mining

MDPI

Article

Predicting Blasting-Induced Ground Vibration in Mines Using Machine Learning and Empirical Models: Advancing Sustainable Mining and Minimizing Environmental Footprint

Nafiu Olanrewaju Ogunsola [1,2,*] and Hendrik Grobler [1]

[1] Department of Mining Engineering and Mine Surveying, University of Johannesburg, Doornfontein Campus, Johannesburg 17011, South Africa
[2] Sibanye-Stillwater Center for Sustainable Mining, University of Johannesburg, Doornfontein Campus, Johannesburg 17011, South Africa
* Correspondence: nogunsola@uj.ac.za

Abstract

Blasting-induced ground vibrations, typically quantified by peak particle velocity (PPV), pose one of the most critical environmental challenges in surface mining and can damage nearby structures and disrupt surrounding ecosystems. Consequently, the development of reliable and accurate predictive models is essential for designing safe, environmentally responsible, and sustainable blasting operations. This study develops a robust predictive framework using a harmonized database of 506 blasting events, from which 386 high-quality records were retained after preprocessing to model PPV as a function of charge per delay (Q), monitoring distance (R), and rock mass rating (RMR). Several machine learning (ML) algorithms, including artificial neural networks trained using the Levenberg–Marquardt algorithm (ANN-LM), adaptive neuro-fuzzy inference systems (ANFIS), Gaussian process regression (GPR), and decision trees (DT), were evaluated alongside conventional empirical models such as the USBM, Ambraseys–Hendron, Langefors–Kihlstrom, and BIS. To further enhance predictive capability, two optimization strategies, Bayesian optimization (BO) and differential evolution (DE), were applied to the GPR model, producing optimized BO-GPR and DE-GPR variants. Model performance was assessed using the correlation coefficient (r), variance accounted for (VAF), mean absolute error (MAE), and relative root mean square error (RRMSE). Results indicate that the BO-GPR model achieved the best predictive performance during testing for both the two-input (Q, R) and three-input (Q, R, RMR) configurations, with r values of 0.97426 and 0.98381, respectively, and VAF values exceeding 94%. SHAP analysis revealed monitoring distance as the dominant attenuating factor controlling PPV. The optimized framework provides an accurate, interpretable tool for vibration prediction and precision blast design, supporting environmentally responsible, sustainable mining operations.

Keywords: peak particle velocity (PPV); blasting-induced ground vibration (BIGV); Bayesian optimization (BO); differential evolution; rock mass rating (RMR); sustainable mining; shapley additive explanations (SHAP)

Check for updates

Academic Editors: María de la Luz García Lorenzo, José María Esbrí, Oscar Andreu Sánchez and Francisco J. Martínez Moreno

Received: 17 March 2026
Revised: 17 April 2026
Accepted: 5 May 2026
Published: 7 May 2026

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

1. Introduction

Blasting is an essential and widely employed technique in mining operations, primarily used for breaking rock in both surface and underground mines. The detonation of

explosives releases significant energy, effectively fracturing rock masses for mineral extraction, but also produces undesirable side effects, such as noise, dust, flyrock, and ground vibrations. The quality of rock fragmentation achieved during blasting fundamentally influences the performance and efficiency of all mining subsystems, including loading, hauling, and crushing operations [1]. Achieving an optimal size distribution of fragmented rock facilitates smoother material handling and transport, significantly enhancing overall economic performance by minimizing equipment wear, reducing energy consumption, and improving throughput rates [2]. Among the side effects of blasting, ground vibration, also commonly known as peak particle velocity (PPV), is particularly significant because it can cause structural damage, harm ecosystems, and lead to social conflicts in nearby communities [3,4]. Effective blast design and control, including controlled and sequential blasting techniques such as pre-splitting and cushion blasting, are implemented to mitigate these impacts. Understanding these interlinked factors is vital for maximizing mining efficiency, ensuring operational safety, and promoting sustainable practices in the mining industry.

PPV has been identified as a critical parameter for assessing the impact of blasting-induced ground vibrations on structures, machinery, and the natural environment in mining areas. Accurate prediction of these vibrations is essential for effective risk management and maintaining operational safety. Elevated PPV levels can cause both temporary (reparable) and permanent (irreparable) damage to built infrastructure, particularly when the vibration frequency aligns with a structure's natural frequency, resulting in resonance and an increased risk of damage. Even PPV levels below structural damage thresholds can induce significant psychological distress among occupants of affected buildings [5]. Such vibrations often cause annoyance and concern among residents near mining sites, sometimes leading to legal and regulatory challenges that delay mining operations and increase costs. Furthermore, high-intensity ground vibrations can have a detrimental impact on the local ecology, contributing to habitat disruption and long-term environmental degradation. It is widely recognized that issues related to PPV are a major source of civil litigation and disputes arising from blasting activities in both civil construction and mining sectors worldwide [4,6]. These vibrations propagate as seismic waves through the Earth, and their magnitude and frequency depend on several factors, including the charge weight, blasting technique, distance from the blast, and the area's geological characteristics, including weathering [7,8]. Accurate prediction capabilities enable mining engineers to design blasting plans that minimize adverse impacts on nearby structures and ecosystems. This, in turn, helps maintain social licenses to operate by addressing community concerns through regulated vibration limits and informed blast scheduling.

In recent years, sustainable and responsible mining has become a primary objective for researchers and industry stakeholders, driven by growing awareness of the environmental consequences of mining, particularly surface mining [9–11]. Sustainable mining practices aim to minimize ecological harm, optimize resource use, reduce carbon emissions, and promote positive social and economic outcomes. These approaches ensure that mining operations align with broader environmental and societal expectations while meeting the global demand for minerals. Key strategies include adopting cleaner energy sources, efficient resource use, land reclamation, biodiversity conservation, and reducing pollutant emissions. Effective management of blasting impacts, such as mitigating noise pollution, controlling dust and flyrock, and minimizing ground vibrations, directly supports these sustainability goals by preserving natural habitats and reducing the environmental footprint of mining operations. Consequently, minimizing adverse blasting effects not only protects local ecosystems and communities but also advances the overarching principles of responsible resource extraction that respect ecological boundaries and social licenses to operate [12,13].

Field measurements or observations using calibrated equipment, such as seismographs and sensors, are the most accurate means of assessing the environmental impacts of PPV. However, measuring PPV using seismographs and sensors requires specialized training and skills to set up the equipment, and it is also time-consuming. In place of direct field measurements, researchers have developed empirical equations to quantify PPV generation from blasting [14–18]. Traditional empirical models for predicting ground vibration are based on simplified relationships between blast parameters and the resulting vibration effects. While useful, these models often lack the flexibility to capture the nonlinear and site-specific nature of blasting phenomena [19–21]. This limitation can lead to inaccurate predictions in complex geological settings. Machine learning (ML) methods offer a promising alternative for estimating PPV from blasting by enabling the analysis of extensive datasets and parameter sets, as well as the identification of intricate patterns within the input variables [22–24]. Integrating machine learning with empirical models combines the strengths of both approaches, improving prediction accuracy and robustness. These hybrid methods can adapt to local conditions, optimize blast designs in real time, and facilitate advanced monitoring and control of blasting operations, ultimately enhancing both environmental and operational outcomes.

This study aims to develop an integrated approach to predict blasting-induced ground vibration by combining empirical formulas with advanced machine learning techniques. The objectives include improving the precision of vibration forecasts, providing actionable insights for blast optimization, and supporting sustainable mining practices by mitigating environmental impacts. The scope encompasses analyzing a comprehensive dataset from mining operations, validating prediction models across diverse conditions, and demonstrating the practical advantages of the combined modeling framework. By addressing these goals, the study contributes to advancing safer, more efficient, and environmentally responsible mining practices that protect both the natural environment and surrounding communities.

This study advances prior PPV prediction research by harmonizing a diverse 506-event database from multiple literature sources for enhanced generalizability, engineering a novel RMR input feature to capture rock mass quality effects beyond traditional Q-R parameters, rigorously comparing four ML models against calibrated empirical benchmarks, and deriving actionable insights for RMR-informed Q/R ratio designs that minimize ecological disruptions and support regulatory compliance.

2. Literature Review

This structured review balances foundational empirical insights, advanced ML innovations, and the imperative for sustainability in blasting-related geomechanical and environmental management literature for mining contexts.

2.1. Summary of Existing Empirical Models for Predicting PPV from Blasting

Empirical models for predicting PPV primarily rely on blast design parameters, such as charge weight, distance from the blasting site, and geological characteristics. Classic models, including the scaled-distance formula, relate PPV to the explosive charge weight per delay and distance using regression equations (Equation (1)). Other empirical models for estimating PPV from blasting are derived using the standard scaled-distance formula in Equation (1), and several are presented in Table 1. These empirical models provide rapid approximations for predicting PPV in blasting operations; however, they often lack flexibility to accommodate site-specific conditions and geological complexities, as they are typically calibrated using field measurements and may oversimplify and be limited in applicability across diverse mining environments.

$$v = kD^{-b} \quad (1)$$

where v represents the PPV (mm/s), and D denotes the scaled distance (m/kg^(1/2)), which is defined as the ratio of the distance from the blast source, R (m), to the square root of the explosive charge mass, Q (kg), measured as the explosive net equivalent charge weight, i.e., Equation (2):

$$D = \frac{R}{\sqrt{Q}} \quad (2)$$

Table 1. Some proposed empirical equations for estimating PPV.

S/N	References	Equations/Estimators
1	[17]	$Q_{max} = k(vD^2)^b$
2	[25]	$v = 0.438D^{-1.52}$
3	[26]	$v = k(R/Q^{1/2})^{-b}e^{-\alpha\beta}$ $v = k(R/Q^{1/3})^{-b}e^{-\alpha\beta}$
4	[27]	$v = 0.828D^{-1.32}$
5	[28]	$v = k(Q^{2/3}/R)^{1.25}$
6	[29]	$v = 0.362D^{-1.63}$
7	[14]	$v = k(R/Q^{1/3})^{-b}$
8	[16]	$v = k(Q/R^{2/3})^{b/2}$
9	[30]	$v = k(R/Q^{1/2})^{-b}$

Note: v: peak particle velocity (mm/s); R: distance from blast source to recording location (m); Q_{max}: net equivalent charge per delay (kg); and k, b, α, and β: attenuation exponents (site-specific constants).

The parameters k and b are site-specific constants determined empirically through field calibration. Together, these constants capture the influence of site characteristics and blasting conditions on vibration levels, allowing the formula to be tailored for accurate prediction of PPV specific to the mining location.

2.2. Review of Machine Learning Applications in Blasting Vibration Prediction and Mining Environmental Management

Machine learning (ML) methods offer notable advancements over traditional empirical models by capturing nonlinear and complex relationships among blasting parameters and vibrational outcomes. ML techniques such as artificial neural networks (ANN), adaptive neuro-fuzzy inference systems (ANFIS), support vector regression (SVR), Gaussian process regression (GPR), and hybrid algorithms have been widely applied to predict PPV with improved accuracy. ML models are trained on large datasets from blast monitoring, enabling better handling of multifactorial influences and site-specific nuances. Beyond prediction, these approaches are integrated into environmental management frameworks for mining to optimize blasting plans, minimize adverse effects on surrounding infrastructure, and ensure compliance with regulatory thresholds, thereby contributing to safer mining operations. Some ML methods reported in the literature for PPV prediction are highlighted in Table 2.

Table 2. Selected ML techniques for predicting PPV in literature.

Citations	ML Methods	Predictor Variables	Evaluation Results
[31]	ANN-KNN	Q, R	$R^2 = 0.88$
[22]	CART	Q, R	$R^2 = 0.95$
[19]	ICA	Q, R	$R^2 = 0.94$
[20]	ANN, KNN, CART, SVM	Q, R	$R^2 = 0.98$

Table 2. *Cont.*

Citations	ML Methods	Predictor Variables	Evaluation Results
[24]	ANN, MLR	Q, R	R^2 = 0.98
[23]	GEP, ANFIS, SCA-ANN	Q, R, ρ, SRH	R^2 = 0.99
[32]	MARS, CART, SVR, MLR	H, D, N, B, S, SL, R, CPH, TC, MCPD	R^2 = 0.95
[33]	GPR, ELM, ANN, MARS	D, B, S, R, PF, SL, MCPD	R^2 = 0.99
[34]	GPR, DT, SVR	N, B, S, R, SD, BH/B, B/H	R^2 = 0.94
[35]	RF	MCPD_Em, MCPD_ANFO, Qv_Em, Qv_ANFO, R, c_m, c_r, ρ, η	r = 0.95
[36]	TSO-SVR, WOA-SVR, CS-SVR	R, Q, PF, S/B, BH/B, B/H, SL/B, U/B	R^2 = 0.95
[37]	PINN, XGBoost, PINN-XGBoost	N, B, S, Q, R, D, SL	R^2 = 0.92
[38]	GPR, MLR	B, S, SL, Q, R, D	R^2 = 0.91
[39]	GWO-SVR, PSO-SVR, HHO-SVR	B, S, SL, L, PF, R, MCPD, Q	R^2 = 0.92
[40]	MLR	B, S, D, Q, R, TC, L, NR	R^2 = 0.82
[41,42]	XGBoost, RF, DF	Q, R	R^2 = 0.78

Note: *R*: distance from blast source to recording location (m); *Q*: charge per delay; *ρ*: rock density; SRH: Schmidt rebound hardness; GEP: gene expression programming; ANFIS: adaptive neuro-fuzzy inference system; SCA-ANN: sine cosine algorithm optimized artificial neural network; MARS: multivariate adaptive regression splines; CART: classification and regression tree; SVR: support vector regression; MLR: multilinear regression; H: hole diameter; D: hole depth; N: number of holes; B: burden; S: spacing; SL: stemming length; CPH: charge per hole; TC: total charge; MCPD: maximum charge per delay; GPR: Gaussian process regression; ELM: extreme learning machine; PF: powder factor; DT: decision tree; SD: scaled distance; BH: bench height; BH/B: bench height to burden ratio; B/H: burden to hole diameter ratio; RF: random forest; MCPD_Em: maximum explosive charge fired per delay for emulsion; MCPD_ANFO: maximum explosive charge fired per delay for ANFO; Qv_Em: explosion heat of explosive at constant volume for emulsion; Qv_ANFO: explosion heat of explosive at constant volume for ANFO; c_m: longitudinal wave velocity of the rock mass; c_r: longitudinal wave velocity of the rock; η: integrity coefficient of the rock mass ($\eta = (c_m/c_r)^2$); TSO-SVR: tuna swarm optimization optimized-SVR; WOA-SVR: whale optimization algorithm optimized-SVR; CS-SVR: cuckoo search optimized-SVR; S/B: spacing to burden ratio; SL/B: stemming length to burden ratio; U: subdrilling; U/B: subdrilling to burden ratio; PINN: physics informed neural networks; XGBoost: LSTM: long short-term memory; PINN-XGBoost: extreme gradient boosting; hybrid PINNs + XGBoost; GWO-SVR: grey wolf optimization optimized-SVR; PSO-SVR: particle swarm optimization optimized-SVR; HHO-SVR: Harris hawks optimization optimized-SVR; L: charge length; TC: total charge per round; NR: number of rows; R^2 = coefficient of determination; *r*: correlation coefficient; RF: random forest; DF: deep forest.

2.3. Discussion on Sustainable Mining Practices Related to Blasting and Vibration Control

Environmental considerations, alongside the physical and geological characteristics of the deposit, ground conditions (including the hanging wall, footwall, and ore zone), and operational factors such as mining and capital costs, labor availability, and other costs, constitute critical determinants in selecting mining methods [43]. Among these, environmental factors hold paramount importance due to their direct impacts on human health and ecosystems. Blasting operations in open-pit mining induce varied environmental consequences, including land degradation, ecosystem disruption, ground vibration, airblast overpressure, noise, dust generation, the release of noxious gases, and water and soil pollution [44].

Sustainable resource management, articulated under Sustainable Development Goal (SDG) 12, prioritizes minimizing adverse environmental impacts and fostering responsible resource consumption. In mining, sustainable practices integrate operational improvements that enhance environmental and social outcomes relative to traditional methods, while ensuring worker safety and community well-being [45]. Key strategies involve reducing

resource inputs, limiting land disturbance, mitigating harmful emissions, and ensuring responsible reclamation and rehabilitation of mined sites [46]. Sustainable mining, particularly in blasting and vibration control, encompasses precision blasting techniques that aim to minimize ground vibration and flyrock, thereby protecting local ecosystems and neighboring communities.

The adoption of predictive modeling and machine learning tools further facilitates proactive environmental management by enabling optimized blasting designs that reduce energy consumption and material waste. Continuous compliance with environmental standards, rigorous monitoring of noise and vibration pollution, and transparent stakeholder engagement are integral to sustainable mining operations. These efforts encompass broader environmental, social, and governance (ESG) commitments that align mining practices with global sustainability objectives while balancing economic viability and ecological integrity.

3. Field Blasting Database Description

The experimental dataset used in this study for training and developing machine learning models to predict blasting-induced peak particle velocity (PPV) was compiled from 12 peer-reviewed sources [47–58] (Table 3), selected for methodological rigor. This yielded 506 data points reported in the literature, with the raw data rigorously verified and harmonized to ensure consistency in units and measurement standards across all sources. PPV values were retained as reported (mm/s); Q was used as reported (kg) without TNT-equivalent conversion; and R was retained in meters (m). Only the International Organization for Standardization (ISO) standard for seismographs in blasting vibration monitoring compliant measurements (seismographs capturing maximum vector PPV) were retained, while screening excluded incomplete or outlying records, ensuring statistically significant sample sizes, appropriate statistical distributions, and data integrity. Multi-source harmonization enhances generalizability beyond site-specific models, enabling cross-geology PPV prediction for global mining applications. This consolidated database includes two primary input variables derived from rock blasting in open-pit mining contexts, charge per delay (Q) and monitoring distance (R), and a single target output, the PPV values measured in field blasting operations (ranging from 0.6 to 28.30 mm/s). Additionally, an engineered input parameter, rock mass rating (RMR), was derived by integrating blasting parameters with site-specific constants. The RMR was selected as an input variable because it serves as a proxy for rock mass. RMR directly quantifies the mechanical characteristics of the propagation medium. The robustness of this consolidation methodology across diverse datasets directly enhances the reliability and predictive performance of the developed ML models.

Table 3. Database compilation source, sample size, and range.

S/N	Citations	Sample Size	Measured PPV Range (mm/s)
1	[47]	43	2.03–28.44
2	[48]	32	0.19–19.60
3	[49]	49	1.10–5.80
4	[50]	20	0.95–92.30
5	[51]	108	0.60–45.20
6	[52]	54	1.02–69.00
7	[53]	30	0.14–5.46
8	[54]	12	4.06–186.00
9	[55]	18	2.92–41.40
10	[56]	41	1.80–33.10
11	[57]	88	2.40–32.10
12	[58]	11	0.19–8.60
Overall	–	506	0.14–186.00

Notably, the initial compiled database comprised 506 data points from 12 peer-reviewed sources. After data preprocessing, including the removal of incomplete records and outliers using the IQR method, a total of 386 data points were retained for model development. This refined dataset was used for all subsequent machine learning analyses.

3.1. Data Preprocessing and Outlier Detection and Removal Process Using the IQR Method

Outliers, i.e., data points significantly different from other observations, can heavily influence statistical analyses and machine learning model performance, leading to biased results. Therefore, identifying and appropriately handling these anomalies is a crucial step in data preprocessing. In this study, the Interquartile Range (IQR) method, a robust statistical technique for detecting and removing outliers from a dataset, was employed to ensure the integrity and reliability of subsequent machine learning analyses.

The IQR method is a widely used nonparametric technique for identifying outliers, as it is less sensitive to extreme values than methods based on the mean and standard deviation [59]. The process involves calculating the first quartile (Q1), the third quartile (Q3), and the Interquartile Range (IQR) itself. Q1 represents the 25th percentile of the data, meaning 25% of the data falls below this value. Q3 represents the 75th percentile, with 75% of the data falling below it. The IQR is the range between Q1 and Q3 (IQR = Q3 − Q1), encompassing the middle 50% of the data.

Outliers are then defined as data points that fall outside specific lower and upper bounds. These bounds are calculated using a multiplier (typically 1.5) applied to the IQR:

* ** Lower Bound: ** Q1 − 1.5 × IQR

* ** Upper Bound: ** Q3 + 1.5 × IQR

Any data point below the lower bound or above the upper bound is classified as an outlier.

The removal of 120 data points was not based solely on high PPV magnitudes but followed the Interquartile Range (IQR) method applied consistently across PPV, Q, and R. Data points were classified as outliers only when falling outside Q1 − 1.5 × IQR or Q3 + 1.5 × IQR, ensuring exclusion based on statistical inconsistency rather than absolute values. Outliers typically showed inconsistencies across multiple variables. Given the integration of 12 independent sources with varying geological and measurement conditions, such anomalies are likely due to inter-source variability rather than reproducible physical behavior, thereby justifying IQR filtering to improve dataset homogeneity and model generalization. While high-PPV events matter, this study focuses on typical operational blasting conditions.

To visually inspect the distribution of data and the presence of outliers before removal, box plots were generated for each numerical column in the original data frame (Figure 1). Box plots graphically display the five-number summary of a data set: minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum. Potential outliers are typically shown as individual points extending beyond the 'whiskers' of the box plot, which correspond to the calculated lower and upper bounds. These initial visualizations clearly indicated numerous data points outside the expected range for several variables, confirming the need for outlier treatment.

After identifying outlier bounds for each numerical column, a systematic approach was used to remove these anomalous data points. The original data frame was iterated through column by column. For each numerical column, rows containing values that fell outside their respective lower and upper IQR bounds were filtered out. This iterative filtering process ensured that only data points within the acceptable range for all numerical

features were retained. The box plot of the new input and output features after outliers' removal is presented in Figure 2.

