ENGINEERING Agriculture

Engineering Journals

RESEARCH ARTICLE

2026 / Volume 13 / Issue 4 / 26674

A hybrid prediction model based on decomposition algorithm and sample entropy-weighted attention mechanism for mobile robot agricultural environment monitoring

Jiangxun LIU[1], Zijun LIU[2], Hui LIU (✉)[1]

1 Institute of Artificial Intelligence and Robotics (IAIR), Key Laboratory of Traffic Safety on Track of Ministry of Education, School of Traffic & Transportation Engineering, Central South University, Changsha 410075, Hunan, China.

2 Dundee International Institute of Central South University, Changsha 410075, Hunan, China.

KEYWORDS

Agricultural environment monitoring, attention mechanism, pollutant concentration prediction, sample entropy

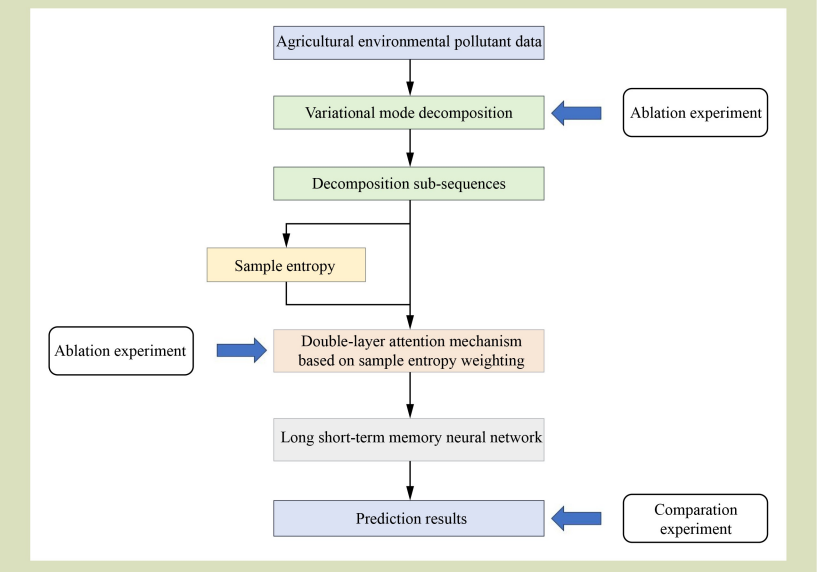
HIGHLIGHTS

- A new hybrid model is proposed that can be used by robots to predict changes in pollutant concentrations in agricultural environments.
- A double-layer attention mechanism based on sample entropy weighting is designed and combined with the prediction model.
- A hybrid prediction model is constructed using variational mode decomposition, the attention mechanism, and long short-term memory neural network.

Received July 23, 2025;
Accepted November 3, 2025.

Correspondence: csuliuhui@csu.edu.cn

GRAPHICAL ABSTRACT



ABSTRACT

The accumulation and spread of agricultural environmental pollutants pose a serious threat to the ecological environment and crop growth. Accurately predicting changes in pollutant concentrations is of great significance for achieving sustainable agricultural development. In response to the challenges of predicting pollutant concentrations in agricultural environments, this paper proposes a novel hybrid deep learning model. The variational mode decomposition algorithm is used to process raw data, reducing nonlinearity and enhancing feature distinguishability. A double-layer attention mechanism based on sample entropy evaluates sub-sequences and focuses on key regions, further improving the predictive performance of the model. Finally, a long

short-term memory neural network is used to obtain prediction results. In time series prediction experiments involving multiple pollutants, the proposed method demonstrated the needed stability and accuracy. Experimental results indicate that, compared to existing methods, this approach achieves a minimum improvement of 4.8% in mean absolute error and 23.5% in mean absolute percentage error for predicting concentrations of three pollutants. Also, the root mean square error of predictions is reduced by up to 29.1%. This study provides reliable technical support for agricultural environmental pollutant monitoring. With mean absolute errors of 5.92, 6.85, and 2.38 for CO, non-methane hydrocarbons and NO_2 predictions respectively, it accurately predicts pollutant variation risks. In the future, it can be deployed on mobile robot platforms to achieve automatic monitoring and early warning, thereby promoting the development of smart agriculture.

