

PAPER

Enhanced Water Quality Prediction in the Yellow River Basin: The Application of the HHO-LSTM Model

Minning Wu^{1,2}(✉), Eric B. Blancaflor¹, Fei Ren¹, Yong Wang¹, Ting Dong¹

¹School of Information Technology, Mapua University, Manila, Philippines

²School of Information Engineering, Yulin University, Yulin, China

pangmao716@163.com

ABSTRACT

In the pivotal water resource region of the Yellow River Basin in China, precise prediction of water resources is essential for their effective and rational management. This study introduces a novel approach to water resource prediction by employing the Harris Hawks Optimization-Long Short-Term Memory (HHO-LSTM) model. This method overcomes the constraints faced by traditional techniques in processing time series data and various variable factors. It encompasses a comprehensive description of the multi-source hydrological data collection process within the Yellow River Basin, followed by meticulous data preprocessing. The data set for this study includes estimates of four critical water quality parameters, and the efficacy of the model is gauged through the mean squared error (MSE) and root mean squared error (RMSE) metrics. This facilitates the projection of future water quality trends in specific areas by leveraging historical water quality data. The HHO-LSTM model has demonstrated outstanding accuracy and robustness in predicting water quality across diverse temporal scales and water resource variables, marking a significant advancement in water resource management within the Yellow River Basin. This approach not only enhances current management strategies but also contributes valuable insights for ongoing water resource research and decision-making processes.

KEYWORDS

water quality, prediction model, Harris Hawks optimization, long short-term memory, Yellow River Basin

1 INTRODUCTION

The Yellow River Basin, a vital water resource region in China, plays a pivotal role in supporting the livelihoods and economic activities of its populace [1–5]. Recent rapid economic growth and urbanization in China have intensified water pollution in the Yellow River, posing significant risks to the ecological environment and the sustainability of the basin's water resources, thereby impacting resident lives [6, 7]. This situation highlights the critical need for precise water quality prediction in the

Wu, M., Blancaflor, E.B., Ren, F., Wang, Y., Dong, T. (2024). Enhanced Water Quality Prediction in the Yellow River Basin: The Application of the HHO-LSTM Model. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(5), pp. 4–14. <https://doi.org/10.3991/ijoe.v20i05.48225>

Article submitted 2023-12-07. Revision uploaded 2024-01-25. Final acceptance 2024-01-28.

© 2024 by the authors of this article. Published under CC-BY.

Yellow River Basin as a cornerstone for the development of effective water resource management strategies and the ability to meet future challenges.

Historically, water quality prediction methodologies largely depended on statistical models, including the autoregressive integrated moving average model (ARIMA), regression analysis, grey systems, and least squares support vector machines. These approaches are known for their strong interpretability and high computational efficiency, yet they demonstrate inherent limitations in processing complex non-linear and temporal relationships [8–14]. Additionally, the effectiveness of traditional methods is often diminished due to the geographical complexity and climatic variability characteristic of the Yellow River Basin [15]. With the evolution of artificial intelligence, machine learning techniques have increasingly been recognized for their potential in water quality prediction. This includes the use of neural networks, decision trees, random forests, and support vector machines [16]. These advanced methods are aptly suited for intricate non-linear systems and are applicable for medium- to long-term water resource predictions. However, they require substantial training data and are susceptible to overfitting in instances of high model complexity [17, 18].

The inherent geographical complexity and climatic variability of the Yellow River Basin present substantial challenges to the efficacy of traditional methods in forecasting future water resource variations. This scenario has escalated the emphasis on and research into neural network technology for water quality prediction, as these technologies demonstrate a heightened capacity for processing complex data sets [19]. Neural network models, particularly when fed with historical water quality data, have been recognized for their potential in accurately predicting water quality scenarios [20]. However, conventional neural network models frequently encounter issues such as gradients vanishing or exploding, particularly when dealing with extended data sequences. In this context, LSTM networks, equipped with memory units, are adept at capturing long-term dependencies. LSTM has a wide range of applications in predicting water resources by analyzing and discerning potential long-term trends and cyclic variations in water quality data [21, 22].