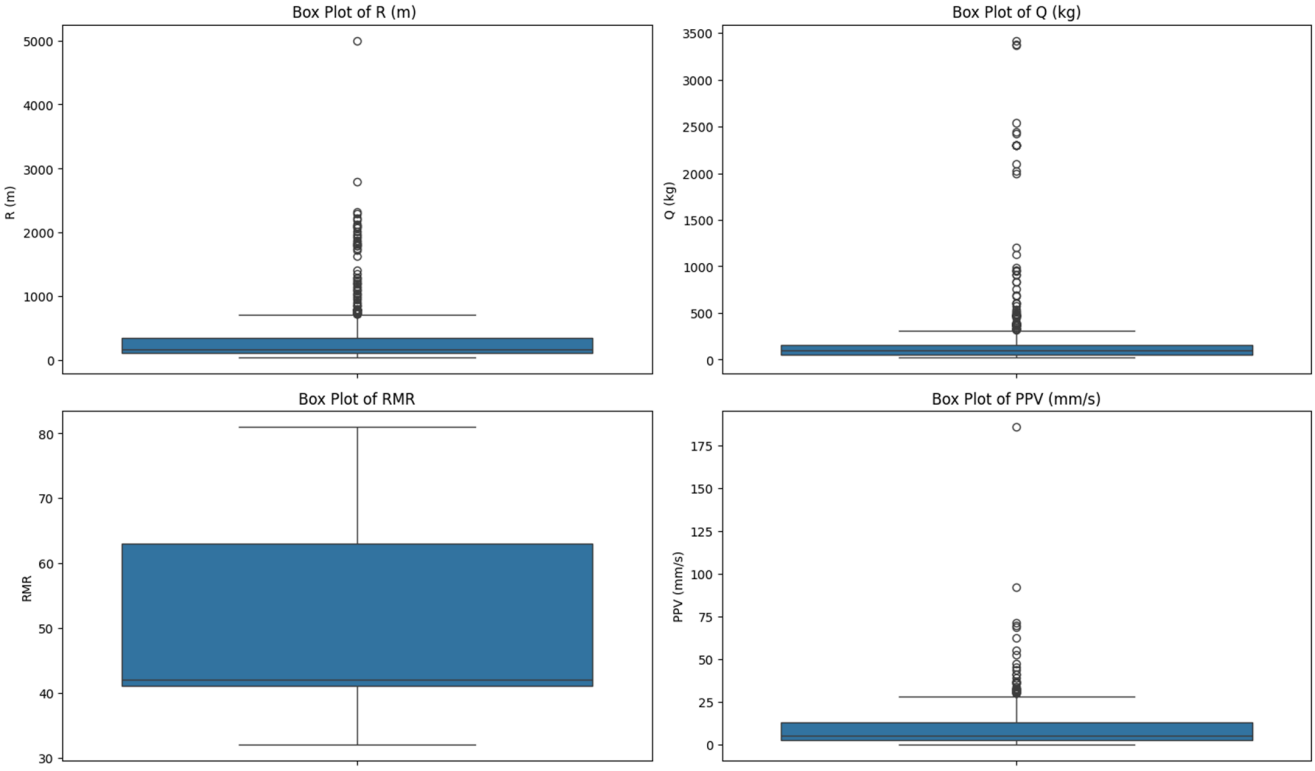


Figure 1. Box plot of the original model input and output parameters before the removal of outliers.

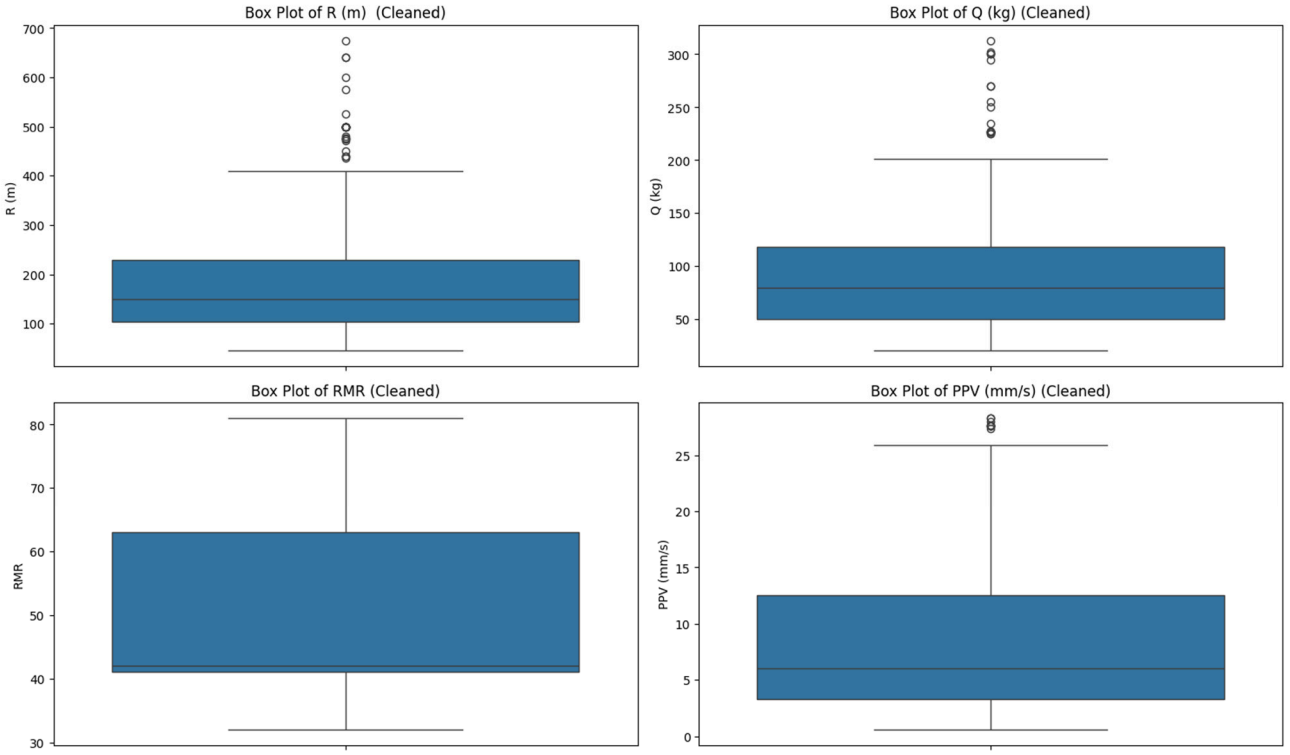


Figure 2. Box plot of the new model input and output parameters after the removal of outliers.

3.2. RMR Derivation Approach

It is important to clarify that RMR values were neither directly reported in the 12 source datasets nor derived from conventional geomechanical indices such as the Geological Strength Index (GSI), Rock Quality Designation (RQD), or Uniaxial Compressive Strength (UCS). Instead, RMR was approximated using a simplified adaptation of Equation (3) [60], enabling the incorporation of rock mass characteristics into the analysis despite the absence of directly reported geomechanical classification data in the original studies.

The attenuation parameter (β) was obtained for each source from reported PPV attenuation relationships through regression fitting of the general vibration prediction equation. These β values, representing site-specific attenuation behavior, were subsequently used to estimate a single representative RMR value for each dataset. This resulted in 12 distinct RMR values corresponding to the 12 independent literature sources.

It is important to emphasize that RMR was not computed at the individual data point level. All observations from a given source were assigned the same RMR value; thus, RMR functions as a site-level descriptor of rock mass conditions rather than a sample-specific variable. Consequently, the model does not receive any transformation of the target PPV for individual observations, and no sample-wise information leakage occurs during model training.

However, it is acknowledged that RMR is indirectly derived from site-level attenuation behavior, as reflected in the empirical relationship:

$$\beta = 0.0058 + 0.0289\ \mathrm{RMR}\ (R^2 = 0.88) \quad (3)$$

This introduces a degree of dependency at the dataset level, since β itself is obtained from PPV attenuation trends within each source. Therefore, RMR should be interpreted as an approximate proxy for rock mass characteristics rather than a strictly independent input variable.

Despite this limitation, the use of RMR enables the incorporation of site-specific geomechanical variability into the modeling framework in a consistent, reproducible manner. Furthermore, assigning a single RMR value per dataset reduces the risk of overfitting, as the variable does not capture point-specific variations in PPV. Additionally, SHAP-based interpretability analysis indicates that RMR contributes meaningfully but not dominantly to the model predictions, confirming that the model does not rely disproportionately on this variable and reducing concerns of indirect target encoding. Detailed descriptive statistics of the datasets are provided in Table 4. Figure 3 presents a correlation matrix based on Pearson correlation coefficients, showing minimal dependency among variables and no evidence of multicollinearity between input parameters or between inputs and outputs. This guarantees robust model performance on unseen data and prevents overfitting. Frequency distribution histograms for input and output parameters are shown in Figure 4, confirming wide variability that supports highly accurate PPV prediction models.

Table 4. The variability intervals of the model input and output parameters.

Model Parameters	Symbol	Units	Class	Statistics			
				Min	Max	Mean	STD
Charge per delay	*Q*	Kg	Input	20.50	313.00	93.04	54.34
Monitoring distance	*R*	m	Input	45.00	675.00	184.76	117.06
Rock mass rating	RMR	-	Input	32.00	81.00	50.00	12.90
Ground vibration	PPV	mm/s	Output	0.60	28.30	8.63	6.99

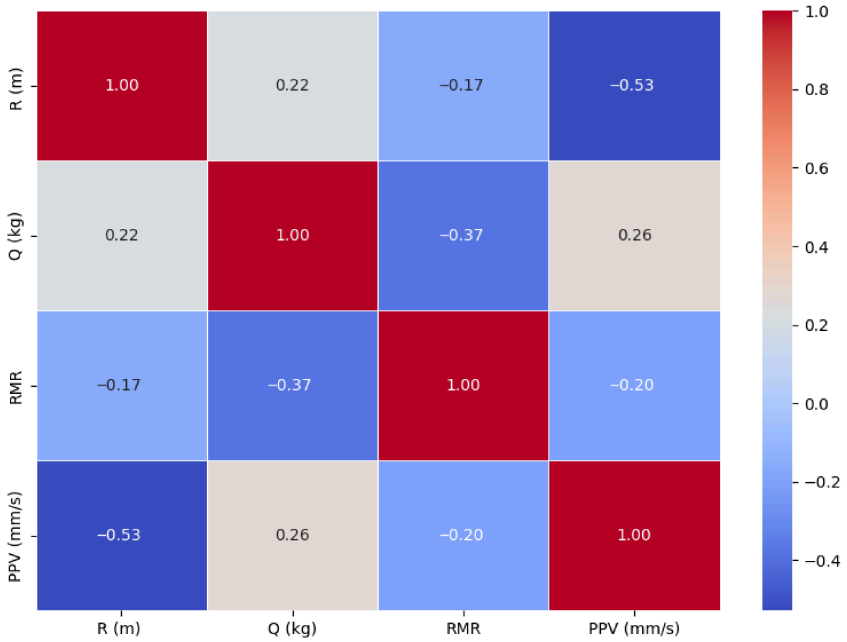


Figure 3. Input and output parameters correlation matrix based on Pearson correlation coefficients.

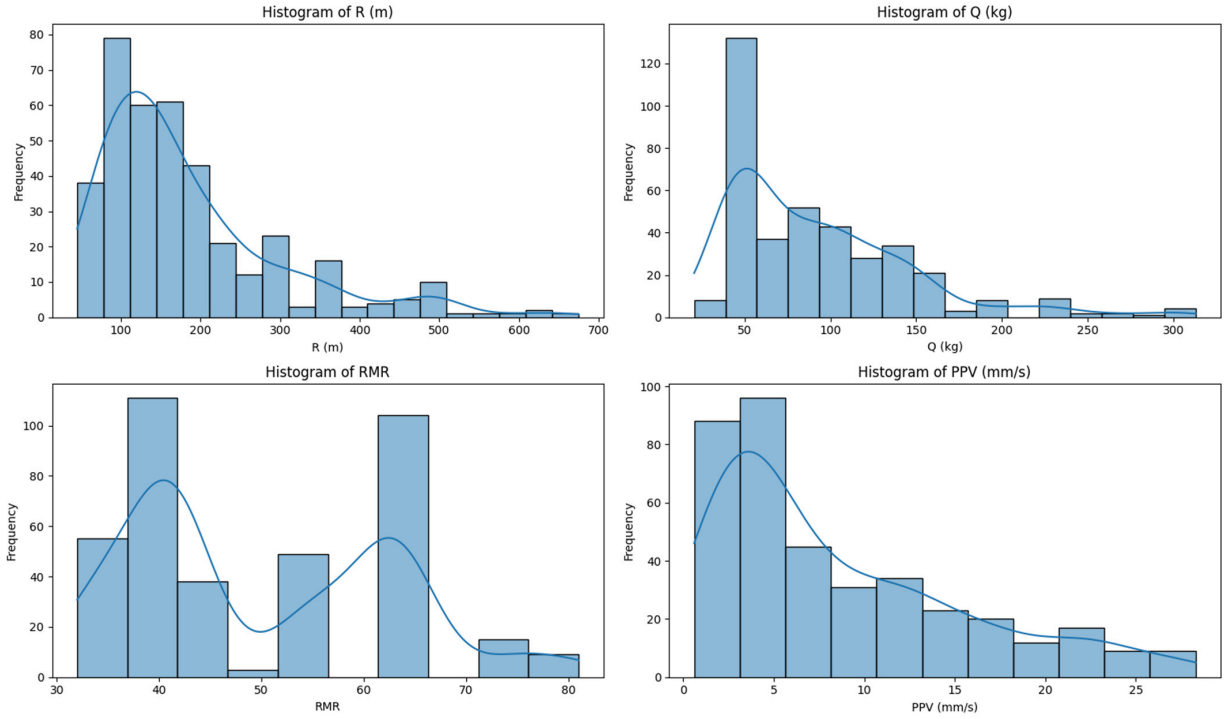


Figure 4. Frequency distribution histograms of input parameters (Q [kg], R [m], RMR [-]) and output PPV [mm/s] showing dataset variability supporting robust model training.

4. Methodology

This section discusses the theoretical background of the ML techniques and the conventional empirical equations used to develop PPV prediction models in this study. Following this, the various PPV model development processes were discussed. The section concludes with a discussion of the various evaluation metrics used to assess the model's performance.

4.1. Machine Learning Techniques

The theoretical background of artificial neural networks (ANNs), adaptive neuro-fuzzy inference systems (ANFIS), Gaussian process regression (GPR), and decision trees (DTs) is discussed in this subsection.

4.1.1. Artificial Neural Network (ANN)

ANNs are computational models inspired by the structure and function of the human brain [61]. They consist of interconnected nodes, called artificial neurons, organized in layers that process information to generate an output. ANNs are a core component of machine learning, capable of learning and modeling complex, nonlinear relationships in data [62].

ANNs comprise an input layer, one or more hidden layers, and an output layer [63]. Each neuron receives signals, processes them using an activation function, and passes the results to other neurons. Connections have adjustable weights, refined during training via optimization algorithms like gradient descent to minimize prediction errors [64]. The architecture (Figure 5) is mathematically summarized in Equation (4). This approach excels in engineering due to its ability to learn data relationships without predefined forms, adapt to various outputs, and straightforward implementation. Backpropagation neural networks (BPNNs) predominate due to their simple topology and their efficacy in modeling nonlinear phenomena [64–66].

$$G = f_0\{w_0 \times [f_1(w_i \cdot h + b_i)] + b_0\} \quad (4)$$

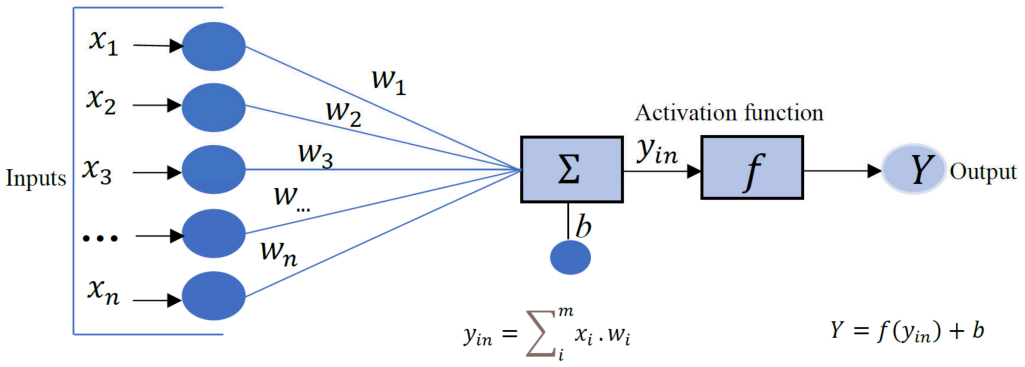


Figure 5. ANN layout schematic with layers and connections.

Here, w_0 and w_1 are weights connecting the intermediate output and the input-intermediate layers, respectively. b_0 and b_1 are biases for the output and intermediate layers, respectively. h and G represent normalized input/output parameters, while f_1 and f_0 are activation functions (hyperbolic tangent sigmoid) applied to both layers in this study.

4.1.2. Adaptive Neuro-Fuzzy Inference Systems (ANFIS)

ANFIS integrates neural network learning with fuzzy-logic reasoning for complex nonlinear systems [67,68]. It employs Takagi–Sugeno fuzzy inference within a 5-layer feed-forward network: (1) fuzzification via membership functions; (2) rule firing strengths; (3) normalization; (4) linear consequent computation; (5) final output aggregation [61,67].

ANFIS employs hybrid learning, combining least-squares (consequent parameters) and gradient descent (premise parameters) for rapid convergence [67]. It captures fuzzy rule expert knowledge and neural data-driven learning, excelling at ambiguous engineering inputs common in nonlinear problems.

ANFIS outperforms traditional ML models in nonlinear prediction through its interpretability and adaptive training, which iteratively minimizes errors. Its versatility

integrates human expert knowledge with data-driven learning, making it valuable for modeling complex systems.

4.1.3. Gaussian Process Regression (GPR)

GPR is a Bayesian nonparametric technique that infers probability distributions over function space, unlike parametric models that assume fixed forms such as $f(x) = mx + c$ [69,70]. GPR allows data to directly dictate the flexibility of $f(x)$ via multivariate Gaussian priors [Equation (5)], enabling exact uncertainty quantification. Maximum entropy principles constrain the solution space for meaningful inference, while pragmatic kernel assumptions enable robust interpolation. This makes GPR powerful for modeling complex nonlinear relationships and uncertain phenomena in engineering applications.

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K & K_T^* \\ K^* & K^{**} \end{bmatrix}\right) \tag{5}$$

Here, T denotes matrix transpose. GPR focuses on conditional probability p($y*|y$), which follows a Gaussian distribution consistent with GPR assumptions (Equation (6)) [66].

$$y^*|y \sim N\left(K^*K^{-1}y, K^{**} - K^*K^{-1}K_T^*\right) \tag{6}$$

The mean of this distribution constitutes the optimal estimate for $y*$ and is represented by Equation (7).

$$\overline{y^*} = K^*K^{-1}y \tag{7}$$

Furthermore, the variance (Equation (8)) quantifies the uncertainty inherent in the estimation and reflects confidence in the predicted values.

$$var(y^*) = K^{**} - K^*K^{-1}K_T^* \tag{8}$$

Based on the preceding explanation, the predicted value $y*$ is obtained at the 10th point (where $x* = 0.2$). This prediction procedure can be systematically extended and applied to multiple points distributed across a designated segment of the x-axis.

4.1.4. Decision Tree (DT)

DTs are nonparametric supervised learning algorithms that construct predictive models by recursively partitioning data based on homogeneity criteria [71,72]. They offer intuitive interpretability of variable interactions and robustness against multicollinearity. DTs form inverted trees with the most influential predictor at the root, descending via binary or multiway splits to terminal leaf nodes.

At each node, splitting decisions compare variable values against thresholds, partitioning the data into two or more subsets using variance-reduction criteria [Equation (9)]. This recursive process, guided by if-then rules, continues each branch until stopping conditions are met, such as homogeneous instances or a minimum sample size. Final predictions aggregate values from terminal leaf nodes.

$$I_V(N) = \frac{1}{|Y|^2}\sum_{i\in S}\sum_{j\in S}\frac{1}{2}(x_i - x_j)^2 - \left(\frac{1}{|Y_t|^2}\sum_{i\in S_t}\sum_{j\in S_t}\frac{1}{2}(x_i - x_j)^2 + \frac{1}{|Y_f|^2}\sum_{i\in S_f}\sum_{i\in S_f}\frac{1}{2}(x_i - x_j)^2.\right) \tag{9}$$

where Y is the set of pre-split sample indices, Y_t is the set of sample indices for which the split test is true, and Y_f is the set of sample indices for which the split test is false.

4.1.5. Metaheuristic Algorithms (MAs)

The two metaheuristic algorithms (MAs) used in this study to improve the performance of the optimum or best-performing ML model are Bayesian optimization (BO) and differential evolution (DE).

The BO, used to optimize black-box functions [73], has a primary advantage: it can find the global optimum with very few experimental iterations, making it particularly suitable for tasks that require many experiments, such as training PPV models. The BO algorithm achieves this capability by constructing a Gaussian process model to describe the probability distribution of the function to be optimized. Then, it balances exploration and exploitation by selecting the next experimental point and iteratively approaching the global optimum. The advantages of the BO algorithm include considering prior parameter information, requiring fewer iterations, and being robust.

The DE algorithm is a vector-based global optimization method that seeks the global optimum of the objective function via differentiation, mutation, and selection of candidate solution vectors [74]. As a metaheuristic, DE avoids rigorous mathematical proofs in optimization; instead, it relies on heuristic rules and strategies to search effectively. At its core, DE mirrors genetic algorithms by iteratively refining the population through mutation, crossover, and selection until converging on the best solution. In the mutation stage, the DE uses differences between two random individual vectors to mutate the current population, creating new offspring from these mutation vectors. After population initialization, individuals get mutated to produce mutation vectors. In the crossover stage, DE applies a probability threshold to select components from the mutated offspring vector, forming trial individuals that enhance population diversity and promote differentiation. In the selection stage, DE follows a greedy selection rule, picking individuals based on their fitness function values and choosing the best from offspring and parents alike.

4.2. Conventional Empirical Equations

The PPV is the primary metric used to assess the impact of blasting on the surrounding environment. Blasting vibrations are influenced by numerous complex factors, making it impractical to develop a universal prediction formula for PPV that accounts for all variables. However, extensive monitoring in blasting engineering reveals that the maximum charge per delay and distance from the blast face are the two most significant factors affecting PPV. Consequently, various empirical PPV prediction formulas have been proposed internationally [14–16,28].

The USBM, AH, LK, and BIS models were selected for this study due to their widespread use and applicability for blasting vibration prediction across diverse conditions. These empirical attenuation formulas (Table 5), which use k and b to represent rock properties, blasting operation parameters, and terrain conditions, provide a robust benchmark for evaluating the developed machine learning models. Field blasting vibration measurements are used to perform regression analysis based on these formulas, enabling the estimation of unknown parameters. Using actual site data, site-specific constants are calibrated to tailor the prediction model for this project. The resulting PPV prediction formulas incorporate the maximum charge per delay and the blast-face distance, enabling accurate PPV calculations for given blasting scenarios at the mining site.

Table 5. Selected empirical equations for estimating PPV in this study.

