ISBOA-Transformer Coal Mine Roadway Gas Concentration Prediction with Biological Intelligence Improvement

Yan Chai*, Zhen Li

College of Science, Liaoning Technical University, Liaoning, China

**Corresponding Author.*

Abstract:

In the domain of coal mine safety, precisely forecasting the gas concentration in roadways holds utmost significance. It serves as a crucial safeguard for miners' lives and well - being and also plays a pivotal role in augmenting the economic performance of mining enterprises. To achieve this accurate prediction of coal mine roadway gas concentration, this research centers around the Bultai coal mine for in - depth investigation. A state - of - the - art concentration monitoring sensor is utilized to perform real - time gas concentration measurements within the coal mine roadways. Moreover, a novel ISBOA - Transformer prediction model for coal mine roadway gas concentration, enhanced through bio - intelligence techniques, is introduced. Initially, improvements are made to the traditional Secretary Bird Optimization Algorithm (SBOA). Latin hypercubic sampling is adopted to ensure a more uniform population initialization, and a fixed - point recombination and mutation strategy is implemented. This strategy aims to boost the evolutionary probability of species and diversify the population. Finally, the optimized algorithm is utilized to precisely tune the hyperparameters associated with the number of attention heads in the Transformer model. Through this process, a highly effective gas concentration prediction model is developed. When validating the model with a training dataset and a testing dataset, the Mean Absolute Error (MAE) of the proposed model is determined to be 0.00062531 for the training dataset and 0.0005832 for the testing dataset. The Root Mean Squared Error (RMSE) registers at 0.0008848 for the training dataset and 0.00082342 for the testing dataset. The Goodness of Fit (R2) values are 0.9712 for the training dataset and 0.9800 for the testing dataset. Clearly, these evaluation metrics vividly demonstrate that the ISBOA - Transformer model performs better than other models in the comparison. Subsequently, the model is implemented on extra gas concentration datasets obtained from the Bultai coal mine.

Keywords: coal mine roadway gas concentration prediction, ISBOA, Transformer, generalization capability

INTRODUCTION

As an important energy industry in China, the safety production in coal mines has always been the focus of extensive attention from all walks of life. In the process of coal mine production, one of the most significant hidden perils endangering the safety of coal mines is gas disaster. The dynamic change of coal mine roadway gas concentration is complex and difficult to grasp accurately, once the gas concentration exceeds the safety threshold, it is very likely to cause gas explosion, combustion and other serious accidents, which will not only cause huge casualties and property losses, but also adversely affect the sustainable development of the coal industry. Consequently, achieving precise and dependable prediction of the gas concentration in coal mine roadways holds crucial practical significance for guaranteeing the safe production of coal mines, mitigating the accident risk, and enhancing the economic benefits of enterprises. Over In the past few decades, new ideas and methods have been provided for the predicting coal mine roadway concentration level of gas. Various types of sensors are able to obtain the concentration immediately and accurately, while intelligent algorithms can deeply mine and analyze these massive data to establish accurate prediction models.

For example, Tutak et al. [1] developed a quantitative assessment method for coal mine gas explosions based on Bayesian networks (BN) and computational fluid dynamics (CFD). The methodology uses BN for probabilistic risk analysis and CFD for simulating gas dispersion and explosion scenarios, providing a comprehensive framework for identifying key risk factors and enhancing safety management in underground coal mines. Nie Z et al. [2] conducted an in - depth exploration of the gas diffusion theory and the environmental characteristics of coal mines. They employed the Gaussian plume model to simulate the gas diffusion within coal mine roadways and proposed an optimization model based on the genetic algorithm and BP neural network. This research concentrated on the gas diffusion theory and the environmental features of coal mines. By simulating the gas diffusion process in coal mine roadways through the Gaussian plume model, an optimized gas prediction model based on the genetic algorithm and BP neural network was constructed. This model has the advantages of small computational volume and high prediction accuracy. Moreover, by integrating continuous machine learning with daily monitoring, its reliability can be further enhanced. This provides a brand - new technical solution for the intelligent monitoring of the coal mine roadway environment, contributes to the innovation and development of the coal mine industry in the field of safety monitoring, and is expected to promote the improvement of the intelligent level of safe coal mine production. Olatomiwa et al. [3] researched the gas diffusion theory as well as coal mine environment, simulated gas diffusion by Gaussian model, and proposed an novel model on the basis of genetic algorithm and BP neural network. The model, characterized by a minimal computational footprint and remarkable prediction precision, can have its reliability augmented when continuous machine learning is incorporated into daily monitoring routines. This combination provides a fresh technological approach for the intelligent surveillance of the coal mine roadway environment. It empowers more exact and effective monitoring within the intricate coal mine setting, thus improving the overall safety and operational efficiency of the mining process. Gan et al. [4] proposed a gas concentration prediction approach grounded. This approach processes the original gas concentration time series data by decomposing the signals using VMD to extract essential features. Subsequently, normalization and wavelet threshold noise reduction are applied to enhance data quality. The VMD-LSTNet-Attention model demonstrates higher prediction accuracy and robustness compared to traditional methods, achieving significant improvements in the prediction of gas concentrations. LUO et al. [5] put forward a prediction model rooted in multi-feature and XGBoost, by integrating the characteristics of historical gas concentration, temperature, and wind speed data, and leveraging the gradient boosting algorithm of XGBoost, the training process of the decision tree is accelerated. This approach not only fully utilizes the multi - dimensional data but also takes advantage of XGBoost's algorithmic efficiency to enhance the training speed of the decision tree, which is expected to contribute to more accurate gas - related predictions in coal mine roadways. Experiments show that the prediction error of this method is lower than the existing deep learning models and the training speed is faster. By simulating the behavior of groups, intelligent optimization algorithms open up completely new paths for attacking complex problems [6-9]. Compared with traditional algorithms, intelligent optimization algorithms achieve an excellent balance between avoiding local optimal solutions and guaranteeing convergence to the global optimal solution by virtue of their subtle design. This makes it show the advantages of high efficiency and accuracy when solving practical problems, especially when dealing with complex and variable problems with lack of information. In this process, scholars have innovatively proposed Particle Swarm Algorithm [10], Sparrow Search (SSA) Model [11], Emperor Penguin Optimization (EPO) Algorithm [12], Harris Hawk Optimization (HHO) Algorithm [13], Sailfish Optimization (SFO) Algorithm [14], Gray Wolf Optimization (GWO) Algorithm [15] and Snake Bird Optimization Algorithm (SBOA) [16]. A series of optimization algorithms. These algorithms persistently conduct exploration activities within the search space, thoroughly delving into the inherent characteristics of the problem itself. This enables a more comprehensive and in - depth understanding of the problem, without being restricted by preconceived assumptions about the model, thus providing a more flexible and effective approach to problem – solving, so as to gradually converge to the optimal solution. Thanks to the gradient-free mechanism, they are able to cope with problems that do not have definite gradient information or sequential derivatives, and have shown greater flexibility and robustness in practical applications. However, intelligent optimization algorithms are not perfect and face many challenges. Among them, the problems of falling into local optimality and premature convergence are more prominent [17-20]. To solve these problems, many scholars have improved the algorithms from different perspectives.