Despite LSTM's achievements across various domains, its limitations become pronounced when addressing complex hydrological systems, such as those exemplified by the Yellow River Basin. To mitigate these limitations, the present study integrates HHO with LSTM, formulating the innovative HHO-LSTM water resource prediction model. We use the HHO algorithm to globally search and optimize LSTM parameters. This is crucial for finding the best parameter settings and improving the effectiveness of the water quality prediction model [23]. The HHO-LSTM model combines the benefits of managing temporal and non-linear relationships with global optimization, which improves both the accuracy and generalizability of water quality predictions. The model's versatility and adaptability further empower it to deliver more precise forecasts of water resource fluctuations in the Yellow River Basin.

2 WATER QUALITY PREDICTION MODEL UTILIZING HHO-LSTM

2.1 LSTM model

The LSTM model, a variant of recurrent neural networks, is specifically engineered to overcome the challenge of long-term dependencies inherent in traditional recurrent neural networks (RNNs). This model is structured with four distinct neural network layers, each functioning interactively in a complex manner rather than as isolated, simplistic layers. LSTM processes information from preceding moments, involving two key types: the cell state and the hidden layer state. The model employs

a triad of gate mechanisms, namely, a forget gate, an input gate, and an output gate. These gates are instrumental in regulating both the transmission and updating processes of the cell state and hidden layer state information [24]. Figure 1 delineates the LSTM hidden layer structure, where C_{t-1} and C_t signify the cell state information at times $t - 1$ and t , respectively, and \tilde{C}_t is the candidate update information at time t . Similarly, h_{t-1} and h_t represent the hidden layer state information at times $t - 1$ and t , respectively, with X indicating the input value at time t . The sigmoid function σ is applied here, with f_t , i_t , and o_t functioning as the control coefficients for the respective gates.

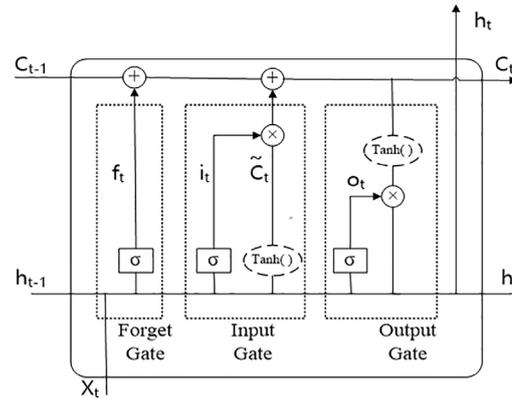


Fig. 1. Hidden layer structure of LSTM

The role of the forget gate is primarily to ascertain the degree to which cell state information C_{t-1} from time $t - 1$ is retained, which is contingent upon the value of f_t . This value, ranging between 0 and 1, is computed using the inputs X_t and h_{t-1} . A value of f_t closer to 0 implies a greater exclusion of information from C_{t-1} , whereas a value nearing 1 suggests a higher retention of information [25]. The operations of control coefficients i_t and o_t are analogous to that of f_t . The input gate is tasked with determining the information to be incorporated into C_t , while the output gate governs the release of hidden layer state information h at time t . The mathematical representation of LSTM involves parameters such as weights (W_f, W_i, W_c, W_o) and biases (b_f, b_i, b_c, b_o).

$$f_t = \sigma(W_f [h_{t-1}, X_t], b_f) \tag{1}$$

$$i_t = \sigma(W_i [h_{t-1}, X_t], b_i) \tag{2}$$

$$\tilde{C}_t = \tanh \sigma(W_c, [h_{t-1}, X_t], b_c) \tag{3}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

$$o_t = \sigma(W_o, [h_{t-1}, X_t], b_o) \tag{5}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{6}$$

2.2 Principle of the HHO algorithm

Researchers designed the HHO algorithm as a metaheuristic approach for optimizing function values, inspired by the distinctive group hunting behavior of Harris's Hawks [26]. It stands out for its operational simplicity, minimal need for parameter

adjustment, and robust convergence capabilities [27]. The HHO algorithm operates through three sequential stages.