Empirical Estimators	References	Equations
United States Bureau of Mines (USBM)	[30]	$v = k(R/Q^{1/2})^{-b}$
Ambraseys–Hendron (AH)	[14]	$v = k(R/Q^{1/3})^{-b}$
Langefors–Kihlstrom (LK)	[16]	$v = k(Q/R^{2/3})^{b/2}$
Bureau of Indian Standards (BIS)	[28]	$v = k(Q/\sqrt[3]{R^2})^{-b}$

Note: v: peak particle velocity (mm/s); R: distance from blast source to recording location (m); Q: net equivalent charge per delay (kg); and k and b: site-specific constants.

4.3. PPV Model Development

This section provides a detailed discussion of the modeling procedures used for both machine learning algorithms and traditional empirical equations, as well as the development of each model applied to PPV prediction. It also outlines the evaluation metrics used to assess model performance.

4.3.1. Model Development Techniques for ANN-LM and Other Machine Learning Frameworks

The procedures adopted in this study for ANN-LM, ANFIS, GPR, and DT modeling involve identifying optimal configurations and parameters for each algorithm to ensure reliable PPV predictions. The procedure followed five distinct stages:

Stage 1. Data compilation, wrangling, processing, and deployment: The data for this study comprised field-measured PPV, drilling and blasting design parameters, and rock mass rating for each of the 386 blasting events. These data were first collected and then processed prior to model development. Data processing involved normalizing the datasets to the range −1 to 1 for the ANN-LM, GPR, and DT models, and to 0 to 1 for the ANFIS model, using the generalized normalization equation (Equation (10)). This was done to prevent overfitting and achieve equal scaling of variables, thereby supporting algorithm convergence. Subsequently, feature engineering techniques were applied to create another model parameter (RMR).

$$Y_{norm} = \frac{(M_{max} - M_{min})(N_i - N_{min})}{(N_{max} - N_{min})} + M_{min} \quad (10)$$

where Y_{norm} represents the normalized parameter, N_i represents the field-compiled data point to be normalized, M_{max} and M_{min} represent the upper and lower limits of the normalization scale, respectively, and N_{max} and N_{min} represent the maximum and minimum values in the field-compiled dataset.

Stage 2. Model training and validation: For ANN-LM, the training process involved defining the network architecture by varying the number of hidden layers (1 or 2) and the number of neurons per layer (1–20). The Levenberg–Marquardt algorithm was employed as the training method, with mean square error used as the cost function. A hyperbolic tangent sigmoid activation function was applied to the neurons in both the hidden and output layers to enable nonlinear mapping.

For ANFIS, training focused on generating fuzzy if-then rules via a hybrid learning algorithm that combines least-squares and backpropagation, adjusting membership functions for both the fuzzification and defuzzification layers. For GPR, modeling entailed selecting appropriate kernel functions and optimizing hyperparameters to represent the data's probabilistic distribution. For DT, model training involved constructing decision trees using recursive partitioning and pruning strategies to reduce overfitting.

Stage 3. Evaluation and selection of the superior model setup: Each trained model was first evaluated using the training dataset. The configuration demonstrating the best performance was then selected based on key metrics, including the correlation coefficient (*r*) and mean squared error (MSE). These metrics provided a quantitative basis for identifying the optimal model setup.

Stage 4. Model validation via test dataset: The chosen configurations of the ANN-LM, ANFIS, GPR, and DT models were evaluated on an unseen test dataset to assess their predictive accuracy and generalization.

Stage 5. Implementation of the top-performing model for prediction: The model with the highest performance during evaluation was recommended for practical application to predict blasting-induced peak particle velocity (PPV) in mines.

4.3.2. ANN-LM PPV Prediction Model

In this study, an ANN-LM model was developed to predict mine blasting-induced PPV. Grid search spanned 20 architectures: 1–2 hidden layers, 1–20 neurons/layers (Table 6). Initially, these were tested using two input parameters: explosive charge per delay (Q) and distance (R), followed by testing with three input parameters: rock mass rating (RMR), Q, and R. Each ANN-LM model was trained using 270 (70%) randomly selected datasets from 386 total field-measured PPV and geological parameters, with the remaining 116 (30%) used for validation and testing. Model performance was evaluated using 5-fold cross-validation to ensure robustness and minimize bias due to random data partitioning. Training was conducted over 1000 epochs with early stopping (max_fail = 100) to avoid overfitting. The early stopping mechanism automatically terminated training when validation performance failed to improve for 100 consecutive epochs, retaining the network weights corresponding to the best validation performance. Fixed settings included the Levenberg–Marquardt trainer, tan-sigmoid activation, MSE cost, a momentum factor (mu) of 0.001, and a convergence threshold (min_grad) of 1×10^{-7}.

Table 6. Sets of ANN tuning parameters for model development.

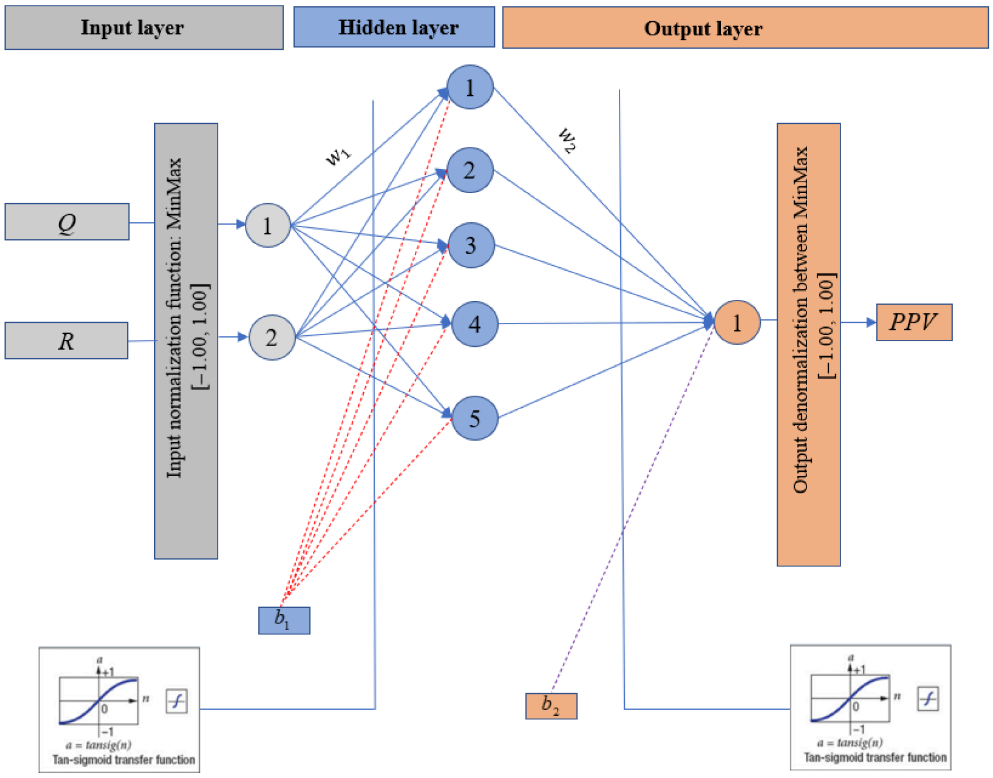
Hyperparameters	Values
Training algorithm	Levenberg–Marquardt (LM)
Network type	Feed-forward backpropagation neural network (BPNN)
Normalization function	Minmax [−1, 1]
Number of hidden layers	1 or 2
Number of neurons per hidden layer	1 to 20 by step 1
Train:test: validation ratio	70:30
Hidden layer activation function	Hyperbolic tangent sigmoid (tan-sigmoid)
Output layer activation function	Hyperbolic tangent sigmoid (tan-sigmoid)
Epochs	1000 with early stopping
Cost function	Mean Square Error (MSE)
Momentum (mu)	0.001
Convergence threshold (min_grad)	1×10^{-7}
Early stopping checks (max_fail)	100

The models' performances were evaluated to select the most reliable structure, with Table 7a,b ranking the 20 network topologies by their correlation coefficients (*r*-values) for two- and three-input parameters, respectively. Notably, network performance did not correlate with the size of the hidden layer; some smaller networks outperformed larger ones. The best model for two parameters (Q and R) was ANN-LM 2-5-1 (overall R = 0.85763), with hyperbolic tangent-sigmoid activation in both the hidden and output layers. For three parameters (Q, R, RMR), the top model was ANN-LM 3-10-1 (training R = 0.88186) with the

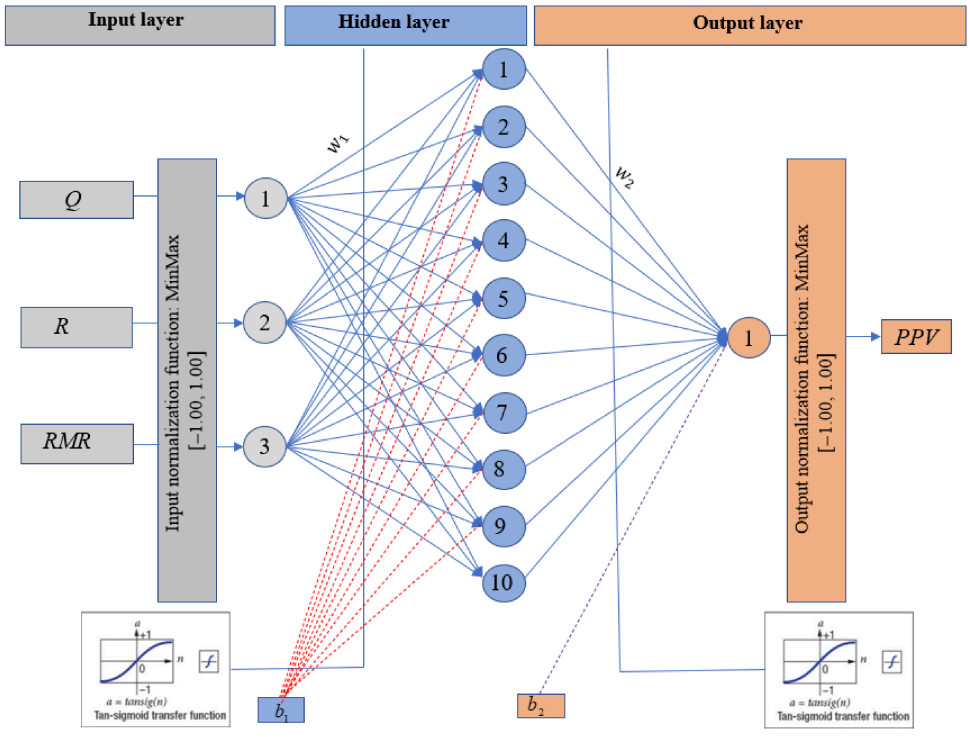
same activation functions. For the optimal ANN-LM 3-10-1 architecture, training converged well before the 1000-epoch limit, with the training and validation MSEs remaining closely aligned, confirming proper convergence and the absence of overfitting. Figure 6a,b illustrate the network topologies of these two optimal ANN-LM models. Figures 7a and 7b present scatter plots and line diagrams comparing measured versus predicted PPV values for the training and testing datasets of the 2-input (Q, R) and 3-input (Q, R, RMR) ANN-LM models, respectively.

Table 7. (**a**). Coefficients of correlation (*r*) for the different tested ANN architectures (2-input). (**b**). Coefficients of correlation (*r*) for the different tested ANN architectures (3-input).

Network	Training	Testing	Validation	Overall
		(**a**)		
2-2-1′	0.73151	0.81693	0.71484	0.73875
2-3-1′	0.77593	0.82233	0.72381	0.76847
2-4-1′	0.76333	0.78318	0.89729	0.81825
2-5-1′	**0.78907**	**0.82989**	**0.86684**	**0.85763**
2-6-1′	0.82390	0.69483	0.793478	0.81235
2-7-1′	0.796071	0.84575	0.79521	0.78414
2-8-1′	0.79718	0.75768	0.80617	0.78472
2-9-1′	0.76391	0.85045	0.86905	0.80991
2-10-1′	0.80771	0.77425	0.88347	0.79347
2-11-1′	0.78731	0.81857	0.85519	0.80107
2-12-1′	0.79850	0.82002	0.80235	0.80238
2-13-1′	0.80736	0.83175	0.77566	0.80333
2-14-1′	0.78983	0.84488	0.78805	0.81036
2-15-1′	0.80457	0.76240	0.80889	0.79364
2-16-1′	0.79433	0.83758	0.70302	0.81364
2-17-1′	0.79872	0.79414	0.81022	0.78961
2-18-1′	0.82911	0.67782	0.75659	0.79834
2-19-1′	0.77971	0.71900	0.89197	0.79184
2-20-1′	0.81005	0.78730	0.81244	0.80726
		(**b**)		
3-2-1′	0.79549	0.81853	0.81827	0.80166
3-3-1′	0.83482	0.68842	0.85210	0.81497
3-4-1′	0.84027	0.71287	0.78221	0.81283
3-5-1′	0.82901	0.92870	0.85475	0.85082
3-6-1′	0.89106	0.88805	0.83218	0.85974
3-7-1′	0.83699	0.89715	0.85741	0.84792
3-8-1′	0.84438	0.87611	0.80061	0.84183
3-9-1′	0.86239	0.84192	0.89443	0.86234
3-10-1′	**0.89181**	**0.95297**	**0.92866**	**0.88186**
3-11-1′	0.84977	0.87455	0.92670	0.86367
3-12-1′	0.81775	0.88892	0.86418	0.86602
3-13-1′	0.82387	0.85019	0.89228	0.83532
3-14-1′	0.85806	0.91647	0.71652	0.85199
3-15-1′	0.85862	0.90002	0.82112	0.87351
3-16-1′	0.84915	0.83360	0.81918	0.85214
3-17-1′	0.83283	0.83141	0.87799	0.85586
3-18-1′	0.82148	0.83863	0.83026	0.84024
3-19-1′	0.82861	0.82844	0.81811	0.83324
3-20-1′	0.85521	0.87406	0.83665	0.84252



(**a**)



(**b**)

Figure 6. Network topologies of the optimal ANN-LM models (**a**) ANN-LM 2-5-1; (**b**) ANN-LM 3-10-1.

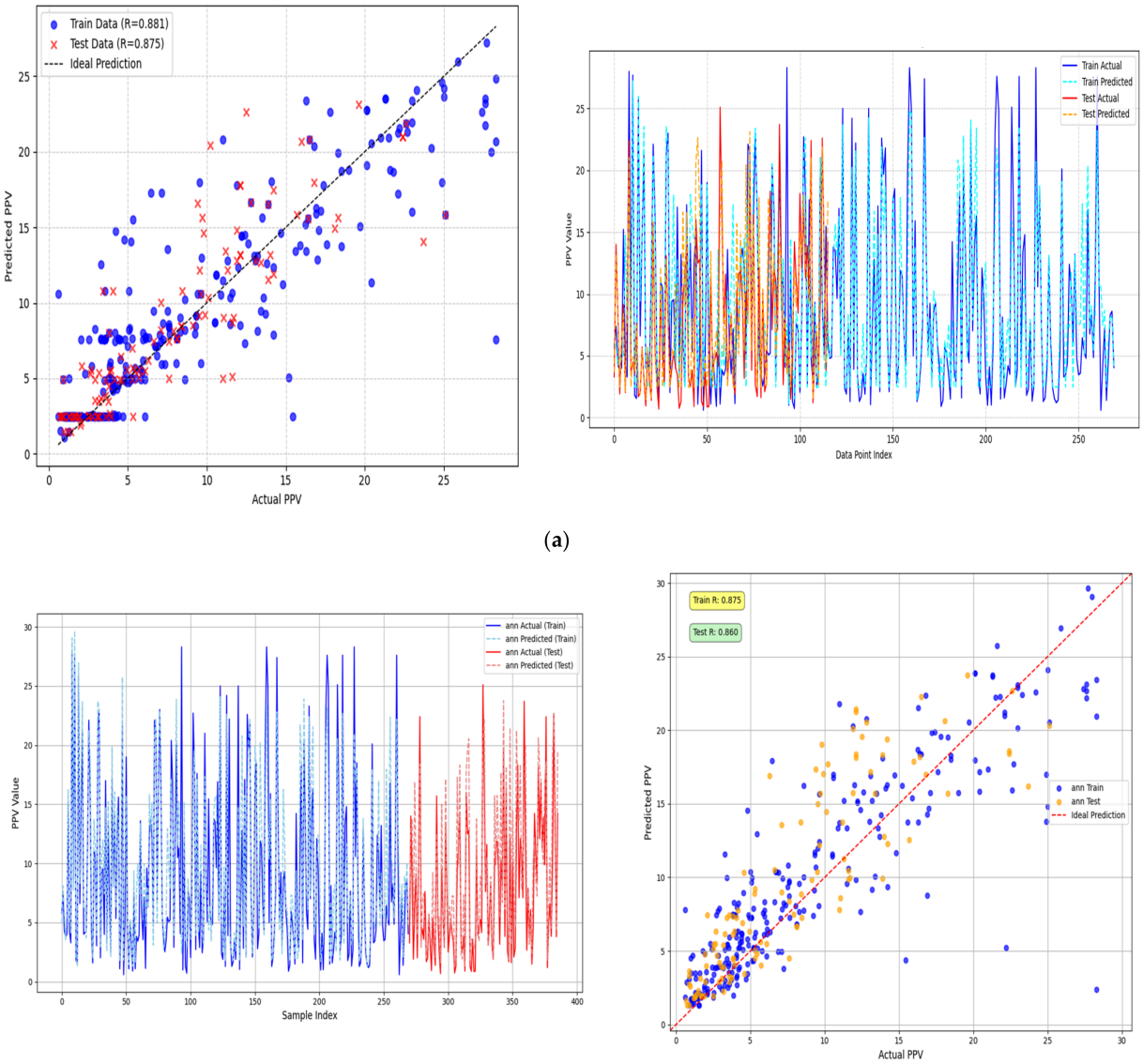


Figure 7. Comparison of the field-measured and ML model-predicted PPV values for training and testing datasets using scatter plots and line diagrams: (**a**) ANN-LM 2-5-1 model; (**b**) ANN-LM 3-10-1 model.

4.3.3. ANFIS PPV Prediction Model

The ANFIS PPV prediction framework was developed using a first-order Sugeno-type fuzzy network initialized with Genfis2 and trained on min–max normalized input–output datasets to ensure numerical stability and efficient convergence. The fuzzy inference system employed Gaussian membership functions for all input variables and linear consequents for the output, providing a smooth yet flexible mapping that can capture the nonlinear relationship between blasting parameters and PPV. For both models, the fuzzy rule base was generated via grid partitioning of the normalized input space. Grid partitioning generated fuzzy rules with Gaussian membership functions (input), linear output; hybrid optimizer (least-squares + backprop); 100 epochs. Varied membership functions per input (2–5); train/test = 70/30.

A hybrid learning strategy combining least-squares estimation of consequent parameters with gradient-based tuning of premise parameters was adopted. The dataset was split into 70% for training and 30% for testing, and training proceeded for 100 epochs.

Model performance was evaluated using 5-fold cross-validation to ensure robustness and minimize bias due to random data partitioning. The full set of hyperparameters for both the 2-input and 3-input configurations, including network architecture, initialization strategy, normalization function, membership function type, FIS generation method, optimization algorithm, train–test ratio, number of epochs, number of membership functions, and the RMSE-based cost function, is summarized in Table 8. The 2-input ANFIS model employed 65 membership functions (best via RMSE minimization), yielding a moderately dense rule base that balances accuracy and computational efficiency. The 3-input ANFIS model used 877 membership functions, providing finer rule granularity to represent higher-order interactions among the three predictors, albeit at the expense of increased computational cost. In both cases, minimizing the Root Mean Squared Error (RMSE) through guided parameter updates produces robust, data-driven neuro-fuzzy models suitable for reliable prediction of PPV in blasting design and vibration control. Figures 8a and 8b present scatter plots and line diagrams comparing measured versus predicted PPV values for the training and testing datasets of the 2-input (Q, R) and 3-input (Q, R, RMR) ANFIS models, respectively.

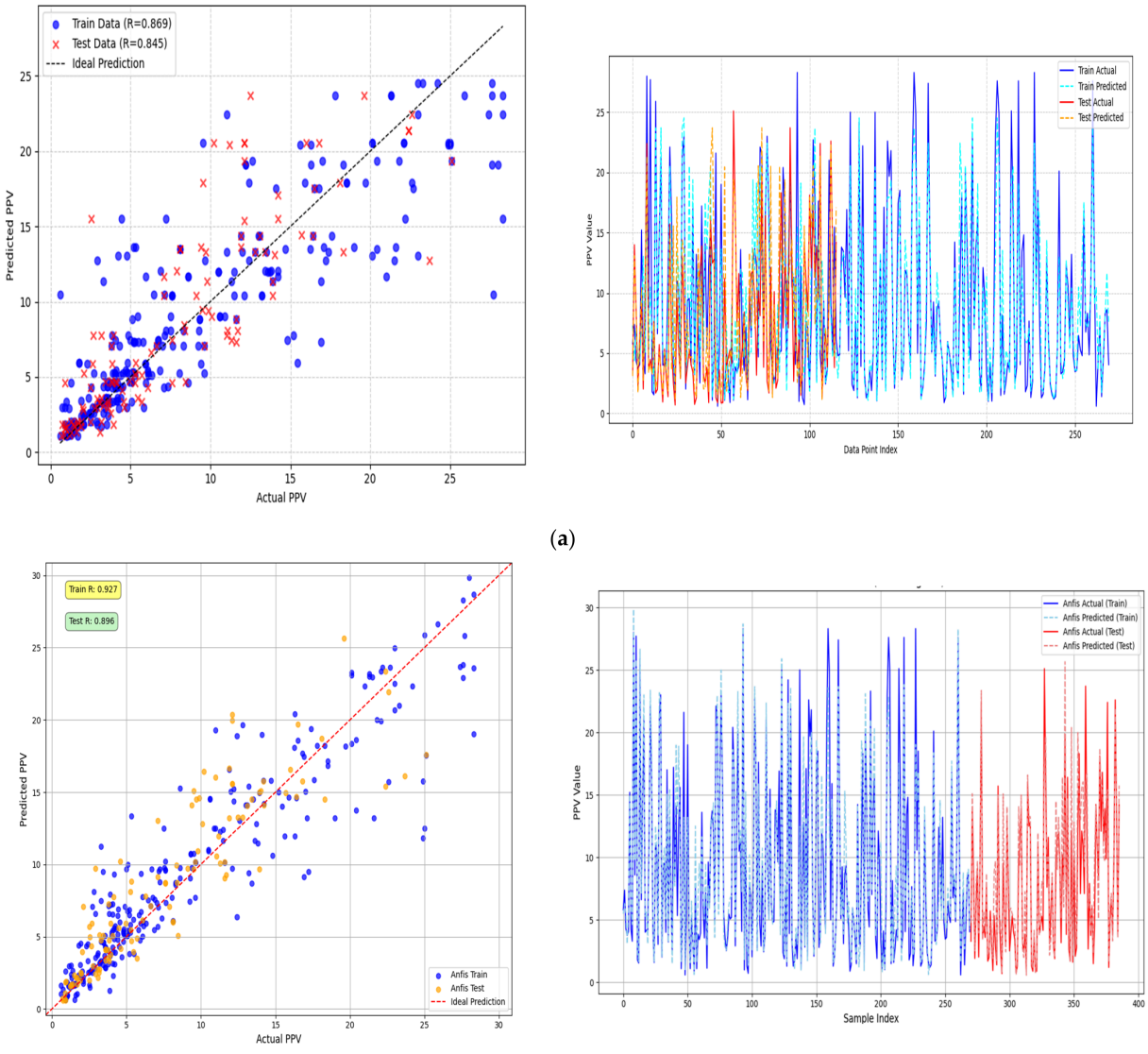


Figure 8. Comparison of the field-measured and ML model-predicted PPV values for training and testing datasets using scatter plots and line diagrams: (**a**) ANFIS (2-input) model; (**b**) ANFIS (3-input) model.