In view of this, this report develops a bio-intelligent improved ISBOA-Transformer coal mine roadway gas concentration prediction model. The Secretary Bird Optimization Algorithm (SBOA) is improved by introducing Latin hypercubic sampling during population initialization to make it homogeneous, and the bio-intelligent fixed-point recombination and mutation strategy is introduced during updating the optimal individuals to increase the evolution probability of the species and enrich the diversity of the population. Finally, it was combined with the Transformer model to find the hyperparameters of the number of optimal attention heads, which was modeled according to the different gas concentration data collected from the Bultai coal mine.

ISBOA-TRANSFORMER COAL MINE ROADWAY GAS CONCENTRATION PREDICTIVE MODELING

Snake and Heron Optimization Algorithm (SBOA)

The snake heron, a remarkable African raptor, is renowned for its unique appearance and behavior. It inhabits the grasslands, savannas, and open river regions of sub - Saharan Africa. Its typical habitats include tropical open grasslands, sparsely wooded savannas, and open areas filled with tall grasses. Additionally, it can also be spotted in semi - deserts or wooded areas with open clearings. Characterized by grayish - brown plumage on its back and wings, a pure white breast, and a dark black belly, the snake heron stands out in its natural environment. Inspired by the snake heron's hunting and escape behaviors, Youfa Fu et al. put forward the SBOA. This algorithm is composed of the snake heron's hunting strategy and escape strategy. These intervals respectively correspond to the three phases of searching for prey, consuming prey, and attacking prey. In the escape phase, the snake heron has two main strategies. First, it camouflages itself into its surroundings. By blending in with the environment, it can avoid potential threats. However, when the bird realizes that the surroundings are not suitable for camouflage, it opts to fly or run quickly as a means to rapidly escape from predators.

The initial phase of the algorithm, which randomly distributes each snake heron in the search space, can be depicted as the below equation:

$$X_{i,j} = lb_j + r \times (ub_j - lb_j) \tag{1}$$

ub_j , lb_j delegate the upper and lower boundaries of the algorithm's search range, respectively. r is a constant number:i is an individual snake heron; j is the dimension of this research.

The hunting phase of the snake heron has three different steps: searching, consuming, and attacking the prey, and each phase accounts for 1/3 of the algorithm iterations, and the position updates in different phases can be represented by Eqs. (2) to (4), respectively.

$$X_{i,j}^{newP1} = x_{i,j} + (x_{random\_1} - x_{random\_2}) \times R_1 \tag{2}$$

$$X_{i,j}^{newP1} = x_{best} + \exp((\frac{t}{T})^\wedge 4) \times (RB - 0.5) \times (x_{best} - x_{i,j}) \tag{3}$$

$$X_{i,j}^{new,P1} = x_{best} + ((1 - \frac{t}{T})^\wedge (2 \times \frac{t}{T})) \times x_{i,j} \times RL \tag{4}$$

x_{random_1} , x_{random_2} is the random position in this phase of the iterative process; and the random number is between (0, 1) and between (0, 0.5); is the optimal position in the current phase being represented by the iterative update Eq. (5).

$$X_i = \begin{cases} X_i^{new,P1} & F_i^{new,P1} < F_i \\ X_i & else \end{cases} \tag{5}$$

F^{newP1} denotes the value of the objective function fitness.

