Physicochem. Probl. Miner. Process., 59(6), 2023, 172096

Physicochemical Problems of Mineral Processing

http://www.journalssystem.com/ppmp

ISSN 1643-1049

Prediction and optimization of tower mill grinding power consumption based on GA-BP neural network

Ziyang Wang [1], Ying Hou [1], Ahmed Sobhy [2,3]

[1] School of Mining Engineering, University of Science and Technology Liaoning, Anshan 114051, China

[2] School of Resources and Environmental Engineering, Shandong University of Technology, Zibo, 255049, China

[3] Minerals Technology Department, Central Metallurgical R&D Institute, Helwan, Cairo, 11421, Egypt

Corresponding authors: houying@ustl.edu.cn (Ying Hou), a.sobhy81@gmail.com (Ahmed Sobhy)

Abstract: Grinding is commonly responsible for the liberation of valuable minerals from host rocks but can entail high costs in terms of energy and medium consumption, but a tower mill is a unique power-saving grinding machine over traditional mills. In a tower mill, many operating parameters affect the grinding performance, such as the amount of slurry with a known solid concentration, screw mixer speed, medium filling rate, material-ball ratio, and medium properties. Thus, 25 groups of grinding tests were conducted to establish the relationship between the grinding power consumption and operating parameters. The prediction model was established based on the backpropagation "BP" neural network, further optimized by the genetic algorithm GA to ensure the accuracy of the model, and verified. The test results show that the relative error of the predicted and actual values of the backpropagation "BP" neural network prediction model within 3% was reduced to within 2% by conducting the generic algorithm backpropagation "GA-BP" neural network. The optimum grinding power consumption of 41.069 kWh/t was obtained at the predicted operating parameters of 66.49% grinding concentration, 301.86 r/min screw speed, 20.47% medium filling rate, 96.61% medium ratio, and 0.1394 material-ball ratio. The verifying laboratory test at the optimum conditions, produced a grinding power consumption of 41.85 kWh/t with a relative error of 1.87%, showing the feasibility of using the genetic algorithm and BP neural network to optimize the grinding power consumption of the tower mill.

Keywords: tower mill, grinding power consumption, energy saving, genetic algorithm, BP neural network

1. Introduction

Under the guidance of the "double carbon" goal and the global zero carbon future action plan, the work of "energy saving, reducing cost, and increasing efficiency" in mining industry is essential (de Bakker 2013; Liu et al., 2022). The most energy-consuming operation in the mining industry is comminution which consists of crushing and grinding stages (Tromans, 2008). Crushing is more energy-efficient than grinding, but grinding is becoming an increasingly essential operation in mineral processing to liberate valuable minerals from host rocks (Fuerstenau and Abouzeid, 2002). Thus, grinding can entail high costs of more than 50% in terms of energy consumption and media use (de Bakker, 2013; de Carvalho and Tavares, 2013; Gupta and Sharma, 2014). Thus, it is essential to investigate saving and consumption reduction of energy (Shi and Xie, 2016; Yuan and Zhang, 2022).

A tower mill is a kind of wet vertical stirred grinding equipment utilizing steel balls or pebble grinding media (Stief et al., 1987). It has a unique design for fine grinding, and provides power savings of up to 50% over traditional grinding equipment due to the use of much smaller grinding media (Stief et al., 1987; Valery and Jankovic, 2002). It was established for many years as a superior to ball mills (Danielle et al., 2017). For example, a tower mill can be employed for a fraction of the cost of an equivalent ball mill. The grinding kinetics of a pilot scale unit were investigated, and unlike normal grinding systems, it can be fitted with a simple first-order breakage model (Austin and Schneider, 2022).

DOI: 10.37190/ppmp/172096

In tower mills, there are many operating parameters affecting the grinding performance, such as the amount of slurry with a known solid concentration, screw mixer speed, medium filling rate, material-ball ratio, and medium properties. Tower mill has some other unique features such as less installation cost, less operation cost, less floor space, less noise, and less vibration (Stief et al., 1987). The tower mill was used in a comminution circuit instead of the energy-inefficient ball mill in the copper-lead-zinc ore processing plant after the primary jaw crusher, secondary cone crusher, and tertiary high-pressure grinding rolls. The circuit reduced the size from 12 mm to 85% finer than 75 µm, with up to 20.16% energy savings (Li et al., 2023). This type of grinding mill has been reviewed, and the advantages of these mills and their potential to revolutionize industrial processes for achieving the energy-efficient ultra-fine grinding of particles were highlighted (Kumar et al., 2023).