In the first stage, referred to as the global search phase, a notable dispersion is observed among the Harris Hawks within their group. During this phase, individual hawks engage in prolonged periods of waiting and observation, employing dual strategies for prey detection [28].

$$X(t+1) = \begin{cases} X_{\text{rand}}(t) - r_1 |X_{\text{rand}}(t) - 2r_2 X(t)| & q \geq 0.5 \\ [X_r(t) - X_m(t)] - r_3 [l_b + r_4 (u_b - l_b)] & q < 0.5 \end{cases} \quad (7)$$

$$X_m(t) = \frac{1}{N} \sum_{i=1}^N X_i(t) \quad (8)$$

The computational formula for this stage incorporates a randomly selected individual X_{rand} from the group, the optimal individual X_r , and the mean position X_m of the population, utilizing random values q, r_1, r_2, r_3, r_4 within the $[0, 1]$ range. The parameters u_b and l_b represent the upper and lower bounds of the population size, denoted by N .

The second stage marks the transition from global search to localized exploitation. This phase is governed by the formula (9), where E_0 , a random number within the range of $[-1, 1]$, signifies the prey's escape energy. The variable t represents the current iteration, while T indicates the maximum number of iterations.

$$E = 2E_0 (1 - t/T) \quad (9)$$

The third stage, known as the local exploitation phase, sees the Harris Hawks adopt four distinct attack strategies. These strategies are formulated based on the prey's escape maneuvers and the Hawks' pursuit tactics. They encompass soft besiege, hard besiege, progressive rapid dive soft besiege, and progressive rapid dive hard besiege [29].

2.3 Development of the HHO-LSTM water quality prediction model

The HHO-LSTM model represents a sophisticated enhancement and optimization of the conventional LSTM model. We use the HHO algorithm, inspired by biomimetics and emulating the hunting behavior of Harris's Hawks, to optimize parameters within the LSTM framework [30]. The application of the HHO algorithm in the HHO-LSTM model notably augments the training speed and overall performance of the LSTM model.

Comprising three integral components—the HHO algorithm, the construction of the LSTM network, and the amalgamation of these elements—the HHO-LSTM model for water quality prediction adheres to a structured framework. Figure 2 delineates this framework, outlining the subsequent steps:

Step 1: Data acquisition stage. In this initial stage, datasets pertinent to water quality prediction are amassed. These datasets typically include historical data related to water quality monitoring, along with additional data concerning variables that may influence water quality.

Step 2: Data preprocessing stage. Subsequently, the acquired data undergoes a comprehensive preprocessing phase. This phase encompasses data cleaning, addressing missing and outlier values, and executing feature engineering. The primary objective of this stage is to refine the data to conform to the input requirements of the model and to optimize its predictive accuracy.

Step 3: Feature engineering stage. The final stage involves the extraction and selection of relevant features for water quality prediction. This process is meticulously tailored to align with the specific requirements of the task at hand and the unique characteristics of the data.

Step 4: Construction of the HHO-LSTM model.

