Paper:

PSO-SVM Optimized Kriging for Geological Modeling of Coal-bearing Formation

Mingdi Ma*,,***, Luefeng Chen*,**,***,†, Min Li*,**,***, Min Wu*,**,***, Witold Pedrycz**** and Kaoru Hirota*****

*School of Automation, China University of Geosciences, Wuhan 430074, China
**Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China
***Engineering Research Center of Intelligent Technology for Geo-Exploration, Ministry of Education, Wuhan 430074, Chaina
****Department of Electrical and Computer Engineering University of Alberta Edmonton, Canada
*****Tokyo Institute of Technology, Japan
†Corresponding author: Luefeng Chen
E-mail: chenluefeng@cug.edu.cn
[Received 00/00/00; accepted 00/00/00]

The accuracy of kriging interpolation mainly depends on the selection of the variogram. In actual coal mine engineering, the number of drilling data in the same coal working face is small, and the amount of data that can be used as training is small, which leads to the low degree of fitting of the traditional variogram model. Therefore, this paper proposes an automatic fitting method for the variogram of the support vector regression machine based on particle swarm optimization (PSO-SVR). Using the powerful fitting ability of SVR when dealing with small samples and the powerful parameter optimization ability of PSO, the variogram can be reconstructed from actual data. Through the traditional kriging spherical model, PSO optimized spherical model (PSO-Sphercial), and SVR model, a comparative analysis of the fitting correlation with the method in this paper is carried out. The experimental results show that the PSO-SVR reconstructed variogram has a higher fitting degree. The proposed method can provide support for coal seam thickness prediction and mining.

Keywords: Coal-bearing formation, geological modeling, PSO-SVR, variogram

1. Introduction

At present, China's coal industry is moving towards a new stage of intelligent mining. Intelligent mines take intelligent mining as the core and integrate big data applications as the support to build a new type of green, safe and efficient mine [1]. However, the uncertainty of geological conditions restricts the precise mining of intelligent mines. For this reason, the establishment of a high-precision geological model of coal-bearing formation is a necessary condition for the realization of intelligent mine construction. Coal seam thickness variation is the most common geological phenomenon in coal mines [2]. If its prediction error is large, it will affect coal mining and reduce work efficiency. Therefore, reflecting changes in coal seam thickness has become one of the main goals of coal-bearing formation geological models [3].

Spatial interpolation is one of the main methods of geological modeling. The function between the attribute value and the spatial distribution is obtained through the discrete sampling point data, and the function is used to optimally approximate all the adopted point data, and then the relevant attributes of any point in the distribution area are calculated [4]. At present, the commonly used spatial interpolation algorithms include Inverse Distance Weight (IDW), Radial basis function and kriging interpolation. The traditional kriging interpolation constructs a spatial interpolation model by fitting the existing variogram [5,6]. In the choice of the theoretical variogram model, there is a strong human subjectivity, which leads to the construction of the kriging interpolation accuracy which needs to be improved. As a more commonly used theoretical model in the field of coal mines, the spherical model itself requires more drilling data as training data. In the actual project, when the amount of drilling data in the coal working face is small, the fitting degree of the spherical model is poor. As coal mine detection projects continue to advance, coal workers continue to increase their requirements for coal seam detection accuracy. Higher-precision detection results can not only save exploration time, but also reduce costs. Therefore, the research to improve the modeling accuracy of a small sample of coal-bearing formation exhibits high practical engineering significance. With the development of artificial intelligence algorithms, the application of the SVR proposed in [7-9] has achieved good results in emotion recognition, providing ideas for modeling small samples of coal-bearing formation.

This paper proposes a Particle Swarm Optimization to optimize the support vector regression machine method (PSO-SVR) to automatically fit the experimental variogram. When using SVR for small samples, it has strong generalization ability, low computational complexity, and can overcome the characteristics of Back Propagation neural network (BP) overfitting. Through the strong group optimization ability of PSO, the penalty factor c and the kernel parameters g in the SVR are optimized to further improve the fit of the variogram. The method proposed in this paper can reduce the subjectivity of the variogram model selection in practical engineering applications with a small amount of drilling data, improve the accuracy of kriging interpolation, and establish a more accurate geological model of coal-bearing formation.At present, this method has not been used to predict coal seam thickness in small data sets.

