remote sensing

MDPI

Article

Combining the Back Propagation Neural Network and Particle Swarm Optimization Algorithm for Lithological Mapping in North China

Yanqi Dong [1], Zhibin Ma [1], Fu Xu [1,2], Xiaohui Su [1,2,*] and Feixiang Chen [1,2]

[1] School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China; yanqidong@bjfu.edu.cn (Y.D.); mmazb@bjfu.edu.cn (Z.M.); xufu@bjfu.edu.cn (F.X.); bjfxchen@bjfu.edu.cn (F.C.)

[2] Engineering Research Center for Forestry-Oriented Intelligent Information Processing, National Forestry and Grassland Administration, Beijing 100083, China

* Correspondence: suxhui@bjfu.edu.cn

Abstract: Lithological mapping is a crucial tool for exploring minerals, reconstructing geological formations, and interpreting geological evolution. The study aimed to investigate the application of the back propagation neural network (BPNN) and particle swarm optimization (PSO) algorithm in lithological mapping. The study area is the Beiliutumiao map-sheet (No. K49E011021) in Inner Mongolia, China. This area was divided into two parts, with the left side used for training and the right side used for validation. Fifteen geological relevant factors, including geochemistry (1:200,000-scale) and geophysics (1:50,000-scale), were used as predictor variables. Taking one lithology as an example, the lithological binary mapping method was introduced in detail, and then the complete lithology was mapped. The model was compared with commonly used spatial data mining methods using the E-measure, S-measure, and Weighted F-measure values. In diorite testing, the accuracy and kappa of the optimized model were 92.11% and 0.81, respectively. The validation results showed that our method outperformed the traditional BPNN and weights-of-evidence approaches. In the extension of the complete lithological mapping, the accuracy, recall, and F1-score were 82.66%, 74.54%, and 0.76, respectively. Thus, the proposed method is useful for predicting the distribution of one lithology and completing the whole lithological mapping at a fine scale. In addition, the trained network can be extended to an adjacent area with similar lithological features.

Keywords: geological mapping; geochemical data; geophysical data; Inner Mongolia; BP neural network

Citation: Dong, Y.; Ma, Z.; Xu, F.; Su, X.; Chen, F. Combining the Back Propagation Neural Network and Particle Swarm Optimization Algorithm for Lithological Mapping in North China. *Remote Sens.* **2023**, *15*, 4134. https://doi.org/10.3390/rs15174134

Academic Editor: Andrzej Stateczny

Received: 8 August 2023
Revised: 19 August 2023
Accepted: 21 August 2023
Published: 23 August 2023

1. Introduction

Lithological mapping refers to the mapping of geological distribution and stratigraphic boundaries on the geological base map according to the selected scale on the basis of a field survey [1,2]. The lithological maps produced by this work lay the data foundation for the exploration and research of regional lithology, minerals, and other resources [3]. The traditional work of lithological mapping is to collect scattered rocks or observe exposed bedrocks to determine the lithology [4]. It can be seen that the traditional work is time-consuming, labor-intensive and inefficient, and presupposes expert knowledge [5]. Thus, some researchers use machine learning for lithological mapping, which is a geographic information system (GIS)-based method involving correlation of lithology with geochemical data [3,6], geophysical data [7,8], or remote sensing data [9,10] to map geological formations. Machine learning applied to lithological mapping is divided into two main types: data-driven and knowledge-driven [11]. Knowledge-driven methods, such as k-means [12] and principal component analysis [13], do not require previously known lithological samples. Data-driven methods include neural network [14,15], random forest

(RF) [16], weights of evidence (WofE) [17], and so on. Previous studies have shown that the knowledge-driven methods require in-depth knowledge of the lithologic phenomena in the study area [18]. In contrast, data-driven methods make it easier to deal with the complex geological environment and immeasurable similarities between lithological samples. Cracknell et al. showed that the neural network method had the shortest constructing time and the best performance for a lithological prediction model [19]. Using a back propagation neural network (BPNN) and support vector machine (SVM), Zhang et al. performed gold mapping and demonstrated that the BPNN had superior predictive capability over the SVM [15]. From the above research, it is found that the BPNN describes the relationship between layers by an activation function and trains the network based on back propagation error [20,21]. The BPNN includes the following steps. Firstly, the network topology needs to be determined. Secondly, the output values of each node in the hidden and output layers and the error between them and the expected value are calculated. Thirdly, the error is propagated forward layer by layer to obtain the error value of each layer. It is indicated that the BPNN possesses the merits of high fault tolerance, self-learning, and adaptability in geology-related research. However, the BPNN's shortcomings include difficult parameter setting. The parameter setting of the BPNN can further significantly impact on the learning process and model performance [22,23]. To minimize the impact of parameter setting and prevent the overfitting problem, metaheuristic optimization algorithms (such as particle swarm optimization (PSO) [24], imperialist competitive [25], whale optimization [26], etc.) are employed to optimize the network structure determination. The PSO algorithm completes the search through cooperation and competition among individuals and has strong global optimization ability [24,27]. Some studies demonstrated that machine learning optimized by PSO outperformed those using machine learning alone [24,27]. However, there is little research on whether PSO can enhance the lithological prediction ability of the BPNN. Therefore, we improved the BPNN with PSO and applied it to lithological mapping.

In mainland China, about one-third of the territory can be divided into overburden areas [7], most of which are covered by vegetation and Cenozoic sediments [3]. Inner Mongolia in northern China is a typical coverage area [28]. Therefore, it is challenging to obtain information about the underlying bedrock in these areas using remote sensing data. Several studies have shown that geochemical and geophysical data are related to the lithology of the underlying bedrock. [29]. Geochemical sampling has focused on stream sediments and soils [30]. Stream sediments are products of rock weathering, formed by erosion and transport by water flow. Their composition integrates the chemical composition characteristics of the underlying bedrock [31]. The geophysics data for this study are geomagnetic data. Each rock has a unique magnetic property, influenced by its type and age. Using these data, we can infer regional geological information and solve the problem of the poor quality of lithological mapping due to few geological outcrops and lack of geological information. To date, more than 6 million square kilometers (km^2) of regional geochemical surveys (1:200,000 scale) have been completed in China. In contrast, it is difficult to obtain large-scale, high-precision geophysical data due to economic and confidentiality reasons [3]. Therefore, we selected geochemical data (1:200,000 scale) and geophysical data (1:50,000 scale) for lithological mapping of the coverage area to ensure the feasibility and applicability of the method.

In summary, we improved the BPNN based on the PSO algorithm and applied it to lithological mapping with the aim of predicting the distribution of the underlying bedrock at a fine scale. Using the geochemical and geophysical data of Inner Mongolia, taking one lithology as an example, the above optimization method was applied to map the lithology distribution, and we finally completed the lithological mapping of the study area. In addition, we compared the improved method with the traditional BPNN and WofE methods.

2. Materials and Methods

2.1. Geological Setting

The study area, the Beiliutumiao map-sheet, extends between the coordinates 113°00′–113°15′N and 42°10′–42°20′E in Inner Mongolia, China (Figure 1a,b) [32]. Its tectonics are the Bainaimiao arc and the Ondor Sum subduction–accretion complex [33]. The elevation ranges from 900 m to 1400 m and belongs to a layered high plain of denudation formed by the uplift of an ancient lake basin. The topography of the study area is dominated by grasslands and mountains.

