2025 9th International Workshop on Materials Engineering and Computer Sciences (IWMECS 2025)

Research on Artificial Intelligence Driven Environmental Pollution Traceability and Dynamic Risk Prediction Method

Moxuan Wang

Hefei Wonwit Senior Hight School, Hefei, 230601, Anhui, China

wmx17555816927@outlook.com

Keywords: Artificial intelligence; Traceability of environmental pollution; Dynamic risk prediction

Abstract: This article focuses on AI (Artificial intelligence) in the field of environmental pollution traceability and dynamic risk prediction. At present, the situation of environmental pollution is severe, and traditional methods can not meet the needs of accurate traceability and prediction, so AI technology is introduced. By collecting the data of pollutant concentration, meteorological and geographical information of several monitoring stations in a typical industrial area for three consecutive years, an algorithm model of environmental pollution traceability based on CNN (Convolutional neural network) and a dynamic risk prediction algorithm model based on LSTM (Long-term and short-term memory network) are constructed. The experimental results show that, in terms of environmental pollution traceability model, the model has a traceability accuracy rate of 85%, a recall rate of 82% and a F1 value of 83.4%, and an accuracy rate of 82%, a recall rate of 80% and a F1 value of 80.9% for organic pollutants. It is not difficult to find that the model not only performs well in environmental pollution traceability and dynamic risk prediction, but also brings help in environmental protection decision-making. However, in the face of sudden environmental incidents, the model performance still has room for improvement.

1. Introduction

In today's global scope, the acceleration of industrialization and urbanization makes the environmental pollution problem more and more serious [1]. Environmental pollution not only causes serious damage to the ecosystem and threatens the survival of many organisms, but also poses a great challenge to human health and the sustainable development of social economy [2]. In the face of increasingly complex pollution sources, traditional environmental monitoring methods often lack accurate pollution source tracking ability, which leads to unsatisfactory treatment effect. Therefore, the introduction of pollution source traceability monitoring technology has become an important means to solve this problem [3]. Tracing environmental pollution sources aims to accurately explore the sources of pollutants and clarify the initial location and causes of pollution, which is very important for fundamentally solving pollution problems [4]. Through the collection, analysis and synthesis of pollution source data, pollution source traceability monitoring technology can realize accurate positioning and tracking of pollution sources and ensure timely identification and treatment of pollution sources [5]. However, in the face of complex and changeable environmental systems, traditional tracking and forecasting methods often expose many limitations and are difficult to meet the actual needs.

By monitoring the emission of pollutants in the atmosphere, environmental detection technology can help the government accurately understand the pollution situation in various regions, as well as the types, emissions and distribution of pollution sources [6]. Introducing AI into the field of environmental pollution traceability and dynamic risk prediction provides a brand-new idea to solve the above problems [7]. AI can process a large amount of environmental data and mine the hidden rules behind the data [8]. This paper focuses on the traceability of environmental pollution and dynamic risk prediction method driven by AI. The purpose of this study is to construct an efficient and accurate environmental pollution traceability and dynamic risk prediction algorithm model by using AI technology. At the same time, it provides strong technical support for environmental protection decision-making through experimental verification.

2. Traceability of environmental pollution and dynamic risk prediction

Through strict environmental monitoring, pollution sources can be found in time, pollution levels can be monitored, and ecological environment conditions can be evaluated, providing strong data support and scientific basis for government decision-making. In the theory of environmental pollution, the theory of pollutant migration and transformation is one of the core contents [9]. This theory describes the movement law of pollutants in air, water, soil and other environmental media and the transformation process between them. The key role of ecological environment monitoring is to identify environmental risks at an early stage and take timely measures to avoid irreversible damage caused by pollution to the ecosystem and public health.

Through advanced monitoring equipment, sensor networks and big data analysis, pollution source traceability technology can track the specific location of pollution sources and the diffusion path of pollution in real time, providing brand-new data support for ecological environment monitoring. Understanding the chemical characteristics of pollutants is helpful to judge their sources, such as pollutants with unique chemical characteristics produced in specific industrial production processes [10]. On the technical level, GIS technology provides a powerful means of spatial analysis for environmental pollution traceability and dynamic risk prediction. Traceability monitoring data of pollution sources can help environmental protection departments to fully understand the types, quantity, distribution and emission characteristics of pollution sources, making the supervision work more targeted and effective. Through satellite or aerial remote sensing, air pollution and water pollution can be monitored in real time, providing timely support for dynamic risk prediction [11]. Data mining technology plays a key role in processing massive environmental data. It can extract valuable information from complicated data, discover the potential correlation between the data, and lay the foundation for building an accurate environmental pollution traceability and dynamic risk prediction model.

