Back-propagation neural network and support vector machines for gold mineral prospectivity mapping in the Hatu region, Xinjiang, China

Nannan Zhang^{1,2,3} · Kefa Zhou^{1,2,3} · Dong Li^{4,5}

Received: 24 November 2017 / Accepted: 16 April 2018 / Published online: 27 April 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Machine Learning technologies have the potential to deliver new nonlinear mineral prospectivity mapping (MPM) models. In this study, Back Propagation (BP) neural network Support Vector Machine (SVM) methods were applied to MPM in the Hatu region of Xinjiang, northwestern China. First, a conceptual model of mineral prospectivity for Au deposits was constructed by analysis of geological background. Evidential layers were selected and transformed into a binary data format. Then, the processes of selecting samples and parameters were described. For the BP model, the parameters of the network were 9–10–1; for the SVM model, a radial basis function was selected as the kernel function with best $C = 1$ and $\gamma = 0.25$. MPM models using these parameters were constructed, and threshold values of prediction results were determined by the concentration-area (C-A) method. Finally, prediction results from the BP neural network and SVM model were compared with that of a conventional method that is the weight-of-evidence (W-of-E). The prospectivity efficacy was evaluated by traditional statistical analysis, prediction-area (P-A) plots, and the receiver operating characteristic (ROC) technique. Given the higher intersection position (74% of the known deposits were within 26% of the total area) and the larger AUC values (0.825), the result shows that the model built by the BP neural network algorithm has a relatively better prediction capability for MPM. The BP neural network algorithm applied in MPM can elucidate the next investigative steps in the study area.

Keywords Mineral prospectivity mapping · Back propagation neural network · Support vector machines · Weight of evidence

Introduction

Mineral exploration aims to identify areas that may include new mineral deposits (Yue et al., 2017). However, finding

new mineral deposits is becoming increasingly difficult, with the most obvious already having been identified, and requires sustained research and development of new methods (Carranza 2017). Mineral prospectivity mapping (MPM) can be a key procedure in mineral exploration; the approach is concerned with quantifying and mapping the likelihood that a mineral may be found by exploration in a study area. Original MPM models can be traced to the works of mathematical geologists such as Harris (1965, 1969), Sinclair and Woodsworth (1970), Agterberg (1971, 1974), and Bonham-Carter (1994).

In MPM, different evidential layers are extracted from known deposit-types as training input data (Bonham-Carter 1994) in order to find a spatial correlation between known deposits and various layers based on numerous computational algorithms (Carranza 2008). The integration of different evidential layers and the identification of favorable areas are key tasks in the MPM process (Zuo et al. 2014; Zhang and Zhou 2015), and mathematical models play an important role in this procedure. In general, mathematical models for MPM can be categorized into knowledge- or data-driven constructs,

Communicated by: H. A. Babaie

✉ Nannan Zhang
znn_0802@163.com

¹ State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, Xinjiang 830011, China

² Xinjiang Research Center for Mineral Resources, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, Xinjiang 830011, China

³ Xinjiang Key Laboratory of Mineral Resources and Digital Geology, Urumqi, Xinjiang 830011, China

⁴ Key Laboratory of Coastal Zone Environmental Processes and Ecological Remediation, Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003, Shandong, China

⁵ University of Chinese Academy of Sciences, Beijing 100049, China

although hybrid models also exist (Cheng and Wu 2017). In knowledge-driven techniques, few known mineral deposits are predicted in areas of interest; the approach is commonly used in poorly explored terrains, and so expert experience and judgment are required (Zhang and Zhou 2015). Knowledge-driven methods include constructs incorporating index overlay (Carranza et al. 1999; Sadeghi and Khalajmasoumi 2015), the fuzzy analytical hierarchy process (Najafi et al. 2014), fuzzy logic (Abedi et al. 2013; Molan and Behnia 2013), wildcat mapping (Carranza 2010; Carranza and Hale 2002a), Dempster-Shafer belief theory or EBFs (e.g., Moon 1990;

Hashemi Tangestani and Moore 2002; Abedi et al. 2017), and the outranking method (e.g., Abedi et al. 2013, 2012a, b). Data-driven methods are suitable for study areas with a reasonably large number of known deposits that are in regions of interest are used as “training points” to recognize and establish spatial relationships among deposits with particular exploration evidential features (Najafi et al. 2014). Data-driven methods include weights-of-evidence (Agterberg and Bonham-Carter 1999; Carranza and Hale 2002b), fuzzy weights-of-evidence (Cheng et al. 2007) logistic regression (e.g., Agterberg and Bonham-Carter 1999; Carranza and Hale 2001; Mejia-

Fig. 1 Schematic geological map of Western Junggar, Xinjiang (modified after 1:200,000 geological maps.).

Herrera et al. 2015), and some machine learning methods such as Bayesian networks classifiers (Porwal et al.; 2006; Abedi and Norouzi 2012), neural networks (Harris and Pan 1999; Brown et al. 2000; Oh and Lee 2010), random forests (Breiman 2001), Boltzmann machine (Chen et al. 2014; Chen 2015), and support vector machines (Zuo and Carranza 2011; Abedi et al. 2012a).