In terms of the excellent grinding performance of the tower mill, it is necessary to maintain the grinding energy consumption of the tower mill at a minimum, and the reasonable selection of key operation parameters is an important part. Although the power of the tower mill can be monitored in real-time, the detection of particle level distribution of grinding products is often obtained by offline manual sorting, which not only costs a lot of manpower and material resources but also the results to lag behind the site situation, which is of little significance to real-time control guidance. With the development of computer technology, the artificial neural network technology has been gradually applied in the mining engineering. For instance, the discrete element method has been used to simulate and optimize the operating factors of the tower mill, and the results indicated that, under the critical screw speed, the larger the speed the better the grinding performance, and the best screw speed was 210 r/min. Also, the results showed that the lower the filling rate the lower the grinding strength, and the larger the filling rate the lower energy consumption, and the best filling rate was 60% (Zhengming et al., 2016). Also, the backpropagation "BP" neural network algorithm was employed to model and optimize the grinding processes of alumina with diamond wheels (Warren Liao and Chen, 1994). From the simulation data obtained, it is indicated that the used neural network produces a more accurate operational model than the regression method. It is also illustrated that dissimilar to the conventional backpropagation network, proper use of the Boltzmann factor with BP can effectively avoid local minima and generate the global best solution (Warren Liao and Chen, 1994).

In this study, the relationship between grinding power consumption and operating parameters was investigated to establish the prediction model of grinding power consumption based on the "BP" neural network algorithm and to optimize the extreme value of the neural network genetic algorithm. The BP neural network algorithm is one of the most widely applied neural network models, and it is a multi-layer feedforward network performed to minimize the mean squared error "MSE" to adjust the model's parameters. The established prediction model can be utilized to guide the optimization process of grinding power consumption, which has great guiding significance for the selection of grinding operation parameters of tower mills.

2. Materials and methods

2.1. Materials

The ore sample used in the test is the rough separation concentrate of high-pressure grinding roll products in an iron mine in Hebei Province. The rough separation process of high-pressure grinding roll products is shown in Fig. 1. After the raw ore with particle size finer than 20 mm is crushed by a high-pressure grinding roll, products are wet-screened using a sieve with a sieve size of 3 mm. The under-screened product (-3 mm) is separated using a wet permanent cylinder magnetic separator to obtain the test ore sample used in this study, and its size distribution is shown in Table 1. Then the test sample which is the concentrate of the rough separation process is fed to the tower mill to investigate predict and optimize its performance.

The tower mill used in the test is TM 200-2.2. The effective volume of the mill is 13 L. The drive motor power is 2.2 kW. The size of the screw mixer is 140 mm. The filling ball medium is 5 mm and 8 mm. The output power of the tower mill is read directly from the electrical control cabinet. The tower mill has a reasonable structure, it occupies a small area, and it has a stable operation and small noise.

In the grinding process, the tower mill chamber fills with water. Then the ball media is released into the chamber from the top. Then all fresh material is fed from the bottom of the grinding chamber. The

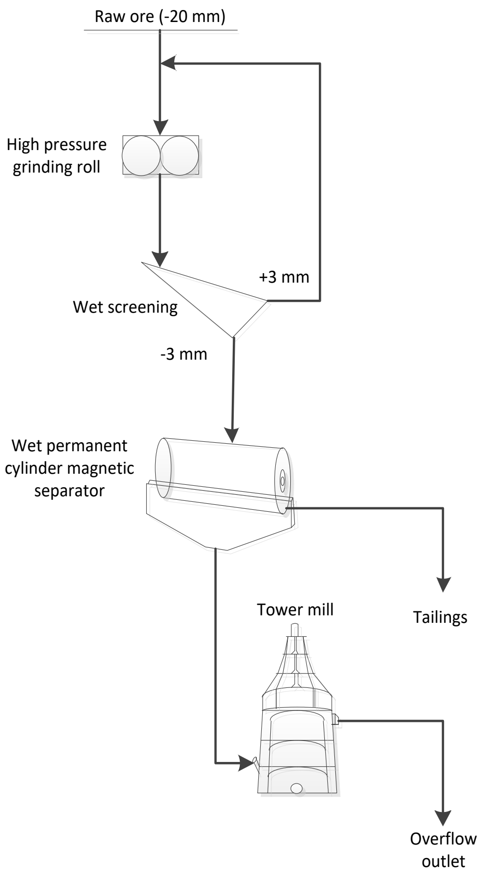

Fig. 1. Tower mill grinding after rough separation process of high-pressure grinding roll products in an iron mine

Table 1. Particle size analysis results

Size fraction (mm)	Yield (%)	Passing cumulative yield (%)
+2.0	2.45	100.00
-2.0+1.0	13.15	97.55
-1.0+0.63	15.91	84.40
-0.63+0.15	29.23	68.49
-0.15+0.074	19.98	39.26
-0.074	19.28	19.28
Total	100.00	

over-grinding is minimized because the finer particles are discharged by the upper-flow and exit through the upper-flow outlet, while the large particles remain in the lower part of the grinding chamber until the target size is produced. The ball media and the feed material can realize an orderly motion cycle and macroscopic forces balance due to the action of centrifugal, gravity, and friction forces. Thus, the material breakage is performed by attrition and not by impact, providing less grinding energy and lower media consumption. In addition, the media is easily refilled during the grinding operation from the top of the grinding chamber.