When the snake heron encounters a threat, it opens an escape strategy, which is categorized into two types: quick escape or camouflage. This can be represented by equations (6) to (7).

$$X_{i,j}^{new,P2} = \begin{cases} C_1: & x_{best} + (2 \times RB - 1) \times ((1 - \frac{t}{T})^\wedge 2) \times x_{i,j} \\ C_2: & x_{i,j} + R_2 \times (x_{random} - K \times x_{i,j}) \end{cases} \tag{6}$$

$$X_i = \begin{cases} X_i^{new,P2} & F_i^{new,P2} < F_i \\ X_i & else \end{cases} \tag{7}$$

The switching conditions for 2,1 are.

$$\begin{cases} C_1 & r < 0.5 \\ C_2 & else \end{cases} \tag{8}$$

Both R_2 and r are number between (0, 1); x_{random} is a random position value in the current iteration; K is a random value 1 or 2.

Improved Snake and Heron Optimization Algorithm (SBOA)

Latin hypercube sampling (LHS) represents a technique for uniformly sampling within a multidimensional space. Its objective is to produce a set of samples that are as evenly distributed as possible, with minimal repetitions, within a specified parameter space.

Latin hypercubic sampling function was used to obtain lhs, which outputs a matrix of $1 \times D$. Thanks to the full - coverage and equal - probability sampling characteristics of the LHS, the samples obtained are more uniformly distributed compared to those generated through random sampling. As a result, when LHS is employed for population initialization, it is beneficial in preventing the population from getting trapped in a local optimal solution during the initialization phase. This, in turn, enhances the algorithm's global search capability. Moreover, a more extensive initial distribution of the population enables the algorithm to explore the advantageous regions within the search space more rapidly. This effectively improves the algorithm's efficiency in identifying the optimal solution. As shown in Figure 1, the generation number is set to 100, which represents the random initialization Yulin hypercubic sampling initialization of individual distribution, respectively. The above results indicate that the values initialized by means of LHS are more evenly distributed.

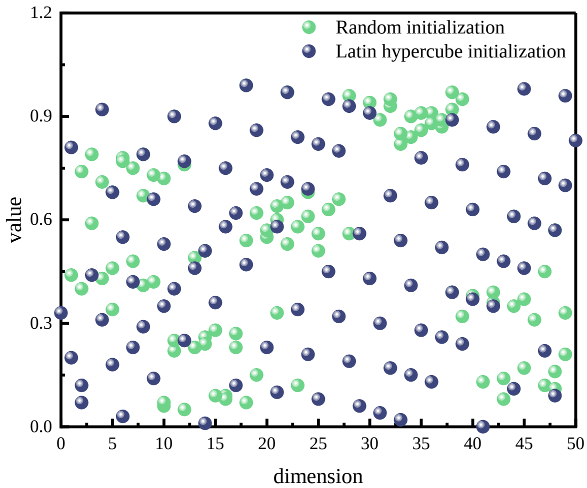


Figure 1. Random initialization vs. Latin hypercube sampling initialization

Targeted recombination and mutagenesis strategies are two powerful techniques in molecular biology that are used to precisely manipulate and introduce DNA variation, respectively. Targeted recombination allows scientists to swap or insert gene fragments at specific locations in the genome. This technique shows great promise including gene editing, functional genomics research, and transgenic technology, and its precision allows the study of gene function and regulation of biological processes in cells. Unlike targeted recombination, mutagenesis strategies aim to introduce random or targeted genetic variants through chemical mutagens, radiation, or molecular tools such as TALENs, ZFNs, and CRISPR/Cas9. These variants can be used to create mutation libraries for forward genetics screening or to directly introduce specific mutations to study their effects on organismal function. Mutagenesis strategies are critical for drug development, crop improvement, and understanding the root causes of genetic diseases, and they provide a means to explore genetic diversity and biological complexity. In this paper, we propose fixed-point recombination and mutagenesis strategies to optimize the SBOA in the capture and escape phases, where the algorithm increases with the number of iterations leading to a decrease in population diversity and falling into local optimal solutions and slow convergence. The mathematical formulation is expressed as follow:

$$X_i(t+1)=\begin{cases} X_{i,recombination}^{r,d} & ,f(X_{i,recombination}^{r,d})\ is\ \ better\ \ than\ \ f(X_i(t)) \\ X_i(t) & ,otherwise \end{cases} \tag{9}$$

$$X_i(t+1)=\begin{cases} X_{i,mutagenesis}^{d} & ,f(X_{i,mutagenesis}^{d})\ is\ \ better\ \ than\ \ f(X_i(t)) \\ X_i(t) & ,otherwise \end{cases} \tag{10}$$

$$d=q\times D \tag{11}$$

$$p_r=\frac{Fitness(i)-Fitness_{\min}}{Fitness_{\max}-Fitness_{\min}} \tag{12}$$

$$X_{i,mutagenesis}^{d}=X_i^d\times rand \tag{13}$$

$X_{i,recombination}^{r,d}$ is the position after the reorganization of the information in d dimensions of the ith individual and the rest of the individuals in the population; $X_{i,mutagenesis}^{d}$ is the position after the mutation of the information in d dimensions of the ith individual; q is the control factor, D is the dimension of the problem to be solved; $Fitness(i)$, $Fitness_{\min}$ and $Fitness_{\max}$ are the current fitness of the individual, the worst individual fitness and the optimal individual fitness within the population, respectively; p_r is the coefficient of mutation for determining whether an individual performs mutation, and takes the value of 0.4; rand is number between [0,1]. When $p_r<0.4$, the individual performs the full cycle fixed-point recombination and mutagenesis strategy.

