

Research on the purification mechanism of heavy metal pollution by biochar composites driven by degree learning

Anran Dai^a, Luxia Jia^b, Aoli Zhan^c, Xiuze Zhang^d

Wuhan University of Technology School of Resources and Environmental Engineering, Wuhan, 430070, China

Abstract. This paper proposes an innovative approach by integrating deep learning technology, specifically employing the GRU recurrent neural network model based on the Seagull optimization algorithm, to enhance the accuracy of predicting biochar performance. The Seagull optimization algorithm, inspired by seagull predatory behavior, is adept at efficiently identifying optimal model parameters, thereby improving the model's generalization ability and robustness. The GRU recurrent neural network, designed for sequence data processing, proves to be instrumental in capturing dynamic and nonlinear interactions between biochar and heavy metals. This, in turn, contributes to heightened prediction accuracy and model interpretability. The article unfolds in a structured manner, beginning with an introduction to the biochar preparation method and its characteristics. It then delves into an analysis of the sources and hazards of heavy metal pollution. Following this, the paper explains the principles and advantages of deep learning technology, providing a comprehensive foundation for the subsequent discussion. The construction and verification process of the proposed model is then detailed, concluding with the presentation of experimental results and in-depth analysis. In essence, this research introduces a pioneering idea and methodology for optimizing biochar design and effectively controlling heavy metal pollution, presenting a fresh perspective on addressing these environmental challenges.

1 Introduction

Due to its unique learning ability, neural networks play a crucial role in biochar performance analysis and prediction [1]. Heavy metal pollution [2] refers to the phenomenon that due to human activities or natural processes, heavy metal elements in the environment exceed normal levels, causing adverse effects on the ecosystem and human health. However, traditional methods usually have shortcomings such as high cost, low efficiency, many by-products, and poor sustainability. Biochar can not only be used as a heavy metal pollution control material, but also as a soil conditioner. In addition, biochar can also be used as a renewable energy source to replace fossil fuels, reduce greenhouse gas emissions, and alleviate the problem of global warming. Therefore, biochar is considered an environmental remediation material with multiple functions and advantages.

The efficacy of biochar and heavy metal removal efficiency are affected by many factors, including biochar preparation conditions, raw material type, surface modification, pollutant concentration, solution pH, temperature, etc. The intricate relationship between these factors brings challenges to traditional mathematical models to accurately describe and predict. To address this complexity and optimize biochar design to enhance heavy metal pollution mitigation, this paper proposes a GRU recurrent neural network model based on Seagull optimization algorithm to accurately predict biochar

performance and heavy metal removal rate. Deep learning is a machine learning technology with artificial neural networks as its core. It is good at learning abstract and complex features and rules from massive data sets, and provides powerful data fitting and prediction capabilities.

The Seagull optimization algorithm, an intelligent optimization approach inspired by seagull predatory behavior, proves effective in identifying optimal model parameters, thereby enhancing the model's generalization ability and robustness. Meanwhile, the GRU recurrent neural network, designed for processing sequence data, adeptly captures the dynamic and nonlinear interactions between biochar and heavy metals[3]. This capability significantly improves the model's prediction accuracy and interpretability. In summary, the integration of these technologies presents a promising avenue for advancing the accurate prediction and optimization of biochar performance in heavy metal remediation scenarios.

Hence, to address the existing challenges in predicting the performance of biochar and enhance its optimal design and application efficacy, this paper introduces a prediction model based on the GRU recurrent neural network and the Seagull optimization algorithm. The proposed model not only achieves high accuracy and interpretability but also demonstrates robustness and generalization capabilities[4]. Its adaptability to diverse scenarios of biochar application and heavy metal contamination in water makes it a valuable tool for advancing biochar research and development.

^adar@whut.edu.cn, ^b1328000870@qq.com, ^c354347179@qq.com, ^d576312457@qq.com

This model offers a new and effective approach to tackle the complexities associated with biochar performance prediction. Its high accuracy ensures reliable predictions, while its interpretability facilitates a better understanding of the underlying dynamics. Moreover, the model's robustness and generalization ability equip it to handle various situations related to biochar and water contaminated with heavy metals. Overall, this paper presents a novel method and tool that contribute to the advancement of biochar research, offering valuable insights and practical applications. The feasibility of the model has been successfully validated through experimentation.