3. AI-driven algorithm model of environmental pollution traceability and dynamic risk prediction

With the increasing complexity of environmental problems, the limitations of traditional methods in environmental pollution traceability and dynamic risk prediction are increasingly prominent. With its excellent data processing and pattern recognition capabilities, AI technology has opened up a new path for building a more accurate and effective algorithm model.

For the algorithm model of environmental pollution traceability, data collection is undoubtedly the key initial link. With the help of environmental monitoring stations distributed in different regions, various environmental data, such as pollutant concentration, meteorological conditions and geographical information, are widely collected. These data will become important materials for model training. On this basis, this article chooses CNN algorithm in deep learning. CNN has achieved great success in the field of image recognition, and its unique convolution layer structure can automatically extract local features from data. In the scene of environmental pollution traceability, environmental data is regarded as a special "image". Taking two-dimensional convolution as an example, assuming that the input data is X and the convolution kernel is K, the convolution operation can be expressed as:

$$Y_{ij} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} X_{i+m,j+n} K_{mn} \tag{1}$$

Where Y_{ij} is the output result after convolution, and features are extracted from the data through this convolution layer. The concentration distribution of pollutants in different regions has specific characteristic patterns in space, and CNN can keenly capture these patterns, and then analyze the possible transmission paths and sources of pollutants. The pooling layer of CNN can reduce the dimension of data. Taking maximum pooling as an example, assuming that the input data is Z, the output is P, and the pool window size is $s \times s$, there are:

$$P_{ij} = \max_{m=0}^{s-1} \max_{n=0}^{s-1} Z_{i \times s+m, j \times s+n} \tag{2}$$

Reduce the amount of data while retaining key features, and improve the operational efficiency of the model. After multi-layer convolution and pooling operation, the extracted features are input into the fully connected layer for classification, so as to judge the source of pollutants.

In the aspect of dynamic risk prediction algorithm model, RNN (Recurrent neural network) and its variant LSTM constitute the core algorithm adopted in this article. In view of the obvious time series characteristics of environmental data, the past pollution situation is of great significance for future risk prediction. RNN can process sequence data and retain historical information in time series through the feedback mechanism of hidden layer. Traditional RNN has the disadvantages of gradient disappearance or explosion when dealing with long-term dependence. LSTM solves this problem by introducing gating mechanism. The gating unit of LSTM includes an input gate i_t, a forgetting gate f_t and an output gate o_t, and their calculation formulas are as follows:

$$i_t = \sigma\left(W_{ii} x_t + W_{hi} h_{t-1} + b_i\right) \tag{3}$$

$$f_t = \sigma\left(W_{if} x_t + W_{hf} h_{t-1} + b_f\right) \tag{4}$$

$$o_t = \sigma\left(W_{io} x_t + W_{ho} h_{t-1} + b_o\right) \tag{5}$$

Where σ is the activation function, W is the weight matrix, b is the bias vector, x_t is the current input, and h_{t-1} is the hidden state at the last moment. These gating units can accurately control the inflow, outflow and retention of information.

In order to further improve the accuracy and stability of the model, it is necessary to optimize the model. Using regularization methods, such as L2 regularization, the expression is:

$$J_{regularized} = J(\theta) + \frac{\lambda}{2m} \sum_{j=1}^{n} \theta_j^2 \tag{6}$$

Where $J(\theta)$ is the original loss function, λ is the regularization parameter, m is the number of samples, and θ is the model parameter, so as to prevent the model from over-fitting. The AI algorithm is deeply integrated with the professional knowledge in the field of environmental pollution, and an efficient algorithm model suitable for environmental pollution traceability and dynamic risk prediction is constructed.

4. Experimental verification and result analysis

In this section, a series of experiments are carried out to verify the effectiveness of the AI-driven environmental pollution traceability and dynamic risk prediction algorithm model. A typical industrial area is selected as the research object in the experiment. There are many types of industrial pollution sources in this area, and the environmental data is rich and representative. The experimental data come from many environmental monitoring stations in this area, covering three years of pollutant concentration data, meteorological data and geographic information data. Divide the collected data into 70% for training model, 20% for verifying model and 10% for testing model.

For the environmental pollution traceability model, accuracy, recall and F1 value are used as evaluation indicators. During the experiment, different types of pollutants were tested for traceability. Table 1 shows the performance of the model in the traceability of different pollutants. As can be seen from the table, for heavy metal pollutants, the accuracy of the model reached 85%, the recall rate was 82%, and the F1 value was 83.4%. For organic pollutants, the accuracy rate is

82%, the recall rate is 80%, and the F1 value is 80.9%. This shows that the model has high reliability in the traceability of different types of pollutants.