During each iteration of the hunting and escape session of the SBOA algorithm, some elite individuals have large deviations between the variables in some dimensions and the theoretical optimal solution, resulting in poor adaptation. However, the variables of these individuals in other dimensions may have approached or reached the optimal solution. Considering the differences in fitness among individuals within the population, there may be information in which an individual's information on the dimensional components that fall into the neighborhood of the optimal solution can be effectively recombined with another individual's information on the poorer dimensions.

To achieve this goal, we employ a fixed-point recombination and mutation strategy. Specifically, we reorganize the latter's information on the worse dimension with the former's information on the corresponding dimension, and at the same time perform a mutation operation on it. The advantage of this strategy is that it maximizes the retention of good solution information and enhances the quality of individuals, thereby boosting the evolutionary probability of species and enriching population diversity.

The steps of the operation in detail are as follow:

Step1: In this step, the population size is determined. By leveraging the Latin Hypercube Sampling (LHS) method, the dimension of the search space is defined based on the number of variables in the optimization problem. The LHS method can cover the entire exploration space more scientifically and evenly, thereby precisely defining the exploration boundary of the algorithm. For each variable's range of values, the Latin hypercube sampling technique is utilized. It randomly creates the starting positions of every individual within the population, specifically within the relevant dimension of the search space. This method ensures a more uniform distribution of initial points compared to simple random sampling. By doing so, it helps in exploring the search space more comprehensively and efficiently when conducting optimization algorithms or simulations. These sampled initial positions will serve as the candidate solutions for the problem. Additionally, the maximum number of algorithm iterations and relevant parameters required for subsequent strategies such as hunting and escape are set.

Step2: Evaluate the fitness, and based on the objective function of the optimization problem, the fitness value of each individual is computed. Introducing the fixed-point reconstruction and mutation strategy to optimize the SBOA in the hunting and escape phases, the algorithm increases with the number of iterations resulting in the reduction of population diversity and falling into the local optimal solution, slow convergence speed and other problems.

Step3: Prey hunting strategy in three stages. The prey search phase uses differential evolution strategy to update individual positions according to the specialized formula, expand the search space by using individual position differences, and decide whether to update by comparing the old and new position adaptations to promote the population to the optimal development. The consuming prey phase incorporates the global optimal position and the Brownian motion principle to guide individuals to explore their neighborhood, continuously evaluate the fitness, and update the new position if it is better, so as to bring the population closer to the possible optimal solution area. The prey attack phase introduces a Levy flight strategy with a nonlinear perturbation factor, allowing individuals to move randomly over long distances to enhance their exploration ability, and comparing fitness after updating to determine whether to update the position or not.

Step4: Escape strategy, when certain pre-conditions are met, such as determining that the current individual may be trapped in the local optimal dilemma, which is just like the snake heron needs to escape when it encounters a danger in reality, then the escape strategy should be enabled. By using the dynamic perturbation factor (where is the current iteration number and is the maximum iteration number), combined with the corresponding position update rule, the position of the individual is updated. Such an update helps the individual to go beyond the limitation of local optimization to explore other regions in the solution space that have not yet been penetrated, thus further improving the probability of the algorithm to find the global optimal solution, and allowing the algorithm to keep moving towards the best result in the continuous adjustment and optimization.

Transformer model

Transformer breaks the limitation of the traditional sequence model relying on the loop structure, and with the self-attention mechanism as the core, it is able to process the input sequence in parallel, which greatly improves the computational efficiency. It is principally constituted by the encoder and the decoder. The encoder conducts feature extraction and representation learning on the input sequence. Through multi- layer multi-head self-attention and feed-forward neural network modules, it captures the intricate dependencies among elements within the sequence. When generating the output sequence, the decoder not only employs self - attention but also utilizes the output of the encoder for cross-attention computation, so as to ensure the relevance between the generated result and the input. The structure is illustrated Figure 2.

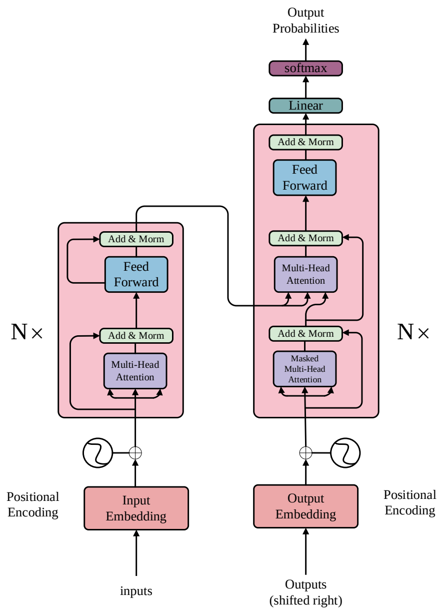


Figure 2. Transformer structure

To take into consideration the order of the input sequence, the Transformer model makes use of a specific position vector, namely position encoding (PE), which ascertains the position of the current data.