2. PSO-SVR Optimized Kriging

The overall structure of the PSO-SVR optimized kriging algorithm proposed in this paper is shown in **Fig. 1**.

2.1. Variogram in Kriging

Kriging interpolation obtains the data quality by evaluating the error of the estimated prediction value, and uses the variogram to characterize continuous random variables. It analyzes the spatial distribution and correlation between interpolation points and neighboring points to determine the neighboring areas that affect the attribute values of the interpolation points. Kriging interpolation point for each neighboring region given a certain weight, the interpolation point is estimated by a weighted average value of the attribute [10,11].

The general formula of kriging interpolation is

$$Z^*(x)=\sum_{i=1}^{n}\lambda_i Z(x_i) \quad (i=1,2,\cdots,n) \quad . . . . \quad (1)$$

where $Z^*(x)$ is the estimated value of the interpolation point attribute, The $Z(x_i)$ is the i-th attribute values at neighboring points x_i. λ_i is the weight of the i-th neighboring point, and it needs to meet the two conditions of unbiasedness and optimality, which are as follows

$$\begin{cases} \sum_{i=1}^{n}\lambda_i=1 \\ \sigma^2=E\left[Z^*(x)-Z(x)\right]^2=\min \end{cases} \quad . . . . . \quad (2)$$

where σ is the standard deviation between the estimated value of the interpolation point attribute $Z^*(x)$ and the true value $Z(x)$.In kriging, the variogram is used to measure the degree of spatial correlation between sample points.

The spatial distribution of coal seam thickness has the autocorrelation of regionalized variables, and the spatial variation structure of coal seam thickness can be expressed according to the characteristics of coal seam thickness distribution revealed by drilling. Because the change of coal seam thickness conforms to the characteristics of regionalized variables, regionalized variables are often characterized by variable difference functions in geostatistics. Therefore, the experimental variogram of coal thickness can be calculated with limited observation points and observation values, and then theoretically fitted.

The sample's variogram (experimental variogram) formula can be represented as

$$\gamma^*(h)=\frac{1}{2N_h}\sum_{i=1}^{n}\left[Z(x_i+h)-Z(x_i)\right] \quad . . . . . \quad (3)$$

where $\gamma^*(h)$ represents experimental variogram. N_h is the number of all point pairs separated by h, h represents separation distance, $Z(x_i)$ represents the attribute value at position x_i. In this study, $\gamma^*(h)$ only depends on the size of the separation distance and not its direction.

Usually, the theoretical variogram model is selected to fit the experimental variogram obtained from the sample points. Common theoretical models are: linear model, spherical model, gaussian model, etc [12,13]. In coal-bearing rock formation modeling, the spherical model is the most widely used. The spherical model formula can be written as

$$\gamma(h)=\begin{cases} C_0 & h=0 \\ C_0+C\left(\frac{3h}{2a}-\frac{h^3}{2a^3}\right) & 0<h\leq a \\ C_0+C & h>a \end{cases} \quad . . \quad (4)$$

where C, C_0, A are undetermined parameters. The selection of different parameters will determine the accuracy of model fitting.

2.2. PSO-SVR Optimized Variogram

Support Vector Machine (SVM) was originally to solve the two-classification problem. Related scholars extended SVM and established SVR to solve the regression problem of function estimation [14,15]. SVR uses the decision boundary of the optimal hyperplane in support vector classification to build a regression model. x_i is the sample input, y_i is the sample output, $\phi(x_i)$ represents the feature vector after x_i is mapped to the high-dimensional feature space, and the corresponding optimal hyperplane formula is obtained as

$$f(x_i)=\omega^T\phi(x_i)+b \quad . . . . \quad (5)$$

where ω is the normal vector and b is the displacement term. The essence of the SVR model training process is to find the optimal ω and b so that $f(x_i)$ is close to y_i, and the convex optimization function is obtained as

$$\min_{\omega,b}\frac{1}{2}\|\omega\|^2+c\sum_{i=1}^{m}(\xi_i^*+\xi_i) \quad . . . . \quad (6)$$