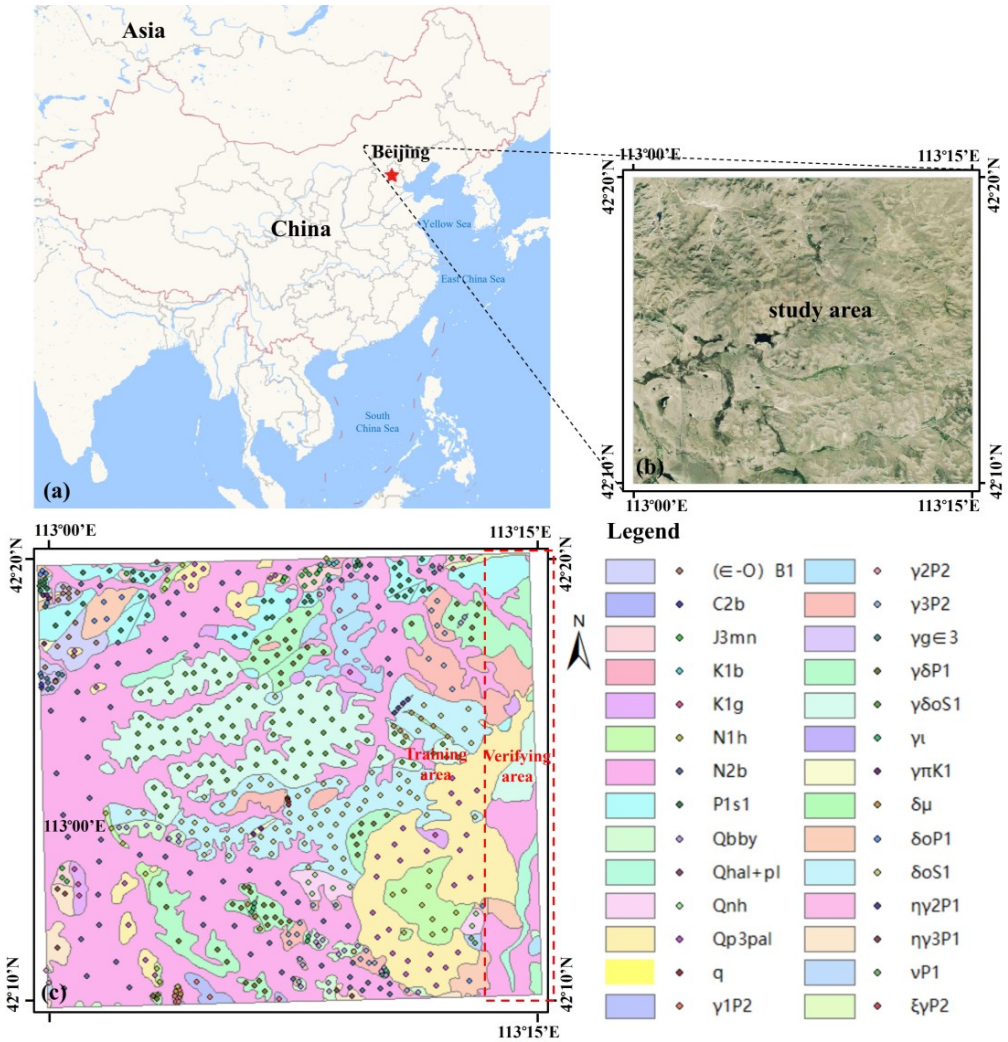
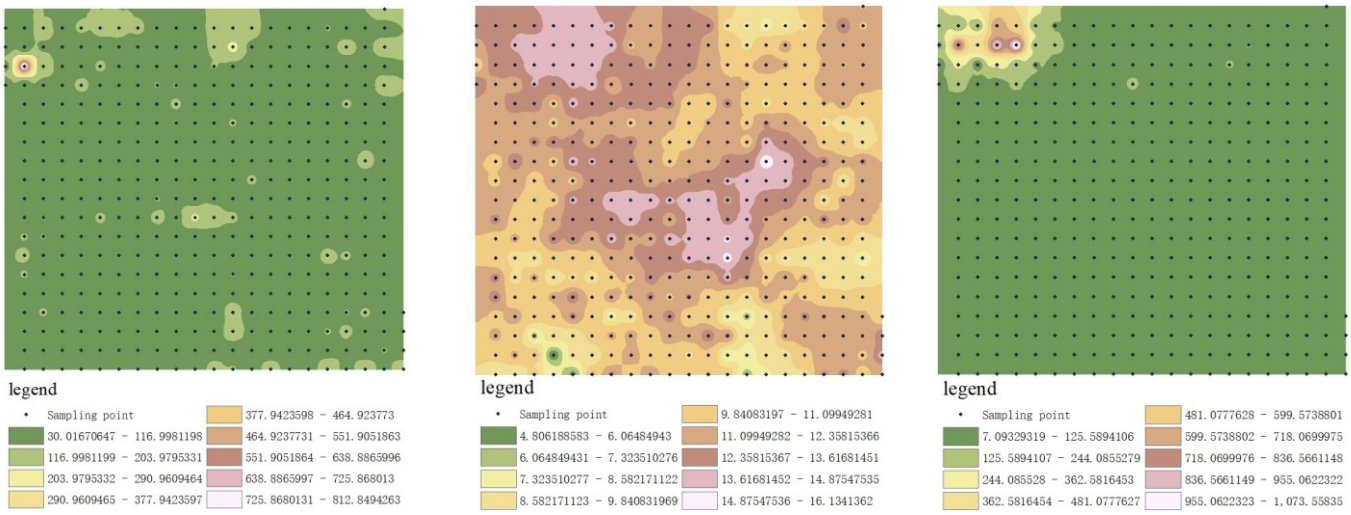


Figure 1. The location (**a**,**b**) and regional lithological map (**c**) of Beiliutumiao map-sheet at 1:50,000 scale. Abbreviations: (Є-O)B1 = Bainaimiao Group; C2b = Benbatu Formation; J3mn = Manitu Formation; K1b = Baiyingaolao Formation; K1g = Guyang Formation; N1h = Hannuoba Formation; N2b = Baogedawula Formation; P1s1 = Sanmianjin Formation; Qbby = Baiyinbaogela Formation; Qhal+pl = Quaternary Holocene alluvium-diluvium gravel bed; Qnh = Hujiertu Formation; Qp3pal = Upper Pleistocene gravel, sandy soil, gravelly sand; q = Permian quartz vein; γ1P2 = Middle Permian fine-granodiorite; γ2P2 = Middle Permian medium-fine-granodiorite; γ3P2 = Middle Permian medium-granodiorite; γgЄ3 = Early silurian fine-pomegranate dolomite granite; γδP1 = Early permian granodiorite; γδoS1 = Early silurian tonalite; γι= Permian granite aplite vein; γπK1 = early Cretaceous granitic porphyry; δμ = Permian diorite porphyrite vein; δoP1 = Early permian quartz diorite; δoS1 = Early silurian quartz diorite; ηγ2P1 = Early permian medium-fine-adamellite; ηγ3P1 = Early permian porphyric biotite adamellite; vP1 = Early permian medium-fine-gabbro; ξγP2 = Middle permian moyite.

The lithological map of the Beiliutumiao map-sheet was divided into two parts (Figure 1c). The left side was used for training and the right side was used for validation. Lithology points were selected in the training area based on the existing lithological map and the actual field lithology points. According to the spatial distribution of the points, 80% of the lithology points were randomly selected as the training dataset, and the remaining 20% as the test dataset. In this study, we selected field-investigated diorite as an example of the underlying bedrock for lithological mapping, which specifically included early silurian quartz diorite (δoS1), early silurian tonalite (γδoS1), and early permian granodiorite (γδP1). The distribution of the training area, validation area, and lithology points in this study is shown in Figure 1c.

2.2. Geochemical and Geophysical Data

The geophysical data, i.e., the geomagnetic data, were collected at a scale of 1:50,000. The geochemical samples were collected at a scale of 1:200,000. Each geochemical layer contains approximately 389 survey points. Geochemical surveys are typically sampled every 5 km^2. These survey points cover the entire study area. Each geochemical sample was analyzed for the concentrations of 14 trace elements (Ag, Al_2O_3, B, CaO, Cu, F, Fe_2O_3, Hg, K_2O, MgO, Na_2O, Ni, SiO_2, Zn). These elements were not heavily reactivated by hydrothermal activity and mineralization and are mainly limited by the lithological distribution. Moreover, these elements were components of the rocks and geology of the study area. Continuous coverage of the data was essential for inferring the lithological distribution throughout the study area. The original data were interpolated using the inverse distance exponential weighting (IDW) method [34] to obtain 15 raster layers with a pixel size of 152 m × 152 m (Figure 2). The IDW method was suitable for data with a uniform distribution of sample points and covering the entire interpolation area. The data were provided by the Development Research Center of the China Geological Survey.