© The Author(s) 2026. Published by Higher Education Press. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0)

1 Introduction

With global population growth and climate change, agricultural production is facing increasingly severe environmental challenges. Soil degradation, water pollution and the spread of air pollutants are among the challenges to be addressed. These pollutants not only threaten crop yields and food safety but may also pose risks to human health. Therefore, achieving real-time, precise monitoring of agricultural environment pollution and predictive prevention and control of pollution have become key research priorities. Standard agricultural environment monitoring methods typically rely on manual sampling and laboratory analysis, which are characterized by low efficiency, high costs and insufficient spatiotemporal resolution, making them inadequate to meet the needs of sustainable agricultural development in the modern era.

In recent years, the integration of robotics and environmental monitoring has brought new solutions to the field. Intelligent robots equipped with multimodal sensors can autonomously or semi-autonomously perform pollutant detection tasks, enabling high-throughput, in-situ detection of pollutants. Lu and Han[1] designed a cooperative receding horizon temporal logic control approach for mobile robot networks to address environmental monitoring issues. Notomista et al.[2] proposed a control theory method for trajectory optimization of mobile robots in environmental monitoring, enabling continuous environmental monitoring. Wu et al.[3] studied a robotic wireless monitoring system based on wireless communication technology and intelligent control technology, achieving real-time monitoring and centralized remote control of underwater debris removal.

In agricultural environments, the dynamic changes in pollutants and their ecological impacts exhibit high complexity and uncertainty. Existing environment monitoring methods primarily rely on post-event detection, which can identify current pollution levels but struggle to provide timely warnings of potential risks, leading to delayed pollution control measures. In recent years, many studies have utilized machine learning methods based on historical data and real-time monitoring information to predict the diffusion trends and cumulative effects of pollutants. Chen et al.[4] combined a deep learning model based on residual neural network, spatiotemporal attention mechanism and convolutional long short-term memory (LSTM) neural networks to extract spatiotemporal correlations in feature sequences for accurate prediction of future pollutant concentrations. Liu et al.[5] established a hybrid prediction model using information gain, wavelet decomposition transformation technology and long short-term memory neural network, achieving higher prediction accuracy and stability for future atmospheric pollutant concentrations. Zhang et al.[6] used residual neural network to deeply extract spatial distribution features of pollutant concentrations and meteorological data, and convolutional long short-term memory network to further extract preliminary spatial distribution features, thereby accurately predicting $PM_{2.5}$ concentrations over a period of time. Bai et al.[7] developed a new deterministic and interval

prediction system for pollutant concentrations, which can provide effective interval predictions for air pollutant concentrations. However, the aforementioned methods are susceptible to noise interference during data processing, and the decomposition results exhibit poor interpretability. Therefore, this paper adopts variational mode decomposition as the data preprocessing method, which enables clear, adaptive separation of modes in the frequency domain and avoids the issue of mode confusion.

For time series with highly nonlinear characteristics, the predictive performance of single models often fails to meet the demands of practical engineering applications. Most studies combine data preprocessing, decomposition algorithms, and optimization algorithms to further enhance model performance. Among these, the attention mechanism enables neural networks to focus on relevant parts of the input data during processing. By incorporating the attention mechanism, neural networks can automatically learn and selectively focus on important information in the input, thereby improving model performance and generalization capabilities. Li et al.[8] proposed a model based on big data association principles and deep learning technology, using spatiotemporal attention and temporal self-attention to improve the accuracy of pollutant concentration predictions. Wan et al.[9] incorporated the attention mechanism into a hybrid model, proposing a method for predicting pollutant emissions from combined heat and power systems. An et al.[10] proposed a convolutional neural network and bidirectional long short-term memory model that integrates spatiotemporal attention mechanism and residual learning, demonstrating strong predictive performance. However, all these methods fail to account for high-efficiency attention analysis under high-latitude data input conditions and neglect the inherent information within the time series itself. Consequently, this paper proposes an attention mechanism based on information entropy weighting.