2.2. Test method

Using a fixed grinding time, a certain quantity of ore sample is weighed each time, and the grinding test is conducted under a corresponding condition. After the grinding process, the products are filtered, dried, screened, and analyzed. The work index of the grinding operation is represented by the grinding power consumption E of the newly generated -0.074 mm product. This index indicates the level of energy consumption efficiency. The equation expression is as follows:

$$E = \frac{P \cdot T}{M} \tag{1}$$

where: E is the energy consumption of the newly generated-0.074 mm product, kwh/t; M is the quantity of the newly generated -0.074 mm product, t; T is the grinding time, h; P is the power of the mill, kW.

The ratio of medium refers to the mass of 8 mm steel balls and the total mass of the steel balls, which is expressed as a percentage. The equation is expressed as follows:

$$D = \frac{M_1}{M_1 + M_2} \times 100\% \tag{2}$$

where: D is the medium ratio,%; M1 is the mass of an 8 mm ball in the mill, kg; M2 is the mass of a 5 mm ball in the mill, kg.

2.3. Test data

First, the grinding concentration, screw mixer speed, medium filling rate, medium ratio, and material-ball ratio were determined as the influencing factors. Under the premise of the grinding time of 260 s, the change trend of each test factor was obtained. A total of 25 groups of grinding tests were conducted, and the test data were processed and displayed in the form of operating parameters corresponding to the grinding power consumption. The data are shown in Table 2.

Table 2. Grinding test data sheet

#	Experimental factor					Milling power consumption (kW·h/t)
	Grinding concentration (%)	Speed (r/min)	Medium filling rate (%)	Media ratio (%)	Material/ball ratio	
1	45	360	20.77	100	0.16	59.76
2	50	360	20.77	100	0.16	52.45
3	55	360	20.77	100	0.16	49.87
4	60	360	20.77	100	0.16	48.43
5	65	360	20.77	100	0.16	44.77
6	70	360	20.77	100	0.16	47.86
7	75	360	20.77	100	0.16	50.32
8	65	340	20.77	100	0.16	44.06
9	65	320	20.77	100	0.16	43.63
10	65	300	20.77	100	0.16	44.39
11	65	280	20.77	100	0.16	45.61
12	65	260	20.77	100	0.16	47.62
13	65	320	20.77	100	0.15	43.19
14	65	320	20.77	100	0.14	44.11
15	65	320	20.77	100	0.13	45.41
16	65	320	20.77	100	0.12	47.81
17	65	320	19.23	100	0.15	46.27
18	65	320	17.69	100	0.15	49.95
19	65	320	16.15	100	0.15	54.34
20	65	320	14.62	100	0.15	60.60
21	65	320	20.77	80	0.15	44.58
22	65	320	20.77	60	0.15	47.25
23	65	320	20.77	40	0.15	50.44
24	65	320	20.77	20	0.15	55.49
25	65	320	20.77	0	0.15	61.75

2.4. Modeling approach

Genetic algorithm "GA" improvement of BP neural network prediction and optimization research is realized through MATLAB, which is mainly divided into three parts: BP neural network prediction model parameter determination, GA optimization of BP neural network prediction model, and neural network genetic algorithm function extreme value optimization.

Five sets of experimental data were randomly selected from the 25 sets of grinding test data in **Błąd! Nie można odnaleźć źródła odwołania.** as test data, with test numbers 3, 11, 15, 19, and 23 as test data, and the remaining 20 sets were used as training data.

The various operating parameters in the training data are taken as the network input, and the network parameters are constantly adjusted during the training process to influence the weights and thresholds of the network nodes so that the output of the network is close to the corresponding grinding energy consumption. the prediction grinding energy consumption values are compared with the actual values, while the correlation coefficient and mean square error are used as the evaluation indexes, to select the appropriate network parameters to determine the prediction model. Then, the genetic algorithm is used to optimize the BP neural network prediction model. The genetic algorithm initializes the population and determines the genetic operator, to determine the parameters of the GA-BP neural network prediction model. Based on the setting of GA-BP neural network prediction model parameters, the upper and lower limits of each influencing factor in the grinding data table are set to optimize the extreme value of the neural network genetic algorithm function.

3. The BP neural network prediction model

3.1. Determination of the BP neural network parameters

3.1.1. Data normalization processing

Data normalization is a method of scaling the sample data to the 0-1 range according to specific proportions. This method can transform the dimension into the dimensionless, so that the sample data is not limited by units, and make the training and prediction of the model more accurate. Scholars usually normalize the data by using the mean-variance method and the maximum minimum method to normalize the sample data with a function of mapminmax.

