Bayesian and LSTM Based Landfill Leakage Risk Study

Zhaozhi Jia[1], Changxin Nai[2]*

jiazhoazhi1209@163.com[1], ncx_lab@163.com[2]*

School of Information and Electronic Engineering Shandong Technology and Business University, Shandong China

Abstracts. The problem of landfill leakage risk has garnered significant attention in the field of environmental protection. Accurate prediction of landfill leakage is crucial for environmental management and risk assessment. This thesis integrates traditional parametric rate-setting methods with deep learning, highlighting the application of modern machine learning in predictive modeling. Initially, a landfill leakage model is established. Subsequently, the Bayesian optimization algorithm is employed for parameter rate-setting to enhance the model's predictive accuracy. Following this, leakage prediction is conducted using an LSTM neural network based solely on measured data. Comparative analysis reveals that LSTM achieves higher prediction accuracy, unveiling the advantages and limitations of both methods. This provides a vital reference for landfill leakage risk assessment. The study analyzes the application of traditional intelligent optimization algorithm machine learning methods in environmental science and introduces new ideas for future research on similar problems.

Keywords: landfill, leakage risk, Bayesian algorithm, LSTM, model prediction

1 Introduction

With industrialization and urbanization, hazardous waste is on the rise, presenting an urgent challenge for effective treatment. Landfills, owing to their reliable technology, straightforward operation, and ample treatment capacity, are widely employed for hazardous waste disposal. However, landfill leakage stands out as a major environmental risk.[1] The complexity and significance of this issue necessitate ongoing research to enhance our understanding and prediction of leakage risks and to implement effective control measures.

Traditional methods for assessing landfill seepage risk primarily rely on predicting design parameters and groundwater environmental risks. [2] Nevertheless, these approaches fall short in analyzing landfill leakage processes comprehensively. Furthermore, the parameters in these models are often either not directly measurable or possess unreasonable values, contributing to increased model uncertainty. [3] Consequently, the accuracy and reliability of existing models are constrained.

To tackle these challenges, our study aims to enhance landfill seepage risk assessment on multiple fronts. Firstly, we will employ a Bayesian optimization approach to rate the parameters of the landfill leakage model, thereby improving the accuracy of parameter values. This will facilitate a more accurate simulation of the landfill leakage process and mitigate

IDEE 2023, November 24-26, Zhengzhou, People's Republic of China

DOI 10.4108/eai.24-11-2023.2343436

model uncertainty. Secondly, we will employ a deep learning approach to construct a predictive model for landfill leakage risk based on analyzed measured data. This model is expected to better capture spatial and temporal variations in leakage risk, consequently enhancing the accuracy of leakage predictions.

In summary, this study promises to bring fresh perspectives and methodologies to the field of landfill leakage risk research. It is poised to guide the enhancement of operation and management practices, improve overall performance, meet industry demands for advancing landfill technology, and stimulate further development in environmental risk management.[4]

2 Model and Methods

2.1 Bayesian optimisation

The Bayesian optimization algorithm is a global optimization method grounded in Bayesian statistical theory, designed to explore a high-dimensional parameter space and identify a globally optimal solution for an objective function. At its core is the Gaussian process (GP), which constructs a mapping of the objective function based on observed data, offering insights into both expectation and uncertainty. [5] The sampling strategy is employed to choose the next parameter setting, striking a balance between exploration and exploitation to swiftly converge toward the global optimal solution, as depicted in Fig. 1.[6]

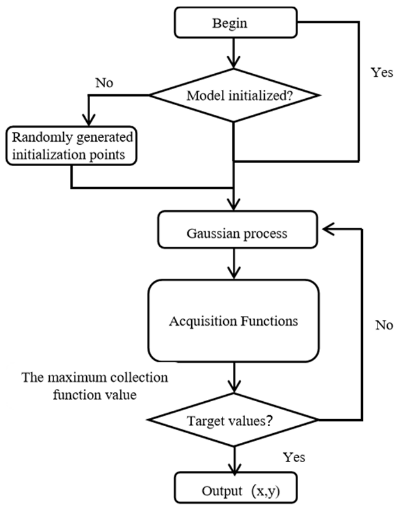

Fig. 1. Flowchart of Bayesian optimization algorithm

2.2 LSTM Model

Long Short-Term Memory (LSTM) is a deep learning neural network particularly well-suited for handling time-series data and sequence modeling tasks. The LSTM architecture is illustrated in Figure 2. The key elements of the LSTM network comprise a Cell State and three gating units: Forget Gate, Input Gate, and Output Gate.[7]The formulate1 to 4 for these components are as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f \quad (1)$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$
$$o_t = \sigma(Wo \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$
$$h_t = o_t * \tanh(C_t) \quad (5)$$