Table 8. Hyperparameter settings for the 2-input and 3-input ANFIS PPV prediction models.

Hyperparameters	Values
Fuzzy network	Sugeno type
Initial FIS for training	Genfis2
Normalization function	Minmax [0, 1]
Membership function type	Input: Gaussmf; Output: linear
FIS generation method	Grid partitioning
Optimization method	Hybrid
Train:test ratio	70:30
Number of epochs	100
Number of membership functions	65 (2-input); 877 (3-input)
Cost function	Root Mean Squared Error (RMSE)

4.3.4. GPR PPV Prediction Model

Table 9 presents the hyperparameter configurations obtained from grid search used for the two GPR models developed in this study: the two-input and three-input models. The table summarizes the selected basis function, kernel structure, and associated kernel hyperparameters used during model training.

Table 9. Hyperparameter settings for the 2-input and 3-input GPR PPV prediction models.

Hyperparameter	2-Input GPR Model	3-Input GPR Model
	Values	Values
Basis function	constant	constant
Kernel	1 ** 2 * Rational Quadratic (alpha = 1, length_scale = 1) + White Kernel(noise_level = 0.1)	Radial Basis Function (RBF)
Constant kernel variance	0.1	10.00
RQ_length scale	0.1	1.00

Note: + denotes kernel addition, * denotes kernel multiplication, and ** denotes exponentiation (kernel raised to a power).

Both models employ a constant basis function, implying that the prior mean of the Gaussian process is assumed to be constant across the input domain. This assumption is commonly adopted in GPR when no strong prior information regarding the global trend of the response variable is available.

For the two-input GPR model, the covariance function is defined as a combination of a Rational Quadratic kernel and a White Kernel. The Rational Quadratic kernel ($\alpha = 1$, length scale = 1) enables the model to capture variations occurring at multiple characteristic length scales, thereby providing flexibility in representing complex relationships between the predictor variables and the response. The White Kernel component (noise level = 0.1) is incorporated to account for noise in the measurements and improve the numerical stability of the regression model. Additionally, the constant kernel variance and the RQ length scale are both set to 0.1. During model development, the training process evaluated 120 hyperparameter candidates, including variations in the Rational Quadratic and White Kernel parameters, such as length scale, alpha, and noise level. Model performance was evaluated using 5-fold cross-validation, yielding 600 total model fits (5 folds $\times$ 120 candidates) to identify the optimal configuration.

In contrast, the three-input GPR model utilizes a Radial Basis Function (RBF) kernel as its covariance function. The RBF kernel assumes a smooth and stationary relationship between the input variables and the output and is widely used in Gaussian process modeling due to its ability to produce smooth function approximations. For this model, the constant

kernel variance is set to 10.00, while the RBF length scale is set to 1.00, indicating a broader covariance structure than in the two-input model. The training process evaluated nine hyperparameter combinations, including variations in the RBF kernel parameters. Using 5-fold cross-validation, this resulted in 45 total model fits for the three-input model.

Figures 9a and 9b present scatter plots and line diagrams comparing measured versus predicted PPV values for the training and testing datasets of the 2-input (Q, R) and 3-input (Q, R, RMR) GPR models, respectively.

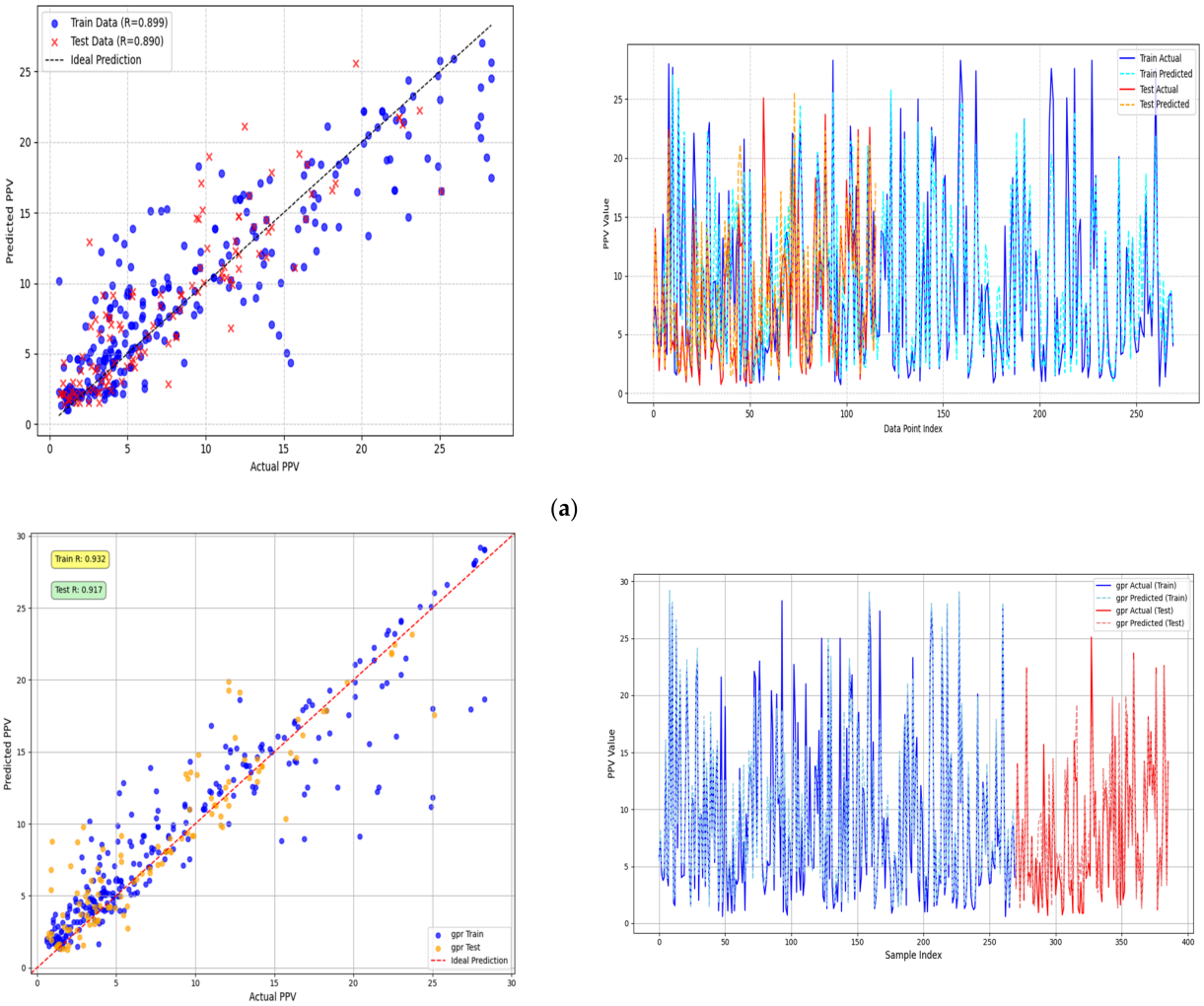


Figure 9. Comparison of the field-measured and ML model-predicted PPV values for training and testing datasets using scatter plots and line diagrams: (**a**) GPR (2-input) model; (**b**) GPR (3-input) model.

4.3.5. DT PPV Prediction Model

The DT PPV prediction model was implemented for both 2-input and 3-input configurations using the hyperparameters summarized in Table 10. A 70/30 data split was used, and model performance was evaluated using 5-fold cross-validation to ensure robustness and minimize bias from random data partitioning. Grid search spanned maximum splits (100–200), minimum leaf/parent sizes (5–20/10–20), MSE split criterion, and all-splits predictor selection, with medium tree complexity as the base across configurations. For each configuration, a medium-complexity regression tree ("Medium Tree") was employed to balance predictive accuracy and model interpretability, which is crucial for explaining PPV behavior in blasting design. Both models used the "allsplits" predictor selection algorithm, which exhaustively evaluates potential split points, thereby improving the identification

of influential predictors for PPV. Additionally, the split criterion was set to mean squared error (MSE), ensuring that each node split minimized variance within its child nodes and enhanced overall regression performance.

Table 10. Hyperparameter settings for the 2-input and 3-input DT PPV prediction models.

Hyperparameter	2-Input GPR Model	3-Input GPR Model
	Values	Values
Model type	'Medium Tree'	'Medium Tree'
'PredictorSelection'—Algorithm used to select the best split predictor	'allsplits'	'allsplits'
SplitCriterion—Split criterion	'mse'	'mse'
'MaxNumSplits'—Maximal number of decision splits	124	118
'MinLeafSize'—Minimum number of leaf node observations	4	6
'MinParentSize'—Minimum number of branch node observations	2	2

For the 2-input model, the best configuration featured a maximum of 124 splits, with a minimum of 4 observations per leaf and 2 per parent node. This yielded a relatively deeper tree, allowing for finer partitioning of the input space and potentially capturing subtle nonlinear interactions between the two input parameters. In contrast, the 3-input model used the optimal setup of 118 splits while maintaining the same minimum parent size of 2 and a smaller minimum leaf size of 6 observations. The reduced leaf size and comparable depth provided additional flexibility to model the increased dimensionality introduced by the third input variable, without excessive overfitting. Figures 10a and 10b present scatter plots and line diagrams comparing measured versus predicted PPV values for the training and testing datasets of the 2-input (Q, R) and 3-input (Q, R, RMR) DT models, respectively.

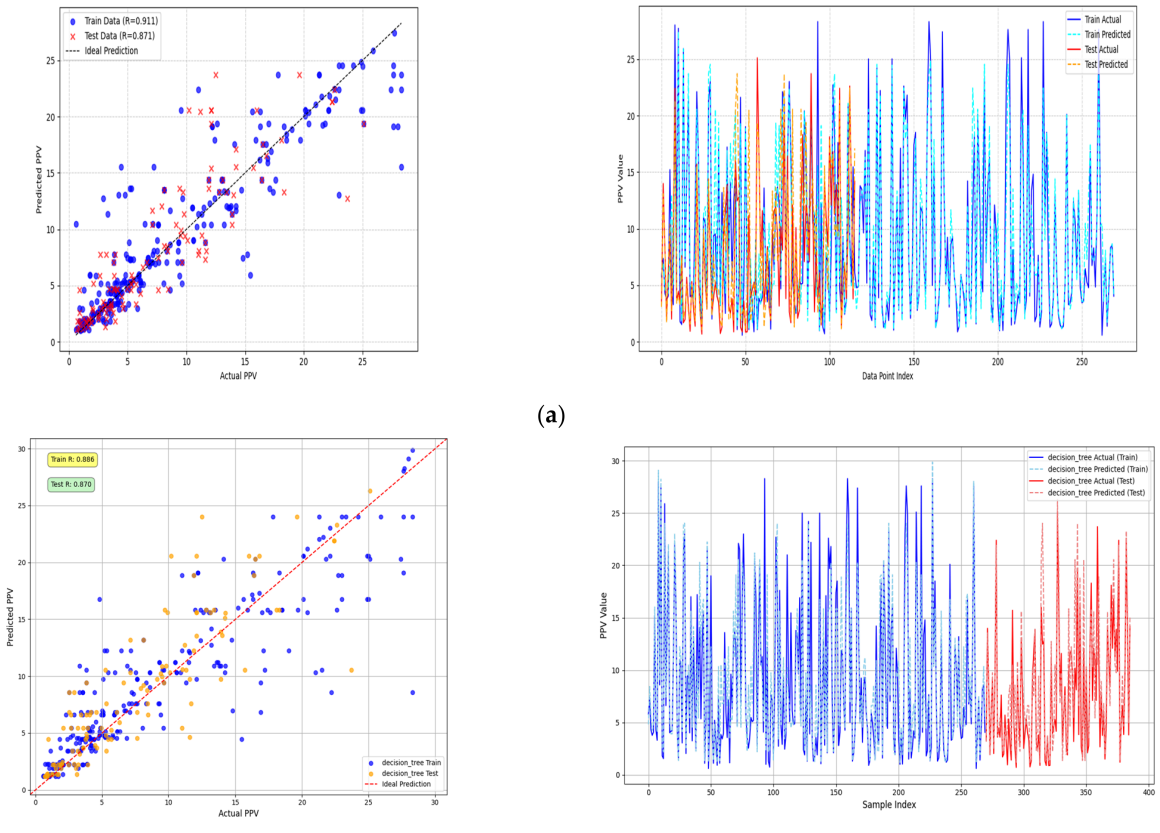


Figure 10. Comparison of the field-measured and ML model-predicted PPV values for training and testing datasets using scatter plots and line diagrams: (**a**) DT (2-input) model; (**b**) DT (3-input) model.

4.3.6. BO-GPR and DE-GPR PPV Prediction Model

The two metaheuristic algorithms, BO and DE, are employed to optimize the hyperparameters of the best-performing stand-alone model (GPR) for the 2-input and 3-input PPV prediction models. Tables 11 and 12 present the optimized hyperparameters and search space bounds for the GPR models developed for PPV prediction using two- and three-input configurations, calibrated via BO and DE algorithms. Model performance was evaluated using 5-fold cross-validation to ensure robustness and minimize bias due to random data partitioning.

Table 11. Hyperparameter settings for the 2-input and 3-input BO-GPR and DE-GPR PPV prediction models.

Hyperparameter	BO (2-Input)	DE (2-Input)	BO (3-Input)	DE (3-Input)
RQ variance	191.2743	653.2977	0.1	0.1
RQ length_scale	20.4912	773.9433	0.5	1.0
RQ alpha	0.3241	332.3814	-	-
Noise level	0.00004	87.1645	0.01 (alpha)	0.1 (alpha)
Constant Kernel variance	-	-	0.05	0.1

Table 12. Hyperparameter search space bounds for BO and DE optimization of the 2-input and 3-input GPR models.

Hyperparameter	2-Input Lower Bound	2-Input Upper Bound	3-Input Lower Bound	3-Input Upper Bound
RQ kernel variance	1×10^{-1}	1×10^{3}	1×10^{-3}	1×10^{1}
RQ length scale	1×10^{-1}	1×10^{3}	1×10^{-2}	1×10^{1}
RQ alpha	1×10^{-1}	1×10^{3}	—	—
Noise level	1×10^{-5}	1×10^{2}	1×10^{-3}	1×10^{1}
Constant kernel variance	—	—	1×10^{-2}	1×10^{1}
Number of candidates evaluated (BO)	120	—	9	—
Population size (DE)	50	—	50	—
Maximum iterations (DE)	1000	—	1000	—
Crossover probability (DE)	0.7	—	0.7	—
Mutation factor (DE)	0.5	—	0.5	—

For the 2-input model, BO produced moderate kernel parameters: an RQ variance of 191.2743, a length scale of 20.4912, and an alpha of 0.3241, along with a very low noise level (0.00004). These values indicate that the BO-based model captures localized variations in the input–output relationship while assuming minimal observational noise. In contrast, the DE-optimized model yielded considerably larger hyperparameter values, including an RQ variance of 653.2977, length scale of 773.9433, alpha of 332.3814, and a higher noise level (87.1645), suggesting a smoother functional approximation with greater tolerance to data variability.

For the 3-input model, both optimization methods produced more compact and consistent hyperparameters, incorporating a Rational Quadratic kernel combined with a constant kernel component. BO resulted in an RQ variance of 0.1, a length scale of 0.5, and a noise level of 0.01, while DE produced a similar RQ variance (0.1) with a slightly larger length scale (1.0) and a slightly larger noise level (0.1). Overall, the results indicate that BO favors tighter data fitting, whereas DE yields smoother models that better accommodate noise.

Figures 11a,b and 12a,b present scatter plots and line diagrams comparing measured versus predicted PPV values for the training and testing datasets of the 2-input (Q, R) and 3-input (Q, R, RMR) BO-GPR and DE-GPR models, respectively.

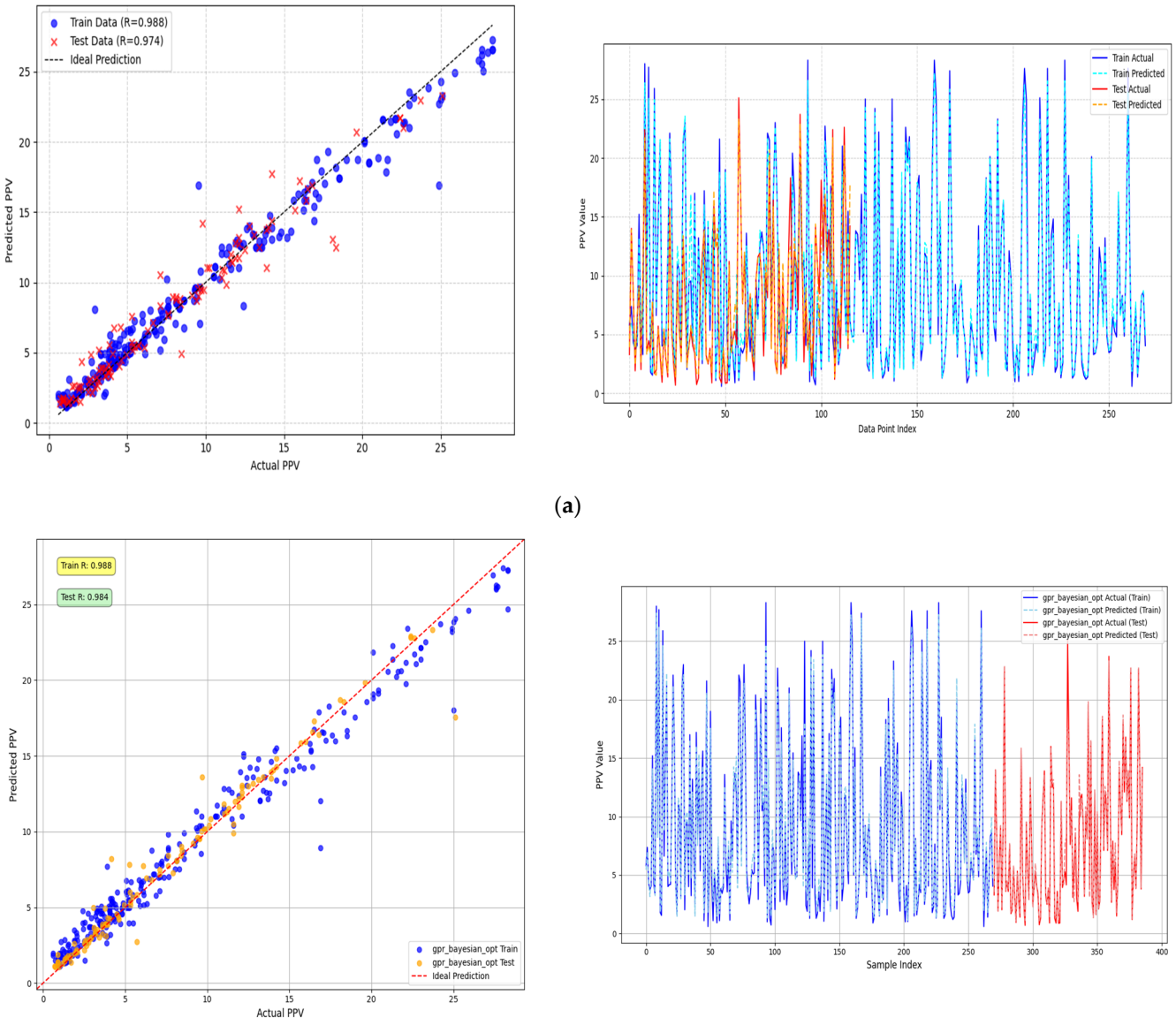


Figure 11. Comparison of the field-measured and ML model-predicted PPV values for training and testing datasets using scatter plots and line diagrams: (**a**) BO-GPR (2-input) model; (**b**) BO-GPR (3-input) model.

4.3.7. Model Development Techniques for Empirical Conventional Equations

The conventional PPV empirical equations for this study were developed by calibrating four widely used predictive models, USBM, AH, LK, and BIS, against the measured field data to obtain site-specific constants that reflect the local geological and blast design conditions. As summarized in Table 13, each model is expressed in a general power-law form that relates the peak particle velocity v to a scaled distance parameter derived from the maximum charge per delay and the distance from the blast source. Table 13 presents the final calibrated equations, along with the corresponding constants k and b, obtained by regression analysis of the recorded vibration data for each model.

It is important to note that the site-specific constants k and b for all four empirical models were derived exclusively through regression analysis on the same 70% training dataset used to train all ML models, and their predictive performance was subsequently evaluated on the same held-out 30% test dataset. This ensures strict consistency and fairness

in the benchmarking process between the empirical formulations and the ML algorithms, with no data leakage occurring across the training and testing phases.