3.1.2. Determination of the network structure

The number of nodes in the input layer is 5 and that in the output layer is 1. This study has only 5 inputs and 1 output, which is not a complex solving problem. Therefore, the prediction model established in this study is a single-hidden layer, so the network structure is a 3-layer BP neural network.

3.1.3. Determination of the number of hidden layer nodes

The number of nodes in the hidden layer is the key factor in constructing the BP neural network prediction model. There is no clear selection method for selecting the number of nodes in the hidden layer. Usually, we refer to the previous empirical formulas for experimental exploration:

$$L = \sqrt{M+N} + A \tag{3}$$

where L is the number of nodes in the hidden layer; M is the number of input nodes; N is the number of output nodes; and A is the adjustment constant between 1 and 10.

In the prediction model, the number of input nodes M=5 and the number of output nodes N=1, so the number of hidden layer nodes L should be selected between 3 and 12. In order to select the best number of hidden layer nodes, the effect of the prediction model of different hidden layers is compared and analyzed. The preset parameters are shown in Table 3, starting the exploration test on the best number of hidden layer nodes.

Based on Table 3, the number of nodes in the hidden layer 3-12 are selected for prediction simulation respectively. The mean square error and sample correlation coefficient generated by the model fitting of the number of nodes in different hidden layers are shown in Table 4.

Table 4 shows that as the number of nodes in the hidden level increases from 3 to 9, the correlation

Table 3. The preset parameters of optimal hidden layer node number training

Factor	Parameter	Factor	Parameter
Number of input layer nodes	5	Training function	trainlm
Number of output layer nodes	1	Implied layer activation function	logsig
Frequency of training	300	Output-layer activation function	purelin
Learning rate	0.0001	Learning target	0.00001

Table 4. Mean square error and sample correlation coefficient of different hidden layer nodes

The number of hidden layer nodes	Training sample		Test sample	
	Correlation coefficient	Mean square error	Correlation coefficient	Mean square error
3	0.89753	14.6286	0.77552	10.9673
4	0.93315	8.5851	0.95853	7.4027
5	0.96771	13.024	0.96321	1.0346
6	0.93846	2.9909	0.88991	6.3567
7	0.9317	2.1631	0.96392	1.673
8	0.97841	1.8664	0.91504	5.1594
9	0.99977	1.6578	0.96289	1.9048
10	0.99301	1.2425	0.9886	2.3495
11	0.99273	2.3705	0.98177	4.4602
12	0.99332	1.3756	0.97886	4.2264

coefficient of the training sample increases, while the mean square error of the training sample decreases; the correlation coefficient of the test sample is basically rising, and the mean square error of the training sample also decreases. As the number of nodes in the hidden layer increases from 9 to 12, the correlation coefficient of the training sample is unchanged, and the mean square error of the training sample is also very small; the correlation coefficient of the test sample is basically unchanged, and the mean square error of the training sample also increases. When the number of nodes in the hidden layer is 9, the correlation coefficient reaches the maximum value is 0.99977, the mean square error of training samples is 1.6578; the correlation coefficients of test samples is 0.96289, and the mean square error of test samples is 1.9048. Whereas for 10 nodes in the hidden layer, for the training and test samples, the correlation coefficient are 0.99301 and 0.9886, and the mean square error are 1.2425 and 2.3495 respectively. Thus, for a number of nodes in the hidden layer of 9 and 10, the difference in the correlation coefficient of the training samples is smaller, but the mean square error of the training samples is smaller when the number of nodes in the hidden layer is 10. Overall, the number of hidden layer nodes of the model takes a value of 10, and its structure is shown in Fig. 2.

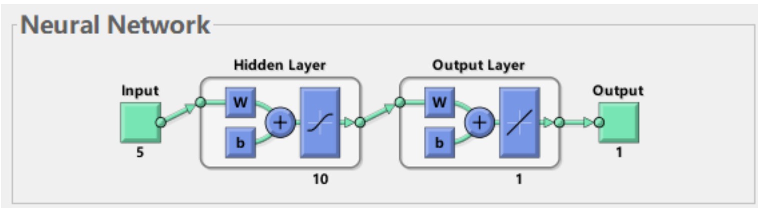

Fig. 2. The BP neural network structure diagram

3.1.4. Determination of the training function

The training function can enable the BP neural network to realize the adaptive adjustment threshold and weight based on the error, then affect the whole network structure, and play an important role in improving the accuracy of the model. In MATLAB, the commonly used training functions are trainoss, trainbfg, trainscg, traincgb, traincgp, traincgf, trainlm, trainrp and trainda. In order to select the best training function, the prediction model effects of the above training functions are used for comparison and analysis. The preset parameters are shown in Table 5, and the best training function is explored.

Based on Table 5, different training function prediction simulations are selected, and the mean square error and sample correlation coefficient generated by the model fitting of different training functions are shown in Table 6.