In the field of MPM, comprehensive metallogenic information has multi-source characteristics; thus, the relationship between this information and mineralization is very complex and mostly nonlinear (Rodriguez-Galiano et al. 2014). Traditional linear mathematical methods struggle to reveal these complex relationships when applied to MPM. Indeed some studies have shown that machine learning algorithms are more advantageous than traditional techniques (Piccini et al. 2012; Hamid et al. 2011). Machine learning can identify and build complicated nonlinear relationships between mineral deposits and evidence. Therefore, the application of machine-learning-based non-linear technology to MPM is helpful as regards obtaining detailed understanding of nonlinear dynamic behavior of metallogenic systems. In addition, the prediction of mineralization system evolution can be incorporated into metallogenic prediction more easily using machine-learning-based non-linear technology; hence, more accurate prediction results can be achieved (Rodriguez-Galiano et al. 2014).

The objective of this study was to compare predictive modeling for MPM based on BP neural network and SVM machine learning algorithm. The MPM model was constructed and tested using evidence layers from the Hatu region of Xinjiang, China. The prospectivity efficacies of the BP neural network, SVM, and a weight-of-evidence (classical method) were contrasted and evaluated using traditional statistical analysis, prediction-area (P-A) plots, and the receiver operating characteristic (ROC) technique. The best results obtained from the proposed methods can be used to direct the next stages of MPM work in the study area.

Geological setting

The western Junggar area is located on the western margin of the Junggar Basin, northern Xinjiang, China. In terms of administrative division, the western Junggar area belongs to Tacheng prefecture, Ili Kazakh Autonomous Prefecture (Zhang and Zhou 2015; Liu et al. 2017). It extends to Sawuer Mountain in the north, Ebinur Lake in the south, Buerkesidai-Karamay in the east, and Omin-Tuoli-Alataw pass in the west (Fig. 1). The area hosts the best gold deposits in Xinjiang to date, with over 200 gold deposits or mineralization occurrences. Western Junggar is situated in the convergence belt between the Siberian Plate and Tarim Plate, and in the Hercynian back-arc basin. Owing to the regional tectonic stress field, the study area has relatively

developed folds and fault structures (Shen et al. 2009). The strata cropped out in the western Junggar area include mainly Ordovician-Silurian epimetamorphic rock series of the lower Paleozoic, Devonian–Carboniferous marine volcanic rocks, turbidite formations of the upper Paleozoic, and Permian-Triassic terrestrial volcano-molasses formation. The Tailegula Formation, Baogutu Formation, and Xibeikula Formation are the main gold-bearing strata in the metallogenic districts; Devonian Kulumudi Formation is also a gold-bearing stratum. Neopaleozoic post-collisional plutonic rocks composed of intermediate–acidic–intrusive rocks are extensively developed in the metallogenic belt; they are divided into two main categories. (Abedi et al. 2017) A: huge acidic batholith dominated by alkali-feldspar granite, which is found on the two sides of the Darabut fault and constitutes the Darabut alkali-rich igneous rock belt (Han et al. 2017). Examples in this category include a pluton under Miao’ergou, Akbastau, Karamay, and Hongshan. (Abedi et al. 2013) A granodiorite-quartz diorite, which is mainly in the form of small stock and distributed southeast of the Dalabut fault (Zhao et al. 2006).

The Hatu region is located on the northwestern (NW) margin of the Junggar Basin (Fig. 1) and is a part of the northern wing of the east section of the Zhayier-Dalabute synclinorium of the western Junggar Variscan fold belt (David et al. 1993; Wang and Zhu 2015; Shen et al. 2016). The formation outcrops of this region are relatively simple since they primarily consist of carboniferous rocks, and include the Lower Carboniferous Xibeikulasi Formation (C1x), Upper Carboniferous Baogutu Formation (C2-3b), and Tailegula Formation (C2-3 t). The Middle Devonian Baerleike Formation (D2b) and Kayierbayi Formation (D2k) were observed only in the northern end of this region. The host rocks are dominantly basalt with minor crystal-vitric tuff and tuffaceous siltstone. Regional structures are generally oriented towards the northeast (NE). The NE fault of the Anqi fault (tensile fault) and the Hatu fault (torsional fracture) together constitute a narrow rift basin, which comprises the basic structural framework in the study area and controls the gold deposit distribution. The evolution of the regional geological structure was conducive to gold formation, and it has the best gold mineralization in Xinjiang (Wang and Zhu 2015). Most of the gold mines are distributed along the northeast Anqi fault, including Hatu, Qi-II, Qi-III, Qi-IV, and Qi-V (Fig. 2). In addition, intrusive magmatic activity in this area was strong, and provided a favorable tectonic environment for gold mineralization. The Hatu Au mineralization is hosted in quartz veins and altered rocks within Lower Carboniferous basalt and minor tuff. As in most gold mines, wallrock alteration is highly developed in the Hatu mining area. The most commonly observed alterations are pyritization and arsenopyritization, followed by carbonitization, silification, sericitization, and chloritization (Shen et al. 2016). A conceptual model of mineral prospectivity for Au deposits in the study area was established on the basis of the anatomy of known deposits and on the analysis of available data (Table 1).