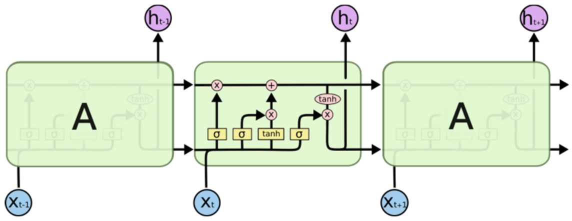

Fig.2 LSTM unit structure diagram

3 Case Study

3.1 Landfill leakage modeling

Utilizing the hydrological model structure of the landfill, we formulated a leachate seepage model grounded in the configuration of a single liner layer. This model encompasses three sub-models: vertical infiltration,[8] lateral conduction and drainage,[9] and geomembrane liner.[10] These sub-models collectively depict the processes of leachate generation, conduction and drainage, and seepage. The aim is to establish a foundation for the automatic determination of model parameters.

Building upon this model, we conducted a case study at a hazardous waste landfill site in Chengdu City, Sichuan Province, China. The daily measured data for the landfill included the cumulative water level value minus the cumulative value of the previous day's leakage. The fundamental parameters of the model are detailed in Table 1.

Table 1 Parameter value table of landfill leakage model

Submodel	Parameter	Value	Unit	Distribution
Vertical permeability Model	Thickness	7	m	Single
	Density	0.6	t/m	Single
	Total porosity	0.671	vol/vol	Single
	Field capacity	0.292	vol/vol	Single
	Withering point	0.077	vol/vol	Single
	Sat.hydr.conductivity	0.864	m/day	Single
Lateral drainage layer	Thickness	0.3	m/day	Single
	Total porosity	0.437	vol/vol	Single
	Field capacity	0.062	vol/vol	Single
	Withering point	0.024	vol/vol	Single
	Sat.hydr.conductivity	-	m/d	Normal(5,0.5)
Geomembrane liner	Geomembrane thickness	2	mm	Single
	Total porosity	0.85	vol/vol	Single
	Field capacity	0.01	vol/vol	Single
	Withering point	0.005	vol/vol	Single
	Sat.hydr.conductivity	1e-10	m/s	Single
	Pingole density	-	#/ha	Normal(2,0.5)
	Installation defects	-	#/ha	Normal(5,0.5)

3.2 Coupling of Bayesian optimization algorithms with leakage models

A Bayesian optimization algorithm is employed in the parameter rate-setting process of the leakage model to enhance its accuracy and reliability. Initially, the leakage model is defined along with the parameters to be optimized, such as the density of holes in the geomembrane of the landfill and the permeability coefficient of the conductive drainage layer. Subsequently, the prior distributions of these parameters are specified, commonly assumed to be normal distributions. Following this, a Bayesian global optimization algorithm computes the posterior distributions of the parameters by comparing the model with observed data to identify the best estimates. The algorithm iteratively generates candidate parameter values, executes the hydrologic model, updates the parameter posterior distributions, and ultimately furnishes a method for optimizing the hydrologic model to better align with the observed data.

3.3 Analysis of results

The model fitting accuracy is the error between the model simulation value and the measured data, and the evaluation indexes we used are mean square error MSE, root mean square error RMSE, mean absolute error MAE, and coefficient of determination R^2.[11]

3.3.1 Parameter rate determination and model prediction results

Through the parameter rate determination, we obtained the best set of parameter values for the leakage model as follows: the density of pinhole holes is 2/ha, the installation holes is 5/ha, and the permeability coefficient of the conductive drainage layer is 5.0112m/day; The division between the rate period and the validation period is 8:2. After the rate determination, we verified the performance of the model by comparing the fit between the model simulation values and the measured values, and the fitting accuracies of the simulation cases for the rate determination period and the validation period are shown in Table 2, and the fitting effects are shown in Fig. 3.