Table 5. The preset parameters of the best training function training

Factor	Parameter	Factor	Parameter
Number of input layer nodes	5	The number of hidden layer nodes	10
Number of output layer nodes	1	Implied layer activation function	logsig
Frequency of training	300	Output-layer activation function	purelin
Learning rate	0.0001	Learning target	0.00001

Table 6. Mean square error and sample correlation coefficient of different training functions

Training function	Training sample		Test sample	
	Correlation coefficient	Mean square error	Correlation coefficient	Mean square error
trainlm	0.99301	1.2425	0.9886	2.3495
traingda	0.89015	8.4422	0.9976	2.9565
trainrp	0.90691	9.8605	0.84941	3.7883
traincgf	0.97	4.5742	0.97919	3.801
traincgp	0.94487	3.0913	0.9155	1.9051
traincgb	0.9709	2.6715	0.91645	1.0452
trainscg	0.95799	3.1137	0.95799	4.0626
trainfg	0.97716	3.0013	0.99874	1.4918
trainoss	0.96422	2.5254	0.99491	0.98381

According to Table 6, the correlation coefficient of the training functions trainoss, trainbfg, trainscg, traincgb, traincgp, traincgf, trainrp and trainda is lower than 0.98, and the mean square error of the training sample is higher than 2.5; the correlation coefficient of the training function trainlm is 0.99301, and the mean square error of the training sample is 1.2425. Each training function selects the training function trainlm; When the training function is trainlm, the correlation coefficient of the test sample is 0.9886, the mean square error of the test sample is 2.3495, and the test sample is also effective. Taken together, the training function of this model is trainlm.

3.1.5. Determination of the activation function

In the absence of an activation function, the BP neural network can only represent a linear relationship, and its learning ability will be greatly reduced. The activation function plays an important role in the neural network by adding non-linear factors to the network to determine whether a neuron is activated and whether the neuron should be discarded or retained. In practice, log S type function logsig, tangent S type function tansig, and linear function purelin are the three common activation functions. Function purelin is a linear function, which is detrimental to the training of the BP neural network, so function purelin is often used as the output layer activation function. In conclusion, all three functions can be used in the output layer activation function, and the functions logsig, and function tansig can be used as the hidden layer activation functions.logsig-purelin, logsig-tansig, logsig-logsig, tansig-purelin, tansig-tansig, tansig-tansig and tansig-logsig as activation functions of the implied and output layer, respectively. In order to select the best activation function, the prediction model effect of the above activation function is used for comparative analysis. The preset parameters are shown in Table 7, and the best activation function is explored.

Based on Table 7, different activation function prediction simulations are selected respectively, and the mean square error and sample correlation coefficient generated by the model fitting of different activation functions are shown in Table 8.

Table 7. The preset parameters of the best activation functions

Factor	Parameter	Factor	Parameter
Number of input layer nodes	5	The number of hidden layer nodes	10
Number of output layer nodes	1	Learning rate	0.0001
Frequency of training	300	Learning target	0.00001
Training function	trainlm		

Table 8. Mean square error and sample correlation coefficient of different activation functions

Activation function	Training sample		Test sample	
	Correlation coefficient	Mean square error	Correlation coefficient	Mean square error
logsig-purelin	0.99301	1.2425	0.9886	2.3495
logsig-tansig	0.97934	2.6917	0.9885	1.3171
logsig-logsig	0.65177	39.2108	0.89868	22.3158
tansig-purelin	0.99547	1.0603	0.98895	1.1892
tansig-tansig	0.98693	1.294	0.97345	1.8514
tansig-logsig	0.7489	35.2513	0.67636	22.2162

According to Table 8, tansig-logsig and logsig-logsig as activation function have the low correlation coefficient value and a large mean square error of training samples. The activation function tansig-purelin training sample correlation coefficient and mean square error were 0.99547, and 1.0603 respectively, while the test sample correlation coefficient and mean square error were 0.98895 and 1.1892 respectively. The correlation coefficient and mean square error of training and test samples were all optimal for the activation function tansig-purelin. Also, the activation function of the hidden layer and the output layer of the model are tansig and purelin, respectively.

In summary, the above research shows that the number of nodes in the input layer of 5 and 10 in the hidden layer, the activation function of the single output layer is tansig and purelin respectively, the training function is trainlm, the training number is 300 times, the learning rate is 0.0001, and the minimum error of the training target is 0.00001.

3.2. Prediction effect of the BP neural network model

The BP neural network model, determined by the trial, was simulated with 5 randomly extracted trials. The prediction relative error is shown in Table 9, and the correlation coefficient of the BP neural network sample is shown in Fig. 3.