Fig. 2 Geological map of the Hatu area (1: 50,000)

Methodology

BP algorithm

To date, the most commonly used artificial neural network (ANN) is a BP neural network, also known as an error back propagation neural network. This kind of algorithm was first

proposed in 1986 by the parallel distributed processing group of Rumelhart et al. (Hamid et al. 2011). The BP network can learn and store a large number of input-output pattern mappings without revealing the mathematical equations that describe this mapping. The learning rule employed in this network uses the steepest descent method and constantly adjusts the network weights and thresholds via back propagation, so

Table 1 Conceptual model of Au deposits in the Hatu region

Metallogenic factor	Description	
Geological background	Tectonic environment	The northeast, northwest-west, and near east-west-striking faults play the most significant roles in controlling gold mineralization in this region.
	Intrusive rocks	The intermediate-to-acidic rocky outcrops in this region frequently contain gold. Hence, acidic magma activity and gold metallogenesis are closely related in time and space within this region.
	Ore-bearing strata	Regions of gold deposits and mineralization in this area are generally concentrated within the Middle-to-Late Carboniferous strata and Middle Devonian strata. Regions of gold mineralization are especially concentrated in sections where intermediate-to-basic volcanic rocks have developed in the Middle-to-Late Carboniferous Tailegula Formation.
	Wallrock alteration	Altered wallrock sections and shattered fault zones where intense pyritization, arsenopyritization, chloritization, sericitization, silification, and carbonitization overlap each other are the most noticeable indicators for this region.
	Dike rock characteristics	The widely distributed quartz veins in this region often contain gold. Small singular, complex, and networked quartz veins produced in fault alteration zones and secondary faults on either side of these zones generally exhibit excellent gold-bearing properties.
Regional geophysical characteristics	Gravity data	The gravity field in gold deposits is elliptical, and distributed in EW–NE direction. The contours near faults are obviously twisted or linear.
Regional geochemical characteristics		Marked As and Ag geochemical anomalies on the surface, natural heavy concentrates and gold anomaly zones, and beneficial mineral (native gold) diffusion halos are all important characteristics for gold prospecting.

Fig. 3 GIS Database of the Hatu area

that the network sum squared error is at a minimum when the training is completed. The algorithm is usually composed of input, hidden, and output layers, including forward and backward propagation. A neuron of the ANN calculates the weighted sum of its inputs and passes it through a transfer function to generate an output. The most popularly used transfer function is the sigmoid logistic function, which is defined by the following formula (Ma et al. 2017):

$$f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

The error function E is the sum of squares of the difference between the desired output and the actual output (see formula 2), where y_j^m is the actual output of the m layer and T_j is the expected value of the output unit.

$$E = \frac{1}{2} \sum_i (y_i^m - T_j)^2 \quad (2)$$

The input and output layers are calculated sequentially, until the output layers. If some deviation occurs between the network output and its expected output, backward

Table 2 Geological variable values of known deposits

Geological Variables Mine Name	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉
Qiqiu No. I	1	0	0	1	1	1	1	1	1
Qiqiu No. II	0	1	1	1	1	1	0	1	0
Qiqiu No. III	0	1	1	0	1	0	0	1	1
Baobei	0	1	0	0	1	0	0	1	1
Qiqiu No. IV Eastern	0	0	1	1	1	0	1	0	1
Qiqiu No. IV Central	1	1	0	1	1	1	1	0	1
Qiqiu No. IV Western	1	1	0	1	1	1	1	0	1
Sa 14	0	1	0	1	1	1	0	0	1
No. 127	1	1	0	0	1	0	1	0	1

propagation is initiated. In the direction of error propagation, the gradient descent method is adopted; that is, the connection coefficient between neurons is adjusted to reduce the total error to the decreasing direction. The mathematical equation is expressed as:

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}} \quad (3)$$

where $0 < \eta < 1$, η is the learning rate, and the adjustment formula of the weight coefficient is $W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}$; in order to accelerate the convergence speed, the last weight

Fig. 4 Binarization map of the pyrition evidence layer

Fig. 5 Mineral prospectivity mapping using the Weight-of-evidence (W-of-E) method

coefficient is often considered and used as one basis of the improvement, so the weight coefficient adjustment formula is defined as:

$$W_{ij}(t+2) = W_{ij}(t) + \Delta W_{ij} + \alpha[W_{ij}(t+1) - W_{ij}(t)] \quad (4)$$

where $0 < \alpha < 1$, and α is the correction constant of the weight coefficient.

SVM algorithm

The SVM method is based on the concept of the optimal classification hyperplane, in which two types of sample are linearly separable. It is driven from statistical theory and utilizes structural risk minimization principles, inner product kernel functions, and VC dimension theory to solve problems including small sample learning, nonlinear recognition, and high-dimensional pattern recognition (Smola and Schölkopf 2004; Zuo and Carranza 2011). In this study, SVM was used as an alternative tool for mineral prospectivity mapping.