where c is the penalty factor, ξ_i^* and ξ_i are relaxation factors. The constraint conditions corresponding to the convex optimization function are obtained as

$$\text{s.t}\begin{cases} f(x_i)-y_i\leq\varepsilon+\xi_i \\ y_i-f(x_i)\leq\varepsilon+\xi \\ \xi_i,\xi_i\geq 0, i=1,2\ldots,m \\ c>0 \end{cases} \quad . . . . \quad (7)$$

where the relaxation factor ε represents the deviation of

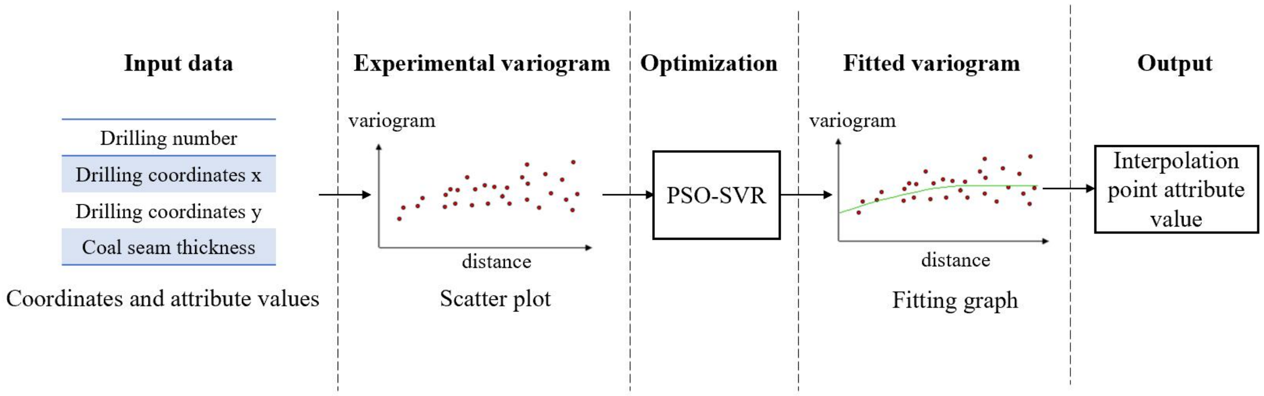

Fig. 1. PSO-SVR optimized Kriging algorithm: an overall structure.

$f(x_i)$ and y_i. Then the Lagrange multipliers a_i and a_i^* are introduced to solve this constrained convex optimization function, and the parameters ω and regression model function $f(x)$ in the SVR model can be obtained as

$$\omega = \sum_{i=1}^{m} (a_i^* - a_i)\,\phi(x_i) \quad (8)$$

$$\begin{cases} f(x) = \sum_{i=1}^{m} (a_i^* - a_i)\,K(x,x_i) + b \\ K(x,x_i) = \exp\left(-g\,\|x - x_i\|^2\right) \end{cases} \quad (9)$$

where g represents the kernel parameter of the kernel function $K(x,x_i)$. The resulting kernel function $K(x,x_i)$ improves the model's ability to deal with nonlinear regression problems. Radial Basis Function (RBF) can effectively improve the fitting effect and prediction performance of the model, so it is often used as a kernel function to optimize the SVR model.

In the SVR model, the penalty factor c controls the error range to avoid over-fitting or under-fitting. In the RBF, the kernel parameters g control the distribution of data after mapping to the new feature space, determine the number of support vectors, and affect the speed of training and prediction. Therefore, choosing the right method to obtain the optimal parameters is a problem we need to solve.

The particle swarm optimization algorithm is another population based optimization algorithm. The algorithm was first proposed by Kennedy and Eberhart in 1995 [16]. Its basic concept originated from the study of artificial life bird predation behavior. It uses information sharing and collaboration between individuals to search for the best location. The solution space of the problem is mapped to the corresponding particle space, and each particle has position and velocity information. For specific problems, adjust in the feasible solution space according to individual fitness and group fitness to search for the optimal value.