2 Model

2.1. Seagull optimization algorithm

In 2019, Himan et al. introduced the Seagull Optimization Algorithm (SOA), inspired by the migration and aggressive behavior observed in seagull groups [5]. This algorithm is designed to simulate the strategic actions of seagulls, allowing them to navigate and avoid collisions, locate optimal positions, and efficiently attack during migration and foraging, all with the ultimate goal of discovering optimal solutions. Compared to alternative optimization algorithms, the Seagull Optimization Algorithm distinguishes itself by requiring fewer parameter adjustments and exhibiting superior global search capabilities. These characteristics contribute to its notable advantages, including a swift convergence speed and high optimization accuracy [6]. The unique emulation of seagull behavior makes SOA a promising and efficient tool for optimization challenges.

The algorithm simulates the migration of seagulls by modeling their positional adjustments during the process of moving from the current position to the next—a phase constituting the global search stage. During this stage, each seagull must adhere to specific conditions [7]. To prevent collisions with other individuals within the seagull group, a variable denoted as A is introduced. This variable is employed to calculate the updated position of the seagull after the transformation, ensuring that individual seagull locations are adjusted to avoid conflicts with others.

Once collision-free positioning is achieved, each seagull then directs its flight towards the location of the seagull with the best fitness value. Upon reaching a position where collisions are safely avoided, the seagull adjusts its trajectory towards the optimal position, ultimately arriving at the new position to complete the update. Figure 1 illustrates the representation of the best position, and the update formula is expressed as follows:

$$X_{ij}^{t+1} = X_{ij}^t + V_{ij}^t \times \text{rand}() \quad (1)$$

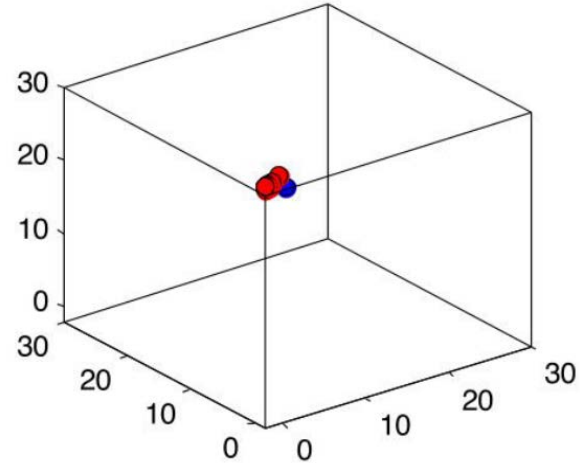


Fig. 1. Best seagull location collection

In the attack stage, seagulls attack food during flight. Specifically, seagulls use their wings and weight to control their flight height, and constantly adjust the angle and speed of attack during the process of hunting food, thus flying in the air. Exhibiting a spiral motion pattern, this process belongs to the local search stage [8]. From the above, the attack position of the seagull can be calculated, namely:

$$\vec{P}_s(x) = (\vec{D}_s \times x' \times y' \times z') + \vec{P}_{bs}(x) \quad (2)$$

Through this updating method, the seagulls in the group gradually search in the solution space, and finally converge to the potential optimal solution of the problem. This makes the Seagull optimization algorithm a swarm intelligence optimization algorithm with global search capabilities.

2.2 GRU recurrent neural network

The Gate Recurrent Unit (GRU) neural network is a simplified structure based on the LSTM recurrent neural network unit proposed by [9]. That is, based on the LSTM recurrent neural network, the gate structure is optimized and the gate structure is optimized. The forgetting gate and the input gate are combined into an "update gate", which can use the output result as a memory state and continuously pass it backwards in a loop without giving out additional memory states. This improves the training speed of the network while ensuring accuracy [10]. Compared with the three gates of the LSTM recurrent neural network, the GRU recurrent neural network only has an update gate and a reset gate. The details are as follows:

Change the input gate, forget gate, and output gate into update gate and reset gate. The unit state C_t and output of the LSTM recurrent neural network are combined into one state h_t . Figure 2 shows the unit structure of the GRU recurrent neural network neuron.

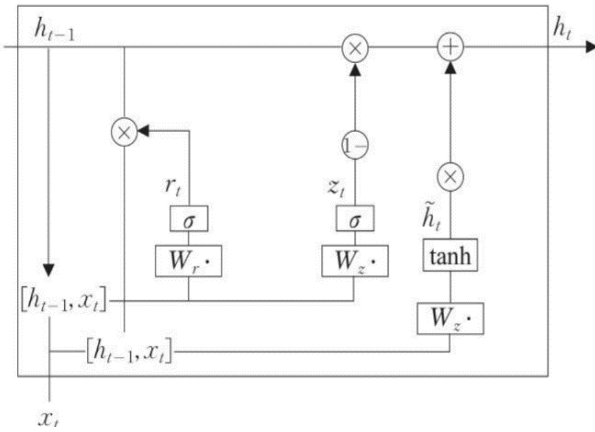


Fig. 2. GRU neuron structure

The workflow of GRU recurrent neural network can be summarized as follows:

The reset gate r_t connects x_t to the hidden layer state at the previous moment, and its weight affects the amount of information to be forgotten in the hidden layer state at the previous moment. The calculation formula is as follows:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (3)$$

In the formula: $\cdot$ represents matrix multiplication, W is the weight matrix of the gate structure, b is the bias; $[h_{t-1}, x_t]$ represents the connection of the matrices, and σ is the sigmoid activation function.