The sample training set contains N vectors of d -dimensional feature space, $x_i \in R^d (i = 1, 2, \dots, N)$, and the target $y_i \in \{-1, 1\}$ is associated with each vector x_i . The problem of

finding the optimal hyperplane is transformed into the problem of solving the convex quadratic equation, as follow:

$$\begin{cases} \text{Minimize : } \frac{1}{2} \|\omega\|^2 \\ \text{Subject to : } y_i(\omega \cdot x_i + b) \geq 1, i = 1, 2, \dots, N \end{cases} \quad (5)$$

where ω, b represents the coefficients of hyperplane equations $f(x) = \omega \cdot x + b$. According to the convex quadratic programming solution in optimization theory, the Lagrange function is constructed as follows:

$$L(\omega, \alpha, b) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^N \alpha_i y_i \times (\omega \cdot x_i + b) + \sum_{i=1}^N \alpha_i \quad (6)$$

where $\alpha_i (i = 1, 2, \dots, N)$ is the introduced Lagrange multiplier.

By adding the error penalty factor C to the objective function and setting the threshold, the above equation is transformed into:

$$\begin{cases} \text{Maximize : } \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (X_i \cdot X_j) K(X_i \cdot X_j) \\ \text{subject to : } \sum_{i=0}^N \alpha_i y_i = 0, 0 \leq \alpha_i \leq c, i = 1, 2, \dots, N \end{cases} \quad (7)$$

The $K(x_i, x_j)$ function is a kind of kernel function, which is conformable to the Mercer theorem. The most common kernel functions include the Linear Function $K(x_i, x_j) = X_i^T X_j$; Polynomial Function $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$; Radial Basis Function, $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$, where $\gamma > 0$; and Sigmoid Function: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)^d$.

GIS database

The geological and mineral space database of the Hatu region was constructed using ArcGIS10.1 (Fig. 3). A geographic coordinate system, namely Beijing 1954,

Fig. 6 Selection of hidden nodes number and training time

Fig. 7 Back Propagation (BP) neural network architecture used in this study

was used (6-degree Gauss-Kruger zone 14, central meridian 81° , and unit m). The database comprises planar elements, linear elements, and dot elements, including: (Abedi et al. 2017) a single dot element representing the Au mineral deposits; (Abedi et al. 2013) a single linear elements representing faults; (Abedi and Norouzi 2012) 7 planar elements representing intrusive rocks, sedimentary (volcanic) strata, quartz veins, the spread halo of natural gold and the anomalies of natural heavy sand. The geological maps were compiled from Xinjiang Bureau of Geology and Mineral Resources mapping at a scale of 1:50,000. The geochemical data comprised the spread halo of natural gold and the anomalies of natural heavy sand, which were at a scale of 1:200,000 and collected from the 1:20 Natural Heavy Sand Database. In this study, according to conceptual model and available data, geological and geochemical datasets were used as sources of evidence for mineral prospectivity mapping.

Data acquisition process

The geological map of the Hatu area was divided into 800 grid units; namely, information statistical units, using the regular grid method. Nine evidence layers according to the conceptual model were selected to build the mineral prospectivity model, including: the spread

halo of natural gold(X1); quartz veins (X2); anomalies of natural heavy sand (X3); fracture intersections (X4); NE-oriented fractures (X5); NW-oriented fractures (X6); EW-oriented fractures (X7); pyrition (X8); and favorable formations (X9). These nine evidence layers were distributed across grids of the same size as the study area, and the grid units were converted to the 0 and 1 data format via binarization classification. That is, if a given evidence layer was present in a certain grid unit, a value of 1 was assigned. Otherwise, a value of 0 was given. Then, evidence layers were initialized using the well-defined regular grid units to form a $m(800) \times n$ (Atkinson and Tatnall 1997) data matrix, where m is the number of grid units and n is the number of evidence layers. As an example, Table 2 shows the values of evidence layers for the nine known exemplary deposits, which were the main deposits in the Hatu region chosen (Fig. 2). Figure 4 shows the nine evidence layers for MPM associated with Au deposits in the area.

Sample selection

Weight-of-evidence (W-of-E) is a traditional and most widely used method; it was first applied to the prediction of mineral deposits by Bonham-Carter et al. (1989) and described in detail by Agterberg and Bonham-Carter (1999). It was used to integrate the nine evidence layers to produce an MPM result (Fig. 5). In total, 23 locations associated with known Au deposits, and 23 ‘non-deposit’ locations in the non-metallogenic district (Fig. 5) were selected. The selection for ‘non-deposit’ locations followed the four criteria defined by Carranza (2008). Finally, the 46 samples were used to learn and set up the models using the BP and SVM methods. Finally, 70% of points from the 46 dots were selected randomly as the training samples; the remainder were used as test samples.

MPM in the study area

Application of the BP algorithm

In this study, MATLAB (2014a) was used to implement the BP algorithm for MPM. The network contains three

Fig. 8 Relationship between target values and output values

Fig. 9 Concentration-area (C-A) model for prospectivity map

parts: one input layer with 9 nodes, one hidden layer with 10 nodes, and one output layer with 1 node representing metallogenic prospectivity. Hyperbolic tangent sigmoid transfer (tansig) and log-sigmoid transfer (logsig) functions were used to describe the transfer functions of the hidden and output layers, respectively. The learning function was the levenberg-marquardt back-propagation function (trainlm), the threshold and weight learning function was the gradient descent weight and bias learning function (learngd), and the performance function was the mean squared normalized error performance function (MSE).