According to statistical ideas, the best fit to the theoretical model is essentially to minimize the variance between the theoretical variogram value and the actual variogram value [17]. So the fitness function is expressed as

$$f_i = \frac{1}{\sum_{i=1}^{N(h)} \left(r^*(h_i) - r(h_i)\right)^2} \quad (10)$$

In this paper, the penalty factor c and the kernel parameter g are used as the particles in the PSO. The procedure of the PSO optimizes SVR parameter in **Fig. 2**. According to the fitness function described above, the parameters are optimized following the PSO algorithm to determine the optimal parameter value.

3. Experiment of Fitting Variogram

This paper uses the drilling data in the actual project to conduct experiments to verify the feasibility of the method. The experimental software used in the reported experiment is MATLAB R2019.

3.1. Data Setting

The data set in this paper comes from a coal working face in Xinjiang. It contains 14 drilling points, recording the x-coordinate, y-coordinate and the thickness of the coal seam. The relative position of its plane and spatial distribution is shown in **Fig. 3, 4**. It can be seen from the drilling data that the thickness of the coal seam in the study area is between 3-9 meters.

3.2. Simulation Experiments

In this part, we analyzed and compared the fitting degree of the theoretical spherical model, PSO-Spherical model, SVR model and PSO-SVR model to the experimental sample variogram.

According to Section 2, calculating the experimental variogram of the sample is the first task of this method. First, calculate the distance matrix of each point according to the drilling data, select the step length $lag = 100$, and obtain the maximum step length distance through calculation. Then find the matching point pair through the loop

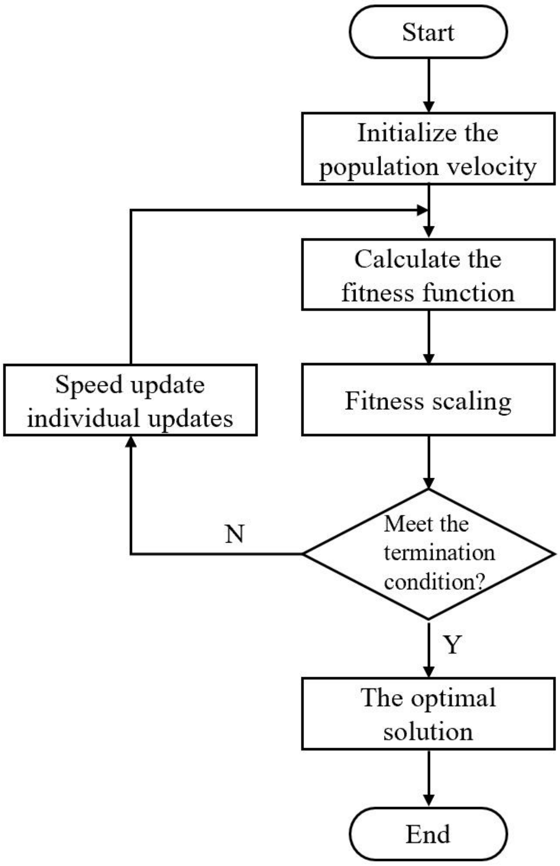

Fig. 2. Procedure of the PSO optimization of SVR parameter.

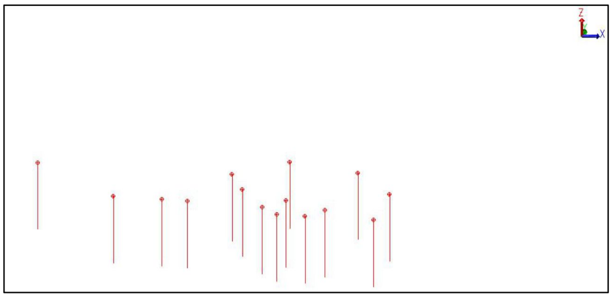

Fig. 4. Relative position distribution of drilling space.

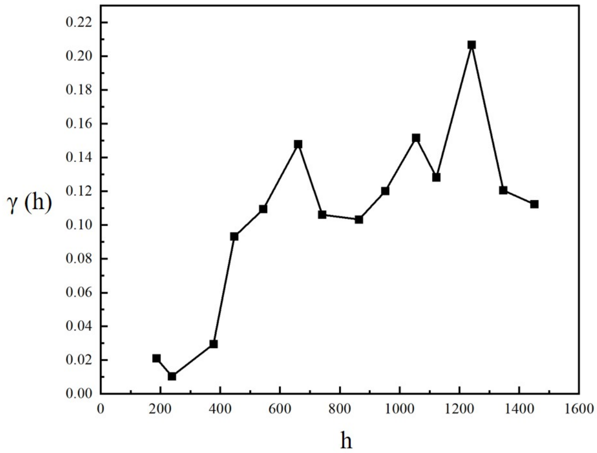

Fig. 5. Sample experimental variogram.