The formula is:

$$p_i^{(2s)} = \sin\left(\frac{i}{10000^{2s/d_{\text{mod}}}}\right) \tag{14}$$

$$p_i^{(2s+1)} = \cos\left(\frac{i}{10000^{2s/d_{\text{mod}}}}\right) \tag{15}$$

i is the input data position, s is the dimension, d_{mod} represents the number of input features, and p_i represents the position encoding.

The multi-head self-attention mechanism in the encoder layer is essentially the splicing of the computation results of multiple self-attention mechanisms, which makes the model go from learning the information of only one representation space to learning the information of multiple different representation spaces. The self-attention mechanism can make the model pay more attention to the important information contained in its own data and reduce the dependence on external information. The calculation method is shown in equation (17).

$$\begin{cases} Q = X_f W^Q \\ K = X_f W^K \\ V = X_f W^V \\ Attention(Q,K,V) = soft\max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \\ h_i = Attention(Q,L,V) \\ MultiHead(Q,K,V) = Concat\left(\sum_{i=1}^{m} h_i\right) \end{cases} \tag{16}$$

Where, X_f denotes the input sequence matrixs, W^Q, W^K and W^V denote the trainable matrix. The d_k is the dimension of Q, K, V; W denotes the weight matrix, m is the number of attention heads, h_i is the result of the ith attention head computation, and the Concat function denotes the result of the computation used to splice the individual attention head.

ISBOA-Transformer Method for Predicting Gas Concentration in Coal Mine Roadway

By virtue of its self-attention mechanism, the Transformer model is able to efficiently process coal mine roadway gas concentration time series data to capture the dependencies between different time steps. Specifically, the historical gas concentration data are used as the input sequence, which is subjected to feature extraction and representation learning by Transformer's encoder, and then the predicted values of future gas concentration are generated by the decoder. The Improved Serpent Heron Optimization Algorithm (ISBOA) is harnessed to carry out the optimization of the number of attention heads within the Transformer framework. The detailed implementation process unfolds as follows: in the initial stage, the methane data obtained from the Bultai coal mine undergoes a random division. Specifically, 70% of the data is allocated to the training set and 30% to the test set, which are concurrently generated for subsequent utilization. Subsequently, by capitalizing on the temporal attributes of the methane data, a preliminary ISBOA-Transformer gas prediction model is erected. This model is designed to leverage the time-series characteristics of the data to capture latent patterns and correlations. Next, ISBOA is applied to optimize the hyperparameters of the model, with the number of attention heads of the model being designated as the key parameter for optimization. During this optimization journey, the root mean square error (RMSE) serves as the objective function. The algorithm iteratively adjusts the parameters to confirm that the error, encompassing the training objective is low (<0.01) and to check if the maximum number of iterations has been attained. This meticulous approach validates the effectiveness of ISBOA in pinpointing the optimal hyperparameters. Finally, the optimized parameters are utilized to construct the final ISBOA - Transformer gas prediction model. This model is then deployed to make predictions on the test data. To validate the model's accuracy and applicability, a comprehensive performance evaluation is executed. By benchmarking the model against other counterparts, the aim is to demonstrate the high-precision prediction capabilities and superiority of the proposed ISBOA-Transformer model. The overall workflow is presented in Figure 3.

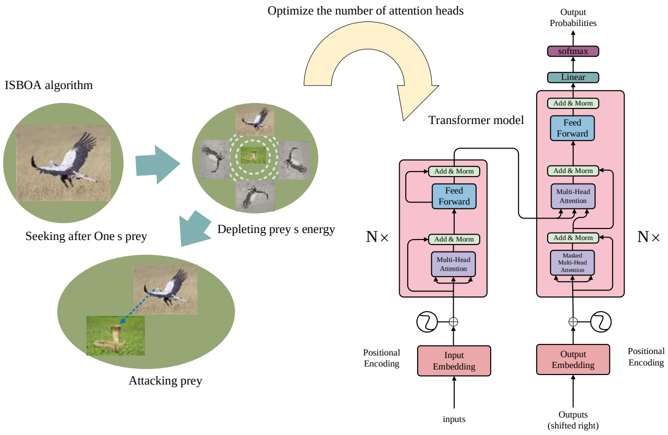


Figure 3. Overall workflow

EXPERIMENTAL ANALYSIS

Source of Data Sets

The Bultai Mine stands as the world's first mega-mine to have attained world-leading benchmarks in terms of mine production capacity, the capacity of the main transportation system, and coal washing and processing capacity. In this mine, monitoring and predicting methane concentrations is particularly critical, not only for environmental protection, but also to ensure worker safety and production continuity.

Current methane sensors have a detection 0.01% accuracy, limiting the precise measurement of methane concentration in the return air [21,22]. Therefore, a laser methane concentration monitoring sensor is used in this study, and the linear arrangement of the sensor can realize continuous measurement along the path, which is applicable to a diverse array of applications within complex environments, and is superior to point-type sensors. The monitoring points are arranged at 1400 m in the 12-coal three-panel area return airway, 1500 m in the 12-coal general return airway, and 460 m in the 42-coal two-panel area return airway in the Bultai coal mine, and the monitoring activities are carried out during the period of April 1, 2024 to July 31, 2024, and the monitoring frequency is set to record the data at intervals of 60 seconds.