The numbers of input- and output-layer nodes are determined by the nature of the actual problem. The choice of the number of hidden nodes is a complex problem determined by many factors, such as the number of training samples, the size of the sample noise,

and the complexity of the problem. Usually, there is no rigorous theoretical guidance to determine the number of neurons in a hidden layer (Atkinson and Tatnall 1997). In this study, a trial and error method using an empirical formula was used to determine the number of nodes in the hidden layer. The number of hidden-layer nodes corresponding to the minimum mean square error (MSE) value can be determined by changing 1 and using the same sample set for the training process. This formula is expressed as (Ma et al. 2017):

$$l = \sqrt{n + m} + a, \quad (8)$$

where l , n , and m are the numbers of hidden-, input-, and output-layer nodes, respectively; and a is a constant adjustable between 1 and 10 (Xu et al. 2015). According to the empirical equation, the number of

Fig. 10 Mineral prospectivity mapping using the Back Propagation (BP) method

Fig. 11 Contrast maps of mean square error (MSE) obtained using four different kernel functions

neurons in hidden layer is between 4 and 13 in this study. As shown in Fig. 6, when the number of hidden-layer nodes was 10, the mean square error (MSE) value of training samples and the training times were smaller than for the other conditions. Hence, the number of hidden layer nodes is set to 10, the maximum number of training times was set to 200, and the error tolerance was set to 10^{-5} . The learning rate was set to just 0.01 because too high a learning rate may lead to system instability, while too small a learning rate may lead to longer training time and slower convergence; however, a lower learning rate can also guarantee that the error value of the network does not jump out of the trough of the error surface, and eventually tends to the minimum error value (Yin et al. 2017). Therefore, smaller learning rates are commonly used to ensure system stability, usually ranging from 0.01 to 0.8. The final network architecture is shown in Fig. 7.

Finding the proper parameters with BP methods can avoid the over fitting phenomenon; the number of hidden layers and the training times, which are two important influencing factors for over fitting, were not larger in this study. In addition, the output tracked the targets

Fig. 12 Grid search method for the selection C and γ values

well for training, testing, and validation, and the R-value is over 0.95 for the total response (Fig. 8). This phenomenon also shows that there is no over fitting in this BP model (Neural Network Toolbox™ 7 User's Guide). After the BP-based prediction model with the above parameters was established, the concentration-area (C-A) method was used to determine the metallogenic prospectivity threshold (Fig. 9). The C-A method was proposed by Cheng et al. (1994), and is presented in detail in Zuo and Wang (2016). The three-classification result for MPM (Fig. 10) could be obtained based on the C-A model (Fig. 9).

Application of the SVM algorithm

MATLAB (2014a) and libsvm3.1 were used to implement the SVM algorithm for MPM. The choice of a kernel function (K) and its parameters for an SVM are crucial for obtaining good results (Zuo and Wang 2016). Four common kernel functions were used to predict the training and test sets. As shown in Fig. 11, the MSE value was closest to the predicted values calculated using the RBF kernel function with the training or test sets; therefore, the RBF function was chosen as the kernel function in this study; the RBF function is widely used in practical application (Foody and Mathur 2004; Oommen et al. 2008).

As shown in formula of the RBF function, C is a penalty value for misclassification error and γ is an inner product coefficient that determines the RBF width and affects classification accuracy. Too large gamma will lead to over fitting. In this study, the grid search method and cross validation were used to determine the values of C and gamma. The results are shown in Fig. 12. In Fig. 12, the x- and y-axes represent the logarithmic values (to the base 2) of C and γ respectively, while the contour line represents the MSE value of the corresponding values of C and γ . The lowest MSE value corresponds to the optimal values of C and γ . Therefore, the best C and γ were obtained ($C = 1$, $\gamma = 0.25$). Since the C and gamma values were small, over learning was avoided. The SVM model was used for MPM, the threshold was determined by the C-A method; the results are shown in Fig. 13.

Discussion

Validation of the predictive spatial model is a vital step in any prospectivity modeling method (Nykänen et al. 2015). In this study, three methods were tested (BP, SVM, and W-of-E). The number of known deposits in high-favorability metallogenic districts and the area

Fig. 13 Mineral prospectivity mapping using the Support Vector Machine (SVM) method

ratios of these districts given by post-classification predictions are used by many scholars to evaluate forecast results. This is because in smaller areas of spatial evidence, it is empirically ‘easier’ to identify undiscovered deposits than in larger areas (Yule 1912; Good 1950; Burger et al. 1989). Using this approach, the statistical results of the BP, SVM and W-of-E methods (Table 3) were obtained. The proportions of high favorability metallogenic districts in the whole study area were found to be 12.875%, 10%, and 9.25% for the three methods, respectively, but all contained 12 known deposits. The proportions of high and medium favorability metallogenic districts were 32.375%, 35.375%, and 30.25% respectively, but they contained 82.61%, 86.96%, and 69.57% of known deposits, respectively. The fewest numbers of known deposits was predicted in the high and medium favorability metallogenic districts by the W-of-E method; however, there was little difference in performance between the BP and SVM methods.

As another means of evaluating prospectivity models, Yousefi et al. (2012, 2013) plotted prospectivity classes against the percentage of known mineral occurrences delineated by (i.e., the prediction rate of) respective prospectivity classes. The same prediction-area (P-A) plots were employed in this study (Fig. 14). Based on this approach, the BP method found that 74% of the

deposits were contained in 26% of the study area; for the SVM method, 72% of the deposits were contained in 28% of the study area; for the W-of-E method, 68% of the deposits were contained in 32% of the study area. Therefore, the BP method exhibited better predictive capability than the SVM and W-of-E methods for the Hatu-region.