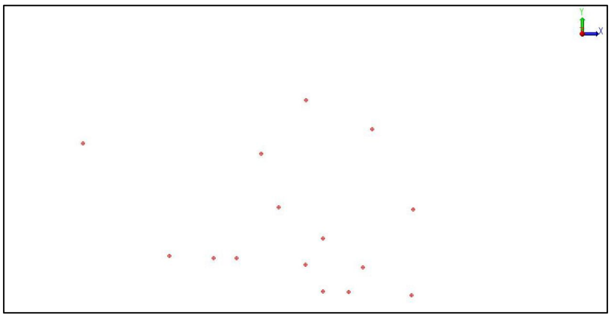

Fig. 3. Relative position distribution of drilling plane.

code. Finally, the experimental variogram of the sample is calculated, as shown in **Fig. 5**.

The spherical model curve is shown in **Fig. 6**. In the spherical model of the variogram theory, range a quantitatively represents the spatial variation range of the regionalization variable, nugget value C_0 represents the randomness of the regionalization variable, and sill value $C+C_0$ represents the magnitude of the spatial variation of the regionalization variable.

In traditional geological modeling software, the parameters of the spherical model are fitted by the Least Square (LS) method. This method has low computational complexity, but could easily fall into a local optimal solution. Therefore, on this basis, using a, C_0 and C as particles, PSO is used to optimize the parameters to improve the fit of the spherical model.

Although the fitting degree of the spherical model optimized by PSO has been improved, when the experimental data is small and the variogram has a peak, its fitting effect needs to be improved, so it is very important to find a model that can be automatically fitted.

Aiming at the experimental variogram curve, this paper uses the SVR model for preliminary fitting.

In order to further improve the fitting degree of SVR, PSO is used to optimize the penalty factor c and the kernel parameter g. This paper sets the learning factor $C_1 = 1.3$, $C_2 = 1.7$, the number of particles $size = 50$, and the maximum number of iterations $max = 500$ in the PSO.

3.3. Result Analysis and Discussion

The experimental variogram of the sample through the traditional kriging spherical model, PSO-Spherical model, SVR model and PSO-SVR model is shown in **Fig. 7**.

It can be seen from **Fig.7** that the value range of the sample experiment variogram is from 0.0105 to 0.2066, and there are three peaks of 0.1479, 0.1518 and 0.2066. The empirical variogram obtained by the theoretical spherical model ranges from 0.0339 to 0.1424, and the

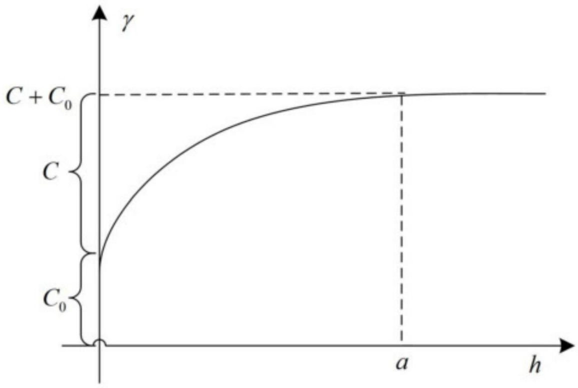

Fig. 6. Spherical model.