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

Figure 2. *Cont.*

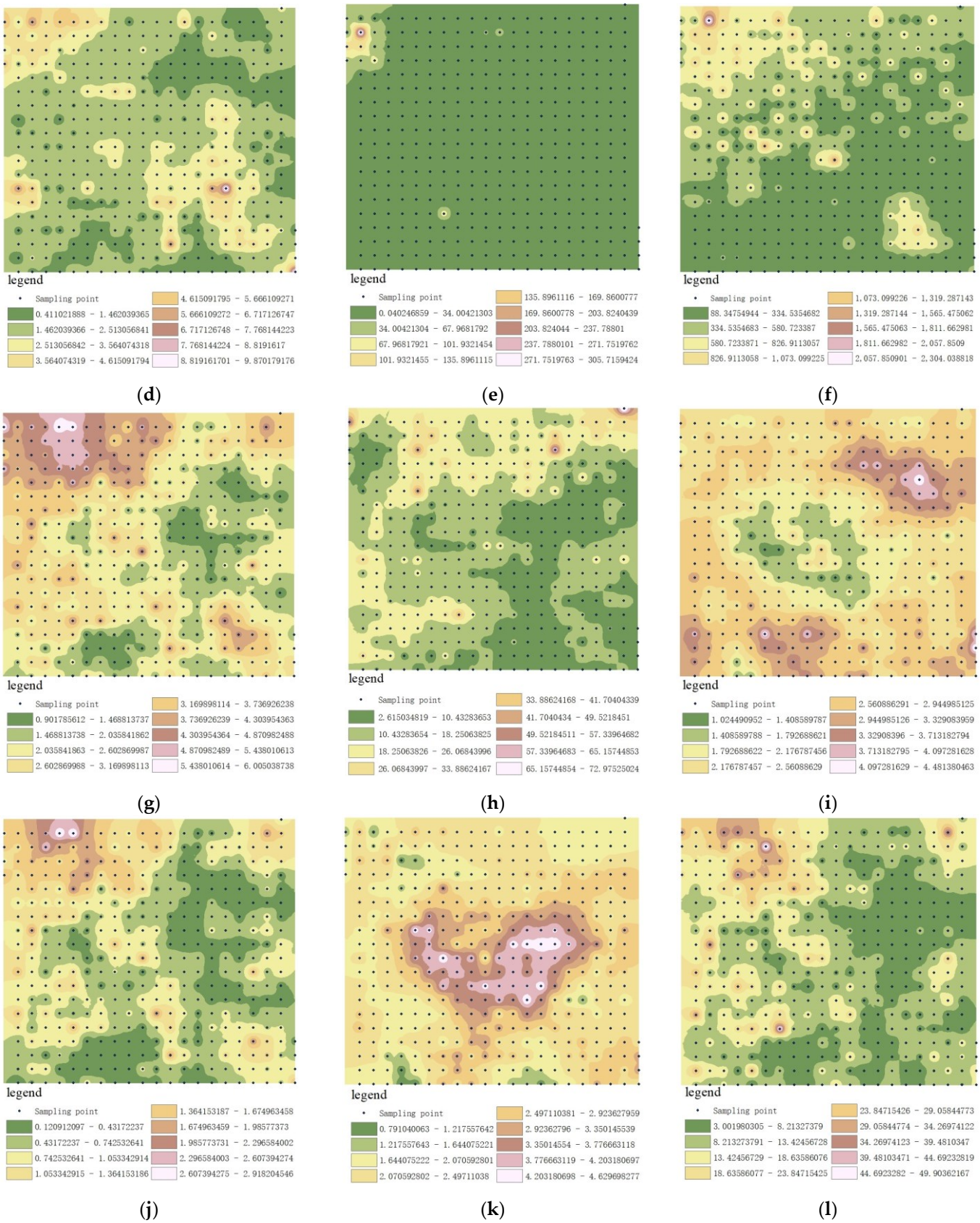
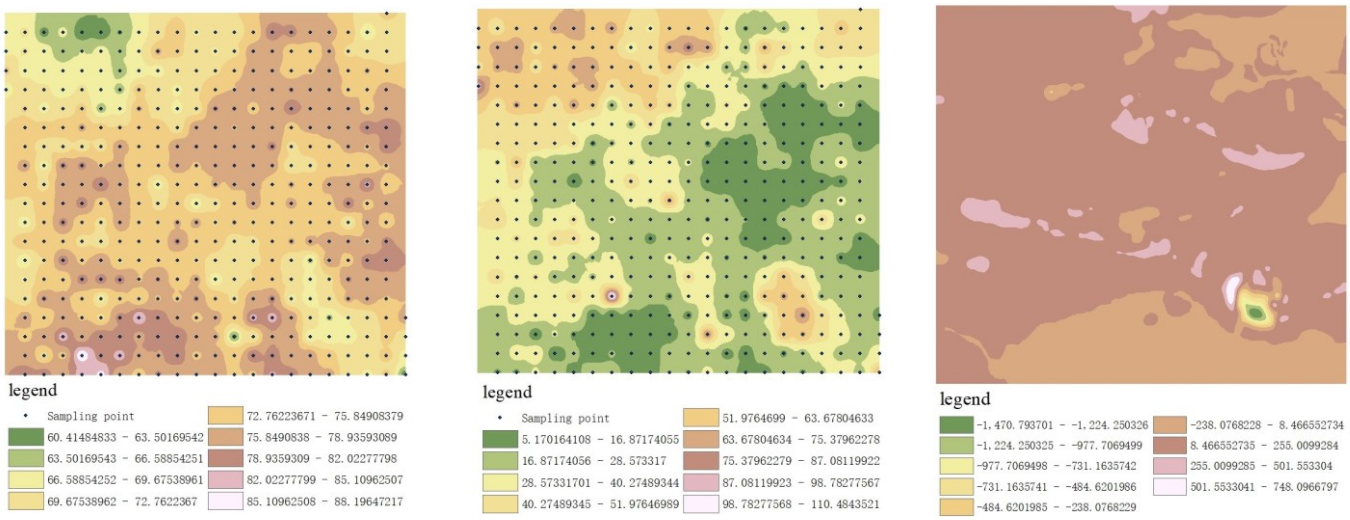


Figure 2. *Cont*.



(**m**) (**n**) (**o**)

Figure 2. The geochemical and geophysical original sampling points and raster layers generated by IDW: (**a**) Ag, (**b**) Al_2O_3, (**c**) B, (**d**) CaO, (**e**) Cu, (**f**) F, (**g**) Fe_2O_3, (**h**) Hg, (**i**) K_2O, (**j**) MgO, (**k**) Na_2O, (**l**) Ni, (**m**) SiO_2, (**n**) Zn, (**o**) geomagnetic layer.

2.3. General Methodology

The proposed method consists of four main steps: data preprocessing, improved BPNN model training, lithological mapping based on the improved BPNN model, and model evaluation (Figure 3).

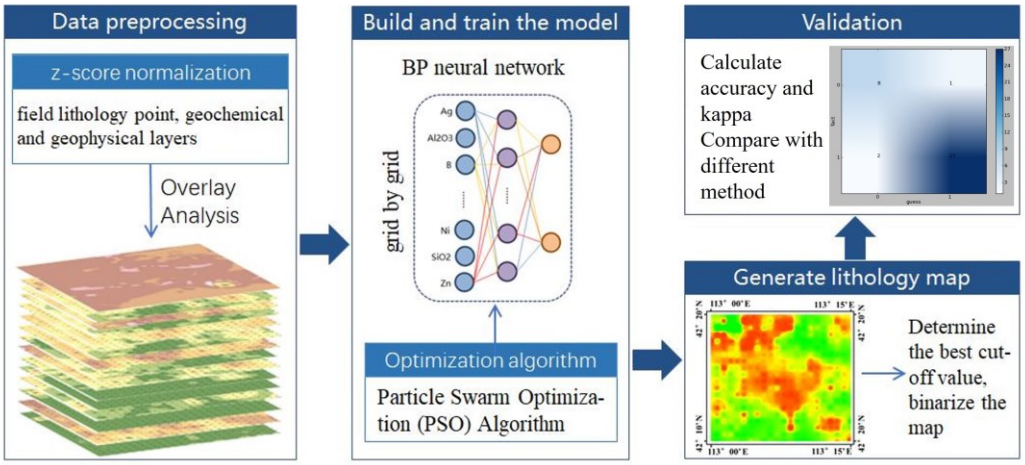


Figure 3. Illustration of the methodology in this study as workflow.