Combined with Fig. 3 and Table 2, it can be seen that the R2 in the rate period is close to 0.95, the MSE and RMSE are very close to zero, and the MAE is very small, which indicates that the model performs very well in the rate period and fits the actual data well. The R2 for the validation period is 0.89 with a larger MSE and RMSE and a larger MAE. This may indicate that the model performs relatively poorly in the validation period and does not generalize well to new data.

Table 2. Model fitting accuracy values by parameterization

	R^2	MSE	RMSE	MAE
Parameter rate periodic	0.95	$4.12e^{-10}$	$2.02e^{-05}$	$1.37e^{-10}$
validation period	0.89	$2.25e^{-8}$	$1.5e^{-4}$	$5.17e^{-8}$

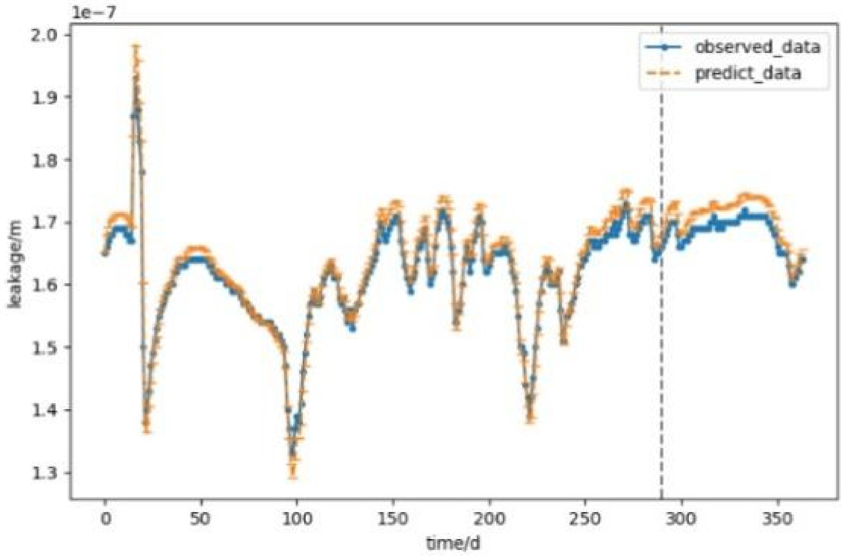

Fig.3. Plot of the effect of fitting the predicted values to the measured values

3.3.2 LSTM neural network prediction results

In the LSTM neural network model structure, it consists of an input layer, three LSTM hidden layers (with 45, 35 and 25 neurons respectively) and an output layer for generating predictions. The model has a batch size of 64 and was trained for 100 rounds with a learning rate of 0.001 and the model parameters were updated using the Adam optimizer. In addition, the data preprocessing phase consisted of data loading, data normalization to scale the data range to [0, 1], and splitting the dataset into training and validation sets in an 8:2 ratio. The accuracy of the model in the training and testing periods is shown in Table 3 and the fit is shown in Figure 4.

Combined with Figure 4 and Table 3, on the training set, the model shows a very high fit with R^2 close to 1, MSE and RMSE very close to zero, and a very small MAE, which indicates that the model performs well on the training data. The model also performs well on the test set. These results show that the LSTM model is able to accurately predict future leakage data with credible and accurate predictions.

Table 3 Model fitting accuracy values in the LSTM

	R^2	MSE	RMSE	MAE
Training period	0.99	$5.65e^{-20}$	2.37^{-10}	$1.26e^{-10}$
Testing period	0.97	$1.17e^{-19}$	$3.42e^{-10}$	$2.17e^{-10}$