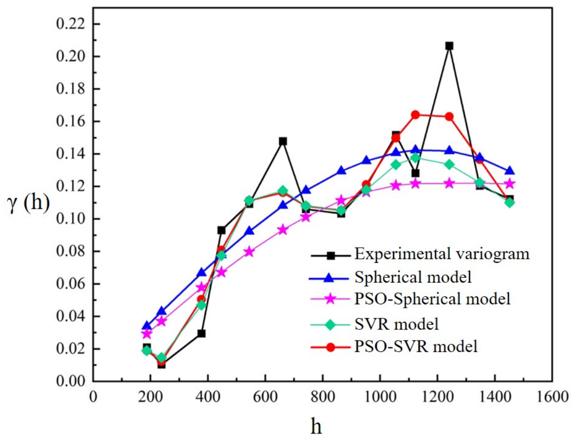

Fig. 7. Fitting plot of experimental variogram.

fitting values for the three peaks are 0.1082, 0.1408 and 0.1419 respectively. Therefore, it can be seen that the model has a poor fitting effect for the numerical mutation of the experimental variogram. The empirical variogram obtained by using the PSO-Spherical model ranges from 0.0293 to 0.1219, and the fitted values for the three peaks are 0.0934, 0.1206, and 0.1219 respectively. This model has the lowest degree of fit for the numerical mutation of the experimental variogram, but its overall curve fit is higher than that of the theoretical spherical model. The empirical variogram obtained by SVR model ranges from 0.0190 to 0.1377, and the fitting values for the three peaks are 0.1175, 0.1335 and 0.1336 respectively. Compared with the previous two models, the model has a significantly improved fit for the numerical mutation of the experimental variogram, and there is a more obvious peak trend change. The PSO-SVR model proposed in this paper obtains the empirical variogram in the range of 0.0190-0.1642, and the fitting values of the three peaks are 0.1163, 0.1499 and 0.1630, respectively.

Due to the small number of drilling in the coal working face in the study area, it is not suitable for cross-validation. Therefore, this article chooses Mean Squared Error (MSE) as the measurement standard. The MSE computed as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad \ldots\ldots\ldots\ldots (11)$$

where n is the number of samples, y_i is the actual data, $\hat{y}_i$ is the fitted data. MSE data can be evaluated the degree of change, the smaller the numerical value is, the experimental data fitting model described better accuracy. The MSE of each model is shown in **Table 1**.

Table 1. MSE of Each Model.

	Mean Squared Error
Spherical model	0.0775
PSO-Spherical model	0.0683
SVR model	0.0519
PSO-SVR model	**0.0362**

It can be seen from the table that $MSE_{(Spherical)} > MSE_{(PSO-Spherical)} > MSE_{(SVR)} > MSE_{(PSO-SVR)}$.

The results show that the PSO-SVR automatic fitting model is higher than other models in terms of peak fitting and overall curve fitting. This is because the PSO algorithm has a unique memory function and can dynamically track the global optimization. Using this advantage of the PSO algorithm, it is reasonable to select the SVR parameters to form the PSO-SVR model, which can make the prediction results more scientific and rigorous. The initial state of the particles in the model can be adjusted autonomously to improve the search accuracy, and finally a fitting model with smaller errors can be obtained.

In practical engineering applications, coal seam thickness prediction and modeling is the focus of geological modeling of coal-bearing rock formations. In the ordinary kriging interpolation process, the PSO-SVR automatic fitting model proposed in this paper is used to form a three-dimensional visualization model of coal seam thickness by writing code, as shown in the **Fig. 8**. In this figure, the coordinate axis Z represents the thickness of the coal seam, and the color change from blue to yellow represents the change in the thickness of the coal seam from thin to thick.

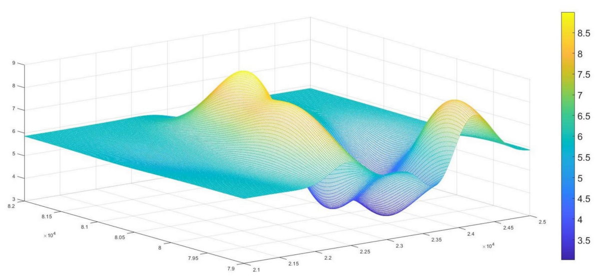

Fig. 8. Coal seam thickness preseated in the study area.