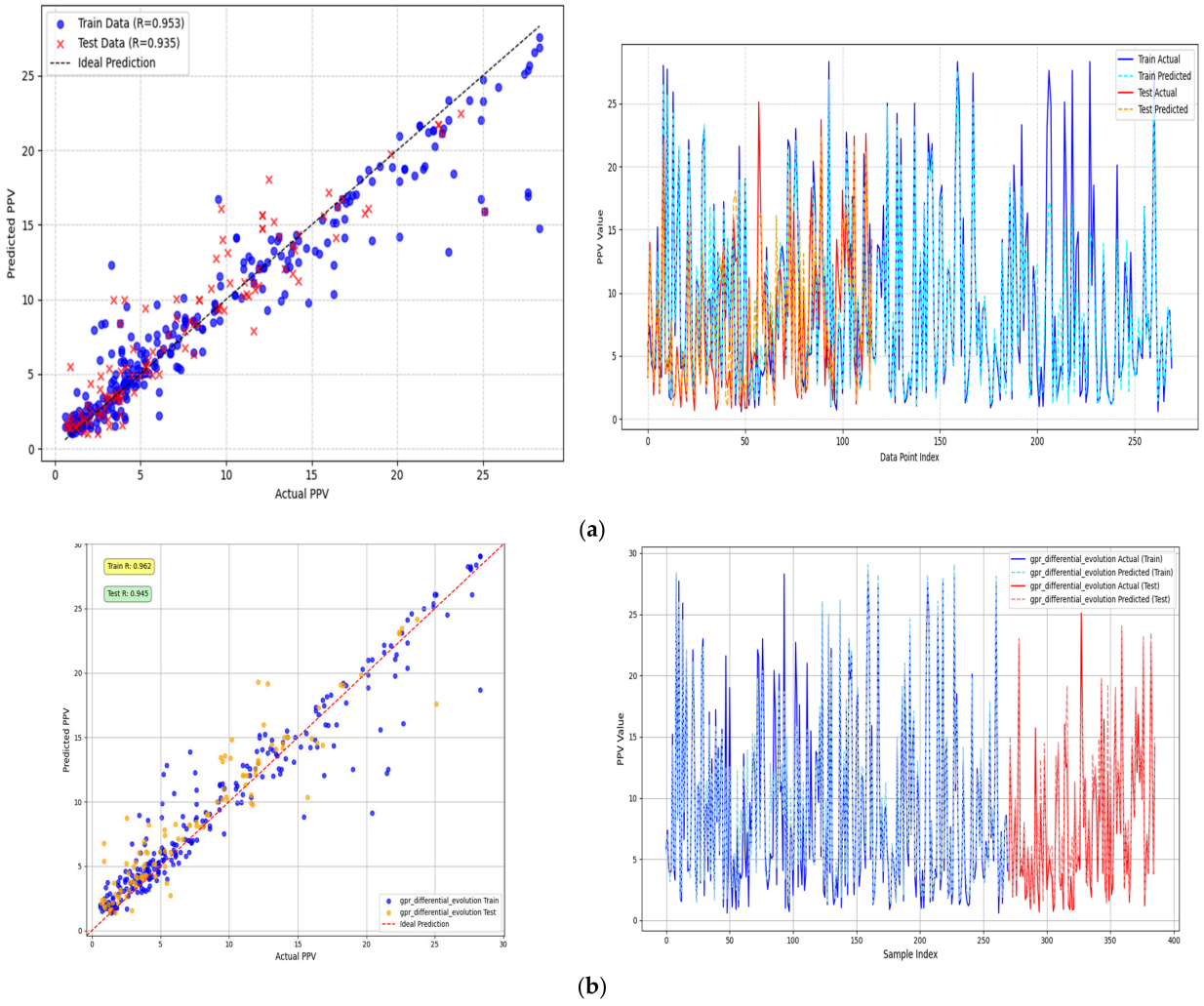


Figure 12. Comparison of the field-measured and ML model-predicted PPV values for training and testing datasets using scatter plots and line diagrams: (**a**) DE-GPR (2-input) model; (**b**) DE-GPR (3-input) model.

Table 13. Site-specific PPV empirical models and derived constants.

Empirical Model	Final Equation	Derived Site-Specific Constants	
		k	b
USBM	$v = 184.0652(R/Q^{1/2})^{-1.1175}$	184.0652	1.1175
AH	$v = 436.9575(R/Q^{1/3})^{-1.1305}$	436.9575	1.1305
LK	$v = 3.3293(Q/R^{2/3})^{1.7329}$	3.3294	1.7329
BIS	$v = 3.3293(Q/\sqrt[3]{^{a}R^{2}})^{0.8664}$	3.3293	0.8664

The curves for generating site-specific constants in the USBM model were obtained by plotting the logarithm of the measured PPV against the logarithm of the USBM-scaled distance, then fitting a regression line to the data points (Figure 13a). The slope and intercept of this line define the exponent and coefficient in the USBM equation, ensuring that the calibrated relationship captures the observed attenuation trend across the monitored range of charge and distance.

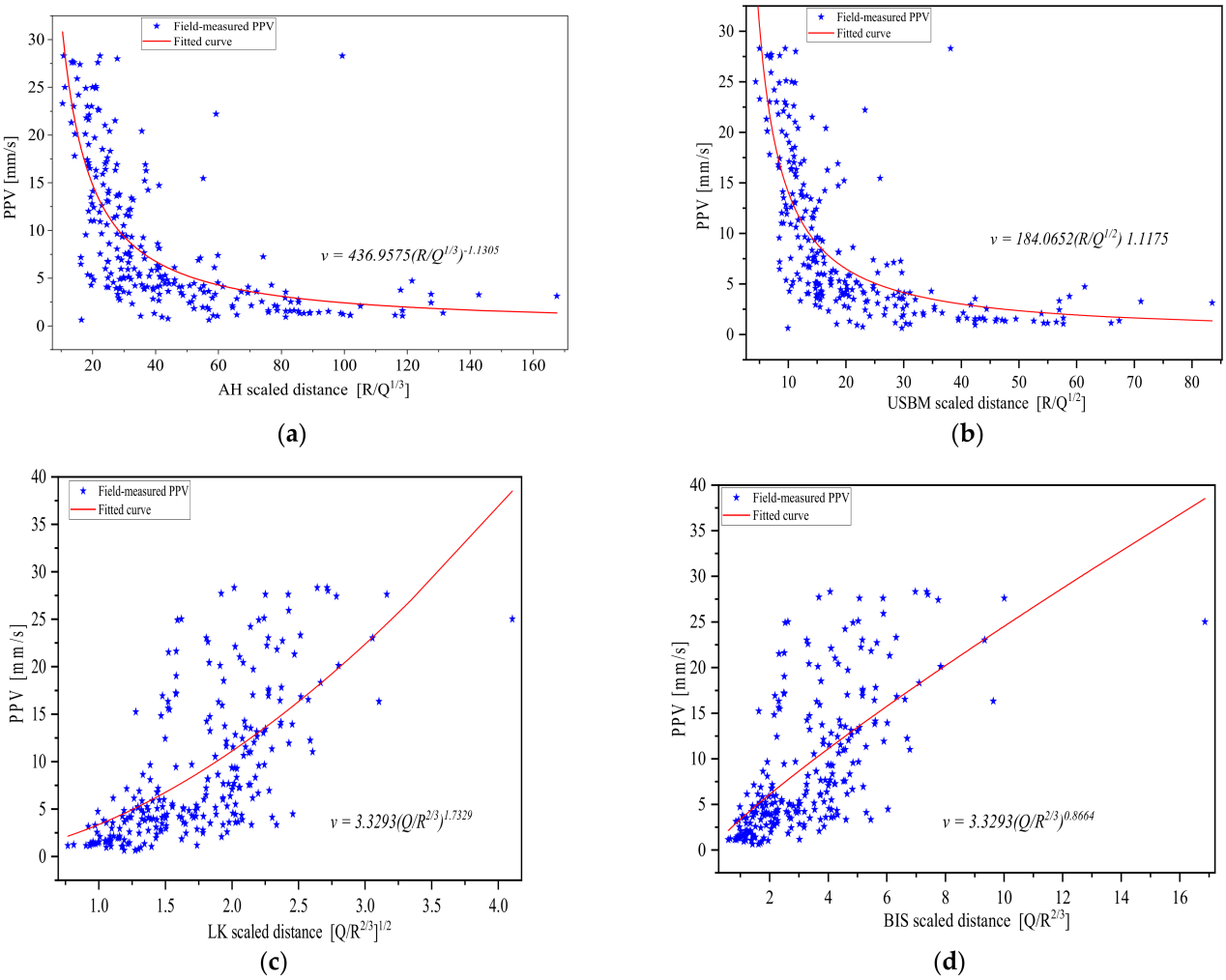


Figure 13. Variation in peak particle velocity (PPV) with scaled distance for four empirical blast vibration prediction models: (**a**) Ambraseys–Hendron (AH), (**b**) USBM, (**c**) Langefors–Kihlström (LK), and (**d**) Bureau of Indian Standards (BIS), showing field-measured PPV data and fitted regression curves.

An identical log–log regression procedure was applied to the AH model (Figure 13b), using its own definition of scaled distance, to derive the corresponding site-specific constants. Recalibration of the AH curve to the same vibration dataset enables a direct comparison of its predictive performance with the USBM form, showing that the AH model can reproduce the local attenuation characteristics.

For the LK model, PPV was plotted against the LK scaled distance, and a regression curve was fitted in log–log space (Figure 13c). From this, the LK constants were obtained and reported in Table 13 as the best-fitting parameters for the site. The BIS model followed the same approach, but with its distinct scaled-distance formulation (Figure 13d), yielding a separate regression line and a distinct set of site-specific constants. Comparison of the four regression curves and their derived constants confirms that each empirical form can be tailored through regression to describe the site-specific vibration behavior, and that Table 13 provides a concise summary of these calibrated PPV prediction models for design and control. Figures 14–17 compare the measured and predicted PPV values for the training and testing datasets of the USBM, AH, LK, and BIS models, using scatter plots and line diagrams, respectively.

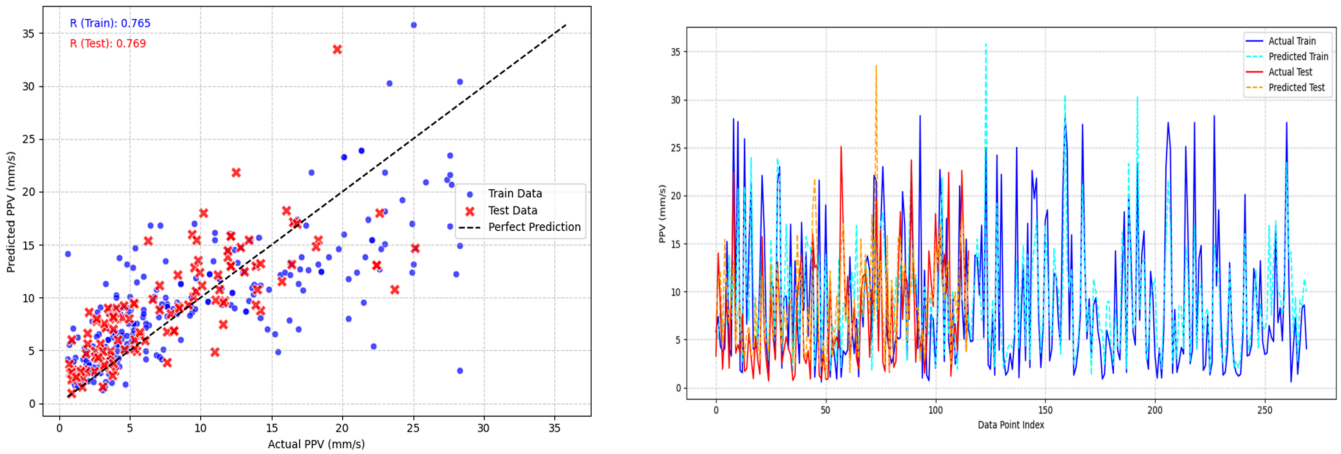


Figure 14. Comparison of the field-measured and USBM model-predicted PPV values for training and testing datasets using scatter plots and line diagrams.

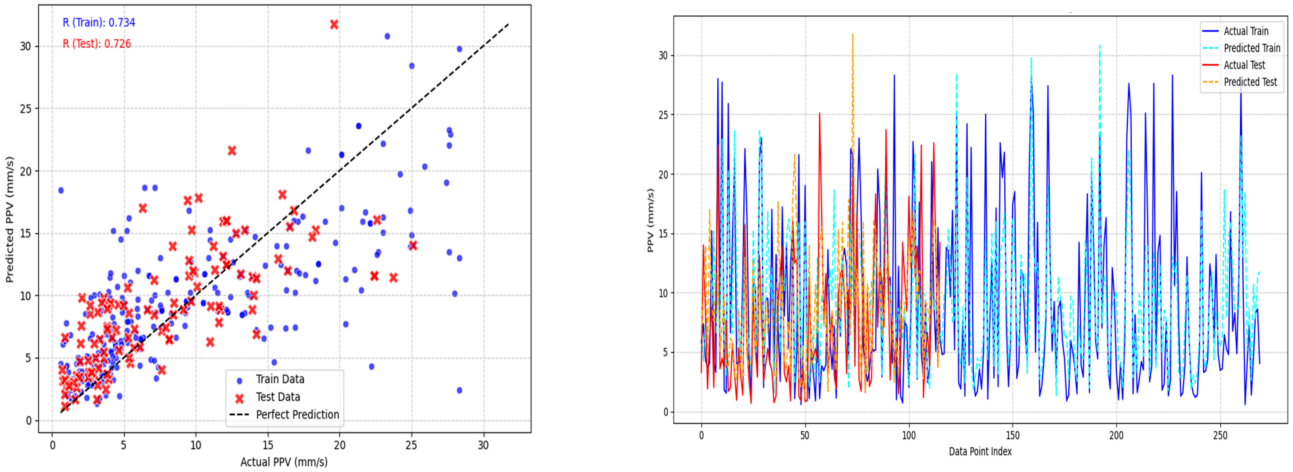


Figure 15. Comparison of the field-measured and AH model-predicted PPV values for training and testing datasets using scatter plots and line diagrams.

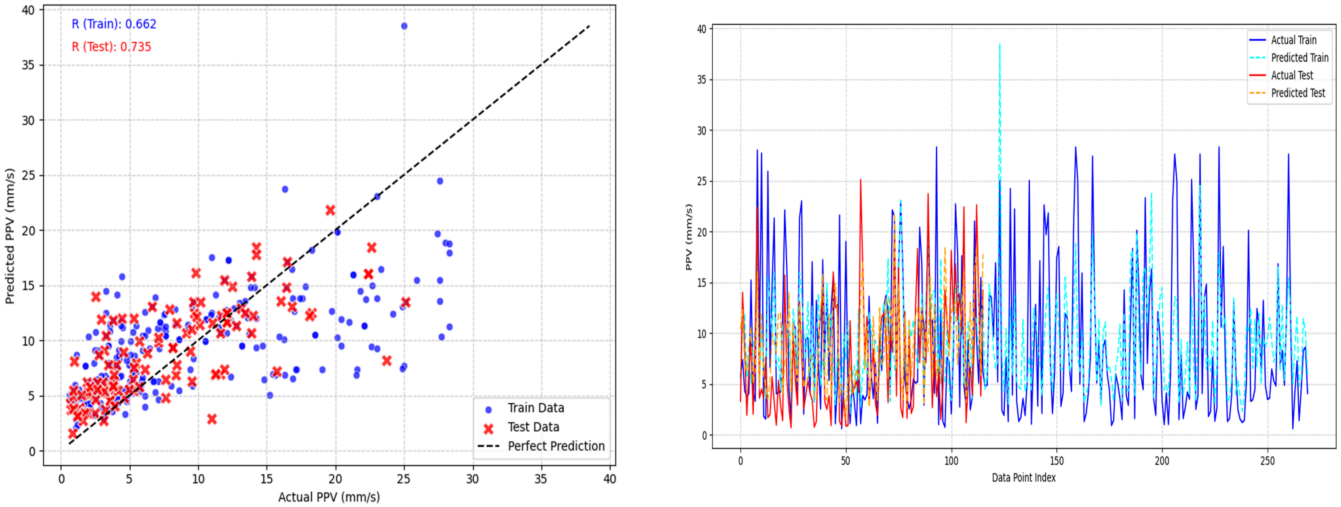


Figure 16. Comparison of the field-measured and LK model-predicted PPV values for training and testing datasets using scatter plots and line diagrams.

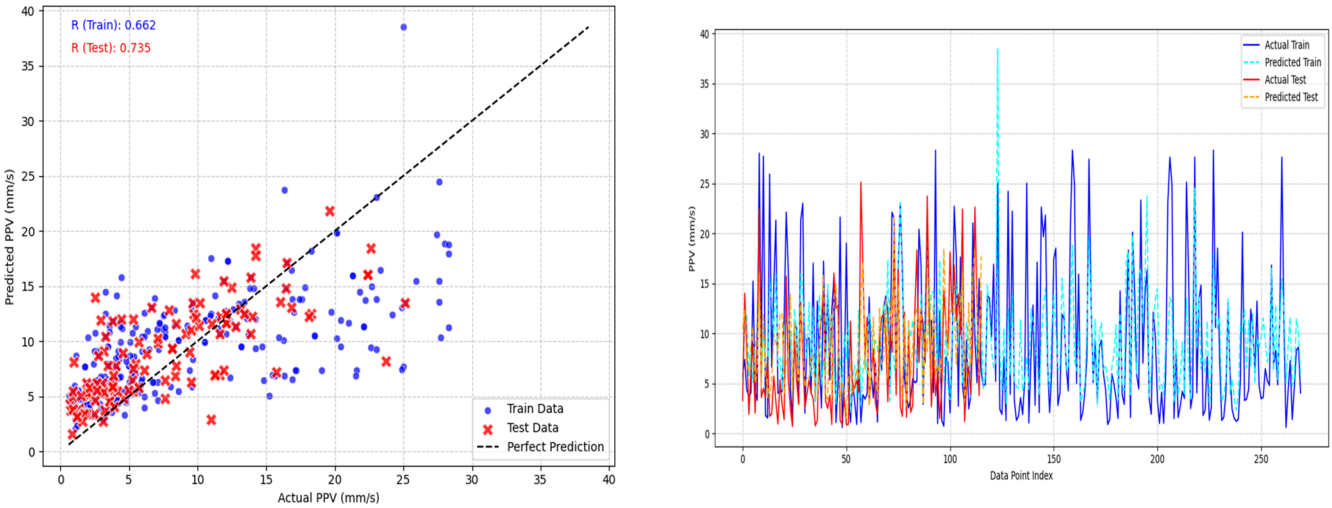


Figure 17. Comparison of the field-measured and BIS model-predicted PPV values for training and testing datasets using scatter plots and line diagrams.

4.4. Performance Metrics Used for Model Evaluation

To evaluate the performance of the blasting-induced PPV model, four standard metrics were employed, including mean absolute error (MAE), relative root mean square error (RRMSE), coefficient of correlation (r), and variance accounted for (VAF). These metrics provide a comprehensive assessment of the model's accuracy, precision, and explanatory power in predicting PPV values from blasting operations.

MAE quantifies the average magnitude of errors in the predicted PPV values without considering their direction, calculated as Equation (11):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{11}$$

where y_i are the observed PPV values, $\hat{y}_i$ are the predicted values, and n is the number of observations. Lower MAE values indicate superior model performance.

RRMSE normalizes the root mean square error by the mean of the observed values, offering a scale-independent measure of relative prediction accuracy, calculated as Equation (12):

$$\mathrm{RRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\frac{1}{n}\sum_{i=1}^{n} y_i} \times 100\% \tag{12}$$

This metric is particularly useful for comparing models across datasets with varying PPV scales, with values closer to 0% signifying a better fit.

The coefficient of correlation (r) assesses the linear relationship strength between observed and predicted PPV, derived from Pearson's correlation coefficient, calculated as Equation (13):

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}} \tag{13}$$

An r-value approaching 1 indicates a strong positive correlation and high predictive reliability.

Finally, VAF measures the proportion of total variance in observed PPV explained by the model, calculated as Equation (14):

$$\text{VAF} = \left(1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right) \times 100\% \quad (14)$$

VAF values exceeding 90% typically indicate excellent model performance in capturing data variability.

5. Results and Discussion

5.1. Performance Evaluation Results of Two-Input and Three-Input PPV Prediction Models

The performance evaluation of the blasting-induced PPV prediction models based on both two-input and three-input configurations (Table 14a,b) reveals a clear superiority of data-driven machine learning approaches over traditional empirical formulations. The consistently high performance across all folds confirms model stability and indicates results are not dependent on a specific data split. For the two-input models (Table 14a), the optimized Gaussian process regression models achieve the best training performance. In particular, BO-GPR achieves the highest correlation coefficient (r = 0.98823), the lowest error indices (MAE = 0.77492; RRMSE = 0.13521), and the largest variance accounted for (VAF = 97.31%), resulting in the best total rank (40) for the training dataset. DE-GPR performs strongly (total rank = 36), while the conventional ML models (DT, GPR, ANN, and ANFIS) show moderate predictive capability, with total ranks of 32, 28, 24, and 20, respectively. In contrast, the empirical formulations (USBM, AH, LK, and BIS) exhibit significantly lower correlation coefficients, larger errors, and substantially lower VAFs, highlighting their limited ability to capture the nonlinear relationship between blasting parameters and PPV when only two inputs are used.

The testing results in Table 14a further reinforce this distinction. The BO-GPR model maintains the highest predictive accuracy (r = 0.97426, VAF = 94.86%, total rank = 40), followed by DE-GPR (total rank = 36), confirming the effectiveness of optimization algorithms in improving GPR performance and enhancing model generalization. Among the conventional ML techniques, GPR, ANN, DT, and ANFIS achieve comparable predictive accuracy, with total ranks of 32, 28, 24, and 20, respectively, indicating reasonable robustness to unseen data. Conversely, the empirical models exhibit noticeable deterioration in predictive accuracy, as evidenced by reduced correlation coefficients and increased prediction errors. Notably, LK and BIS exhibit nearly identical statistical performance, reflecting the similarity in their scaled-distance attenuation formulations (k: 3.3294 vs 3.3293; b: -1.7329 vs -0.8664), yet both models remain among the poorest performers, underscoring the limitations of empirical scaling laws for reliable PPV prediction within the two-input framework. The similarity in the performance of the LK and BIS models in this study closely aligned with the findings of Jiang et al. [75].

For the three-input PPV models summarized in Table 14b, the inclusion of the additional predictor (RMR) substantially improves predictive capability across all machine learning techniques. During the training phase, BO-GPR again records the best overall performance (r = 0.98767, VAF = 97.11%, total rank = 24), followed by DE-GPR (total rank = 20). The remaining ML models (GPR, ANFIS, DT, and ANN) achieve moderately high correlation coefficients ranging from approximately 0.87 to 0.93, with VAF values between about 76% and 87%, indicating improved representation of PPV variability relative to the two-input configuration. However, these conventional models receive lower total ranks due to comparatively larger prediction errors.

In the testing phase of Table 14b, BO-GPR maintains the highest predictive performance (r = 0.98381, MAE = 0.50752, RRMSE = 0.13679, VAF = 96.78%), followed by DE-GPR,

which also demonstrates strong predictive reliability. Among the remaining models, GPR and ANFIS achieve better accuracy than DT and ANN, which exhibit slightly higher errors and lower VAFs. These results indicate that optimization-enhanced GPR models provide the best balance between predictive accuracy and robustness when three inputs are used.

