.3390/min11060601>
- [23] M. Khandelwal, P.K. Kankar, Prediction of blast-induced air overpressure using support vector machine. *Arab. J. Geosci.* **4**, 3-4, 427-433 (2011). DOI: <https://doi.org/10.1007/s12517-009-0092-7>
- [24] P. Hajikhodaverdikhan, M. Nazari, M. Mohsenizadeh, S. Shamshirband, K.W. Chau, Earthquake prediction with meteorological data by particle filter-based support vector regression. *Eng. Appl. Comput. Fluid Mech.* **12**, 1, 679-688 (2018). DOI: <https://doi.org/10.1080/19942060.2018.1512010>
- [25] V.A. Temeng, C.K. Arthur, Y.Y. Ziggah, Suitability assessment of different vector machine regression techniques for blast-induced ground vibration prediction in Ghana. *Model. Earth Syst. Environ.* **8**, 1, 897-909 (2022). DOI: <https://doi.org/10.1007/s40808-021-01129-0>
- [26] S. Hosen, R. Amin, Significant of Gradient Boosting Algorithm in Data Management System. *Eng. Int.* **9**, 2 (2021).
- [27] G. Komarasamy, S.C.B. Jaganathan, K. Sridharan, A. Mital, S. Awal, Harmony Gradient Boosting Random Forest Machine Learning Algorithms for Sentiment Classification. In *Proceedings – 2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security, iSSSC 2022*, Institute of Electrical and Electronics Engineers Inc., (2022). DOI: <https://doi.org/10.1109/iSSSC56467.2022.10051210>
- [28] J. Renaud, R. Karam, M. Salomon, R. Couturier, Deep learning and gradient boosting for urban environmental noise monitoring in smart cities. *Expert Syst. Appl.* **218**, May (2022). DOI: <https://doi.org/10.1016/j.eswa.2023.119568>
- [29] A.-L. Boulesteix, S. Janitza, J. Kruppa, I.R. König, Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics pre-review version of a manuscript accepted for publication in *WIREs Data Mining & Knowledge Discovery*. 2012. [Online]. Available: <http://www.stat.uni-muenchen.de>
- [30] M. Savargiv, B. Masoumi, M.R. Keyvanpour, A new random forest algorithm based on learning automata. *Comput. Intell. Neurosci.* **2021**, (2021). DOI: <https://doi.org/10.1155/2021/5572781>
- [31] W. Lin, Z. Wu, L. Lin, A. Wen, J. Li, An ensemble random forest algorithm for insurance big data analysis. *IEEE Access* **5**, 16568-16575 (2017). DOI: <https://doi.org/10.1109/ACCESS.2017.2738069>
- [32] P. Ragam, A.R. Komalla, N. Kanne, Estimation of blast-induced peak particle velocity using ensemble machine learning algorithms: A case study. *Noise Vib. Worldw.* **53**, 7-8, 404-413 (2022). DOI: <https://doi.org/10.1177/09574565221114662>
- [33] <https://github.com/kamleshitb/Ground-Vibration-Prediction-in-blasting-using-Neural-Network/blob/main/Blasting%20dataset.csvp>