After calculating the reset gate, the candidate state $\tilde{h}_t$ of the hidden layer at the current moment is obtained. The calculation formula is as follows:

$$\tilde{h}_t = \tanh(W_z \cdot [r_t \odot [h_{t-1}, x_t] + b_z]) \quad (4)$$

Among them, $\odot$ represents the multiplication of corresponding elements of the matrix. If the value of r_t is larger, it means that the hidden layer saved more information at the previous moment, and then $\tilde{h}_t$ will be affected more.

Z_t is the update gate in the network, which controls the amount of information transferred from the hidden state at the previous moment in the network to the current hidden state h_t . When the weight is larger, it means more information needs to be saved. The calculation formula is as follows:

$$Z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (5)$$

Finally, the hidden layer state h_t at the current moment is output, which represents the amount of information that should be extracted from the information state h_{t-1} at the previous moment and the amount of information that needs to be retained from the candidate state $\tilde{h}_t$. The

output is jointly controlled by Z_t , h_{t-1} and $\tilde{h}_t$. Calculated as follows

$$h_t = (1 - Z_t) \odot h_{t-1} + Z_t \odot \tilde{h}_t \quad (6)$$

It can be seen from the GRU neuron structure diagram and calculation formula that each GRU unit makes a decision on retaining or forgetting information, thus creating dependencies between neural network units. The GRU recurrent neural network can save and discard

information only through update gates and reset gates. It is different from the LSTM recurrent neural network and has fewer parameters, which effectively improves the training speed. In addition, Shewalkar et al. also compared and evaluated three types of recurrent neural networks: RNN, LSTM and GRU, and tested their performance using data sets with long-term dependency characteristics. The experimental results illustrate the performance and accuracy of GRU and LSTM recurrent neural networks. Close, and the error rate is significantly lower than that of the RNN recurrent neural network. When the data set is the same, the GRU recurrent neural network converges faster [11]. Therefore, combining the characteristics of RNN, LSTM and GRU, this paper uses GRU recurrent neural network as the basic model to study the biochar performance prediction model.

2.3 SANN model

Despite its straightforward structure and limited parameters, the GRU recurrent neural network, like other neural networks, requires the manual setting of hyperparameters. Recognizing the potential challenges stemming from the subjective and random nature of manual hyperparameter tuning, this study adopts the Seagull optimization algorithm. The aim is to automate and optimize the adjustment of three pivotal hyperparameters in the GRU recurrent neural network: the number of hidden layer nodes, the initial learning rate, and the L2 regularization coefficient. By leveraging the Seagull optimization algorithm, this approach seeks to improve predictive accuracy by offering a more systematic and data-driven hyperparameter optimization process.

This article delineates the optimization process for the GRU recurrent neural network using the Seagull optimization algorithm (SOA), encompassing the following steps: (1) Initialization of parameters for the Seagull optimization algorithm, defining key aspects like the number of populations, iterations, hyperparameters for optimization, and the upper and lower limits of optimization; (2) Establishment of a GRU recurrent neural network with parameter settings including the number of input and output layer nodes, as well as the number of hidden layers; (3) Application of the Seagull optimization algorithm to optimize hyperparameters of the GRU recurrent neural network; (4) Computation of the fitness value of the seagull and implementation of position updates; (5) Evaluation of whether the Seagull optimization algorithm has reached the maximum iteration count. If conditions are met, consider the obtained optimal parameter values as the global solution; otherwise, return to step (4) until convergence; (6) Incorporation of the optimal results from the Seagull optimization algorithm as hyperparameters for the GRU recurrent neural network. These hyperparameters include the number of hidden layer nodes, initial learning rate, and L2 regularization coefficient. Subsequently, the optimized GRU recurrent neural network is trained for predicting biochar properties.