The receiver operating characteristic (ROC) technique has been widely used to test the validity of spatial predictive models (Bradley 1997; Nykänen et al. 2015) and represents an alternative approach to visualizing, organizing, and selecting classifiers based on their performance. The value of the area under curve (AUC) ranges from 0 to 1.0; the larger the AUC, the better the performance. A ROC curve nearer to (0, 1), leading to a larger AUC value, represents a more efficient classification (Fawcett 2006; Gao et al. 2016). ROC curves and statistical analysis were adopted to assess the prediction performance of the three methods. Figure 15 shows the results of ROC analysis considering both the true positive rate (TPR) and the false positive rate (FPR) according to different probability threshold values of mineral prospectivity. The AUC for the BP, SVM and W-of-E methods were 0.825, 0.811, and 0.790, respectively. The results show that a satisfactory effect was achieved by the application of BP, because when the value of AUC is between 0.7–0.9, higher accuracy is obtained.

Table 3 Statistical results for the BP, SVM and W-of-E methods

	High-favorability metallogenic districts	Medium-favorability metallogenic districts	Non-metallogenic districts
Different methods	BP/SVM/W of E	BP/SVM/W of E	BP/SVM/Wof E
Number of known deposits	12/12/12	7/8/4	4/3/7
Area covered	12.875%/10%/9.25%	19.5%/25.375%/21%	67.625%/64.625%/69.75

Fig. 14 P-A plot for mineral prospectivity. a: Back Propagation (BP) method; b: Support Vector Machine (SVM) method; c: Weight of evidence (W-of-E) method

The P-A plot (Fig. 14) was constructed based on the ability of prospectivity models to predict mineral deposits with respect to the size of predicted target areas. The ROC curves (Fig. 15) were constructed based on whether every grid of the study area was classified correctly or not. In summary, two different validation methods both found that the BP method provided the more accurate results in this study.

Fig. 15 Receiver operating characteristic (ROC) curves for validation

Conclusions

In this study, a method of MPM based on the BP and SVM algorithms was developed and implemented, and compared with the application of a traditional W-of-E method. The following conclusions were obtained:

1. Sample and parameter selection is very important as regards BP and SVM model application. Sample selection must be according to four criteria. Parameter selection is usually according to the MSE value; smaller MSE values result in better fitting performance.
2. P-A and ROC were used to effectively evaluated the three tested methods; and results indicate that MPM using the BP and SVM machine learning methods have better prediction performance than the traditional W-of-E approach. The BP method has relatively best prediction capability for MPM in this area; however, this validation was not obtained using a traditional statistical method.
3. The next investigative step for predictive work in the study area will be to focus on high-favorability metallogenic districts that are not known to contain deposits.
4. The construction and application of the nonlinear MPM model based on machine learning is closely related to the analysis of regional metallogenic regularity and the extraction of ore-forming geological variables. However,

further work is needed to identify the most suitable geological variables for machine learning in MPM models.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 41602339, U1503291), the Western Light Foundation of the Chinese Academy of Sciences (CAS; Grant No. XBBS-2014-19), the Xinjiang Uygur Autonomous Major Project (Grant No. 201330121-3), the National Basic Research Program of China (Grant No. 973Program2014CB440803) and the Strategic Priority Research Program of the Chinese Academy of Sciences(XDA19030204).