Table 14. (**a**) Statistical performance indices and ranking of 2-input PPV prediction models during training and testing phases. (**b**) Statistical performance indices and ranking of 3-input PPV prediction models during training and testing phases.

Phase	PPV Model	Performance/Error Metrics								
		r	Rank	MAE	Rank	RRMSE	Rank	VAF (%)	Rank	Total Rank
					(**a**)					
Training	ANFIS	0.86903	5	2.42871	5	0.40840	5	75.52	5	20
	GPR	0.89946	7	2.29320	7	0.36319	7	80.84	7	28
	DT	0.91053	8	1.87341	8	0.34361	8	82.84	8	32
	ANN	0.88109	6	2.30227	6	0.39211	6	77.62	6	24
	BO-GPR	0.98823	10	0.77492	10	0.13521	10	97.31	10	40
	DE-GPR	0.95345	9	1.36874	9	0.25884	9	90.24	9	36
	USBM	0.76460	4	3.51796	4	0.53246	4	58.41	4	16
	AH	0.73361	3	3.62874	3	0.56104	3	53.80	3	12
	LK	0.66240	1	4.22536	1	0.61950	1	43.72	1	4
	BIS	0.66242	2	4.22528	2	0.61948	2	43.73	2	8
Testing	ANFIS	0.84453	5	2.31504	5	0.45652	5	65.13	5	20
	GPR	0.88981	8	1.98478	8	0.37124	8	77.82	8	32
	DT	0.87062	6	2.01644	6	0.40641	6	71.81	6	24
	ANN	0.87534	7	1.98810	7	0.38969	7	74.75	7	28
	BO-GPR	0.97426	10	0.82750	10	0.17228	10	94.86	10	40
	DE-GPR	0.93529	9	1.37482	9	0.27203	9	87.41	9	36
	USBM	0.76914	4	2.92310	4	0.50977	4	57.88	4	16
	AH	0.72646	1	3.16857	1	0.54562	1	51.08	1	4
	LK	0.73513	2	3.14830	2	0.52968	2	54.04	2	8
	BIS	0.73515	3	3.14811	3	0.52966	3	54.05	3	12
					(**b**)					
Training	ANFIS	0.92679	3	1.85009	3	0.31132	3	85.88	3	12
	GPR	0.93197	4	1.73448	4	0.30370	4	86.75	4	16
	DT	0.88586	2	2.22975	2	0.38333	2	78.47	2	8
	ANN	0.87487	1	2.42038	1	0.40746	1	76.39	1	4
	BO-GPR	0.98767	6	0.94174	6	0.14070	6	97.11	6	24
	DE-GPR	0.96165	5	1.29907	5	0.23074	5	92.40	5	20
Testing	ANFIS	0.89628	3	1.92199	3	0.36158	3	79.03	3	12
	GPR	0.91651	4	1.56782	4	0.31888	4	83.81	4	16
	DT	0.86957	2	2.21724	2	0.42657	2	70.89	2	8
	ANN	0.85981	1	2.78611	1	0.47943	1	68.54	1	4
	BO-GPR	0.98381	6	0.50752	6	0.13679	6	96.78	6	24
	DE-GPR	0.94522	5	1.46304	5	0.27455	5	89.29	5	20

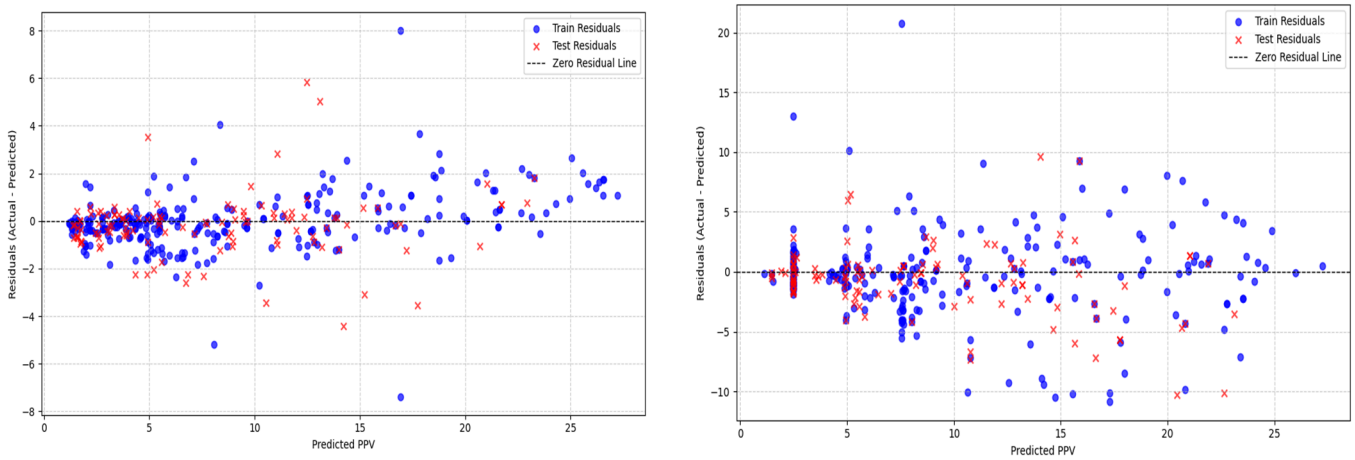
Taken together, the comparisons between Table 14a,b demonstrate that transitioning from a two-input to a three-input configuration systematically improves model stability and predictive reliability. The addition of RMR enhances ML models' ability to capture the complex interactions governing blasting-induced ground vibrations. Moreover, the consistent dominance of BO-GPR and DE-GPR across both training and testing phases highlights the effectiveness of metaheuristic optimization strategies in enhancing GPR model performance. In contrast, the comparatively poor rankings of the empirical models in Table 14a emphasize the limitations of conventional scaled-distance equations for accurate PPV prediction. Overall, these findings underscore the necessity of flexible, nonlinear, and

optimization-enhanced machine learning approaches for reliable prediction of blasting-induced vibrations in sustainable mining operations.

Figures 18 and 19 present residual plots (predicted PPV minus measured PPV) versus predicted PPV for the combined training and test datasets of the two-input and three-input models, respectively, providing further insight into model accuracy. Figure 18, which evaluates the two-input models, shows that the ML-based approaches produce tighter residual clustering around the zero line than the empirical equations, indicating superior predictive precision. Figure 19 illustrates the residual distributions for the three-input ML models, with the optimized BO-GPR and DE-GPR models exhibiting the narrowest residual dispersion, confirming the enhanced predictive stability achieved through the integration of optimization algorithms and the additional RMR parameter in the PPV prediction framework.

5.2. Comparative Performance Analysis of Blasting-Induced PPV Prediction Models

For the two-input models (Table 14a), the comparative evaluation of training and testing phases shows a clear hierarchy in model performance, with the optimized GPR models demonstrating the highest predictive capability. During the training phase, the BO-GPR model achieved the best overall performance with the highest correlation coefficient (r = 0.98823), the lowest MAE (0.77492), and the smallest RRMSE (0.13521), while explaining 97.31% of the variance (VAF) in PPV. This represents a substantial improvement over the conventional ML models (ANN, DT, GPR, and ANFIS) and the empirical equations (USBM, AH, LK, and BIS). For instance, BO-GPR improved the correlation coefficient by approximately 8–32% compared with the classical ML models and empirical formulations, while reducing MAE by roughly 60–80% relative to the empirical predictors. The DE-GPR model ranked second across all metrics (r = 0.95345; MAE = 1.36874; RRMSE = 0.25884; VAF = 90.24%), indicating that evolutionary optimization substantially enhances the predictive capability of the baseline GPR model. Among the conventional ML methods, DT and GPR demonstrated relatively strong calibration performance, whereas ANN and ANFIS performed moderately. In contrast, the empirical models exhibited the weakest predictive capability, with the LK and BIS equations producing the lowest correlations and the highest prediction errors.



(**a**) (**b**)

Figure 18. *Cont.*

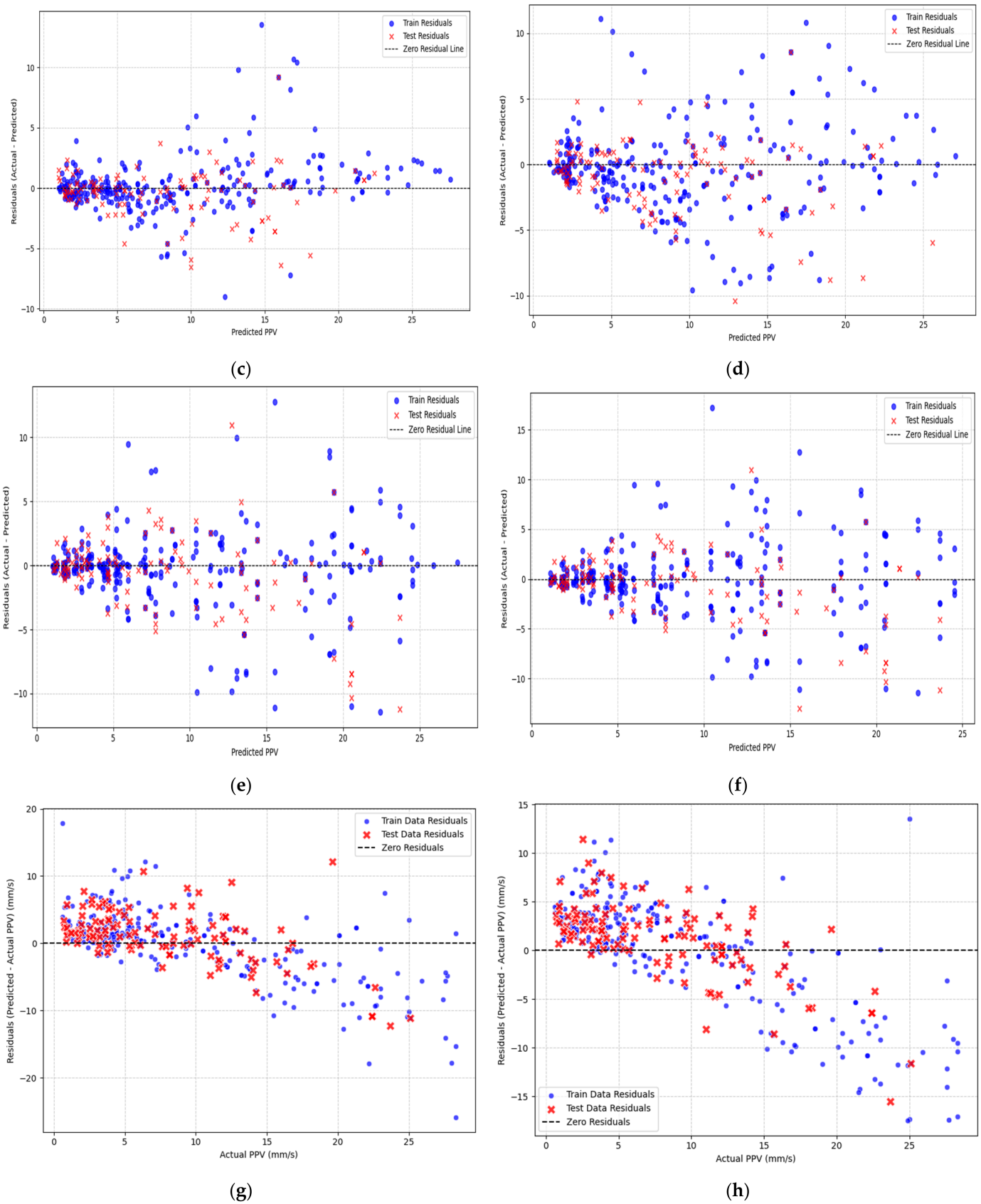
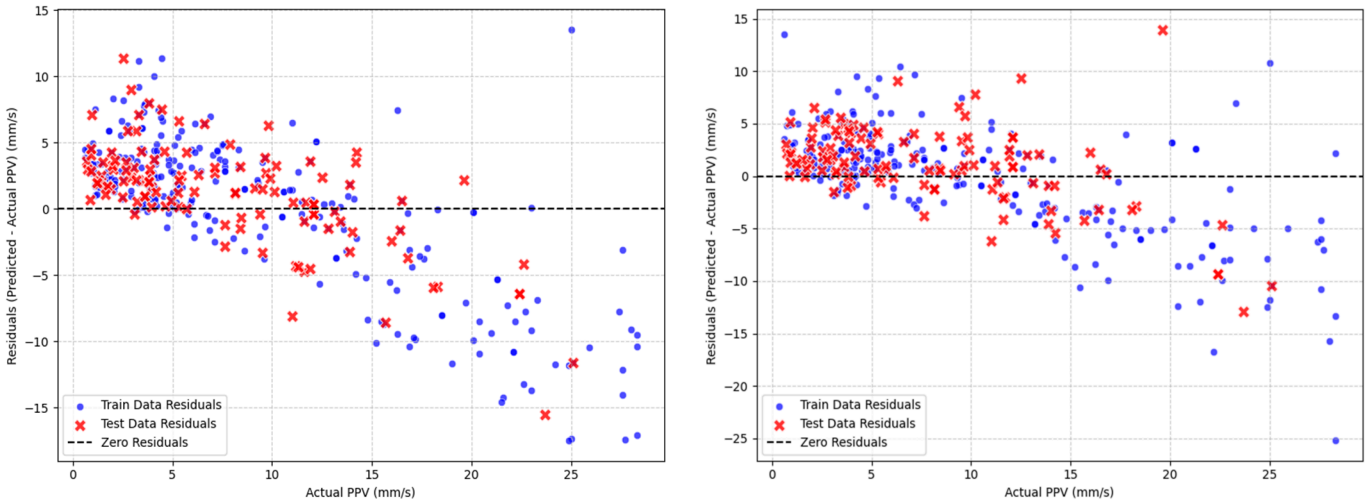
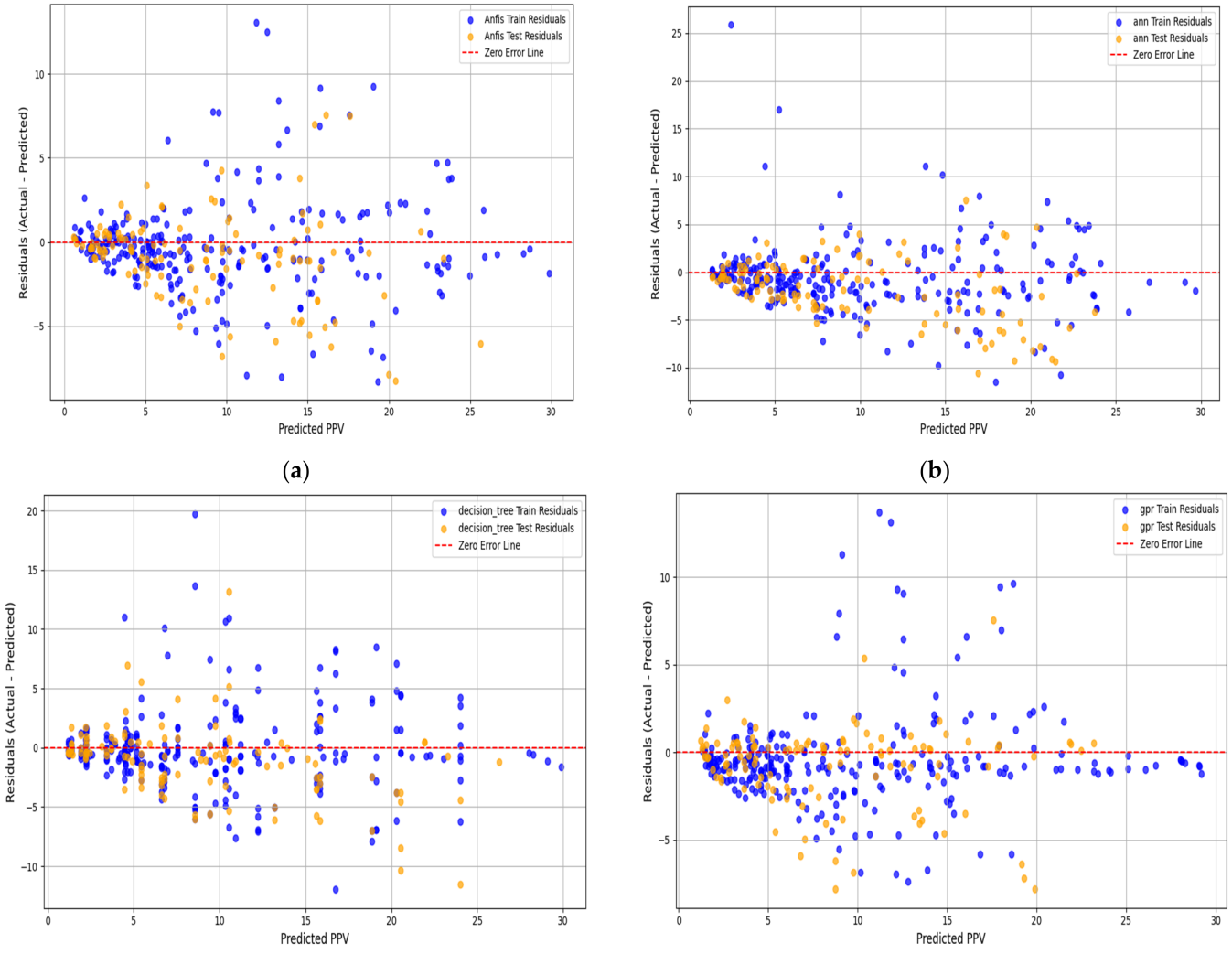


Figure 18. *Cont.*



(**i**) (**j**)

Figure 18. Residuals (predicted-measured PPV) vs. predicted PPV for (**a**) BO-GPR, (**b**) ANN, (**c**) DE-GPR, (**d**) GPR, (**e**) DT, (**f**) ANFIS, (**g**) AH, (**h**) BIS, (**i**) LK, and (**j**) USBM 2-input testing data.



(**a**) (**b**)

(**c**) (**d**)

Figure 19. *Cont*.

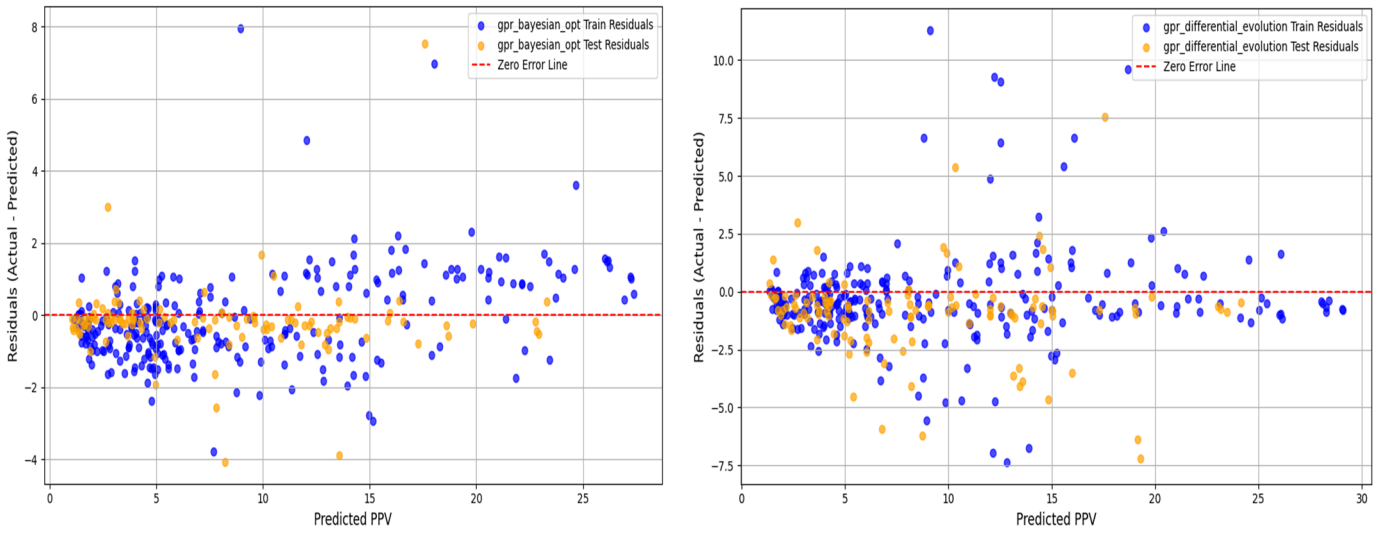


Figure 19. Residuals (predicted-measured PPV) vs. predicted PPV for (**a**) ANFIS, (**b**) ANN, (**c**) DT, (**d**) GPR, (**e**) BO-GPR, and (**f**) DE-GPR 3-input testing data.

A similar trend is observed during the testing phase, confirming the robustness and generalization of the optimized models. The BO-GPR model again produced the best results, with r = 0.97426, MAE = 0.82750, RRMSE = 0.17228, and VAF = 94.86%, maintaining its dominance under unseen conditions. Compared with the empirical models, BO-GPR reduced the average prediction error by approximately 70–75% and improved variance explanation by nearly 40 percentage points. The DE-GPR model also maintained strong predictive capability (r = 0.93529; VAF = 87.41%), while the conventional ML models showed moderate performance. Although ANN and GPR outperformed the empirical equations, their prediction errors remained notably higher than those of the optimized models.

For the three-input models (Table 14b), incorporating rock mass rating (RMR) improved the overall predictive accuracy of all ML models. During the training phase, BO-GPR again emerged as the best-performing model, achieving r = 0.98767, MAE = 0.94174, RRMSE = 0.14070, and VAF = 97.11%, followed by DE-GPR (r = 0.96165; VAF = 92.40%). Among the standard ML models, GPR and ANFIS demonstrated stronger calibration performance than DT and ANN. The optimized models significantly reduced prediction errors compared with the baseline algorithms, highlighting the effectiveness of Bayesian and evolutionary optimization in improving the hyperparameter configuration of the GPR framework.

The testing results further confirm the superiority of the optimized approaches. The BO-GPR model achieved the highest predictive accuracy (r = 0.98381; MAE = 0.50752; RRMSE = 0.13679; VAF = 96.78%), while DE-GPR maintained the second-best performance (r = 0.94522; VAF = 89.29%). In contrast, the conventional ML models exhibited moderate predictive capability, with GPR and ANFIS outperforming DT and ANN. Overall, the inclusion of optimization algorithms markedly enhanced the predictive reliability of the GPR models.