As can be seen from Fig. 3, the training and the test correlation coefficient are 0.99547 and 0.98895 respectively, indicating that the trained neural network has a certain predictive effect. Table 9 shows that the relative error between the actual value of grinding power consumption and the predicted value

Table 9. Prediction error analysis of BP neural networks

Sequence number	Sample actual value (kW·h/t)	Sample predictive value (kW·h/t)	Relative error (%)
1	49.87	48.84	2.07
2	45.61	46.86	2.74
3	45.41	46.77	2.99
4	54.34	53.22	2.06
5	50.44	50.92	0.95

is within 3%, and the mean square error of the predicted sample is 1.19. The prediction effect of this model has certain credibility, indicating that the prediction model of tower grinding power consumption established by the BP neural network is effective. However, there is still a relative error approaching 3%, indicating that the BP neural network model still needs to be further optimized.

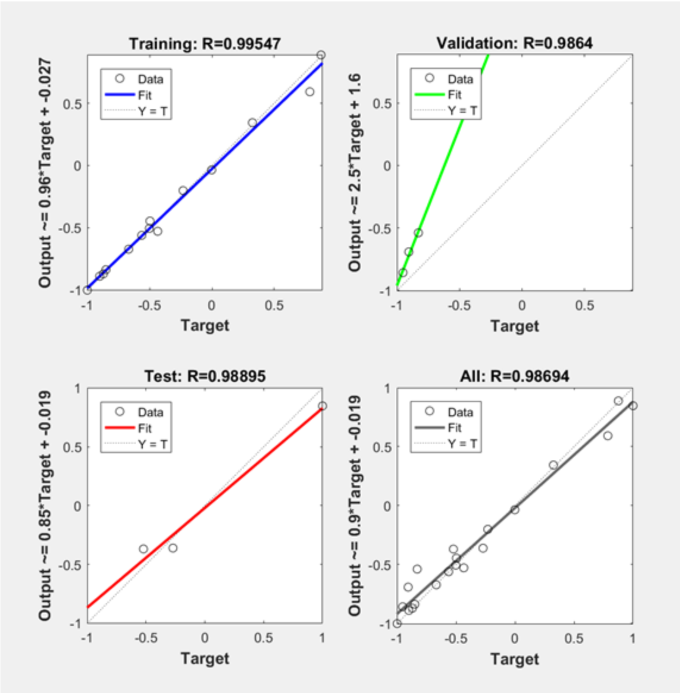

Fig. 3. Sample phase diagram of the BP neural network

4. The GA-BP neural network prediction model

4.1. Determination of the network parameters of the genetic algorithm

4.1.1. Determination of the population initialization

Population initialization includes the design of initial parameters such as population size, number of iterations, and encoding mode. Population size refers to the number of individuals, and the more individuals the easier it is to find the global optimal solution, but the computation time is longer. If the number of individuals in the population is too small, it is easy to fall into the local optimal solution. Taken together, the population size set in this study is 15.

The maximum number of iterations refers to the upper limit of the number of evolution of the population, and when the population iteration reaches the number of iterations, the population stops evolving. If the maximum number of iterations is too small, it will lead to the termination of the genetic algorithm without finding the optimal solution. The large number of iterations will cause serious computing time consuming and reduce the computational efficiency. Thus, the maximum number of iterations set in this study was 40. The real number encoding in this study was chosen to encode individuals into real number strings containing weights and threshold information.

4.1.2. Selection of genetic operators

The selection of genetic operators mainly includes the picking of selection operators, cross operators, and variation operators.

Selection operations are based on the differential fitness values of individuals in the population. The selected probability is proportional to their fitness value. Thus, individuals with higher fitness values have a greater probability of being selected. The selection operator bears the responsibility of introducing the fitness function in the genetic process and also determines whether the individual is selected to breed or to be eliminated by the population. In this study, the most commonly used roulette selection method was chosen as the selection operator.

Crossing operation is an important way of offspring generation in a population, which is of great significance for the continuous evolution of the population, and designing the type of crossing operators becomes an important step in the study of genetic algorithms. Since the real code method is adopted above, the single point real crossover method is chosen as the crossover operator in this study. Usually, the probability of crossing is greater than 0 and less than 1, and the crossing probability is 0.4.

The variation process in the genetic algorithm has some probability to negatively impact the direction of the individual population, but the population must have the variation process because the variation of appropriate probability will make the population diverse and prevent the population development from stagnating. Single-point variants were selected as the variation operator in this study. The variation probability value is set low, and the variation probability is 0.03.

In summary, the above studies found that the GA-BP neural network has 5, 10, and 1 input, hidden, and output layer nodes respectively, the training function is trainlm, and the activation function is tansig and purelin respectively. The training function is trainlm, the training number is 300 times, the learning rate is 0.0001, the minimum error of the training target is 0.00001, the population size is 15, 40 iterations, the crossover probability is 0.4, and the variation probability is 0.03.

4.2. Prediction effect of the GA-BP neural network model

The GA-BP neural network model, determined by trial, is trained and simulated with 5 randomly extracted trial data. The prediction error is shown in Table 10, and the sample correlation coefficient of the GA-BP neural network is shown in Fig. 4.