- a) Initialization of HHO algorithm parameters, including population size, number of iterations, and exploration rate, is systematically conducted.
- b) The structure of the LSTM neural network is established, encompassing the specification of input, hidden, and output layers. Key LSTM hyperparameters, such as the number of neurons in the hidden layers, learning rate, iteration count, batch size, and the number of LSTM layers, are precisely defined.
- c) Integration of the model is achieved by amalgamating the HHO-optimized LSTM with the specific requirements of water quality prediction.
- d) Data segmentation into training and test sets is carried out, typically utilizing a time-series split approach for this purpose.
- e) The training phase involves using the training set to iteratively refine the model parameters.
- f) Monitoring of the training process is conducted, with particular focus on the changes in the loss function, ensuring the model’s gradual convergence on the training data.
- g) Evaluation of the model’s effectiveness is performed on the test set, providing insights into its performance on previously unseen data.
- h) Quantitative analysis of performance metrics, including RMSE and mean absolute error (MAE), is undertaken.
- i) Comparative visualization analysis is executed, juxtaposing predicted results against actual observations to elucidate the model’s predictive capabilities.
- j) Optimization of model parameters is performed based on the outcomes of these assessments to enhance overall performance.
- k) The application of techniques such as cross-validation is employed to thoroughly assess the model’s generalizability.
- l) Finally, the fully trained HHO-LSTM model is applied to real-world water quality prediction tasks, with subsequent generation and analysis of the prediction results.

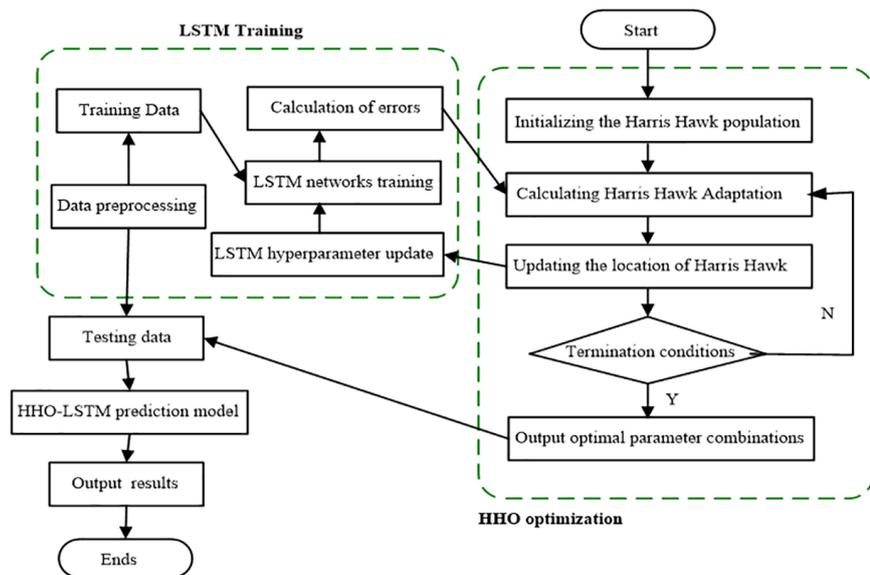


Fig. 2. HHO-LSTM water quality prediction modeling

3 PERFORMANCE ANALYSIS OF PREDICTIONS

3.1 Data source and processing

For this investigation, real-time data obtained from the national surface water quality automatic monitoring system of the China Environmental Monitoring Station was utilized as the primary data source. The methodology employed for data processing comprised the following steps:

Step 1: Acquisition of data. Data for daily water quality monitoring, specifically from the Yuxi River Monitoring Station in Yulin, Shaanxi, within the Yellow River Basin, were collected for the period between June 18, 2021, and December 31, 2023. This dataset, which was pivotal for model testing, included various indicators such as water monitoring data, pH, temperature, and dissolved oxygen levels. Table 1 presents an exhaustive analysis of this dataset.

Table 1. Analysis of Yuxi River water quality dataset

	Temperature	pH	Dissolved Oxygen
Count	426.000000	426.000000	426.000000
Mean	11.456338	8.333756	8.997653
Std	8.003415	0.150905	1.966463
Min	0.600000	7.810000	4.290000
25%	3.200000	8.250000	7.255000
50%	11.050000	8.350000	9.280000
75%	19.000000	8.460000	10.762500
Max	30.800000	8.720000	12.010000

Step 2: Selection of research focus. Upon comparison with the Surface Water Environmental Quality Standards, it was determined that the dissolved oxygen content significantly influenced the water quality classification of the Yuxi River. Consequently, the dissolved oxygen content was identified as the primary parameter for assessing the water quality of the Yuxi River. The dissolved oxygen content of the water quality of the Yuxi River is related to the season, temperature, and climatic factors, and the general trend is that dissolved oxygen is low when the temperature is high and high when the temperature is low, and the change of its dissolved oxygen in the last three years is shown in Figure 3.