2.3.1. Data Preprocessing

Data preprocessing includes standardized geochemical and geophysical raster layers and overlay analysis of the lithology points and raster layers. The Z-score standardization method was used in this study (Formula (1)). This avoids the effect of extreme values of raster layers becoming outliers.

$$x_n = \frac{x_o - \mu}{\sigma}, \quad (1)$$

where μ is the mean value of the original data, σ is the standard deviation of the original data, x_o is the original value of the index factor, and x_n is the standardized result of the index factor.

2.3.2. Improved BPNN Model Training

Based on the BPNN with PSO algorithm, the lithological mapping model is constructed and trained. The output of the model includes the probability of lithology presence (*conf*) and the probability of its absence (*unconf*). Lithology prediction value (*res*) can be calculated:

$$res = \frac{conf + (100 - unconf)}{2}, \quad (2)$$

where the ranges of *res*, *conf*, and *unconf* are 0–100%.

2.3.3. Lithological Mapping Based on Improved BPNN Model

The lithology map of the study area was generated using the optimal model of diorite. Most lithological maps are binary, i.e., 1 (presence) or 0 (absence), so the optimal cut-off value needs to be determined by statistically analyzing the sample number of diorite presence and absence.

2.3.4. Method Comparison and Assessment

We also compared the proposed method with the traditional BPNN and WofE. Accuracy and kappa are used to assess lithological mapping. In lithological mapping, the predicted distribution of a lithology usually results in a binary map. Therefore, to validate the prediction performance, the binary maps of the validation area generated by different methods can be evaluated by the E-measure (E-m) [35], S-measure (S-m) [36], and Weighted F-measure (WF-m) [37]. These can be calculated using Saliency Evaluation Toolbox [38,39] (https://github.com/Mehrdad-Noori/Saliency-Evaluation-Toolbox (accessed on 20 August 2023)).

2.4. Improved BPNN

2.4.1. The Overall Flow of Improved BPNN

The improved BPNN includes the following process (Figure 4). The strategy to combine the BPNN and the PSO algorithm is to encode the weights and biases of each layer as particles according to the BPNN structure. Firstly, the particle swarm is initialized, and the activity range and velocity of particles are limited according to the parameter-taking range of the BPNN. Then, the particles completed by the iterative search are reduced to the corresponding weights and biases of the BPNN. Finally, a local optimization search is performed based on the BPNN.

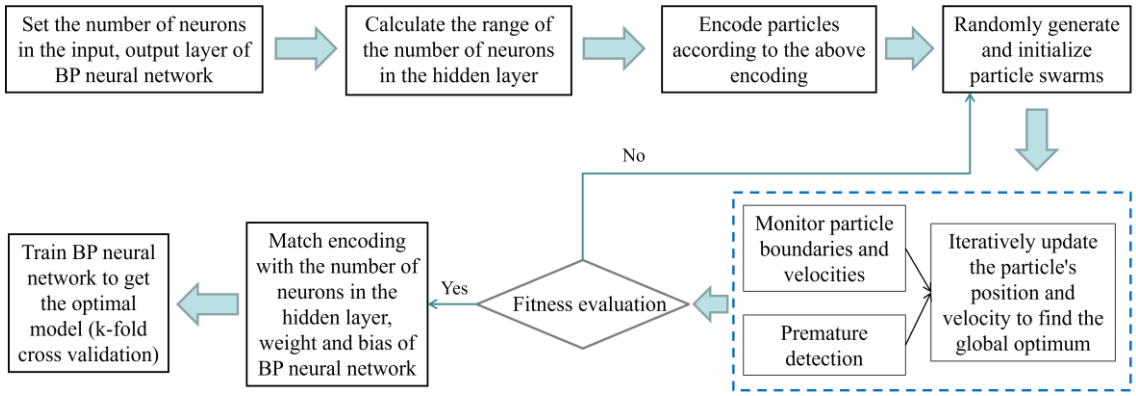


Figure 4. Algorithm flow of improved BPNN.

Determine the Structure of BPNN

The number of potential evaluation index factors is used to determine the number of nodes of the input layer i. The number of output layer nodes o is fixed at two. The range of the number of nodes h in the hidden layer is $[h_{\min}, h_{\max}]$:

$$h = \sqrt{i + o} + a, \tag{3}$$

where a is a constant between [1,10].

Particle Encoding

The hidden layer control interval, weight parameter interval, and output node bias interval of the individual particles are encoded according to the particle encoding approach in Section 2.4.2., combined with the BPNN structure.

Initialize Particle Swarm

We set the parameters such as population size, weight range, and search range of the particle population. The algorithm activated the weight parameter area by hiding the encoding of the control area. We initialized the particle position and velocity and reset the particle learning velocity. The initialized particles are multi-dimensional real vectors.

Iterative Updates

Of the sample data, 80% were randomly selected as the input training sample. The position and velocity of each particle was updated iteratively. When searching for the extremes, the search boundary, velocity, and control code of the particles need to be monitored in real time to ensure that the individual extremes and global extremes are searched.

Premature Particle Detection

It is necessary to determine whether or not the particles converge prematurely. If there is no premature convergence, the fitness evaluation is performed. If there is premature convergence, the particle velocity is mutated and the number of times required to reach the threshold is calculated, then the particle population is reorganized, and finally iterative update is performed.

Determine to Stop the Search

If the fitness evaluation value is within the threshold value or the number of iterations is greater than the maximum set number, the search is stopped. A set of weights and biases are selected to optimize the BPNN as the optimization result. The final BPNN optimal model is obtained by training.

2.4.2. Encoding Based on PSO Algorithm

The PSO algorithm uses particle position vectors and velocity vectors to find the optimal solution. Each particle finds its own optimal solution, namely, the individual extreme value, and shares it with the other particles. The optimal solution of the individual extreme value is the global extremum [40,41]. The algorithm has the characteristics of fast convergence and strong global search ability.

The geochemical and geophysical indicators that affect lithological mapping are different in the study area. The number of nodes in the hidden layer have a range of values based on the number of inputs and outputs. Therefore, we added the number of nodes in the hidden layer in the existing individual particle encoding method. The optimal number of nodes in the hidden layer and the weights and biases of each layer were searched using the PSO algorithm. Our encoding method of the individual particles is as follows:

$$x_i = [T_{hmin} T_{hmin+1} \dots T_{hmax} W_{hmin+1} W_{hmin+2} \dots W_{hmax} B_1 B_2], \tag{4}$$

where the individual particle encoding includes three parts: the hidden layer control interval T_h, the weight parameter interval W_h, and the output node bias interval B_o. The range of T_h is [0, 1]. o takes the values 1 or 2. W_h is coded as follows:

$$W_H = \begin{bmatrix} w^2_{(hmin)1} \cdots w^2_{(hmin)o} w^2_{(hmin+1)1} \cdots w^2_{(hmin+1)o} \cdots w^2_{(hmax)1} \cdots w^2_{(hmax)o}, \\ b^2_{hmin} \cdots b^2_{hmax} w^3_{1(hmin)} \cdots w^3_{1(hmax)} w^3_{2(hmin)} \cdots w^3_{2(hmax)} \end{bmatrix}, \quad (5)$$

where W_h includes the connection weights from the input layer to the hidden layer, the hidden layer bias, and the connection weights from the hidden layer to the output layer. ${w_{ij}}^l$ is the connection weight of the *jth* node at *(l − 1)th* layer and the *ith* node at *lth* layer. ${b_i}^l$ is the bias of *ith* node at *lth* layer. The encoding length of the individual particle x_i is determined by both the range of nodes in the hidden layer and the number of nodes in the input and output layers. The value of the hidden layer control interval determines the number of nodes in the hidden layer. If the value of T_j in the hidden layer control interval of a particle is greater than 0.5, the corresponding weight parameter interval is initialized with it.

3. Results

3.1. Lithological Map of Diorite

With the improved PSO algorithm, we obtained the optimal model of the lithological map. The number of network nodes in the input, hidden, and output layers is 15, 12, and 2, respectively. The learning rate is 0.3. The impulse coefficient is 0.45. The error threshold is 0.005. We obtained the optimal model by comprehensively considering the accuracy and kappa of the models. The accuracy of the optimal model is 92.11% and the kappa is 0.81. Based on this model, 15 related factors were applied to generate a lithological map of diorite (Figure 5a).