Fig. 4 shows that the training and test correlation coefficient of the sample are 0.99353 and 0.99253

Table 10. Prediction error analysis of the GA-BP neural network model

Sequence number	Sample actual value (kWh/t)	Sample predictive value (kWh/t)	Relative error (%)
1	49.87	50.20	0.67
2	45.61	45.80	0.40
3	45.41	45.18	0.51
4	54.34	55.26	1.70
5	50.44	51.06	1.24

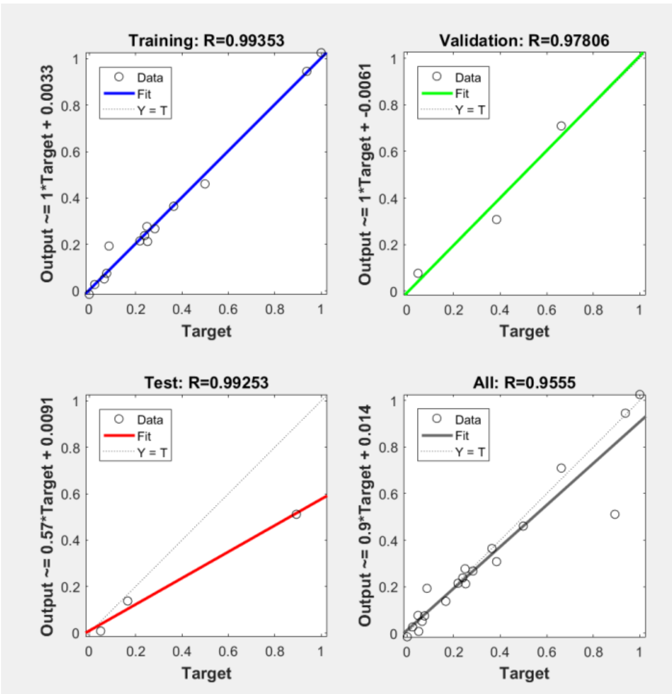

Fig. 4. Sample-phase diagram of the GA-BP neural network

respectively, indicating that the trained neural network has a certain predictive effect. It can be seen from Table 10, that relative error between the actual and predicted values of grinding power consumption is within 2%, and the mean square error of the predicted sample is 0.29.

Comparing Table 9 with Table 10, the predicted value of GA-BP is closer to the actual sample value than that of BP, indicating that the GA-BP neural network is more suitable than the BP neural network for the study of yield prediction of intermediate easy selection level of grinding products. It shows that the tower mill power prediction model of BP neural network is effective and the GA-BP neural network model can be used in the prediction of the tower mill performance.

5. Genetic algorithm neural network function extreme value optimization

5.1. Setting of the neural network parameters of the genetic algorithm

The constructed GA-BP neural network prediction model was used as the fitness function of this study, and its prediction value was set as the fitness value, and used to evaluate the quality of individuals in the population. All other parameters were the same as those of the GA-BP neural network prediction model. The optimization process is random value-taking. If the value range is not set, the optimal grinding parameters obtained by the genetic algorithm are likely too different from the actual situation. In order to prevent this issue, it is necessary to set the value range of the grinding factors represented by each chromosome in advance, and the value range is determined by the upper and lower limits of each operating parameter in the grinding data table.

5.2. Genetic algorithm optimization results and test verification

Based on the setting of GA-BP neural network parameters, the grinding power consumption of the tower mill is optimized. Under the grinding concentration of 66.49%, 301.86%, the filling rate of the medium of 96.61%, the medium ratio of 96.61%, and the material-ball ratio of 0.1394, the grinding power consumption is 41.069 kWh/t. The grinding time was maintained for 260 s, the laboratory tests were performed under these conditions, and the grinding products were subjected to particle size screening, as shown in Table 11.

According to Table 11, the content of -0.074 mm is 65.5%, and the actual grinding power consumption is 41.85 kWh/t. The relative error is 1.87% compared with the predicted grinding power consumption of 41.069 kWh/t. Furthermore, Fig. 5 shows the particle size distribution of the feed to and the product from the tower mill.