To assess that the model can effectively leverage the restricted field measurements for sufficient learning, it was noted in this study that outliers were present in the different datasets. Therefore, a rigorous data preprocessing step was implemented for all datasets. The goal is to protect the model from outliers during the training and prediction phases, thus improving its accuracy and reliability. Figure 4 shows the data set after preprocessing.

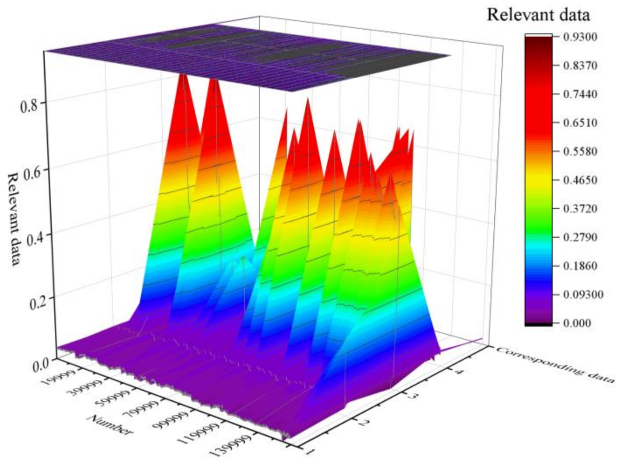


Figure 4. Data set after preprocessing

Model Testing

For the purpose of evaluating the model's ability to generalize to new data in our study, the gas concentration dataset was randomly split into a training subset and a testing subset, with the training subset accounting for 70% and the testing subset for 30%. The model's prediction results for both the training subset and the testing subset are depicted in Figure 5. As analyzed in Figure 5, the distribution of the prediction results on the training set and the test set is close to a straight line, which indicates that the prediction performance of the ISBOA-Transformer gas prediction model on both is stable and close to the real value, and suggests that the model has good feasibility and applicability in practical applications.

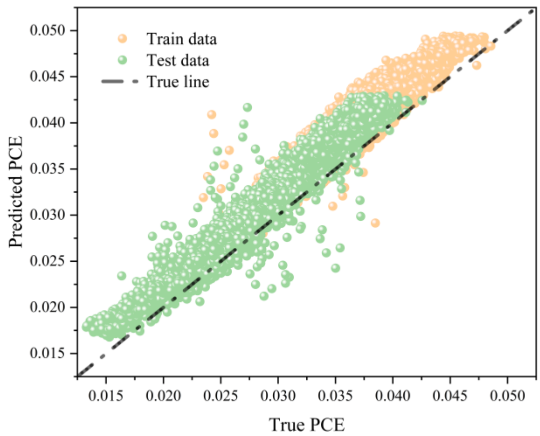


Figure 5. Model training set and test set prediction results

Model Comparison

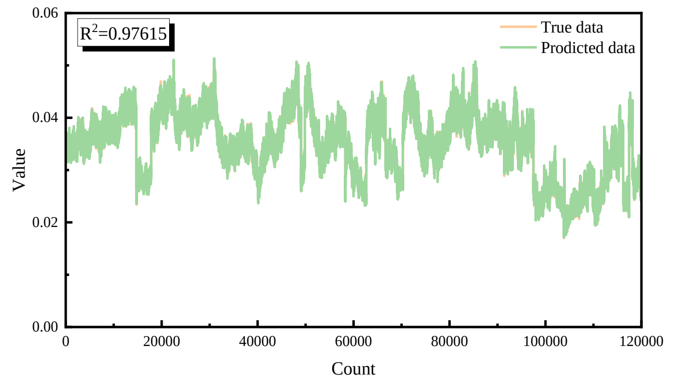
In order to corroborate the superiority of the developed ISBOA-Transformer prediction model, five models were selected for comparison experiments in this study, specifically。 Additionally, four commonly - employed statistical metrics are utilized to quantify the degree of fit between the predicted values and the actual values. These metrics include the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE), the Root Mean Square Error (RMSE), and the goodness - of - fit R2. The expressions for these metrics are obtained [1]. The results of each model evaluation index are displayed in Table 1.