References

- Abedi M, Norouzi GH (2012) Integration of various geophysical data with geological and geochemical data to determine additional drilling for copper exploration. *J Appl Geophys* 83:35–45
- Abedi M, Torabi SA, Norouzi GH, Hamzeh M, Elyasi GR (2012a) PROMETHEEII: a knowledge-driven method for copper exploration. *Comput Geosci* 46:255–263
- Abedi M, Torabi SA, Norouzi GH, Hamzeh M (2012b) ELECTRE III: a knowledge-driven method for integration of geophysical data with geological and geochemical data in mineral prospectivity mapping. *J Appl Geophys* 87:9–18
- Abedi M, Norouzi GH, Fathianpour N (2013) Fuzzy outranking approach: a knowledge-driven method for mineral prospectivity mapping. *Int J Appl Earth Obs Geoinf* 21:556–567
- Abedi M, Kashani SBM, Norouzi GH, Yousefi M (2017) A deposit scale mineral prospectivity analysis: a comparison of various knowledge-driven approaches for porphyry copper targeting in Seridune, Iran. *J Afr Earth Sci* 128:127–146
- Agterberg FP (1971) A in the forward propagation process, index for detecting favourable geological environments. *Can Inst Min Metall* 10:82–91
- Agterberg FP (1974) Automatic contouring of geological maps to detect target areas for mineral exploration. *Math Geol* 6:373–395
- Agterberg FP, Bonham-Carter GF (1999) Logistic regression and weights of evidence modeling in mineral exploration. *Proc. 28th Int Symp app Comput mineral Ind (APCOM)*. Golden, CO, USA, pp 483–490
- Atkinson P, Tatnall A (1997) Introduction neural networks in remote sensing. *Int J Remote Sens* 18:699–709
- Bonham-Carter GF (1994) *Geographic information Systems for Geoscientists: modeling with GIS*. Pergamon Press, Ontario, Canada 398
- Bonham-Carter GF, Agterberg FP, Wright DF (1989) Weights of evidence modelling: a new approach to mapping mineral potential, statistical applications in earth. *Sciences* 89(9):171–183
- Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recogn* 30:1145–1159
- Breiman L (2001) Random forests. *Mach. Learn* 45:5–32
- Brown WM, Gedeon TD, Groves DI, Barnes RG (2000) Artificial neural networks: a new method for mineral prospectivity mapping. *Aust J Earth Sci* 47:757–770
- Burger H, Kirsch C, Skala W (1989). The application of microcomputers in exploration and exploitation of mineral deposits. *Original Research Article Computers & Geosciences* 15(4): 587–591
- Carranza EJM (2008) Geochemical anomaly and mineral prospectivity mapping in GIS. In: *Handbook of exploration and environmental geochemistry*, vol 11. Elsevier, Amsterdam, p 351
- Carranza EJM (2010) Improved wildcat modelling of mineral prospectivity. *Resour Geol* 60:129–149
- Carranza EJM (2017) Natural resources research publications on geochemical anomaly and mineral potential mapping, and introduction to the special issue of papers in these fields. *Nat Resour Res* 26(4): 379–410
- Carranza EJM, Hale M (2001) Logistic regression for geologically constrained mapping of gold potential, Baguio district. *Philipp. Explor Min Geol* 10:165–175
- Carranza EJM, Hale M (2002a) Wildcat mapping of gold potential, Baguio district, Philippines. *Trans Inst Min Metall Appl Earth Sci* 111:100–105
- Carranza EJM, Hale M (2002b) Where porphyry copper deposits are spatially localized? A case study in Benguet province, Philippines. *Nat Resour Res* 11:45–59
- Carranza EJM, Mangaoang JC, Hale M (1999) Application of mineral exploration models and GIS to generate mineral potential maps as input for optimum land use planning in the Philippines. *Nat Resour Res* 8:165–173
- Chen Y (2015) Mineral potential mapping with a restricted Boltzmann machine. *Ore Geol Rev* 71:749–760
- Chen Y, Lu L, Li X (2014) Application of continuous restricted Boltzmann machine to identify multivariate geochemical anomaly. *J Geochem Explor* 140:56–63
- Cheng Y, Wu W (2017) Mapping mineral prospectivity using an extreme learning machine regression. *Ore Geol Rev* 80:200–213
- Cheng Q, Agterberg FP, Ballantyne SB (1994) The separation of geochemical anomalies from background by fractal methods. *J Geochem Explor* 51:109–130
- Cheng Q, Chen ZJ, Khaled A (2007) Application of fuzzy weights of evidence method in mineral resource assessment for gold in Zhenyuan District, Yunnan Province, China. *Earth Sci J China Univ Geosci (In Chinese)* 32:175–184
- David BS, Paul KT, Shen SQ et al (1993) The Hatu gold anomaly, Xinjiang-Uygur autonomous region, China — testing the hypothesis of aeolian transport of gold. *J Geochem Explor* 47:201–216
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recogn Lett* 27:861–874
- Foody GM, Mathur A (2004) A relative evaluation of multiclass image classification by support vector machines. *IEEE Trans Geosci Remote Sens* 42(6):1335–1343
- Gao Y, Zhang ZJ, Xiong YH, Zuo RG (2016) Mapping mineral prospectivity for Cu polymetallic mineralization in southwest Fujian Province, China. *Ore Geol Rev* 75 (2016) 16–28
- Good IJ (1950) Probability and the weighting of evidence. *Griffin, London* 119pp
- Hamid NA, Nawi NM, Ghazali R (2011) Accelerating Learning Performance of Back Propagation Algorithm by Using Adaptive Gain Together with Adaptive Momentum and Adaptive Learning Rate on Classification Problems Computing and Multimedia Applications. *Second International Conference, UCMA 2011, Daejeon, Korea, April 13–15. Proceedings, Part II*
- Han BF, Qinq JJ, Sun B et al (2017) Late Paleozoic vertical growth of continental crust around the Junggar Basin, Xinjiang, China (Part I): Timing of post-collisional plutonism. *Acta Petrol Sin* 2006 22(5): 1077–1086 (in Chinese)
- Harris DP (1965) *An Application of Multivariate Statistical Analysis to Mineral Exploration*
- Harris DP (1969) Alaska's base and precious metals resources: a probabilistic regional appraisal. *Q. J. Colorado Sch. Min* 64:295–327
- Harris D, Pan G (1999) Mineral favorability mapping: a comparison of artificial neural networks, logistic regression, and discriminant analysis. *Nat Resour Res* 8:93–109
- Hashemi Tangestani M, Moore F (2002) The use of Dempster-Shafer model and GIS in integration of geoscientific data for porphyry copper potential mapping, north of Shahr-e-Babak, Iran. *Int J Appl Earth Observ Geoinform* 4:65–74