Table 1: Performance Evaluation Table for Environmental Pollution Source Tracing Models

Pollutant Type	Sample Quantity	Correct Source Tracing Count	Incorrect Source Tracing Count	Accuracy Rate	Recall Rate	F1 Score	False Negative Rate	False Positive Rate
Heavy Metal Pollutants	200	170	30	85%	82%	83.4%	18%	15%
Organic Pollutants	180	148	32	82%	80%	80.9%	20%	17.8%

In the aspect of dynamic risk prediction model, in order to visually show the prediction effect of the model, a line chart is drawn to compare the predicted value with the actual value (see Figure 1). Take the environmental pollution risk index in the next week as the forecast object. In the figure, the abscissa is time (in days) and the ordinate is risk index. It can be clearly seen from the figure that the predicted value of the model is closely related to the actual value. Although there are some deviations at individual time points, the overall trend is highly consistent. From the 3rd day to the 5th day, the actual risk index showed an upward trend, and the predicted value accurately captured this change. The error between the predicted value and the actual value is quantified by calculating RMSE, and the result shows that RMSE is 0.08, which indicates that the prediction accuracy of the model is high.

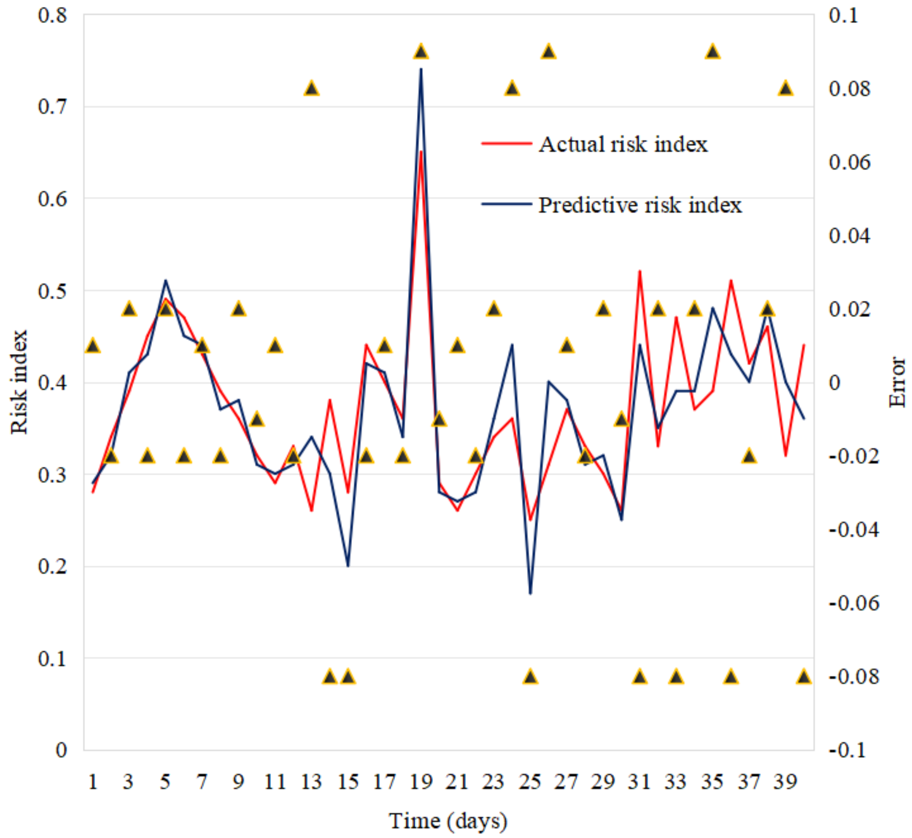

Figure 1 Comparison chart between predicted value and actual value of dynamic risk prediction model

The results show that the proposed AI-driven environmental pollution traceability and dynamic risk prediction algorithm model shows good performance in practical application. The environmental pollution traceability model can accurately identify the source of pollutants and provide a powerful direction for pollution control. Dynamic risk prediction model can accurately predict the changing trend of environmental pollution risk, which is helpful to formulate preventive measures in advance and reduce environmental risks. However, there are still some aspects that can be optimized in the model, such as the timeliness and accuracy of prediction in the face of sudden environmental events. In the future, we can consider introducing more real-time monitoring data and improving the model structure to improve the overall performance of the model.