Table 1. Hyperparameters of each model

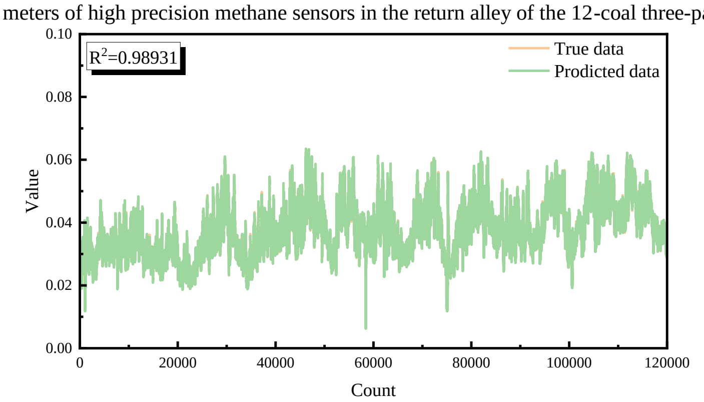
models	MAE		MAPE/%		RMSE		R^2	
	Training set	*Test set*	*Training set*	*Test set*	*Training set*	*Test set*	*Training set*	*Test set*
ISBOA-Transformer	0.00062531	0.0005832	1.9531	2.3819	0.0008848	0.00082342	0.9712	0.9800
Transformer	0.006421	0.0006452	2.2515	2.4817	0.000913	0.0007747	0.94232	0.9311
Autoformer	0.008145	0.0007999	2.4761	2.6491	0.001841	0.0008275	0.9134	0.9511
LR	0.008886	0.0008736	2.4913	2.7923	0.001414	0.0007975	0.924	0.911
XGBoost	0.007642	0.0008123	2.3910	2.4812	0.001241	0.0007436	0.9355	0.9376

From Table 1, it can be found that the ISBOA-Transformer model predicts methane on the dataset collected by the Bultai coal mine through the 460 high-precision methane sensors in the return-air alley of the 42-coal second disc area, and the values of its MAE in the training set and the test set are 0.00062531 and 0.0005832, and the values of its RMSE in the training set and the test set are 0.0008848 and 0.00082342, and the values of R2 are 0.9712 and 0.9800 in the training and test sets. therefore, it shows good performance compared to other models.

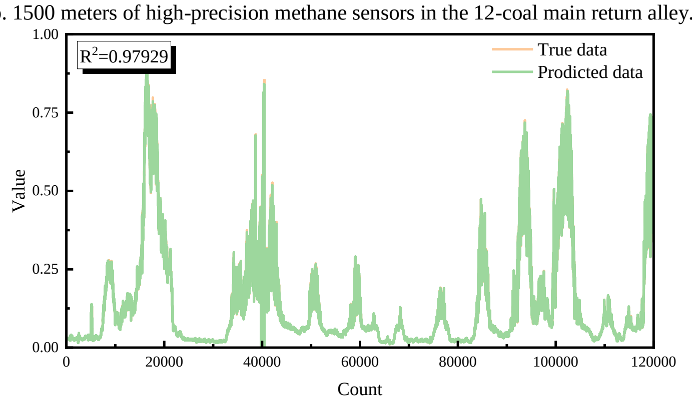
Through the data collected from 1400 m of high-precision methane sensors in the return-air alley of the 12-coal three-panel area, 1500 m of laser methane concentration monitoring sensors in the 12-coal total return-air alley, 2800 m of high-precision methane sensors in the 22-coal total return, and 460 m of laser methane concentration monitoring sensors in the 42-coal two-panel area, the model constructed by this study is verified in terms of its ability to be generalized, and model parameters The model parameters are consistent with the optimization method reference and above. The fitting effect of the test set is shown in Figure 6.



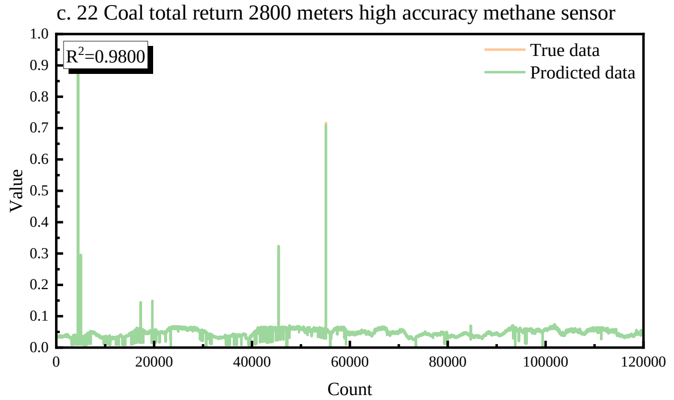
a. 1400 meters of high precision methane sensors in the return alley of the 12-coal three-panel area



b. 1500 meters of high-precision methane sensors in the 12-coal main return alley.



c. 22 Coal total return 2800 meters high accuracy methane sensor



d. 460 high-precision methane sensors in the return-air alley of 42 Coal No. 2 Plate Area

Figure 6. Fitting of gas test sets in different coal mines

As can be observed from Figure 6, the ISBOA - Transformer prediction model for coal mine roadway gas concentration developed in this study showcases outstanding performance when applied to diverse datasets within the same application context and modeling conditions. Significantly, the R2 value exceeds 0.95 for both the training dataset and the test dataset. This clearly manifests its high - precision predictive prowess. Additionally, the values of MAPE, MAE, and RMSE are substantially lower, which implies that the constructed model has a superior performance compared to other models.