3 Experiment

3.1. Data set

We conducted a series of experiments to evaluate and compare the performance of our GRU recurrent neural network prediction model, which is based on the Seagull optimization algorithm. The experiments utilized a data set provided by Wang Min et al. [12] (2022) consisting of 96 samples, encompassing 24 different biochar composites and 4 heavy metals (lead, cadmium, chromium, and mercury). Each sample contains detailed records of the biochar preparation process, including 4 preparation parameters: original material, carbonization temperature, activator, and activation time. Additionally, the data set includes physical and chemical parameters related to heavy metal adsorption, such as initial concentration, solution pH, and temperature. Furthermore, each sample provides information on the adsorption performance of biochar for heavy metals, including two parameters: adsorption capacity and adsorption isotherm. This comprehensive data set offers valuable information for conducting in-depth studies on heavy metal adsorption by biochar composites, serving as a crucial foundation for research and development in related fields.

3.2 Evaluation standard

The evaluation of a predictive model involves considering several key metrics. The root mean square error (RMSE) is a measure of overall prediction accuracy, calculated as the average square root of the differences between the model's predicted values and the actual values. RMSE is particularly sensitive to outliers in the data. Conversely, the mean absolute error (MAE) provides insight into the model's average prediction error by calculating the absolute average differences between predicted and actual values. This metric offers a broader view of the model's performance across the dataset. In addition, the correlation coefficient gauges the linear relationship between the model's predictions and the actual values. With values ranging from -1 to 1, a coefficient closer to 1 signifies a stronger positive linear relationship. Together, these metrics contribute to a comprehensive assessment of the model's capabilities, encompassing accuracy, error, and the degree of linear association between predicted and actual values.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (7)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (8)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (9)$$

These formulas provide calculation methods for root mean square error, mean absolute error and correlation coefficient, and are used to evaluate the performance of

the GRU recurrent neural network prediction model based on the Seagull optimization algorithm in time series prediction tasks.

3.3 Result analysis

The GRU neural network prediction model, optimized through the Seagull optimization algorithm, is meticulously configured for enhanced performance. The Seagull optimization algorithm plays a crucial role in fine-tuning parameters, particularly determining the number of hidden units in the model's hidden layer. With 128 neurons shaping the complexity and learning abilities of the model, this choice is pivotal. To ensure effective learning, a judicious learning rate of 0.001 is set, striking the right balance for appropriate parameter updates during training—neither too rapid nor too slow. The training process unfolds across 50 iterations, a parameter subject to practical adjustments based on the model's convergence behavior. Each training cycle involves the input of 32 samples, optimizing efficiency.

Critical parameter adjustments contribute to refining the model's convergence and generalization capabilities. Experimental validation leverages a dataset provided by Wang Min et al., featuring 24 diverse biochar composite materials and comprehensive data on the adsorption of four distinct heavy metals (lead, cadmium, chromium, and mercury). The dataset, comprising 96 samples, encompasses crucial information such as biochar preparation parameters, physical and chemical characteristics of heavy metals, and parameters elucidating adsorption performance.

This paper undertakes a comparative analysis between the Seagull optimization algorithm-enhanced GRU neural network prediction model (SOA-GRU) and a standard benchmark GRU neural network prediction model. The outcomes, including root mean square error and average absolute error values, are detailed in Table 1 for reference.

Table 1. Model evaluation index value

evaluation standard	GRU	SOA-GRU
RMSE/%	12.020	9.230
MAE/%	8.301	6.230

According to the experimental results, the RMSE of the baseline GRU model is 12.02%, while the RMSE of the improved GRU model is 9.23%. This shows that the improved model has a smaller average prediction error on the data set relative to the baseline model. And the MAE of the baseline GRU model is 8.301%, while the MAE of the improved GRU model is 6.23%. This means that the improved model performs better in terms of mean absolute error.

By comparing RMSE and MAE, we conclude that the improved GRU model is significantly better than the baseline GRU model on the same data set. This means that the improved model more accurately captures the distribution of real data and has better generalization performance. Our improvement is mainly reflected in the use of the Seagull optimization algorithm for parameter optimization, which helps the model better learn the complex patterns of the data.