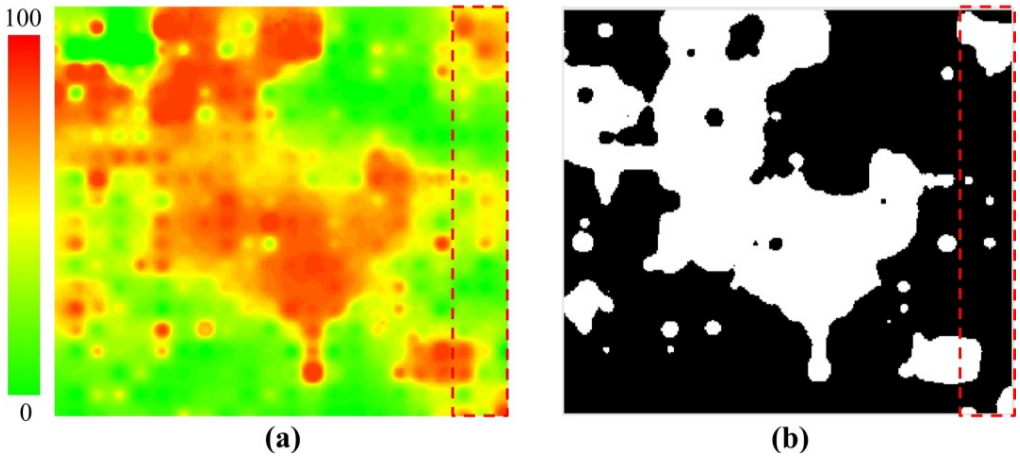


Figure 5. The lithological map (**a**) and binary map (**b**) of diorite based on improved BPNN.

The distribution of a specific lithology type is generally binary. To realize binary mapping, we set a threshold by which the continuous probability was converted to a set of binary data. By comparing the number and accuracy of prediction at different cut-off values (Table 1), we found the optimal cut-off value to be 59%. The lithological map after binarization is shown in Figure 5b. Black indicates that diorite is absent, while white indicates that diorite is present.

Table 1. Partial cut-off value results.

Cut-Off Value (%)	Number of Diorite Presences	Number of Diorite Absences	Presence of Predictive Accuracy (%)	Absence of Predictive Accuracy (%)
49	144	15	97.30	35.71
50	143	16	96.62	38.10
51	141	18	95.27	42.86
52	141	19	95.27	45.24
53	139	21	93.92	50.00
54	137	21	92.57	50.00
55	132	23	89.19	54.76
56	128	23	86.49	54.76
57	122	26	82.43	61.90
58	122	27	82.43	64.29
59	**120**	**28**	**81.08**	**66.67**
60	117	28	79.05	66.67
61	110	29	74.32	69.05
62	98	30	66.22	71.43
63	84	32	56.76	76.19
64	79	33	53.38	78.57
65	73	34	49.32	80.95
66	67	35	45.27	83.33

3.2. Model Evaluation

3.2.1. Model Validation

To verify our method, we extracted the predicted results of the validation area and compared them with the actual distribution. In the lithological map of the validation area generated by the improved BPNN, E-m, S-m, and WF-m are 0.66, 0.44, and 0.25, respectively.

3.2.2. Comparison with Traditional BPNN

The model network structure of the traditional BPNN is set to 15-12-1, the learning rate is 0.3, and the error threshold is 0.005. The prediction accuracy of the traditional BPNN is 71.05%, which is 21.06% lower than the prediction accuracy of the improved BPNN. The kappa of this model is 0.25, which is 0.56 lower than that of the improved BPNN. The optimal cut-off value is 73%. The lithological map and binary map of diorite generated by the traditional BPNN model is shown in Figure 6.

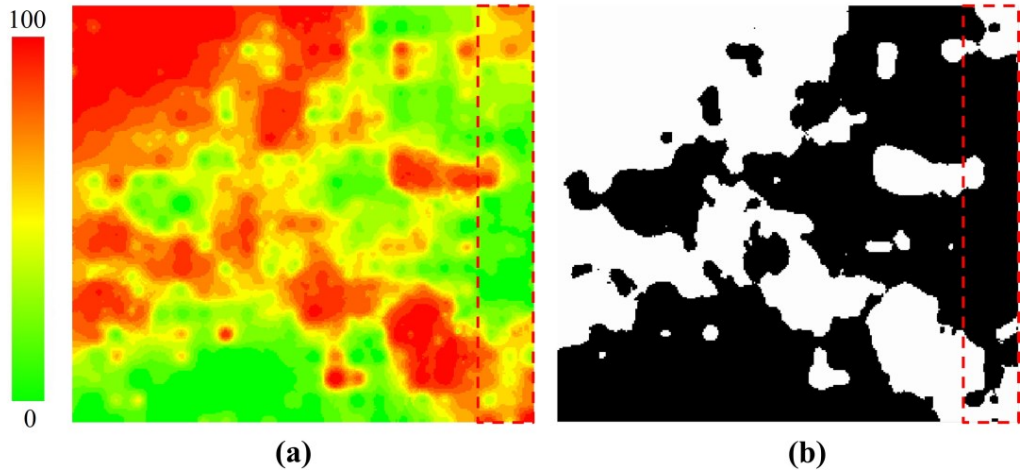


Figure 6. The lithological map (**a**) and binary map (**b**) of diorite based on traditional BPNN.

In the lithological map of the validation area generated by the traditional BPNN, E-m, S-m, and WF-m are 0.60, 0.39, and 0.18, respectively. The E-m of the traditional BPNN is

0.06 lower than that of the improved BPNN. The S-m of the traditional BPNN is 0.05 lower than that of the improved BPNN.

3.2.3. Comparison with WofE

In geological exploration, WofE is a mainstream spatial data mining method [42]. Both WofE and the improved BPNN are supervised methods in spatial data mining methods. So, we compared the geological prediction ability of both methods. WofE requires only a dataset of samples to indicate the presence of diorite. The weight of each evidence layer is calculated to build the model for predicting the target of the unknown area. The posterior probability of each pixel is then calculated based on the prediction model. The probability is positively correlated with the likelihood of occurrence of the target. Based on WofE, we used the same geochemical and geophysical data to construct the prediction model and generated a lithological map of diorite (Figure 7a). The optimal cut-off value is 61%. The binary map of diorite generated by the WofE model is shown in Figure 7b.

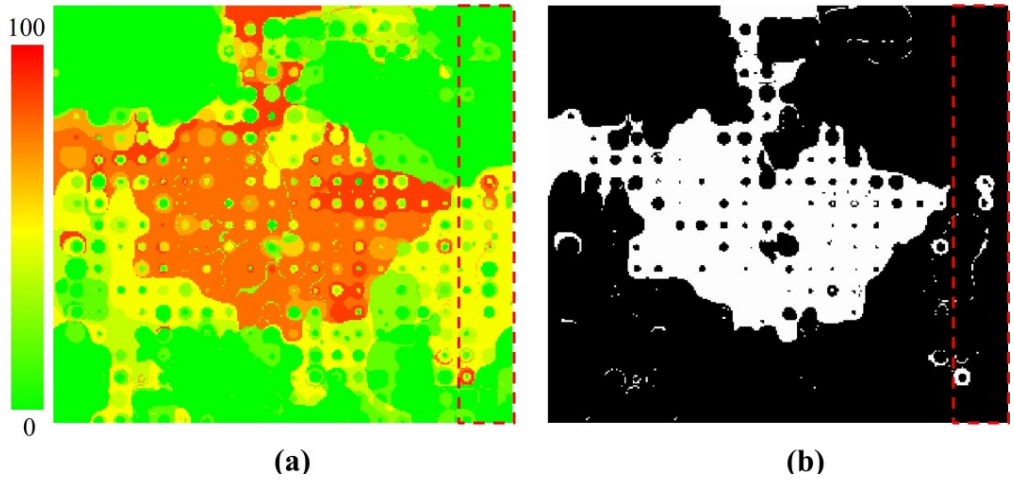


Figure 7. The lithological map (**a**) and binary map (**b**) of diorite based on WofE.