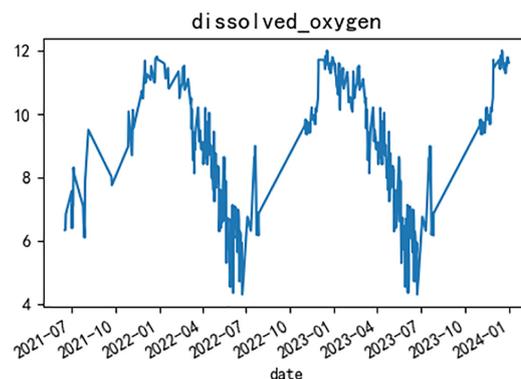


Fig. 3. Changes in dissolved oxygen content in Yuxi River

Step 3: Normalization of data. Normalization of the dissolved oxygen data from the Yuxi River was performed. This step was crucial in mitigating the impact of outliers on model convergence and facilitating both the speed and stability of the HHO-LSTM model's convergence. The normalization adjusted the input values for the HHO-LSTM model to range between $[-1, 1]$. The employed normalization formula, depicted as Equation (10), is formulated as follows:

$$X_n = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (10)$$

where, X denotes the original data, X_n the normalized data, X_{\max} the maximum value within the original data, and X_{\min} the minimum value of the original data.

Step 4: Segmentation of data into training and validation sets. The division of the sample data into training and validation sets was undertaken at a ratio of 6:4. Explicitly, 60% of the sample data was allocated for the purpose of training the model. The remaining 40% of the data was utilized for validation, serving to assess the model's performance.

3.2 Experimental implementation of predictive models

Predictive experiments were conducted using both the LSTM and HHO-LSTM models, which were implemented in Python, to analyze the water quality data of the Yuxi River. Training of these models on the dataset was performed using the Adam optimizer, with a specified learning rate of 0.001 and an epoch setting of 100. This process culminated in the derivation of the final experimental outcomes.

3.3 Fit results of the training set

The performance of both HHO-LSTM and LSTM models was quantitatively assessed using MAE and RMSE as evaluation metrics, with the results detailed in Table 2. It was observed that both models demonstrated relatively low error rates in fitting the training set. Notably, the HHO-LSTM model exhibited superior predictive performance compared to the LSTM model, reflected in a reduction of RMSE by 0.002 and MAE by 0.001.

Table 2. Evaluation of model loss values

Model	RMSE	MAE	Evaluation Value
LSTM	0.183	0.033	0.881621
HHO-LSTM	0.182	0.031	0.882438

3.4 Prediction results of the validation set

In order to elucidate the disparities in the predictive capabilities of the HHO-LSTM and LSTM models for water quality, a comparative analysis was conducted.

Figure 4 presents this comparison, showcasing the alignment of the predictions from both the HHO-LSTM and LSTM models with the actual water quality data. It was observed that both models proficiently forecasted the periodic fluctuations in water quality. However, a higher degree of accuracy in predictions was noted in the case of the HHO-LSTM model, as evidenced by its prediction curve more closely mirroring the actual data trends.