CONCLUSIONS

In this study, an ISBOA-Transformer coal mine roadway gas concentration prediction model with improved biological intelligence was constructed. Algorithmic improvements were made to incorporate Latin hypercubic sampling in the population initialization part to make the initial position of the snake heron's population more homogeneous and to improve the population diversity. In the selection of optimal individuals, the algorithm is inspired by biological intelligence and incorporates the fixed-point recombination and mutation strategies to optimize the SBOA in the hunting and escape phases, which leads to a decrease in population diversity with the increase in the number of iterations, a fall into the local optimal solution, and a slow convergence speed. Finally, ISBOA was used to optimize the number of attention heads of Transformer model. Gas concentration data collected from four distinct geographical areas of the Burdai coal mine were predicted. The prediction results were compared with the outputs generated by the Transformer, Autoformer, LR, and XGBoost algorithms. The comparison results reveal that the model accuracy of the method proposed in this paper is higher. Specifically, the R2 value exceeds 0.95 for both the training set and the test set, thereby validating the generalization ability of the ISBOA - Transformer prediction model.

FUNDING

This research was funded by Youth Program of Ministry of Education, China, grant number 21YJCZH204; Natural Science Foundation of Liaoning Province, China, grant number 2020-MS-301, the Liaoning Provincial Department of Education, China, grant number LJKMZ20220694.

REFERENCES

[1] Tutak M, Brodny J. Predicting methane concentration in longwall regions using artificial neural networks. International Journal of Environmental Research and Public Health, 2019, 16(7): 1406.

[2] Nie Z, Ma H, Zhang Y. Research on Gaussian Plume Model of Gas Diffusion in Coal Mine Roadway Based on BP Neural Network Optimized by Genetic Algorithm//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2020, 526(1): 012158.

[3] Olatomiwa L, Mekhilef S, Shamshirband S, et al. A support vector machine–firefly algorithm-based model for global solar radiation prediction. Solar Energy, 2015, 115: 632–644.

[4] Junwei Zhuo, Xingyu Chen, Huisheng Zhang, et al. Rapid and high-accuracy concentration prediction of gas mixtures based on PMH-TCN, Measurement, Volume 242, Part C, 2025, 116003, ISSN 0263-2241.

[5] Wenchao Gan, Ruilong Ma, Wenlong Zhao, et al. A VMD-LSTNet-Attention model for concentration prediction of mixed gases, Sensors and Actuators B: Chemical, Volume 422, 2025, 136641, ISSN 0925-4005.

[6] Zhiqiang L U O, Hao Z, Jiajun T. Incorporating multi-features and XGBoost algorithm for gas concentration prediction. CHINA MINING MAGAZINE, 2024, 33(S1): 359-363, 370.

[7] Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. Applied Intelligence, 2021, 51(3): 1371-1388.

[8] Kim J, Shin W, Hong S, et al. A novel pathway to construct gas concentration prediction model in real-world applications: Data augmentation; fast prediction; and interpolation and extrapolation. Sensors and Actuators B: Chemical, 2023.

[9] Siarry P and Kulkarni A. Metaheuristics for Sustainable Manufacturing. Engineering Applications of Artificial Intelligence, 2024.

[10] Cheng S, Qin Q, Chen J, et al. Brain storm optimization algorithm: a review. Artificial Intelligence Review, 2016, 46: 445-458.

[11] Gad A G. Particle swarm optimization algorithm and its applications: a systematic review. Archives of computational methods in engineering, 2022, 29(5): 2531-2561.

[12] Müller A J, Michell R M, Pérez R A, et al. Successive Self-nucleation and Annealing (SSA): Correct design of thermal protocol and applications. European Polymer Journal, 2015, 65: 132-154.

[13] Khalid O W, Isa N A M, Sakim H A M. Emperor penguin optimizer: A comprehensive review based on state-of-the-art meta-heuristic algorithms. Alexandria Engineering Journal, 2023, 63: 487-526.
[14] Qiao L, Liu K, Xue Y, et al. A multi-level thresholding image segmentation method using hybrid Arithmetic Optimization and Harris Hawks Optimizer algorithms. Expert Systems with Applications, 2024, 241: 122316.
[15] Zhang Y, Mo Y. Chaotic adaptive sailfish optimizer with genetic characteristics for global optimization. The Journal of Supercomputing, 2022, 78(8): 10950-10996.
[16] Niu P, Niu S, Chang L. The defect of the Grey Wolf optimization algorithm and its verification method. Knowledge-Based Systems, 2019, 171: 37-43.
[17] Abdel-Basset, Mohamed, et al. "Secretary bird optimization algorithm: A new bio-inspired metaheuristic algorithm for global optimization problems." Engineering Applications of Artificial Intelligence 122 (2023): 110136.
[18] Pandey H M, Chaudhary A, Mehrotra D. A comparative review of approaches to prevent premature convergence in GA. Applied Soft Computing, 2014, 24: 1047-1077.
[19] Evers G I. An automatic regrouping mechanism to deal with stagnation in particle swarm optimization. The University of Texas-Pan American, 2009.
[20] Song Z, Ren C, Meng Z. Differential Evolution with perturbation mechanism and covariance matrix based stagnation indicator for numerical optimization. Swarm and Evolutionary Computation, 2024, 84: 101447.
[21] Ye H, Dong J. An ensemble algorithm based on adaptive chaotic quantum-behaved particle swarm optimization with weibull distribution and hunger games search and its financial application in parameter identification. Applied Intelligence, 2024: 1-30.
[22] Conrad B M, Tyner D R, Johnson M R. Robust probabilities of detection and quantification uncertainty for aerial methane detection: Examples for three airborne technologies. Remote Sensing of Environment, 2023, 288: 113499.