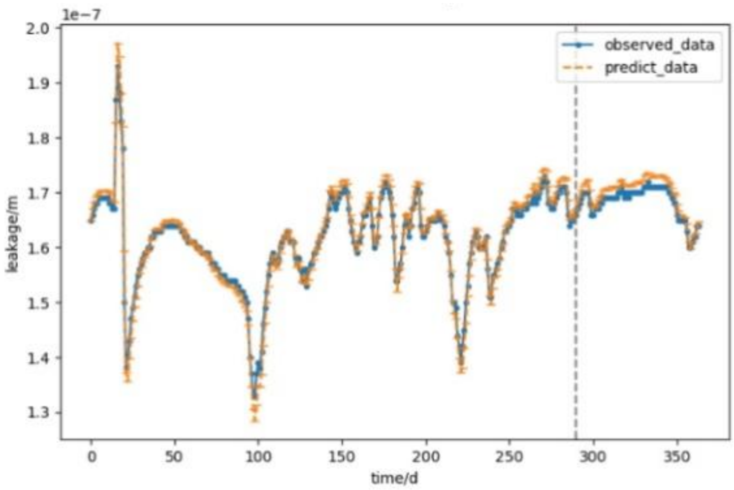

Fig.4. Effect of LSTM model fitting accuracy

3.3.3 Comparison and Analysis

The parametric rate-determined model performs well on the training data, but the performance drops during the validation period, which may be due to overfitting. This suggests that more care needs to be taken in parameter selection and model evaluation to avoid model performance degradation on new data.The LSTM model performed well in predicting the measured leakage data with its strong capability and high generalization performance. This suggests that LSTM is a powerful tool suitable for time series data prediction and can be used to improve the accuracy of leakage prediction.For comparative analysis of the above results, see Table 4.

Table 4 Comparison of prediction results between traditional Bayes and LSTM

Methods	R^2	MSE	RMSE	MAE
Bayesian optimization	0.89	$2.25e^{-8}$	$1.5e^{-4}$	$5.17e^{-8}$
LSTM model	0.97	$1.17e^{-19}$	$3.42e^{-10}$	$2.17e^{-10}$

In conclusion, the parametric rate-determined model performs well in the training period but may not be robust enough in the validation period, whereas the LSTM model shows good generalization ability and is suitable for reliable prediction of real leakage data. It is necessary to choose an appropriate modeling method based on specific needs and data characteristics.

4 Conclusion

The traditional global optimization algorithm Bayesian optimization algorithm is used to determine the parameters of the landfill leachate leakage model, and the optimal set of parameters can be obtained, which shows that the algorithm is feasible.

Compared with the calculation of the simulation value of the leakage amount by using the idea of optimizing the leakage model, the generalization of the numerical model of the leachate leakage process is not considered. According to the measured values, the prediction accuracy of the leakage amount by using LSTM is higher, and the effect of risk assessment is better.

Reference

[1] Tao F Research on the full quantitative treatment of leachate from municipal domestic waste landfills[J]Guangdong Chemical,2022,49(06):144-147.

[2] Tagliaferri F, Invernizzi M, Sironi S, et al. Influence of Model Parameters When Simulating Landfill Odour Emission[J].Chemical Engineering Transactions, 2021.DOI:10.3303/CET2185029.

[3] Wang H, Zhang C, Wang Y N, et al. Simultaneous degradation of refractory organics, antibiotics and antibiotic resistance genes from landfill leachate concentrate by GAC/O3[J].Journal of cleaner production, 2022.

[4] Traxler J M Trends in syntactic parsing: anticipation, Bayesian estimation, and good-enough parsing[J]. Trends in Cognitive Sciences,2014,18(11).

[5] Science - Automation Science; Recent Findings from University of Calgary Advance Knowledge in Automation Science (A steam injection distribution optimization method for SAGD oil field using LSTM and dynamic programming)[J]. Energy Weekly News,2020.

[6] Haidong Y. A parameter rate-setting method for predicting model parameters of sudden water pollution events based on improved Bayesian-MCMC[J]. Control and Decision Making, 2018, 33(4):8.

[7] Meiqin T. Study on pollutant transport model of hazardous waste landfill [D].Shandong Technology and Business University,2021

[8] Guodong J. Study on environmental risk assessment model of leakage in hazardous landfill [D]. Shandong Normal University,2017.

[9] uyu Z. Based on waste landfill leakage risk prediction of deep learning [D]. Shandong business school, 2023.

[10] Shuoyang G.Research on Landfill Leakage Based on BP Neural Network [D]. Shandong Institute of Technology and Technology,2023.

[11] Xiaoyuan Z.Bayesian statistical method for hydrological parameter calibration and influencing factors identification of SWMM model [D]. Chongqing University,2019.