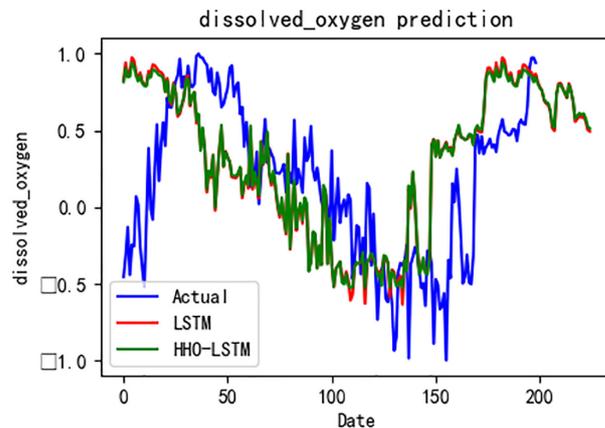


Fig. 4. Comparative analysis of HHO-LSTM and LSTM water quality prediction models

4 CONCLUSION

Combining the HHO algorithm with the LSTM model to create the HHO-LSTM water quality prediction model significantly improves how we optimize the LSTM model's parameters. This advancement has effectively addressed the previously identified limitations in accurately predicting water quality, particularly in terms of precision. The results show that the HHO-LSTM model is very accurate in predicting the dissolved oxygen content of the Yuxi River. It's better at fitting function curves than the traditional LSTM model. Also, compared to the LSTM model, the HHO-LSTM model performs better, with lower MAE and RMSE. In practical scenarios, factors such as outliers, missing values, or inconsistencies in data quality may influence the model's efficacy. Future research endeavors might focus on the development of more sophisticated data cleaning and preprocessing techniques aimed at enhancing data quality and consistency. Furthermore, addressing the constraints associated with small sample datasets is essential. In this context, the exploration of data augmentation methods to broaden the dataset and bolster the model's generalizability is suggested as a viable research direction.

5 ACKNOWLEDGEMENT

This work is partially supported by the Natural Science Basic Research Plan in Shaanxi Province of China (Grant No. 2024NC-YBXM-223), Yulin Key Laboratory of Big Data and Intelligent Decision Making, the Youth Innovation Team of Shaanxi Universities, the National Science Foundation of China Funding Project for Department of Education of Shaanxi Province of China (Grant No. 22JC063), Natural Science and Technology Project Plan in Yulin of China (Grant No. CXY-2022-94, CXY-2022-92, CXY-2022-93), and Thanks for the help.