To comprehensively compare the developed two-input and three-input models and identify the optimal predictor for blasting-induced ground vibration (BIGV), the performance metrics and model rankings obtained from the training and testing datasets were visualized using line graphs and stacked bar charts (Figure 20a–d). Although the present study advances BIGV prediction using machine learning and empirical approaches on a dataset of 386 blasting events, several limitations remain. First, the dataset was compiled

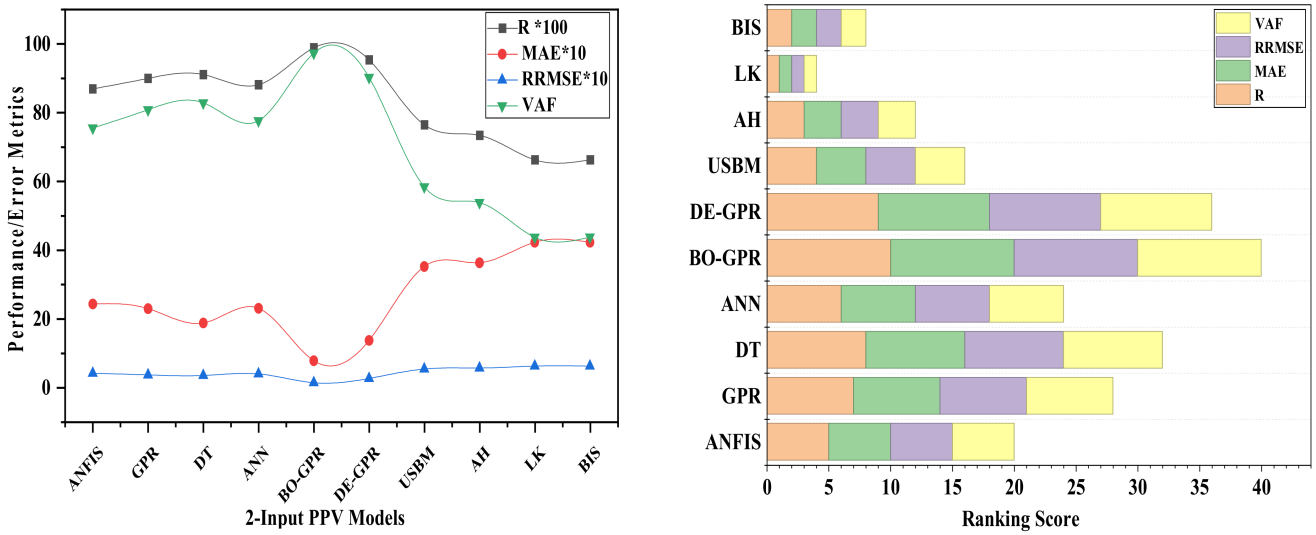
from consolidated literature sources rather than unified site-specific measurements, which may introduce variability in geological conditions and measurement protocols. Second, the model inputs were limited to charge per delay (Q), monitoring distance (R), and rock mass rating (RMR), excluding other influential blast design parameters such as burden, spacing, stemming length, powder factor, and detailed lithological characteristics. Third, the exclusion of statistically identified extreme values during preprocessing. The IQR-based removal reduced the dataset from 506 to 386 points. However, some records removed for high PPV may represent valid extreme events, particularly for structural damage assessment. Consequently, the models are mainly applicable to typical operational conditions and may not reliably predict extreme scenarios; caution is advised beyond the validated range. Also, the random 70/30 split likely places data from the same site in both subsets, potentially inflating generalization metrics. While 5-fold cross-validation was used for hyperparameter optimization, cross-site leakage remains a limitation. Future studies should focus on integrating more comprehensive blast design variables, performing uncertainty and sensitivity analyses, and validating hybrid ML-empirical frameworks using site-specific field datasets to further enhance the generalizability and regulatory applicability of BIGV prediction models. Additionally, future multi-source studies are encouraged to adopt leave-one-site-out (LOSO) cross-validation to provide a more rigorous assessment of cross-site generalization. A key limitation of the predictive framework is its scope. Standards such as DIN 4150-3 [42] and BS 7385-2 [76] assess structural safety using both PPV and the dominant vibration frequency, with limits that vary by frequency band. However, the current models predict only PPV, due to inconsistent frequency data across datasets. Future work should develop integrated PPV-frequency or multi-output models. Beyond overall performance benchmarking, direct comparisons with state-of-the-art models are essential. The proposed optimal models in this study quantitatively compare favorably with several high-performing BIGV predictors reported in the literature, including [19,20,22,24,31,40]. Model efficacy reflects not only algorithmic architecture but also, critically, the input feature vector design, Q and R, and engineered RMR in this study.

5.3. BIGV Input Variable Importance Using SHAP Analysis

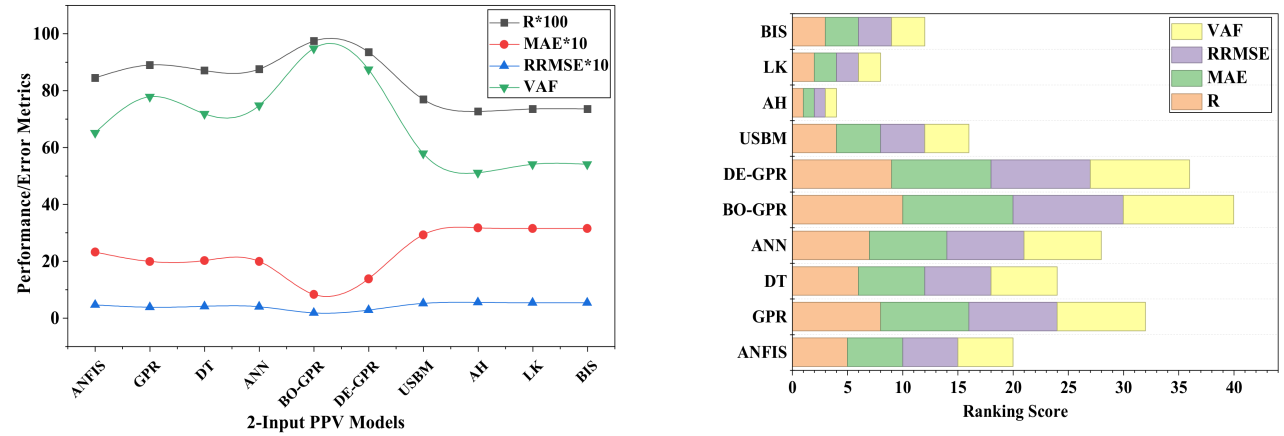
The Shapley additive exPlanations (SHAP) method is generally used to enhance understanding of the relationships between the model's input parameters and the dependent variable. It is a feature-attribution method rooted in cooperative game theory, developed to provide accurate, consistent explanations of individual predictions from complex or highly parameterized models. Its theoretical basis comes from the Shapley value framework [77], which treats model predictions as a collaborative effort among features. The contribution of each feature is determined by calculating the weighted average of its marginal contributions across all possible orders of feature inclusion. SHAP analysis was implemented using the GPR model for both two-input and three-input PPV prediction models.

In this study, all SHAP values are expressed in mm/s, which is the unit of PPV. The parameter strengths for the two-input model are shown in Figure 21. The SHAP global feature importance bar chart (Figure 21a) reveals that R (m) is the dominant predictor for PPV (mm/s), followed by Q (kg). Bar lengths represent average absolute SHAP values, quantifying each feature's overall contribution to prediction magnitude. Notably, R (m) exhibits a substantially higher average absolute SHAP value (5 mm/s) than Q (kg), which exhibited an average absolute SHAP value of 2.98 mm/s, establishing it as the primary driver of model outputs. This finding is highly consistent with conventional blasting practice, indicating that the distance from the point of blast initiation to the seismograph (R, m) is the primary factor influencing the generation and attenuation of PPV. The SHAP summary plot (Figure 21b) further elucidates directional (nonlinear) positive and negative

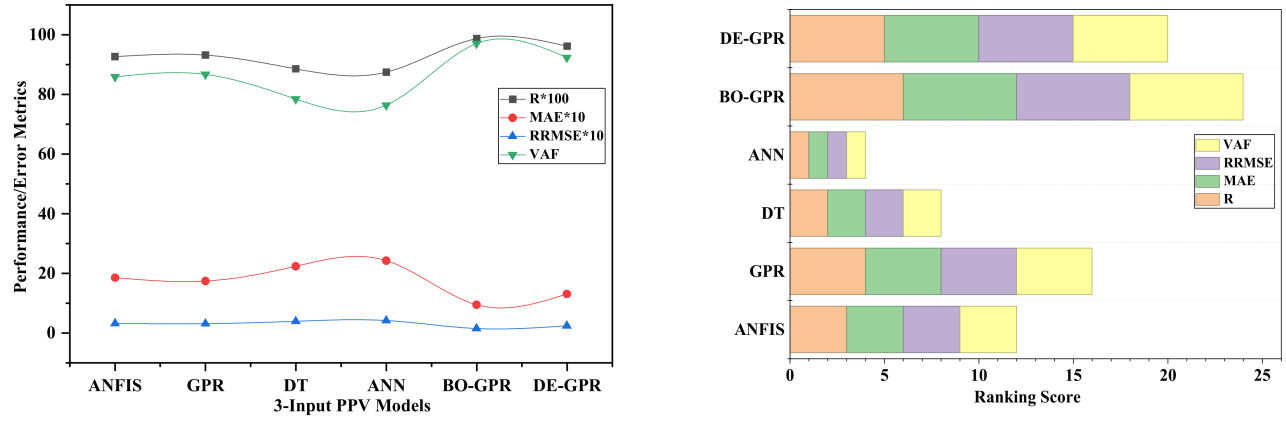
correlations between feature values and the PPV. High Q (kg) values (red dots) yield positive SHAP values, elevating the predicted PPV (mm/s), whereas low Q (kg) values (blue dots) yield negative SHAP values, confirming a direct positive relationship. Conversely, high R (m) values are associated with negative SHAP values, reducing PPV (mm/s), whereas low R (m) values are associated with positive SHAP values, indicating an inverse relationship.



(**a**)

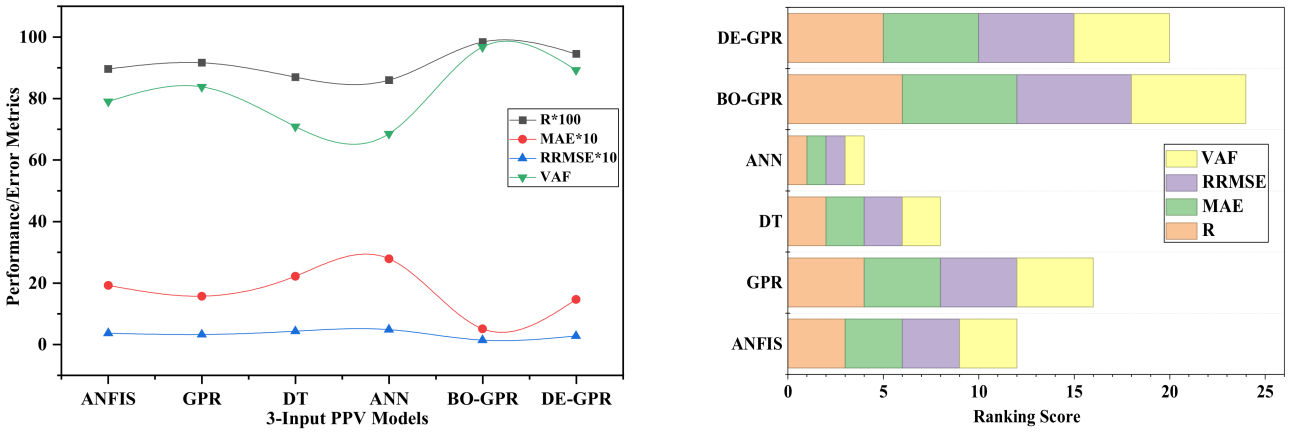


(**b**)



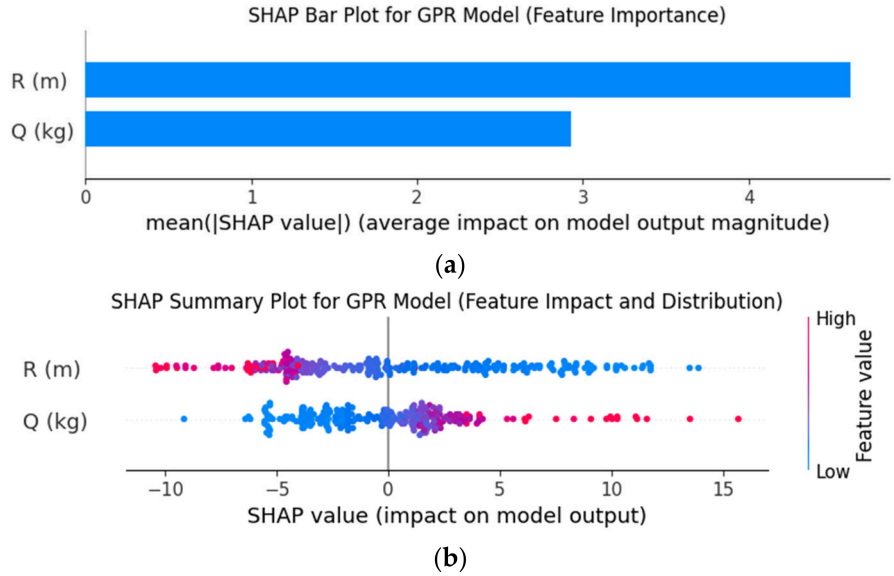
(**c**)

Figure 20. *Cont.*



(**d**)

Figure 20. PPV models comparison (**a**) 2-input training data performance evaluation results and ranking scores, (**b**) 2-input testing data performance evaluation results and ranking scores, (**c**) 3-input training data performance evaluation results and ranking scores, and (**d**) 3-input testing data performance evaluation results and ranking scores. Note: * = multiply by.



(**b**)

Figure 21. (**a**) SHAP variable importance bar chart (2-input model); (**b**) SHAP summary plot displaying the distribution of SHAP values (violin plot) for 2-input model.

For the three-input PPV model, the SHAP global feature importance bar chart (Figure 22a), based on mean absolute SHAP values, ranks R (m) (3.4 mm/s) as the most influential predictor of PPV (mm/s), followed by RMR (1.7 mm/s), with Q (kg) (1.4 mm/s) showing the least impact. This finding also aligns with conventional blasting and rock mechanics practices, indicating that the distance from the point of blast initiation to the seismograph (R, m) and the rock mass inherent strength property greatly influence the generation and attenuation of PPV. This ranking quantifies each feature's contribution to model predictions, highlighting standoff distance as the most significant factor. The SHAP summary violin plot (Figure 22b) reveals directional effects. For R (m), higher values (red dots) yield negative SHAP values, reducing PPV (mm/s), whereas lower values (blue dots) yield positive SHAP values, confirming an inverse relationship consistent with vibration attenuation. Conversely, higher Q (kg) (red) produces positive SHAP values, elevating PPV (mm/s), whereas lower values (blue) decrease it, reflecting a direct charge-weight relationship. RMR shows minor, dispersed effects, with higher values slightly lowering

predictions, suggesting better rock quality dampens vibrations. This finding is consistent with established principles of rock mass wave attenuation. Higher RMR values indicate better-quality rock mass with fewer discontinuities and greater intact strength, enhancing elastic stiffness and promoting more rapid distribution of blast-induced seismic energy over a larger volume, thereby reducing PPV. Reduced discontinuity density further limits wave scattering and resonance amplification. The negative SHAP values associated with higher RMR confirm a geomechanically meaningful and physically interpretable attenuation relationship.

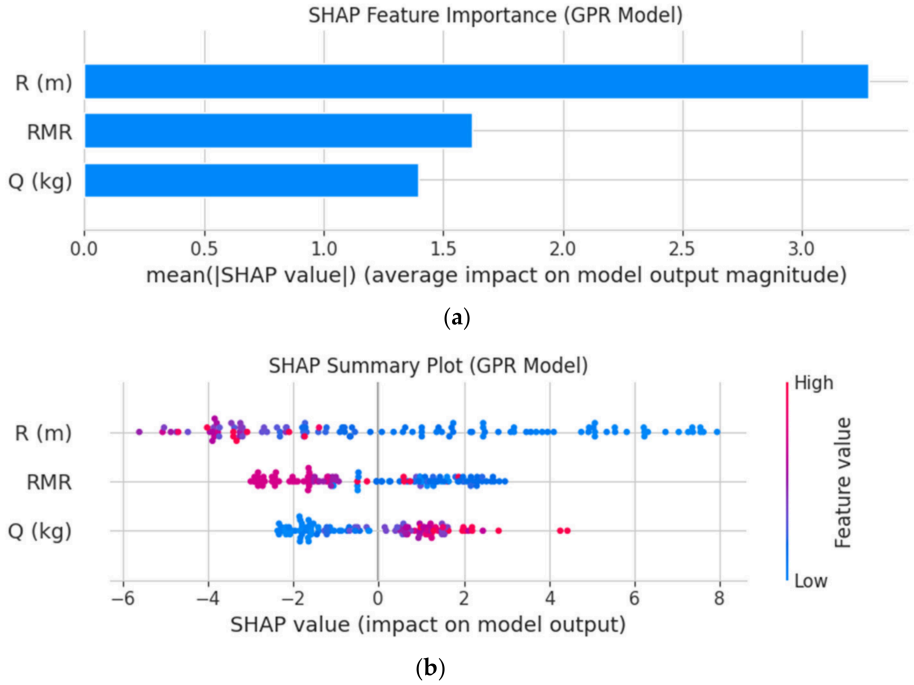


Figure 22. (**a**) SHAP variable importance bar chart (3-input model); (**b**) SHAP summary plot displaying the distribution of SHAP values (violin plot) for 3-input model.

5.4. Relevance to Sustainable Mining and Environmental Footprint Reduction

Accurate prediction of PPV using the developed optimal ANN-LM model, particularly with the three-input configuration (r = 0.907, VAF = 82.3%), enables precision blast design that minimizes environmental disturbance, including PPV. By optimizing charge per delay (Q) and monitoring distance (R), the optimal model can help protect adjacent ecosystems, water bodies, and communities from excessive ground vibrations that could disrupt habitats or contaminate water bodies. The optimum model and other developed models could also substantially reduce noise and airblast overpressure pollution, aligning with established attenuation principles where increased R exponentially lowers PPV propagation.

Enhanced structural integrity and operational safety result from predictions maintained below site-specific damage thresholds, preventing resonance-induced failures in mining infrastructure and nearby buildings while alleviating psychological distress among residents. The models' superior performance over empirical alternatives (e.g., USBM, AH) ensures reliable forecasting, reducing over-blasting that leads to excessive energy consumption, explosive wastage, and associated greenhouse gas emissions.

Regulatory compliance is strengthened through verifiable PPV forecasts that adhere to international standards (e.g., [42,76]), thereby mitigating litigation risks associated with vibration-related disputes prevalent in mining operations. The integration of RMR refines site-specific predictions, providing defensible data for community consultations and permitting processes that strengthen the social license to operate.

To facilitate practical field application, we propose a structured operational workflow for blasting engineers to utilize the BO-GPR model in precision blast design (Figure 23). Users input site-specific R and RMR, along with regulatory/project-specific PPV thresholds. The model iteratively calculates the maximum allowable Q to ensure predicted vibrations remain within safe limits. This approach enhances decision making and the model's applicability to real-world blasting operations.

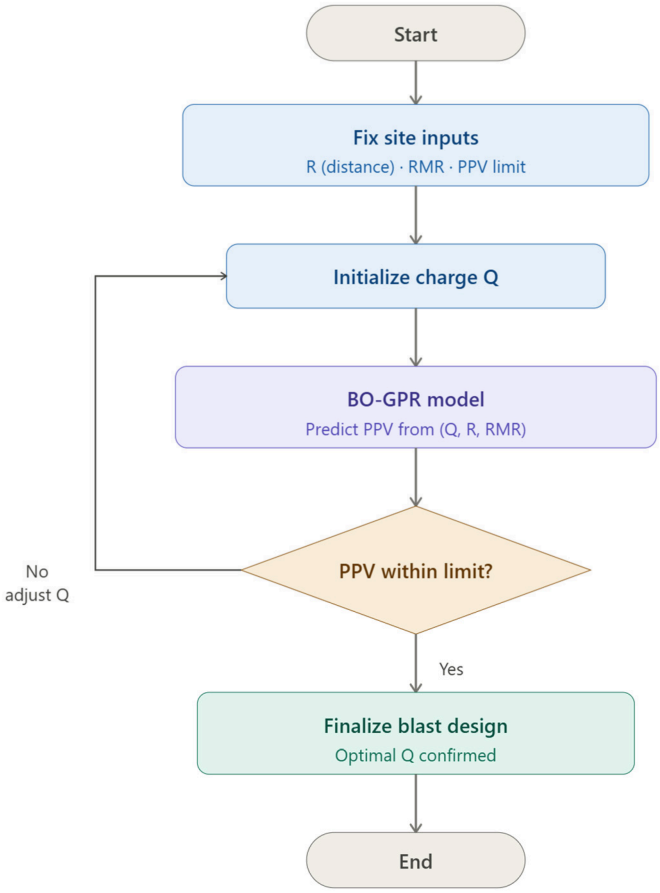


Figure 23. Operational workflow for applying the BO-GPR model in precision blast design.

These predictive tools should be integrated seamlessly into comprehensive environmental management systems, supporting real-time blast optimization aligned with Sustainable Development Goal 12 (responsible consumption and production). Hybrid ML-empirical frameworks adapt to geological variability, enabling precision blasting techniques such as pre-splitting that preserve biodiversity, minimizing flyrock, and enhancing downstream efficiency in loading, hauling, and crushing cycles. Deployment within mine planning software facilitates environmental, social, and governance (ESG) reporting, continuous monitoring via Internet of Things (IoT) seismographs, and proactive mitigation, ultimately balancing critical mineral supply with reduced ecological footprint.