4. Conclusion

The accuracy of Kriging interpolation largely depends on the selection of the variogram. In practical engineering applications, when targeting a certain coal working face, there are characteristics of less drilling data and insufficient available data. Therefore, the commonly used theoretical variogram model has a low degree of fit. In response to this situation, this paper proposes a PSO-SVR model to automatically fit the variogram. By comparing the theoretical spherical model, PSO-Spherical model, and SVR model, it is concluded that this method has the best fit. The method in this paper provides a feasible scheme for the prediction of coal seam thickness in practical engineering applications.In the follow-up work, try to apply the ant colony algorithm and neural network commonly used in emotion recognition and target detection to the study of coal seam thickness prediction of small data sets.

Acknowledgements

This work is supported by National Natural Science Foundation of China under Grants 61973286, 61603356, and 61733016, the 111 project under Grant B17040, and the Fundamental Research Funds for the Central Universities, China University of Geosciences (No. 2021068).

References:

[1] X L Liu. Digital construction of coal mine big data for different platforms based on life cycle. *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, 2017, 456-459.

[2] Y Shi, Y Sun, Q Zhong, et al. Research on Three-Dimensional Reproduction Technology of Coal Face. *2016 International Conference on Multimedia Systems and Signal Processing (ICMSSP)*, 2016, 105-109.

[3] Q Zeng, W Ming, Z Zhang, et al. Construction of a 3D Stratum Model Based on a Solid Model. *IEEE Access*, 2021, 9: 20760-20767.

[4] Q Jia, W Li, D Che. A Triangulated Irregular Network Constrained Ordinary Kriging Method for Three-Dimensional Modeling of Faulted Geological Surfaces.*IEEE Access*, 2020, 8: 85179-85189.

[5] Z Zhang, X Lai, M Wu. Improved Kriging for Drilling Visualization Based on Quantum Genetic Algorithm. *2018 37th Chinese Control Conference (CCC)*, 2018: 10281-10284.

[6] Wikipedia, New York, NY, USA. Variogram.[Online], available: https://en.wikipedia.org/wiki/Variogram, 2019.

[7] L Chen, Y Feng, M Maram, et al. Multi-SVM based Dempster-Shafer Theory for Gesture Intention Understanding Using Sparse Coding Feature. *Applied Soft Computing*, 2019, 85: 105787-105795.

[8] L Chen, M Zhou, M Wu, et al. Three-Layer Weighted Fuzzy Support Vector Regression for Emotional Intention Understanding in Human-Robot Interaction. *IEEE Transactions on Fuzzy Systems*, 2018, 26(5): 2524-2538.

[9] L Chen, M Wu, M Zhou, et al. Dynamic Emotion Understanding in Human-Robot Interaction Based on Two-Layer Fuzzy SVR-TS Model. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, 50(2): 490-501.

[10] H Wang, R Zhang, K Liu, et al. Improved Kriging Interpolation Based on Support Vector Machine and Its Application in Oceanic Missing Data Recovery. *2008 International Conference on Computer Science and Software Engineering*, 2008: 726-729.

[11] H Chen, X Wang and G Chen. Soil phosphorus mapping using Kriging interpolation method. *2010 World Automation Congress*, 2010: 167-170.

[12] GISGeography Semi-Variogram: Nugget, Range and Sill. [Online]. available: https://gisgeography.com/semi- variogram-nugget-range-sill, 2019.

[13] S Garrigues, D Allard and F Baret. Using First- and Second-Order Variograms for Characterizing Landscape Spatial Structures From Remote Sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(6): 1823-1834.

[14] A Khusro, M S Hashmi, A Q Ansari. Enabling the Development of Accurate Intrinsic Parameter Extraction Model for GaN HEMT using Support Vector Regression(SVR). *IET Microwaves Antennas & Propagation*, 2019, 13(9): 1457-1466.

[15] J Zou, C Li, Q Yang, et al. Fault prediction method based on SVR of improved PSO. *The 27th Chinese Control and Decision Conference (2015 CCDC)*, 2015: 1671-1675.

[16] D Pal, P Verma, D Gautam, et al. Improved optimization technique using hybrid ACO-PSO. *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, 2016: 277-282.

[17] F Yulianto, W F Mahmudy, A A Soebroto. Comparison of Regression, Support Vector Regression(SVR), and SVR-Particle Swarm Optimization(PSO) for Rainfall Forecasting. *Journal of Information Technology and Computer Science*, 2020, 5(3): 235.