6 REFERENCES

- [1] Z. Zhao, H. Li, X. Song, and W. Sun, "Dynamic monitoring of surface water bodies and their influencing factors in the Yellow River Basin," *Remote Sensing*, vol. 15, no. 21, p. 5157, 2023. <https://doi.org/10.3390/rs15215157>
- [2] Y. Sun, X. Chen, Y. Luo, D. Cao, H. Feng, X. Zhang, and R. Yao, "Agricultural water quality assessment and application in the Yellow River delta," *Agronomy*, vol. 13, p. 1495, 2023. <https://doi.org/10.3390/agronomy13061495>
- [3] K. Zhou, "Integrated allocation model of water quantity and quality and its application in the Yellow River," *Water Supply*, vol. 20, no. 8, pp. 3768–3778, 2020. <https://doi.org/10.2166/ws.2020.126>
- [4] P. Li, Y. L. Ran, X. M. Li, Z. Han, Y. Zhang, L. Y. Zhang, Z. J. Du, and X. M. Wu, "Thermal equilibrium analysis of small-to-medium river ecosystems in northern China under multi-factor coupling and decision-making for ecological restoration," *International Journal of Heat and Technology*, vol. 41, no. 5, pp. 1291–1300, 2023. <https://doi.org/10.18280/ijht.410519>
- [5] Y. Jin, X. Y. Wang, and Y. P. Dong, "Variation of water quality in Ningxia section of the Yellow River in recent 5 years," *Journal of Chemistry*, vol. 2022, p. 7704513, 2022. <https://doi.org/10.1155/2022/7704513>
- [6] Y. Xu and C. S. Wang, "Ecological protection and high-quality development in the Yellow River Basin: Framework, path, and countermeasure," *Bulletin of Chinese Academy of Sciences (Chinese Version)*, vol. 35, no. 7, pp. 875–883, 2020. <https://doi.org/10.16418/j.issn.1000-3045.20200425002>
- [7] Y. P. Chen, B. J. Fu, Y. Zhao, K. B. Wang, M. M. Zhao, J. F. Ma, J. H. Wu, C. Xu, W.G. Liu, and H. Wang, "Sustainable development in the Yellow River Basin: Issues and strategies," *Journal of Cleaner Production*, vol. 263, p. 121223, 2020. <https://doi.org/10.1016/j.jclepro.2020.121223>
- [8] A. Katimon, S. Shahid, and M. Mohsenipour, "Modeling water quality and hydrological variables using ARIMA: A case study of Johor River, Malaysia," *Sustainable Water Resources Management*, vol. 4, pp. 991–998, 2018. <https://doi.org/10.1007/s40899-017-0202-8>
- [9] X. Yang, Q. Liu, X. Luo, and Z. Zheng, "Spatial regression and prediction of water quality in a watershed with complex pollution sources," *Scientific Reports*, vol. 7, no. 1, p. 8318, 2017. <https://doi.org/10.1038/s41598-017-08254-w>
- [10] G. J. Hasham and M. M. Ramal, "Water quality assessment of Euphrates river within Fallujah city using water quality indices technique," *International Journal of Design & Nature and Ecodynamics*, vol. 17, no. 4, pp. 563–570, 2022. <https://doi.org/10.18280/ijdne.170410>
- [11] V. Obradovic and A. Vulevic, "Water resources protection and water management framework in western balkan countries in drina River Basin," *Acadlore Transactions on Geosciences*, vol. 2, no. 1, pp. 24–32, 2023. <https://doi.org/10.56578/atg020103>
- [12] A. S. Al Chalabi, S. M. Naeem, and R. J. Alkhafaji, "Water Quality Index (WQI) for main water treatment plants in Basra City, Iraq," *International Journal of Design & Nature and Ecodynamics*, vol. 18, no. 6, pp. 1407–1416, 2023. <https://doi.org/10.18280/ijdne.180614>
- [13] M. H. Mohsen and F. M. Al-Mohammed, "Assessment of irrigation water quality using the Canadian Water Quality Index (CWQI) in the Hilla Main Canal, Iraq," *Instrumentation Measure Métrologie*, vol. 22, no. 3, pp. 105–111, 2023. <https://doi.org/10.18280/i2m.220303>