- Liu Y, Zhou KF, Xia QL (2017) A MaxEnt model for mineral Prospectivity mapping. *Nat Resour Res*. <https://doi.org/10.1007/s11053-017-9355-2>
- Ma D, Zhou T, Chen J et al (2017) Supercritical water heat transfer coefficient prediction analysis based on BP neural network. *Nucl Eng Des* 320:400–408
- Mejía-Herrera P, Royer JJ, Caumon G, Cheilletz A (2015) Curvature attribute from surface-restoration as predictor variable in Kupferschiefer copper potentials: an example from the fore-Sudetic region. *Nat Resour Res* 24(3):275–290
- Molan YE, Behnia P (2013) Prospectivity mapping of Pb–Zn SEDEX mineralization using remote-sensing data in the Behabad area, Central Iran. *Int J Remote Sens* 34(4):1164–1179
- Moon WM (1990) Integration of geophysical and geological data using evidential belief function. *IEEE Trans Geosci Remote Sens* 28:711–720
- Najafi A, Karimpour MH, Ghaderi M (2014) Application of fuzzy AHP method to IOCG prospectivity mapping: a case study in Taherabad prospecting area, eastern Iran. *Int J Appl Earth Obs Geoinf* 33:142–154
- Nykänen V, Lahti I, Niiranen T, Korhonen K (2015) Receiver operating characteristics (ROC) as validation tool for prospectivity models — a magmatic Ni–Cu case study from the Central Lapland Greenstone Belt, northern Finland. *Ore Geol Rev* 71:853–860
- Oh H, Lee S (2010) Application of artificial neural network for gold–silver deposits potential mapping: a case study of Korea. *Nat Resour Res* 19:103–124
- Oommen T, Misra D, Twarakavi NKC et al (2008) An objective analysis of support vector machine based classification for remote sensing. *Math Geosci* 40:409–422
- Piccini C, Marchetti A, Farina R, Francaviglia R (2012) Application of indicator kriging to evaluate the probability of exceeding nitrate contamination thresholds. *Int J Environ Res* 6(4):853–862
- Porwal A, Carranza EJM, Hale M (2006) Bayesian network classifiers for mineral potential mapping. *Comput Geosci* 32(1):1–16
- Rodriguez-Galiano V, Sanchez-Castillo M, Chica-Olmo M et al (2014) Machine learning predictive models for mineral prospectivity: Anevaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol Rev* 71:804–818
- Sadeghi B, Khalajmasoumi M (2015) A futuristic review for evaluation of geothermal potentials using fuzzy logic and binary index overlay in GIS environment. *Renew Sust Energ Rev* 43:818–831
- Shen P, Shen YC, Liu TB et al (2009) Geochemical signature of porphyries in the Baogutu porphyry copper belt, western Junggar, NW China. *Gondwana Res* 16(2):227–242
- Shen P, Pan HD, Zhu HP (2016) Two fluid sources and genetic implications for the Hatu gold deposit, Xinjiang, China. *Ore Geol Rev* 73(2):298–312
- Sinclair AJ, Woodsworth GL (1970) Multiple regression as a method of estimating exploration potential in an area near terrace, B.C. *Econ Geol* 65:998–1003
- Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14(3):199–222
- Wang L, Zhu YF (2015) Multi-stage pyrite and hydrothermal mineral assemblage of the Hatu gold district (west Junggar, Xinjiang, NW China): implications for metallogenic evolution. *Ore Geol Rev* 69:243–267
- Xu YJ, You T, Cl D (2015) An integrated micromechanical model and BP neural network for predicting elastic modulus of 3-D multi-phase and multi-layer braided composite. *Compos Struct* 122:308–315
- Yin LB, Liu GC, Zhou JL et al (2017) A calculation method for CO₂ emission in utility boilers based on BP neural network and carbon balance. *Energy Procedia* 105:3173–3178
- Yousefi M, Kamkar-Rouhani A, Carranza EJM (2012) Geochemical mineralization probability index (GMPi): a new approach to generate enhanced stream sediment geochemical evidential map for increasing probability of success in mineral potential mapping. *J Geochem Explor* 115:24–35
- Yousefi M, Carranza EJM, Kamkar-Rouhani A (2013) Weighted drainage catchment basin mapping of stream sediment geochemical anomalies for mineral potential mapping. *J Geochem Explor* 128:88–96
- Yule GU (1912) On the methods of measuring association between two attributes. *J.R. Stat. Soc* 75:579–642
- Zhang NN, Zhou KF (2015) Mineral prospectivity mapping with weights of evidence and fuzzy logic methods. *J Intell Fuzzy Syst* 29:2639–2651
- Zhao ZH, Bai ZH, Xiao XL, Mei HJ (2006) The diagenetic and mineralization of the rich alkali igneous rocks in northern China. Xinjiang Geological Press, Beijing 2006:1–302 (in Chinese)
- Zuo RG, Carranza EJM (2011) Support vector machine: a tool for mapping mineral prospectivity. *Comput Geosci* 37:1967–1975
- Zuo R, Wang J (2016) Fractal/multifractal modeling of geochemical data: a review. *J Geochem Explor* 164:33–41
- Zuo R, Zhang ZJ, Zhang DJ et al (2014) Evaluation of uncertainty in mineral prospectivity mapping due to missing evidence: a case study with skarn-type Fe deposits in southwestern Fujian Province, China. *Ore Geol Rev* 71:502–515