In the lithological map of the validation area generated by the WofE, the E-measure, S-measure, and Weighted F-measure are 0.40, 0.43, and 0.11, respectively. The E-m of the WofE is 0.26 lower than that of the improved BPNN. The S-m of the WofE is 0.01 lower than that of the improved BPNN. The WF-m of the WofE is 0.14 lower than that of the improved BPNN.

3.3. Extended Application of Regional Lithologic Mapping

We extended the above approach to predict lithologies across the study area. By predicting additional rocks in the study area, the lithology was mapped (Figure 8). The prediction accuracy of the lithologic map was 82.66%, the recall was 74.54%, and the F1-score was 0.76. The experimental results show that lithological mapping can be effectively carried out using geochemical and geophysical data, based on the BPNN with the PSO algorithm.

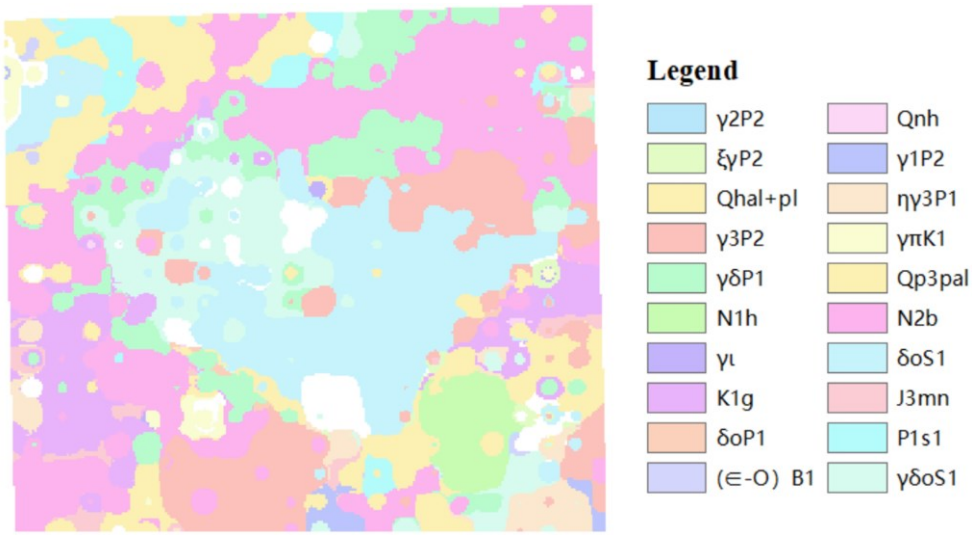


Figure 8. The complete lithological map of Beiliutumiao map-sheet based on improved BPNN.

4. Discussion

We constructed the lithological mapping model of diorite based on the improved BPNN and integrated it into the self-developed platform AoGIS. In this section, sensitivity analysis and error analysis are discussed by analyzing the whole experimental process.

4.1. Optimal Parameters from Sensitivity Analysis

This section focused on the sensitivity analysis of the combination of the number of neurons in the hidden layer, initial learning rate, impulsivity coefficient, and error threshold, finding that the different parameter combinations have a significant influence on the training results of the BPNN [43].

Figure 9a illustrates the result of the sensitivity analysis of the number of neurons in the hidden layer. When the hidden layer dimension is between 4 and 8, the model accuracy is low, and the number of iterations is relatively small. When the number of neurons in the hidden layer is above 10, the accuracy improves but the number of model iterations does not increase significantly. The best result is obtained for a network containing 12 neurons in the hidden layer, whose accuracy and kappa are 92.11% and 0.81, respectively. In summary, the topological structure (15-12-2) of our model is optimal.

Figure 9b,c illustrate the result of the sensitivity analysis of the initial learning rate and impulsivity coefficient. The optimal values are 0.3 and 0.45, respectively. These two parameters generally take values between 0.005 and 0.8. With the increase of the impulse coefficient, the number of iterations showed a downward trend, and the model accuracy showed an overall upward trend. As the initial learning rate increases, the number of iterations and model accuracy generally decline. When the initial learning rate is greater than 0.45, the number of model iterations tends to be stable, and the model accuracy has no prominent value. The impulse coefficient and the initial learning rate are important factors in the adjustment of the network weight value, which directly determine the training time and generalization ability of the model.

Figure 9d illustrates the result of the sensitivity analysis of the error threshold. Its optimal value is 0.005. In this experiment, the error thresholds in the three ranges of 0.005–0.0075, 0.01–0.0275, and 0.03–0.04 are 92.11%, 89.47%, and 86.84%, respectively. The number of iterations varies around 5000 times. With the decrease of the error threshold, that is, the continuous relaxation of the end condition of model training, the accuracy of the model will decrease. However, the model prediction performance does not change too much due to the influence of the error threshold.

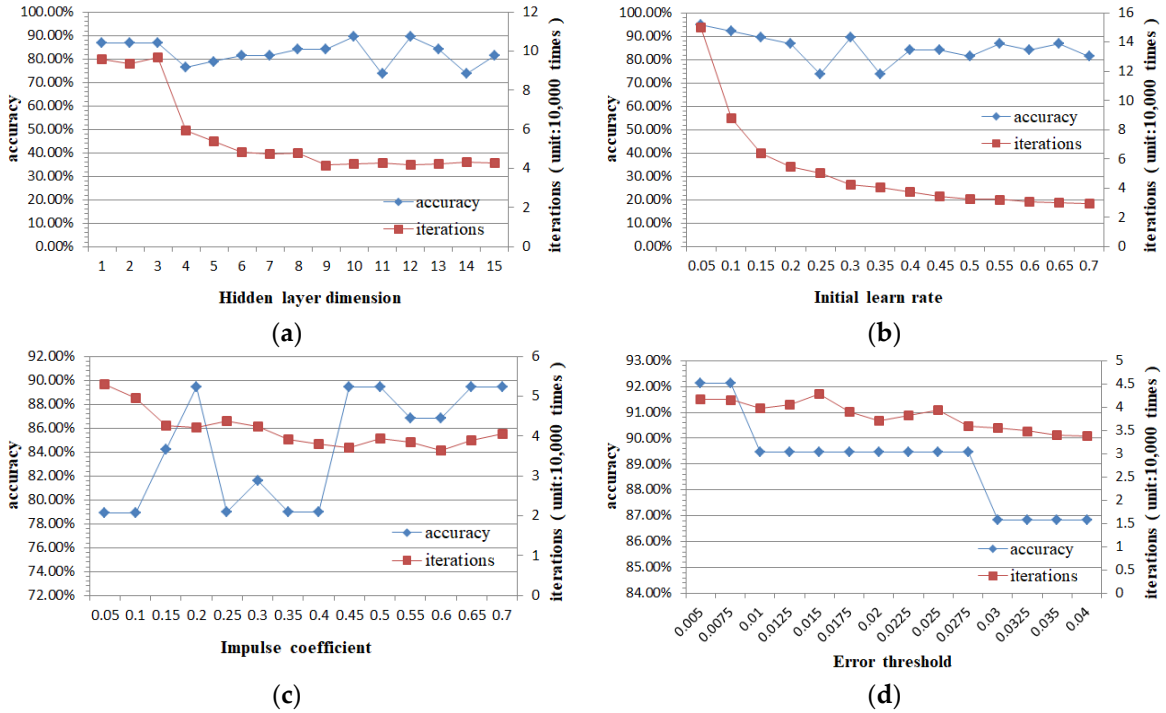


Figure 9. The sensitivity analysis of different parameters: (**a**) hidden layer dimension, (**b**) initial learn rate, (**c**) impulse coefficient, (**d**) error threshold.

4.2. Comparative Analysis

By comparing the accuracy and kappa of the different models, it can be seen that the improved BPNN is better than the other two methods. By analyzing the results, if the traditional BPNN is directly applied to the lithological mapping, the prediction results vary greatly from the actual ones. WofE has poor prediction performance for the regions where lithology points are absent. This can also be seen from the subsequent prediction results within the validation area.

By comparing the E-m, S-m, and WF-m of the validation area based on different areas, it can be seen that the improved BPNN is better than the other two methods. Since there are no training lithology points in the validation area, these three results represent the ability of the method to be extended to other neighboring areas with similar lithological characteristics. The results show that the improved BPNN model can be better extended to other areas for geological mapping at a fine scale.