- [14] B. T. Akinyemi, O. D. Ogundele, and A. B. Afolabi, "Advancements in sustainable membrane technologies for enhanced remediation and wastewater treatment: A comprehensive review," *Acadlore Transactions on Geosciences*, vol. 2, no. 4, pp. 196–207, 2023. <https://doi.org/10.56578/atg020402>
- [15] L. Jiang, Q. Zuo, J. Ma, and Z. Zhang, "Evaluation and prediction of the level of high-quality development: A case study of the Yellow River Basin, China," *Ecological Indicators*, vol. 129, p. 107994, 2021. <https://doi.org/10.1016/j.ecolind.2021.107994>
- [16] U. Ewuzie, O. P. Bolade, and A. O. Egbedina, "Application of deep learning and machine learning methods in water quality modeling and prediction: A review," *Current Trends and Advances in Computer-Aided Intelligent Environmental Data Engineering*, pp. 185–218, 2022. <https://doi.org/10.1016/B978-0-323-85597-6.00020-3>
- [17] M. Drogkoula, K. Kokkinos, and N. Samaras, "A comprehensive survey of machine learning methodologies with emphasis in water resources management," *Applied Sciences*, vol. 13, no. 22, p. 12147, 2023. <https://doi.org/10.3390/app132212147>
- [18] R. Huang, C. Ma, J. Ma, X. Huangfu, and Q. He, "Machine learning in natural and engineered water systems," *Water Research*, vol. 205, p. 117666, 2021. <https://doi.org/10.1016/j.watres.2021.117666>
- [19] Y. Chen, L. Song, Y. Liu, L. Yang, and D. Li, "A review of the artificial neural network models for water quality prediction," *Applied Sciences*, vol. 10, no. 17, p. 5776, 2020. <https://doi.org/10.3390/app10175776>
- [20] A. N. Ahmed, F. B. Othman, H. A. Afan, R. K. Ibrahim, C. M. Fai, M. S. Hossain, M. Ehteram, and A. Elshafie, "Machine learning methods for better water quality prediction," *Journal of Hydrology*, vol. 578, p. 124084, 2019. <https://doi.org/10.1016/j.jhydrol.2019.124084>
- [21] M. Bulut, "Hydroelectric generation forecasting with Long Short Term Memory (LSTM) based deep learning model for Turkey," *arXiv*, preprint, 2021. <https://doi.org/10.48550/arXiv.2109.09013>
- [22] S. Khullar and N. Singh, "Water quality assessment of a river using deep learning Bi-LSTM methodology: Forecasting and validation," *Environmental Science and Pollution Research*, vol. 29, no. 9, pp. 12875–12889, 2022. <https://doi.org/10.1007/s11356-021-13875-w>
- [23] M. A. Mossa, O. M. Kamel, H. M. Sultan, and A. A. Z. Diab, "Parameter estimation of PEMFC model based on Harris Hawks' optimization and atom search optimization algorithms," *Neural Computing and Applications*, vol. 33, pp. 5555–5570, 2021. <https://doi.org/10.1007/s00521-020-05333-4>
- [24] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN-LSTM model for gold price time-series forecasting," *Neural Computing and Applications*, vol. 32, pp. 17351–17360, 2020. <https://doi.org/10.1007/s00521-020-04867-x>
- [25] C. Wu, F. Wu, Y. Chen, S. Wu, Z. Yuan, and Y. Huang, "Neural metaphor detecting with CNN-LSTM model," in *Proceedings of the Workshop on Figurative Language Processing*, New Orleans, Louisiana, 2018, pp. 110–114. <https://doi.org/10.18653/v1/W18-0913>
- [26] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019. <https://doi.org/10.1016/j.future.2019.02.028>
- [27] B. S. Yıldız and A. R. Yıldız, "The Harris hawks optimization algorithm, salp swarm algorithm, grasshopper optimization algorithm and dragonfly algorithm for structural design optimization of vehicle components," *Materials Testing*, vol. 61, no. 8, pp. 744–748, 2019. <https://doi.org/10.3139/120.111379>
- [28] R. Roy, V. Mukherjee, and R. P. Singh, "Harris hawks optimization algorithm for model order reduction of interconnected wind turbines," *ISA Transactions*, vol. 128, pp. 372–385, 2022. <https://doi.org/10.1016/j.isatra.2021.09.019>

- [29] S. Jiao, G. Chong, C. Huang, H. Hu, M. Wang, A. A. Heidari, H. L. Chen, and X. H. Zhao, "Orthogonally adapted Harris hawks optimization for parameter estimation of photovoltaic models," *Energy*, vol. 203, p. 117804, 2020. <https://doi.org/10.1016/j.energy.2020.117804>
- [30] D. Dhawale, V. K. Kamboj, and P. Anand, "An improved Chaotic Harris Hawks Optimizer for solving numerical and engineering optimization problems," *Engineering with Computers*, vol. 39, no. 2, pp. 1183–1228, 2023. <https://doi.org/10.1007/s00366-021-01487-4>

7 AUTHORS

Minning Wu, School of Information Technology, Mapua University, Manila 1002, Philippines; School of Information Engineering, Yulin University, Yulin 719000, China (E-mail: pangmao716@163.com).

Eric B. Blancaflor, School of Information Technology, Mapua University, Manila 1002, Philippines.

Fei Ren, School of Information Technology, Mapua University, Manila 1002, Philippines.

Yong Wang, School of Information Technology, Mapua University, Manila 1002, Philippines.

Ting Dong, School of Information Technology, Mapua University, Manila 1002, Philippines.