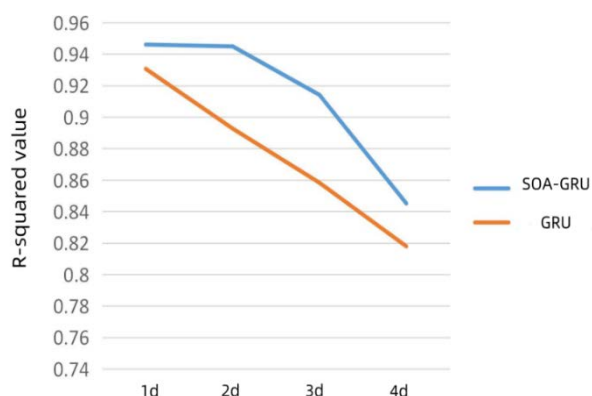


Fig. 3. Comparison chart of R² values of two models

R² is used to measure the fit of the model. When the time steps are at the same moment, the R² of the SOA-GRU model is greater than that of the GRU model. When the prediction time steps are one day, two days, three days and four days, its R² is higher than that of the GRU model. 1.66%, 5.85%, 6.51%, 3.33%. The R² of the SOA-GRU model is above 0.94 when the prediction time is one day, and its R² values are above 0.90 when the prediction time is three days; compared with the R² of the GRU model, only when the time step is one day, its R² value is greater than 0.90, which is 0.93059. In addition, comparing the decline in fitting degree of the two models as the prediction time increases, it can be found that the R² of the SOA-GRU model decreases even more, ranging from 0.94605 to 0.84508, a decrease of 10.67%; The R² of the GRU model decreased by 12.12%, and its range was 0.93059-0.81783. Plotting the R² of the two models into a graph, as shown in Figure 3, it can be found that the R² value of the SOA-GRU model shows an approximately linear downward trend as time increases.

By comparing the evaluation index data of the SOA-GRU model and the GRU model for the same time step, it can be found that the RMSE, MAPE and R² of the SOA-GRU model are better than the GRU model, proving that it has better prediction effects and can achieve biological prediction. Carbon for the purpose of more accurate performance predictions.

4 Summary

In this paper, by adopting the improved GRU neural network model driven by deep learning and using the Seagull optimization algorithm for parameter optimization, we have made significant progress in the study of the purification mechanism of heavy metal pollution by biochar composite materials. Experiments have proven that the improved model shows higher prediction accuracy on the data set, reflecting the effectiveness of the Seagull optimization algorithm in improving model performance. This research lays the foundation for the application of deep learning in the field of biochar composite materials, provides a more accurate prediction tool for future environmental governance and resource management, and has extensive scientific research and practical application value.

References

1. Sim, Sunghyun, Dohee Kim, and Hyerim Bae. "Correlation recurrent units: A novel neural architecture for improving the predictive performance of time-series data." *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023).
2. Choi, Ji Seok, Chang June Lee, and Jung Keun Lee. "A parallel recurrent neural network for robust inertial and magnetic sensor-based 3D orientation estimation." *IEEE Access* (2023).
3. El Abassi, Rida, et al. "Exploring crude oil price movements as a complex time series using recurrent neural networks: Complex time series using recurrent neural networks." *Proceedings of the 6th International Conference on Networking, Intelligent Systems & Security*. 2023.
4. Xu, Jian-You, et al. "Demand Prediction of Shared Bicycles Based on Graph Convolutional Network-Gated Recurrent Unit-Attention Mechanism." *Mathematics* 11.24 (2023): 4994.
5. Kumar, Kaushal. "Reservoir Computing in Epidemiological Forecasting: Predicting Chicken Pox Incidence." *medRxiv* (2023): 2023-04.
6. Eleftheriadis, Panagiotis, et al. "Hyperband optimization of stacked bidirectional long short-term memory neural network for the state of charge estimation." *2023 IEEE International Conference on Environment and Electrical Engineering and 2023 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE, 2023.
7. Kalahasty, Rohan. "Neuro-Inspired Plasticity for Biologically Realistic Self-Adaptation of Neural Network Weights." *2023 IEEE International Conference on Development and Learning (ICDL)*. IEEE, 2023.
8. Ameer, Iqra, et al. "Emotion Classification in Texts over Graph Neural Networks: Semantic Representation is better than Syntactic." *IEEE Access* (2023).
9. Gehrig, Mathias, and Davide Scaramuzza. "Recurrent vision transformers for object detection with event cameras." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.
10. Alqahtani, Fehaid, et al. "A hybrid deep learning model for rainfall in the wetlands of southern Iraq." *Modeling Earth Systems and Environment* (2023): 1-18.
11. Yu, Weihao, et al. "Routing hypergraph convolutional recurrent network for network traffic prediction." *Applied Intelligence* 53.12 (2023): 16126-16137.
12. Wangmin, Yaser A., et al. "Riverside Landslide Susceptibility Overview: Leveraging Artificial Neural Networks and Machine Learning in Accordance with the United Nations (UN) Sustainable Development Goals." *Water* 15.15 (2023): 2707.