In addition, we found that the elemental contents of Na_2O, Al_2O_3, and SiO_2 are low and the remaining elements are high in the predicted diorite locations obtained. Based on Pearson correlation coefficient analysis, there is a negative correlation between Na_2O and diorite [44] and a positive correlation between K_2O and diorite [45]. The combination of these elements can be used as an indicator element for predicting diorite in the Beiliutumiao map-sheet and even in other areas. Ag, Pb, and Cu are typical syngenetic combinations with high content in areas with a high diorite possibility.

4.3. Limitations of the Method

In the experiment of predicting the distribution of diorite in the Beiliutumiao map-sheet, the accuracy of the model was 92.11% and the kappa was 0.81. In the lithological map of the validation area generated by the improved BPNN, the E-m, S-m, and WF-m are 0.66, 0.44, and 0.25, respectively. When extending the trained network to another neighboring region, it is necessary to ensure that the input raster layers are of the same class and order. If the input data have a different original scale, they need to be interpolated or resampled to the same spatial resolution. The spatial resolution of the training raster layers is 152 m $\times$ 152 m in this research.

By analyzing the above experiments, it can be seen that the accuracy is not high when the upper right corner area of the study area is considered. This situation is more obvious in the lithological map generated by the traditional BPNN. Fe_2O_3, MgO, Ni, Zn, and other input layers have a high value distribution in the upper right corner area, which affects the final output map. Therefore, it can be inferred that certain extreme maximal and minimal values in the input layer affect the output layer. We will analyze and discuss the reasons why the predicted results differ from the actual distribution in the study area to improve the method in the future.

5. Conclusions

Lithological mapping needs to determine the type and distribution of lithology. We combine the BPNN and PSO algorithm and apply them to predict the spatial distribution of the underlying bedrock, using diorite as an example. Firstly, the improved PSO algorithm was used to obtain the optimal number of nodes in the hidden layer, and the weights and bias of the BPNN. Then, the BPNN model was constructed for lithological mapping. Lithological mapping was carried out using geochemical data (1:200,000-scale) and geophysical data (1:50,000-scale) of the Beiliutumiao map-sheet. This area was divided into two parts, with the left side used for training and the right side used for validation. The average accuracy and kappa of the optimal model were 92.11% and 0.81, respectively. The diorite distribution map generated by our method was consistent with the actual lithological distribution. The prediction accuracy and kappa of the improved BPNN were, respectively, 21.06% and 0.56 higher than those of the traditional BPNN. The validation results of the E-m, S-m, and WF-m showed that our method outperformed the traditional BPNN and WofE. In addition to this, we extend the method to geological mapping of all lithologies in the study area. The prediction accuracy of the lithologic map was 82.66%, the recall was 74.54%, and the F1-score was 0.76. The research showed that the proposed method could be used to predict one lithology and complete lithological mapping at a fine scale. The trained network could be extended to another adjacent area that shares similar lithological features and can generate lithological maps at a fine scale.

Author Contributions: Conceptualization, X.S. and Y.D.; methodology, Y.D.; software, F.C. and Y.D.; validation, Z.M., Y.D. and X.S.; formal analysis, Z.M.; investigation, Y.D.; resources, F.C.; data curation, Z.M.; writing—original draft preparation, Y.D.; writing—review and editing, F.X., X.S., F.C. and Y.D.; visualization, Z.M.; supervision, X.S.; project administration, F.C. and F.X.; funding acquisition, X.S., F.X. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key R&D Program of China, grant number 2022YFF1302700, The Emergency Open Competition Project of National Forestry and Grassland Administration, grant number 202303, and Outstanding Youth Team Project of Central Universities, grant number QNTD202308.

Data Availability Statement: The authors do not have permission to share data.

Acknowledgments: The authors are grateful to the Development Research Center of the China Geological Survey in China for providing the regional geological map and geochemical and geophysical data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Carosi, R.; Montomoli, C.; Iaccarino, S. 20 years of geological mapping of the metamorphic core across Central and Eastern Himalayas. *Earth Sci. Rev.* **2018**, *177*, 124–138. [CrossRef]
2. Guartán, J.A.; Emery, X. Regionalized Classification of Geochemical Data with Filtering of Measurement Noises for Predictive Lithological Mapping. *Nat. Resour. Res.* **2020**, *30*, 1033–1052. [CrossRef]
3. Ge, Y.; Zhang, Z.; Cheng, Q.; Wu, G. Geological mapping of basalt using stream sediment geochemical data: Case study of covered areas in Jining, Inner Mongolia, China. *J. Geochem. Explor.* **2022**, *232*, 10688. [CrossRef]
4. Clegg, P.; Bruciatelli, L.; Domingos, F.; Jones, R.R.; De Donatis, M.; Wilson, R.W. Digital geological mapping with tablet PC and PDA: A comparison. *Comput. Geosci.* **2006**, *32*, 1682–1698. [CrossRef]