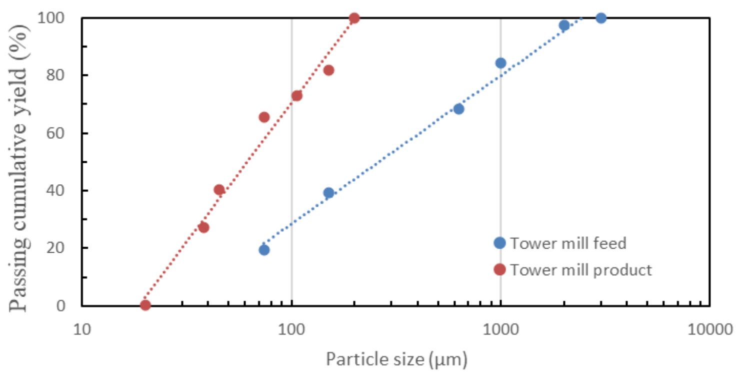

Fig. 5. The particle size distribution of the samples before after grinding using a tower mill

6. Conclusions

In this study, a BP neural network prediction model was established using the grinding concentration, screw speed, medium filling rate, medium ratio, and material-ball ratio as the input; and grinding power consumption as the output. With the BP neural network model prediction, the relative error is within 3%, and the relative error can be reduced to within 2% by further optimizing the prediction model by the genetic algorithm GA. Based on the setting of the prediction model parameters of the GA-BP neural network, the upper and lower limits of each operating parameter in the grinding data table

Table 11. Results of the product granularity analysis

Size fraction (mm)	Yield (%)	Passing cumulative yield (%)
+0.15	18.51	100
-0.150+0.106	8.41	81.94
-0.106+0.074	7.58	73.08
-0.074+0.045	25.16	65.5
-0.045+0.038	13.1	40.34
-0.038+0.020	26.94	27.24
-0.02	0.3	0.3
Total	100	

are set to optimize the grinding power consumption of the tower mill and verified by the grinding test. Thus, the actual grinding power consumption of 41.85 kWh/t with a 1.87% relative error is close to the predicted grinding power consumption of 41.069 kWh/t.

References

AUSTIN, L. G., C. L. SCHNEIDER. 2022. *A Kinetic Model for Size Reduction in a Pilot Scale Tower Mill: Model Verification*. Minerals 12(6), 679.

DANIELLE, C. R., S. ERIK, PATRICK, M. HUGH. 2017. *Prediction of Product Size Distribution of a Vertical Stirred Mill Based on Breakage Kinetics.* International Scholarly and Scientific Research & Innovation 11(11), 1740-1744.

DE BAKKER, J. 2013. *Energy Use of Fine Grinding in Mineral Processing.* Metallurgical and Materials Transactions E 1(1), 8-19.

DE CARVALHO, R. M., L. M. TAVARES. 2013. *Predicting the effect of operating and design variables on breakage rates using the mechanistic ball mill model*. MInerals Engineering 43-44, 91-101.

FUERSTENAU, D. W., A.-Z. M. ABOUZEID. 2002. *The energy efficiency of ball milling in comminution.* International Journal of Mineral Processing 67, 161-185.

GUPTA, V. K., S. SHARMA. 2014. *Analysis of ball mill grinding operation using mill power specific kinetic parameters.* Advanced Powder Technology 25(2), 625-634.

KUMAR, A., R. SAHU, S. K. TRIPATHY. 2023. *Energy-Efficient Advanced Ultrafine Grinding of Particles Using Stirred Mills—A Review*. Energies 16(14), 5277.

LI, L., B. WEI, Q. ZHANG, J. ZHANG, X. ZHANG, C. WANG, N. LI, Z. LIU. 2023. *Evaluating the performance of an industrial-scale high pressure grinding rolls (HPGR)-tower mill comminution circuit*. Minerals Engineering 191.

LIU, S., Z. LIN, Y. JIANG, T. ZHANG, L. YANG, W. TAN, F. LU. 2022. *Modelling and discussion on emission reduction transformation path of China's electric power industry under "double carbon" goal.* Heliyon 8 (9), e10497.

SHI, F., AND W. XIE. 2016. *A specific energy-based ball mill model: From batch grinding to continuous operation*. Minerals Engineering 86, 66-74.

STIEF, D. E., W. A. LAWRUK, L. J. WILSON. 1987. *Tower mill and its application to fine grinding*. Minerals and Metallurgical Processing, 45-50.

TROMANS, D. 2008. *Mineral comminution: Energy efficiency considerations*. MInerals Engineering 21 (8):613-620.

VALERY, W., A. JANKOVIC. 2002. *The Future of Comminution* 34th IOC on Mining and Metallurgy, 30 Sept.-3 Oct.2002, Hotel "Jezero", Bor Lake, Yugoslavia.

WARREN LIAO, T., L. J. CHEN. 1994. *A neural netwark approach for grinding processes: Modeling and optimization.* International Journal of Machining Tools Manufacturing 34, 919-937.

YUAN, Y., Y. ZHANG. 2022. *Improvement of the grindability of vanadium-bearing shale and the direct vanadium leaching efficiency of grinded product via microwave pretreatment with particle size classification.* Colloids and Surfaces A: Physicochemical and Engineering Aspects 647, 128979.

ZHENGMING, X., W. XIN, W. XING, W. LONG. 2016. *Simulation Analyses and Experimental Investigation on Optimum Matching of Operating Parameters of Tower Mill.* China Mechanical Engineering 27(4), 483-487.