5. Conclusions

The purpose of this study is to solve the problems of environmental pollution traceability and dynamic risk prediction, and the corresponding algorithm model is constructed with the help of AI technology and verified by experiments. In the aspect of environmental pollution traceability, the model based on CNN shows a high recognition ability for different types of pollutants. The accuracy of heavy metal pollution traceability is 85%, and the F1 value is 83.4%. The accuracy of organic pollutants is 82%, and the F1 value is 80.9%, which means that the model can accurately trace the pollution source and provide a direction for fundamentally solving the pollution problem. In the aspect of dynamic risk prediction, the predicted value of the model based on LSTM is highly consistent with the actual value, and RMSE is only 0.08. It can capture the changing trend of environmental pollution risk, help to plan prevention strategies in advance and reduce environmental risks.

However, the study also found that the model has some limitations. In the face of sudden environmental events, the timeliness and accuracy of prediction need to be improved. This is due to the abnormality and complexity of data in unexpected situations, and the existing model structure and data processing methods are difficult to adapt quickly. In the future, we can try to introduce more real-time monitoring data, so that the model can get the latest information in time and adjust the forecast. At the same time, the model structure is further optimized to enhance its ability to handle complex and changeable data. Generally speaking, this study provides a useful exploration for the application of AI in the field of environmental pollution, and continuous improvement is expected to bring greater value to the cause of environmental protection.

References

[1] Cheng Zhigao, Li Gang, Wu Bo. Residual characteristics of petroleum hydrocarbons in soils of oilfield areas and prediction of ecological environmental risks[J]. Chinese Journal of Ecology, 2023, 42(8): 1944-1952. DOI:10.13292/j.1000-4890.202308.005.

[2] Wang Yu, Yuan Jiahui, Chen Hao. Evolution characteristics of soil phosphorus pool and environmental risk prediction in typical farmland of Taihu Lake Basin[J]. Acta Pedologica Sinica, 2022, 59(06): 1640-1649. DOI:CNKI:SUN:TRXB.0.2022-06-016.

[3] Zhang Shiyang, Li Qingqing, Yang Jie. Odor risk assessment and control strategies for volatile organic contaminated soil during excavation[J]. Research of Environmental Sciences, 2022, 35(09): 2185-2194. DOI:10.13198/j.issn.1001-6929.2022.05.03.

[4] Liu Han, Cai Wangwei, Xia Zhichang. Risk prediction of important reservoir water sources based on BP neural network[J]. Sichuan Environment, 2020, 39(03): 32-38. DOI:10.14034/j.cnki.schj.2020.03.007.

[5] Xiao Kaiqi, Xu Honggen, Gan Jie. Source analysis and environmental quality assessment of soil heavy metal pollution in Western Hunan[J]. Environmental Science, 2024, 45(03): 1760-1768. DOI:10.13227/j.hjkx.202304221.

[6] Ma Wenjuan, Liu Danni, Yang Fang. Research progress in isotope tracing of pollutants in water environments[J]. Journal of Environmental Engineering Technology, 2020, 10(02): 242-250. DOI:CNKI:SUN:HKWZ.0.2020-02-019.

[7] Wu Xinyi, Zhu Mengyuan, Sun Peipei. Advances in artificial intelligence-driven traceability of

water environmental pollutants[J]. Journal of Ecology and Rural Environment, 2025, 41(09): 1143-1157. DOI:10.19741/j.issn.1673-4831.2025.0330.

[8] Shi Jincheng, Wang Guoyu, Wang Yingchun. Risk identification and governance model of large-scale artificial intelligence models from an ESG perspective[J]. Bulletin of Chinese Academy of Sciences, 2024, 39(11): 1845-1859. DOI:10.16418/j.issn.1000-3045.20240415001.

[9] Jin Xiangyi, Zhang Wenfei. Artificial intelligence and corporate pollution reduction: Environmental effects of intelligent governance[J]. China Population, Resources and Environment, 2023, 33(08): 138-145. DOI:CNKI:SUN:ZGRZ.0.2023-08-013.

[10] Chen Liang, Zhu Junye, Jin Long. Time-series data-driven dynamic early warning for chemical process risks[J]. Journal of Safety and Environment, 2023, 23(10): 3491-3501. DOI:10.13637/j.issn.1009-6094.2022.1165.

[11] Wang Jinglin, Wang Dongsheng, Yuan Hongying. Application of fingerprinting technology and artificial intelligence in traceability analysis of pollutants[J]. Environmental Protection Science, 2022, 48(06): 130-137. DOI:10.16803/j.cnki.issn.1004-6216.2021100040.