5. Shirmard, H.; Farahbakhsh, E.; Muller, R.D.; Chandra, R. A review of machine learning in processing remote sensing data for mineral exploration. *Remote Sens. Environ.* **2022**, *268*, 112750. [CrossRef]
6. Almasi, A.; Jafarirad, A.; Afzal, P.; Rahimi, M. Orogenic Gold Prospectivity Mapping Using Geospatial Data Integration, Region of Saqez, Nw of Iran. *Bull. Miner. Res. Explor.* **2015**, *150*, 65–76. [CrossRef]
7. Yuyan, Z.; Jiang, W.; Zhao, Y.; Hao, L.; Wang, D.; Lu, J. Integrating Geochemical and Geophysical Information to Improve Geological Mapping in Northeast China: A Data Transferring Technology for Characterization and Classification. *Geochem. Int.* **2020**, *58*, 352–362. [CrossRef]
8. Kuhn, S.; Cracknell, M.J.; Reading, A.M. Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia. *Geophysics* **2018**, *83*, B183–B193. [CrossRef]
9. Ge, W.; Cheng, Q.; Tang, Y.; Jing, L.; Gao, C. Lithological Classification Using Sentinel-2A Data in the Shibanjing Ophiolite Complex in Inner Mongolia, China. *Remote Sens.* **2018**, *10*, 638. [CrossRef]
10. Esmaeili, S.; Tangestani, M.H.; Tayebi, M.H. Sub-pixel Mapping of Copper- and Iron-Bearing Metamorphic Rocks Using ASTER Data: A Case Study of Toutak and Surian Complexes, NE Shiraz, Iran. *Nat. Resour. Res.* **2020**, *29*, 2933–2948. [CrossRef]
11. Carranza, E.J.M. Data-Driven Evidential Belief Modeling of Mineral Potential Using Few Prospects and Evidence with Missing Values. *Nat. Resour. Res.* **2014**, *24*, 291–304. [CrossRef]
12. Asante-Okyere, S.; Shen, C.; Ziggah, Y.Y.; Rulegeya, M.M.; Zhu, X. A Novel Hybrid Technique of Integrating Gradient-Boosted Machine and Clustering Algorithms for Lithology Classification. *Nat. Resour. Res.* **2019**, *29*, 2257–2273. [CrossRef]
13. Piippo, S.; Sadeghi, M.; Koivisto, E.; Skytt, P.; Baker, T. Semi-automated geological mapping and target generation from geochemical and magnetic data in Halkidiki region, Greece. *Ore Geol. Rev.* **2022**, *142*, 104714. [CrossRef]
14. Wang, X.; Zuo, R.; Wang, Z. Lithological Mapping Using a Convolutional Neural Network based on Stream Sediment Geochemical Survey Data. *Nat. Resour. Res.* **2022**, *31*, 2397–2412. [CrossRef]
15. Zhang, N.; Zhou, K.; Li, D. Back-propagation neural network and support vector machines for gold mineral prospectivity mapping in the Hatu region, Xinjiang, China. *Earth Sci. Inform.* **2018**, *11*, 553–566. [CrossRef]
16. Wang, Z.; Zuo, R.; Jing, L. Fusion of Geochemical and Remote-Sensing Data for Lithological Mapping Using Random Forest Metric Learning. *Math. Geosci.* **2020**, *53*, 1125–1145. [CrossRef]
17. Jia, W.; Wang, G. Multiple level prospectivity mapping based on 3D GIS and multiple geoscience dataset analysis: A case study in Luanchuan Pb-Zn district, China. *Arab. J. Geosci.* **2019**, *12*, 332. [CrossRef]
18. Magalhães, L.A.; Souza Filho, C.R. Targeting of Gold Deposits in Amazonian Exploration Frontiers using Knowledge- and Data-Driven Spatial Modeling of Geophysical, Geochemical, and Geological Data. *Surv. Geophys.* **2011**, *33*, 211–241. [CrossRef]
19. Cracknell, M.J.; Reading, A.M. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput. Geosci.* **2014**, *63*, 22–33. [CrossRef]
20. Tan, Q.; Huang, Y.; Hu, J.; Zhou, P.; Hu, J. Application of artificial neural network model based on GIS in geological hazard zoning. *Neural Comput. Appl.* **2020**, *33*, 591–602. [CrossRef]
21. Niu, H. Smart safety early warning model of landslide geological hazard based on BP neural network. *Saf. Sci.* **2020**, *123*, 104572.
22. Xu, J.; Zhao, Y. Stability Analysis of Geotechnical Landslide Based on GA-BP Neural Network Model. *Comput. Math. Methods Med.* **2022**, *2022*, 3958985. [CrossRef] [PubMed]
23. Zhu, C.; Zhang, J.; Liu, Y.; Ma, D.; Li, M.; Xiang, B. Comparison of GA-BP and PSO-BP neural network models with initial BP model for rainfall-induced landslides risk assessment in regional scale: A case study in Sichuan, China. *Nat. Hazards* **2019**, *100*, 173–204. [CrossRef]
24. Xi, W.; Li, G.; Moayedi, H.; Nguyen, H. A particle-based optimization of artificial neural network for earthquake-induced landslide assessment in Ludian county, China. *Geomat. Nat. Hazards Risk* **2019**, *10*, 1750–1771. [CrossRef]
25. Tien Bui, D.; Himan, S.; Ataollah, S.; Kamran, C.; Nhat-Duc, H.; Binh Thai, P.; Quang-Thanh, B.; Chuyen-Trung, T.; Mahdi, P.; Baharin Bin, A.; et al. A novel integrated approach of relevance vector machine optimized by imperialist competitive algorithm for spatial modeling of shallow landslides. *Remote Sens.* **2018**, *10*, 1538. [CrossRef]
26. Yang, W.; Xia, K.; Fan, S.; Wang, L.; Li, T.; Zhang, J.; Feng, Y. A multi-strategy whale optimization algorithm and its application. *Eng. Appl. Artif. Intell.* **2022**, *108*, 104558. [CrossRef]
27. Hakim, W.L.; Rezaie, F.; Nur, A.S.; Panahi, M.; Khosravi, K.; Lee, C.W.; Lee, S. Convolutional neural network (CNN) with metaheuristic optimization algorithms for landslide susceptibility mapping in Icheon, South Korea. *J. Environ. Manag.* **2022**, *305*, 114367. [CrossRef]
28. Mu, S.; Yang, H.; Li, J.; Chen, Y.; Gang, C.; Zhou, W.; Ju, W. Spatio-temporal dynamics of vegetation coverage and its relationship with climate factors in Inner Mongolia, China. *J. Geogr. Sci.* **2013**, *23*, 231–246. [CrossRef]
29. Grunsky, E.C.; Mueller, U.A.; Corrigan, D. A study of the lake sediment geochemistry of the Melville Peninsula using multivariate methods: Applications for predictive geological mapping. *J. Geochem. Explor.* **2014**, *141*, 15–41. [CrossRef]
30. Xiao, F.; Wang, K.; Hou, W.; Erten, O. Identifying geochemical anomaly through spatially anisotropic singularity mapping: A case study from silver-gold deposit in Pangxidong district, SE China. *J. Geochem. Explor.* **2020**, *210*, 106453. [CrossRef]
31. Kirkwood, C.; Everett, P.; Ferreira, A.; Lister, B. Stream sediment geochemistry as a tool for enhancing geological understanding: An overview of new data from south west England. *J. Geochem. Explor.* **2016**, *163*, 28–40. [CrossRef]

32. Liu, Y.; Teng, F.; Wang, W.; Yang, Z.; Wang, S.; Hu, X.; Guo, S.; He, P. 1:50,000 Geological Map Database of Beiliutumiao Mapsheet, Inner Mongolia. *Geol. China* **2020**, *47*, 63–76.
33. Xiao, W.; Windley, B.F.; Hao, J.; Zhai, M. Accretion leading to collision and the Permian Solonker suture, Inner Mongolia, China: Termination of the central Asian orogenic belt. *Tectonics* **2003**, *22*, 1069–1090. [CrossRef]
34. Mahmood Ajaj, Q.; Aidi Shareef, M.; Davut Hassan, N.; Falih Hasan, S.; Noori, A.M. GIS Based Spatial Modeling to Mapping and Estimation Relative Risk of Different Diseases Using Inverse Distance Weighting (IDW) Interpolation Algorithm and Evidential Belief Function (EBF) (Case study: Minor Part of Kirkuk City, Iraq). *Int. J. Eng. Technol.* **2018**, *7*, 185–191. [CrossRef]
35. Fan, D.; Gong, C.; Cao, Y.; Ren, B.; Cheng, M.; Borji, A. Enhanced-alignment measure for binary foreground map evaluation. *arXiv* **2018**, arXiv:1805.10421.
36. Fan, D.; Cheng, M.; Liu, Y.; Li, T.; Borji, A. Structure-measure: A new way to evaluate foreground maps. *arXiv* **2017**, arXiv:1708.00786.
37. Ran, M.; Lihi, Z.; Ayellet, T. How to evaluate foreground maps. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 23–28 June 2014.
38. Mohammadi, S.; Noori, M.; Bahri, A.; Majelan, S.G.; Havaei, M. Cagnet: Content-aware guidance for salient object detection. *Pattern Recognit.* **2020**, *103*, 107303. [CrossRef]
39. Noori, M.; Mohammadi, S.; Majelan, S.G.; Bahri, A.; Havaei, M. Dfnet: Discriminative feature extraction and integration network for salient object detection. *Eng. Appl. Artif. Intell.* **2020**, *89*, 103419. [CrossRef]
40. Ofori-Ntow Jnr, E.; Ziggah, Y.Y.; Relvas, S. Hybrid ensemble intelligent model based on wavelet transform, swarm intelligence and artificial neural network for electricity demand forecasting. *Sustain. Cities Soc.* **2021**, *66*, 102679. [CrossRef]
41. Yasin, Q.; Sohail, G.M.; Ding, Y.; Ismail, A.; Du, Q. Estimation of Petrophysical Parameters from Seismic Inversion by Combining Particle Swarm Optimization and Multilayer Linear Calculator. *Nat. Resour. Res.* **2020**, *29*, 3291–3317. [CrossRef]
42. Xie, S.; Huang, N.; Deng, J.; Wu, S.; Zhan, M.; Carranza, E.J.M.; Zhang, Y.; Meng, F. Quantitative prediction of prospectivity for Pb–Zn deposits in Guangxi (China) by back-propagation neural network and fuzzy weights-of-evidence modelling. *Geochem. Explor. Environ. Anal.* **2022**, *2*, 22. [CrossRef]
43. Wang, J.; Xu, D. Artificial Neural Network-Based Microwave Satellite Soil Moisture Reconstruction over the Qinghai–Tibet Plateau, China. *Remote Sens.* **2021**, *13*, 5156. [CrossRef]
44. Tian, J.; Teng, X.; Liu, Y.; Teng, F.; Guo, S.; He, P.; Wang, W. The chronology, geochemistry of the Early Permian granodiorite in Langshan area, Inner Mongolia and its tectonic setting. *Geol. China* **2020**, *47*, 767–781.
45. Dong, H.; Meng, Q.; Liu, G.; Duan, X.; Ti, Z.; Zhu, W.; Xue, P.; Cheng, H.; Ji, T. Geochemical Characteristics of Early Silurian Granite from Biaoshan Area in Beishan, Inner Mongolia and Their Tectonic Implications. *Northwest Geol.* **2018**, *51*, 159–174.