6. Conclusions

This study developed and evaluated a comprehensive framework for predicting PPV using both machine learning algorithms and calibrated empirical equations based on a harmonized database of 386 blasting events. The developed models incorporated charge per delay (Q), monitoring distance (R), and RMR as predictor variables, enabling investigation of both two- and three-input predictive configurations. In addition to conventional machine learning models, including ANN-LM, ANFIS, GPR, and decision trees, two optimization

strategies, Bayesian optimization (BO) and differential evolution (DE), were employed to enhance the predictive performance of the optimum stand-alone GPR model. This integrated methodology advances blast design from empirical approximations to precise, site-adaptive predictions that support operational efficiency and regulatory compliance. From this study, the following key findings and conclusions are drawn:

1. The comparative results demonstrate that optimization-enhanced models significantly outperform both traditional ML algorithms and empirical scaled-distance equations (USBM, AH, LK, and BIS). Among all developed models, the BO-GPR model achieved the best predictive performance during testing for both the two-input and three-input configurations, with r = 0.97426 and r = 0.98381, respectively, together with the lowest error indices (MAE = 0.82750 and 0.50752) and the highest variance explanation (VAF = 94.86% and 96.78%). The DE-GPR model consistently ranked second, confirming that metaheuristic optimization substantially enhances the predictive capability of the baseline GPR model. In comparison, the conventional ML models demonstrated moderate predictive accuracy, while the empirical models produced the weakest results due to their limited ability to represent nonlinear blast dynamics.
2. The incorporation of rock mass rating (RMR) as an additional input parameter further improved the predictive reliability and stability of the models, underscoring the importance of accounting for rock mass characteristics in vibration prediction studies. SHAP analysis revealed that monitoring distance (R) is the most influential factor controlling PPV attenuation, followed by charge per delay (Q). In contrast, RMR provides a stabilizing influence on the three-input models. These findings are consistent with established blasting mechanics while demonstrating the capability of machine learning techniques to quantify complex nonlinear interactions among blasting variables.
3. From a practical perspective, the optimized BO-GPR framework provides a powerful predictive tool for pre-blast PPV forecasting and blast design optimization. The model can assist engineers in determining safe charge-per-delay limits and appropriate standoff distances to maintain vibration levels below regulatory thresholds such as those specified in DIN 4150. By enabling more accurate prediction and control of blasting-induced ground vibrations, the proposed framework contributes to improved operational safety, reduced environmental disturbance, and enhanced community acceptance of mining activities.
4. The BO-GPR and DE-GPR models were trained on a 386-event database spanning Q = 20.5–313 kg (charge per delay) and R = 45–675 m (monitoring distance), with RMR = 32–81 (Table 4). Field engineers should exercise caution when applying these models outside their ranges, particularly for near-field (R < 45 m) conditions or massive blasts (Q > 313 kg), where extrapolation may yield unreliable predictions.
5. Overall, this study demonstrates that optimization-enhanced machine learning models, particularly BO-GPR and DE-GPR, offer a robust and reliable approach for predicting blasting-induced ground vibration in complex geological environments. Their integration into mine planning and environmental monitoring systems can support precision blasting strategies that minimize explosive consumption, reduce ecological disturbance, and advance sustainable mining practices aligned with global sustainability objectives (SDG 12).

Author Contributions: Conceptualization, N.O.O.; methodology, N.O.O.; software, N.O.O.; validation, N.O.O.; formal analysis, N.O.O.; investigation, N.O.O.; resources, N.O.O.; data curation, N.O.O.; writing—original draft preparation, N.O.O.; writing—review and editing, N.O.O.; visualization,

N.O.O.; supervision, H.G.; project administration, H.G.; funding acquisition, H.G. All authors have read and agreed to the published version of the manuscript.

Funding: The authors received no external funding for this research.

Data Availability Statement: The data used in this study will be made available by the corresponding author upon a reasonable request.

Acknowledgments: The first author gratefully acknowledges the Global Excellence and Stature (GES) Fellowship at the University of Johannesburg and the Sibanye-Stillwater Center for Sustainable Mining in the Department of Mining Engineering and Mine Surveying at the University of Johannesburg for providing the postdoctoral research fellowship opportunity and for their support in the preparation of this manuscript.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PPV	Peak Particle Velocity
BIGV	Blasting-Induced Ground Vibration
USMB	United States Bureau of Mines
AH	Ambraseys and Hendron
R	Correlation Coefficient
RRMSE	Relative Root Mean Square Error
VAF	Variance Accounted For
MAE	Mean Absolute Error
LK	Langefors and Kihlstrom
BIS	Bureau of Indian Standards
ML	Machine Learning
ANN	Artificial Neural Network
ANFIS	Adaptive Neuro-Fuzzy Inference Systems
GPR	Gaussian Process Regression
DT	Decision Trees
BO	Bayesian Optimization
DE	Differential Evolution
SHAP	Shapley Additive Explanations
RMR	Rock Mass Rating
SDG	Sustainable Development Goals

References

1. Monjezi, M.; Ahmadi, M.; Sheikhan, M.; Bahrami, A.; Salimi, A.R. Predicting blast-induced ground vibration using various types of neural networks. *Soil Dyn. Earthq. Eng.* **2010**, *30*, 1233–1236. [CrossRef]
2. Michaux, S.; Djordjevic, N. Influence of explosive energy on the strength of the rock fragments and SAG mill throughput. *Miner. Eng.* **2005**, *18*, 439–448. [CrossRef]
3. Ogunsola, N.; Pureun, J.; Cho, S.; Kim, Y. An artificial neural network-based model for blast-induced ground vibration prediction in urban area. In *Rock Dynamics: Progress and Prospect*; CRC Press: Boca Raton, FL, USA, 2023; Volume 2, pp. 229–234. [CrossRef]
4. Ogunsola, N.O.; Shin, C.; Lawal, A.I.; Kim, Y.K.; Cho, S. Blasting-induced ground vibration modeling in tunnel excavation: A comparative study of ANN, hybrid ANNs, and empirical Models. *Rock Mech. Rock Eng.* **2025**, *58*, 2711–2737. [CrossRef]
5. Dowding, C.H. *Blast Vibration Monitoring and Control*; Prentice-Hall Inc.: Englewood Cliffs, NJ, USA, 1985; p. 297.
6. Ogunsola, N.O.; Kim, Y.G.; Cho, S. Application of ANN and metaheuristic algorithm optimized ANNs for the prediction of blasting-induced ground vibration in urban areas. *Geomech. Eng.* **2025**, *43*, 483–505.
7. Frank, A.; Kolapo, P.; Ogunsola, N.; Munemo, P.; Akinola, A. Application of improved blasting techniques in open pit mining for maximum productivity: A case of Oakyam Quarry Limited, Ogun State, Nigeria. *Science* **2022**, *10*, 12–23. [CrossRef]
8. Ogunsola, N.O.; Lawal, A.I.; Saliu, M.A. Variations of physico-mechanical, mineralogical and geochemical properties of marble under the influence of weathering. *Min. Miner. Depos.* **2019**, *13*, 95–102. [CrossRef]

9. Khademian, A.; Bagherpour, R. Environmentally sustainable mining through proper selection of explosives in blasting operation. *Environ. Earth Sci.* **2017**, *76*, 166. [CrossRef]
10. Bakhtavar, E.; Hosseini, S.; Hewage, K.; Sadiq, R. Green blasting policy: Simultaneous forecast of vertical and horizontal distribution of dust emissions using artificial causality-weighted neural network. *J. Clean. Prod.* **2021**, *283*, 124562. [CrossRef]
11. Wang, Z.M.; Zhou, W.; Jiskani, I.M.; Ding, X.H.; Liu, Z.C.; Qiao, Y.Z.; Luan, B. Dust reduction method based on water infusion blasting in open-pit mines: A step toward green mining. *Energy Sources Part A Recovery Util. Environ. Eff.* **2025**, *47*, 6029–6043. [CrossRef]
12. Ogunsola, N.O.; Kim, Y.G.; Cho, S. Machine Learning-Based Prediction Models of Airblast Overpressure in Tunnel Excavation for Sustainable Development. In Proceedings of the ARMA US Rock Mechanics/Geomechanics Symposium, Santa Fe, NM, USA, 8–11 June 2025; p. D022S020R005. [CrossRef]
13. Jiskani, I.M.; Cai, Q.; Zhou, W.; Shah, S.A.A. Green and climate-smart mining: A framework to analyze open-pit mines for cleaner mineral production. *Resour. Policy* **2021**, *71*, 102007. [CrossRef]
14. Ambraseys, N.N.; Hendron, A.J.; Ambraseys, N.R.; Hendron, A.J. Dynamic behavior of rock masses. In *Rock Mechanics in Engineering Practice*; Stagg, K., Wiley, J., Eds.; Wiley: London, UK, 1968; pp. 203–207.
15. Duvall, W.I.; Fogelson, D.E. *Review of Criteria for Estimating Damage to Residences from Blasting Vibrations*; US Department of the Interior, Bureau of Mines: Washington, DC, USA, 1962; Volume 5968.
16. Langefors, U.; Kihlstrom, B. *The Modern Techniques of Rock Blasting*; John Wiley and Sons: New York, NY, USA, 1963.
17. Rai, R.; Singh, T.N. A new predictor for ground vibration prediction and its comparison with other predictors. *Indian J. Eng. Mater. Sci.* **2005**, *11*, 178–184.
18. Roy, P.P. Prediction and control of ground vibration due to blasting. *Colliery Guard.* **1991**, *239*, 215–229.
19. Armaghani, D.J.; Hasanipanah, M.; Amnieh, H.B.; Mohamad, E.T. Feasibility of ICA in approximating ground vibration resulting from mine blasting. *Neural Comput. Appl.* **2018**, *29*, 457–465. [CrossRef]
20. Nguyen, H.; Bui, X.N.; Moayedi, H. A comparison of advanced computational models and experimental techniques in predicting blast-induced ground vibration in open-pit coal mine. *Acta Geophys.* **2019**, *67*, 1025–1037. [CrossRef]
21. Ogunsola, N.O.; Shin, C.; Kim, Y.K.; Lawal, A.I.; Cho, S. Development of an ANN-based closed-form equation for the prediction of airblast overpressure induced by construction rock excavation blasting in urban areas. *Appl. Sci.* **2022**, *12*, 12654. [CrossRef]
22. Hasanipanah, M.; Faradonbeh, R.S.; Amnieh, H.B.; Armaghani, D.J.; Monjezi, M. Forecasting blast-induced ground vibration developing a CART model. *Eng. Comput.* **2017**, *33*, 307–316. [CrossRef]
23. Lawal, A.I.; Kwon, S.; Hammed, O.S.; Idris, M.A. Blast-induced ground vibration prediction in granite quarries: An application of gene expression programming, ANFIS, and sine cosine algorithm optimized ANN. *Int. J. Min. Sci. Technol.* **2021**, *31*, 265–277. [CrossRef]
24. Lawal, A.I. An artificial neural network-based mathematical model for the prediction of blast-induced ground vibration in granite quarries in Ibadan, Oyo State, Nigeria. *Sci. Afr.* **2020**, *8*, e00413. [CrossRef]
25. Nicholson, R.F. Determination of Blast Vibrations Using Peak Particle Velocity at Bengal Quarry. Master's Thesis, Lulea University of Technology, Lulea, Sweden, 2005.
26. Ghosh, A.; Daemen, J.K. A simple new blast vibration predictor. In Proceedings of the 24th U.S. Symposium of Rock Mechanics, College Station, TX, USA, 20–23 June 1983; p. 151e61.
27. Siskind, D.E.; Stagg, M.S.; Kopp, J.W.; Dowding, C.H. *Structure Response and Damage Produced by Ground Vibration from Surface Mine Blasting*; U.S. Department of the Interior, Bureau of Mines: Washington, DC, USA, 1980.
28. *ISI Bull IS-6922*; Criteria for Safety and Design of Structures Subjected to Underground Blast. Bureau of Indian Standard: New Delhi, India, 1973.
29. Nicholls, H.R.; Charles, F.J.; Duvall, W.I. *Blasting Vibrations and Their Effects on Structures*; U.S. Department of the Interior, Bureau of Mines: Washington, DC, USA, 1971.
30. Duvall, W.I.; Petkof, B. *Spherical Propagation of Explosion Generated Strain Pulses in Rock*; U.S. Department of the Interior, Bureau of Mines: Washington, DC, USA, 1959.
31. Amiri, M.; Bakhshandeh Amnieh, H.; Hasanipanah, M.; Mohammad Khanli, L. A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure. *Eng. Comput.* **2016**, *32*, 631–644. [CrossRef]
32. Komadja, G.C.; Rana, A.; Glodji, L.A.; Anye, V.; Jadaun, G.; Onwualu, P.A.; Sawmliana, C. Assessing ground vibration caused by rock blasting in surface mines using machine-learning approaches: A comparison of CART, SVR and MARS. *Sustainability* **2022**, *14*, 11060. [CrossRef]
33. Arthur, C.K.; Bhatawdekar, R.M.; Mohamad, E.T.; Sabri, M.M.S.; Bohra, M.; Khandelwal, M.; Kwon, S. Prediction of blast-induced ground vibration at a limestone quarry: An artificial intelligence approach. *Appl. Sci.* **2022**, *12*, 9189. [CrossRef]
34. Fissha, Y.; Ikeda, H.; Toriya, H.; Owada, N.; Adachi, T.; Kawamura, Y. Evaluation and prediction of blast-induced ground vibrations: A Gaussian process regression (GPR) approach. *Mining* **2023**, *3*, 659–682. [CrossRef]

35. Dzimunya, N.; Besa, B.; Nyirenda, R. Prediction of ground vibrations induced by bench blasting using the random forest algorithm. *J. South. Afr. Inst. Min. Metall.* **2023**, *123*, 123–132. [CrossRef] [PubMed]
36. Huang, Y.; Zhou, Z.; Li, M.; Luo, X. Prediction of ground vibration induced by rock blasting based on optimized support vector regression models. *CMES-Comput. Model. Eng. Sci.* **2024**, *139*, 3147–3165. [CrossRef]
37. Ala, C.K.; Mayaluri, Z.L.; Kaushik, A.; Parveen, N.; Saxena, S.; Zamani, A.T.; Muduli, D. An explainable AI-based framework for predicting and optimizing blast-induced ground vibrations in surface mining. *Results Eng.* **2025**, *27*, 106046. [CrossRef]
38. Ullah, S.; Ren, G.; Ge, Y.; Kinyua, E.M. Evaluation and Prediction of Blast-Induced Ground Vibration Using Gaussian Process Regression at Saindak Copper–Gold Open Pit Mine, Pakistan. *Min. Metall. Explor.* **2025**, *42*, 155–170. [CrossRef]
39. Rong, K.; Xu, X.; Wang, H.; Yang, J. Predicting the blast-induced ground vibration with support vector regression optimized by five swarm algorithms. *Earth Sci. Inform.* **2025**, *18*, 424. [CrossRef]
40. Ravikumar, A.; Vardhan, H.; Sarma, K.V. Prediction of blast-induced ground vibrations in limestone mine using Multiple Linear Regression (MLR) Analysis. *J. Sustain. Min.* **2025**, *24*, 359–368. [CrossRef]
41. Asteris, P.G.; Armaghani, D.J. An empirical-driven machine learning (EDML) approach to predict PPV caused by quarry blasting. *Bull. Eng. Geol. Environ.* **2025**, *84*, 200. [CrossRef]
42. *DIN4150-3*; DIN4150-3 Structural Vibration Part 3: Effects of Vibration on Structures. DIN German Institute for Standardization: Berlin, Germany, 1999.
43. Hartman, H.L.; Mutmansky, J.M. *Introductory Mining Engineering*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2002.
44. Dumakor-Dupey, N.K.; Arya, S.; Jha, A. Advances in blast-induced impact prediction—A review of machine learning applications. *Minerals* **2021**, *11*, 601. [CrossRef]
45. Gorman, M.R.; Dzombak, D.A. A review of sustainable mining and resource management: Transitioning from the life cycle of the mine to the life cycle of the mineral. *Resour. Conserv. Recycl.* **2018**, *137*, 281–291. [CrossRef]
46. Dubiński, J. Sustainable development of mining mineral resources. *J. Sustain. Min.* **2013**, *12*, 1–6. [CrossRef]
47. Ak, H.; Iphar, M.; Yavuz, M.; Konuk, A. Evaluation of ground vibration effect of blasting operations in a magnesite mine. *Soil Dyn. Earthq. Eng.* **2009**, *29*, 669–676. [CrossRef]
48. Kumar, A.; Singh, P.; Sharma, S.K.; Kishore, N.; Singh, C.S. Quantitative assessment of BIGV and structural response based on velocity and frequency around an opencast mine. *Curr. Sci.* **2021**, *121*, 275–285. [CrossRef]
49. Paurush, P.; Rai, P. Evaluation of ground vibrations induced by blasting in a limestone quarry. *Curr. Sci.* **2022**, *122*, 1279–1287. [CrossRef]
50. Xue, X.; Yang, X. Predicting blast-induced ground vibration using general regression neural network. *J. Vib. Control.* **2014**, *20*, 1512–1519. [CrossRef]
51. Agrawal, A.; Choudhary, B.S.; Murthy, V.M.S.R.; Murmu, S. Impact of bedding planes, delay interval and firing orientation on blast induced ground vibration in production blasting with controlling strategies. *Measurement* **2022**, *202*, 111887. [CrossRef]
52. Kahriman, A. Analysis of ground vibrations caused by bench blasting at can open-pit lignite mine in Turkey. *Environ. Geol.* **2002**, *41*, 653–661. [CrossRef]
53. Kuzu, C.; Ergin, H. An assessment of environmental impacts of quarry-blasting operation: A case study in Istanbul, Turkey. *Environ. Geol.* **2005**, *48*, 211–217. [CrossRef]
54. Ozer, U.; Karadogan, A.; Kahriman, A.; Aksoy, M. Bench blasting design based on site-specific attenuation formula in a quarry. *Arab. J. Geosci.* **2013**, *6*, 711–721. [CrossRef]
55. Görgülü, K.; Arpaz, E.; Uysal, Ö.; Durutürk, Y.S.; Yüksek, A.G.; Koçaslan, A.; Dilmaç, M.K. Investigation of the effects of blasting design parameters and rock properties on blast-induced ground vibrations. *Arab. J. Geosci.* **2015**, *8*, 4269–4278. [CrossRef]
56. Hudaverdi, T.; Akyildiz, O. Investigation of the site-specific character of blast vibration prediction. *Environ. Earth Sci.* **2017**, *76*, 138. [CrossRef]
57. Hudaverdi, T. Application of multivariate analysis for prediction of blast-induced ground vibrations. *Soil Dyn. Earthq. Eng.* **2012**, *43*, 300–308. [CrossRef]
58. Pradhan, S.K. A Neural Network Approach for the Prediction of Ground Vibrations Induced Due to Blasting. Doctoral Dissertation, National Institute of Technology Rourkela, Rourkela, India, 2016.
59. Aruna, M.; Vardhan, H.; Tripathi, A.K.; Parida, S.; Raja Sekhar Reddy, N.V.; Sivalingam, K.M.; Li, Y.; Elumalai, P.V. Enhancing safety in surface mine blasting operations with IoT based ground vibration monitoring and prediction system integrated with machine learning. *Sci. Rep.* **2025**, *15*, 3999. [CrossRef]
60. Görgülü, K.; Arpaz, E.; Demirci, A.; Koçaslan, A.; Dilmaç, M.K.; Yüksek, A.G. Investigation of blast-induced ground vibrations in the Tülü boron open pit mine. *Bull. Eng. Geol. Environ.* **2013**, *72*, 555–564. [CrossRef]
61. Ogunsola, N.O.; Quadri, A.I.; Bankole, A.O. Comparative Modeling of Compressive and Tensile Strengths in Thermally Exposed Pozzolanic Mortar Using ANN-LM, ANFIS, and MVRA with Closed-Form Equations. *Iran. J. Sci. Technol. Trans. Civ. Eng.* **2025**, *49*, 1–30. [CrossRef]

62. Ogunsola, N.O.; Oh, S.; Min, G.; Cho, S. Soft computing-based modeling of high strain rate-controlled rock compressive strength: A comparative analysis of ANN and metaheuristic algorithms optimized ANN models. *Multiscale Multidiscip. Model. Exp. Des.* **2025**, *8*, 422. [CrossRef]
63. Tijani, I.A.; Lawal, A.I.; Ogunsola, N.O.; Kwon, S. Prediction of ultimate strength of FRP-confined predamaged concrete using backward multiple regression motivated soft computing methods. *Sci. Iran.* 2023, *in press*. [CrossRef]
64. Ogunsola, N.O.; Grobler, H. Soft computing techniques for predicting mode-II fracture toughness: A cost-effective approach to advancing rock mechanics' design and sustainability. In Proceedings of the SANIRE (AFRIROCK) Symposium SAIMM, Muldersdrift, South Africa, 25–26 June 2025.
65. Asteris, P.G.; Roussis, P.C.; Douvika, M.G. Feed-forward neural network prediction of the mechanical properties of sandcrete materials. *Sensors* **2017**, *17*, 1344. [CrossRef]
66. Lawal, A.I.; Oniyide, G.O.; Kwon, S.; Onifade, M.; Köken, E.; Ogunsola, N.O. Prediction of Mechanical Properties of Coal from Non-destructive Properties: A Comparative Application of MARS, ANN, and GA. *Nat. Resour. Res.* **2021**, *30*, 4547–4563. [CrossRef]
67. Jang, J.S. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [CrossRef]
68. Walia, N.; Singh, H.; Sharma, A. ANFIS: Adaptive neuro-fuzzy inference system-a survey. *Int. J. Comput. Appl.* **2015**, *123*, 32–38. [CrossRef]
69. Ebden, M. *Gaussian Processes for Regression: A Quick Introduction*; The Website of Robotics Research Group in the Department of Engineering Science, University of Oxford: Oxford, UK, 2008; Volume 91, pp. 424–436.
70. Schulz, E.; Speekenbrink, M.; Krause, A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *J. Math. Psychol.* **2018**, *85*, 1–16. [CrossRef]
71. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Wadsworth International Group: Belmont, CA, USA, 1984; Volume 432, p. 9.
72. Tiryaki, B. Predicting intact rock strength for mechanical excavation using multivariate statistics, artificial neural networks, and regression trees. *Eng. Geol.* **2008**, *99*, 51–60. [CrossRef]
73. Muleta, M.K.; McMillan, J.; Amenu, G.G.; Burian, S.J. Bayesian approach for uncertainty analysis of an urban storm water model and its application to a heavily urbanized watershed. *J. Hydrol. Eng.* **2013**, *18*, 1360–1371. [CrossRef]
74. Storn, R.; Price, K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **1997**, *11*, 341–359. [CrossRef]
75. Jiang, T.; Jiang, A.; Zheng, S.; Xu, M. Prediction of Blast-Induced Ground Vibration (BIGV) of Metro Construction Using Difference Evolution Algorithm Optimized Gaussian Process (DE-GP). *Shock. Vib.* **2021**, *2021*, 2847180. [CrossRef]
76. *BS 7385-2*; Evaluation and Measurement for Vibration in Buildings—Part 2: Guide to Damage Levels from Groundborne Vibration. British Standards Institution (BSI): London, UK, 1993.
77. Winter, E. *Handbook of Game Theory with Economic Applications*; Elsevier: Amsterdam, The Netherlands, 2002; Volume 3, pp. 2025–2054.

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