The integration of predictive research with technologies such as intelligent robots can establish a closed-loop system of monitoring-prediction-decision-making, driving the development of smart agriculture. By integrating preprocessing and decomposition methods, attention mechanisms and neural networks to construct hybrid models, and deploying them on mobile robots, autonomous monitoring and predictive control of agricultural environmental pollutants can be achieved. Wang et al.[11] used fully integrated empirical mode decomposition and variational mode decomposition with adaptive noise to perform a secondary decomposition of time series, effectively improving the prediction accuracy of air pollutant concentrations. Xiao et al.[12] developed an adaptive decomposition and integrated model combining independent component analysis, demonstrating the outstanding performance of decomposition methods in early warning of harmful pollutants. Existing research has demonstrated that decomposition algorithms have a satisfactory effect on improving model prediction performance. However, the selection of neural network models is crucial. Wang et al.[13] proposed a multilayer perceptron model based on wavelet decomposition to extract non-stationarity and nonlinear dependencies in time series. Wang et al.[14] used a long short-term memory artificial neural network optimized by the grey wolf optimization algorithm for prediction and modeling, with this hybrid framework serving as an effective tool for air pollutant prediction and early warning. Given the superior performance and low computational cost of long short-term memory neural networks in achieving real-time monitoring and prediction, they have been selected as the neural network model used.

In summary, although scholars have achieved useful results in predicting agricultural environment pollutants, there is still room for exploration in high-precision prediction hybrid models based on mobile robots in complex environments. This study proposes an improved deep learning model for predicting agricultural pollutants, including a variational mode decomposition algorithm, a double-layer attention mechanism (DLAM) with sample entropy (SE) weighting and a long short-term memory neural network. There are three main contributions of this study.

We propose a novel hybrid deep learning model, the long short-term memory neural network based on variational mode decomposition and double-layer attention mechanism with sample entropy weighting (VMD-SE-DLAM-LSTM). By integrating variational mode decomposition for data processing and attention mechanisms for feature processing into a long short-term memory neural network, this constitutes, to our knowledge, an innovative architecture for high-precision forecasting of agricultural environmental pollutant concentrations.

We propose an attention mechanism based on sample entropy weighting. For the decomposed feature subsequences, to select subsequences with higher information value, we calculate their

sample entropy and apply weighting, enabling the model to focus more on the critical parts of data features, reduce redundant information interference, and make the learning and prediction of pollutant sequences more efficient.

Experimental results on the UCI air quality dataset for three pollutants demonstrate the superior performance of the proposed method, enabling accurate monitoring of agricultural environments and showcasing its potential for application on mobile robots.

2 Methodology

2.1 Variational mode decomposition

Variational mode decomposition (VMD) is a signal decomposition estimation method[15]. VMD determines the frequency center and bandwidth of each component by iteratively searching for the optimal solution of the variational mode during the decomposition process, thereby enabling adaptive frequency domain partitioning of the signal and effective separation of each component[16,17].

Assuming that the original signal S is decomposed into K components u, ensuring that the decomposition sequence is a finite bandwidth modal component with a center frequency, and that the sum of the estimated bandwidths of each mode is minimized, with the constraint that the sum of all modes is equal to the original signal, then the constrained variational expression is:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{ \sum\limits_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \dfrac{j}{\pi t} \right) \cdot u_k(t) \right] \mathrm{e}^{-j\omega k t} \right\|_2^2 \right\} \\ s.t. \sum\limits_{k=1}^{K} u_k = S \end{cases} \tag{1}$$

By solving this constrained variational expression, the constrained variational problem is transformed into an unconstrained variational problem, obtaining the augmented Lagrange expression:

$$\begin{aligned} L[\{u_k(t)\},\{\omega_k\},\lambda(t)] &= \alpha \sum_{k=1} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] \mathrm{e}^{-j\omega k t} \right\|_2^2 \\ &+ \left\| S(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), S(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \end{aligned} \tag{2}$$

Parameter values are updated iteratively according to the following equations. When the component satisfies $\sum_k \left\| u_k^{n+1} - u_k^n \right\|_2^2 \Big/ \left\| u_k^n \right\|_2^2 < \varepsilon$, the solution is complete.

$$u_k^{n+1}(\omega) = \frac{f(\omega) - \sum_{i \neq k} u_i(\omega) + \dfrac{\lambda(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \tag{3}$$

$$\omega_k^{n+1} = \frac{\int_0^{\infty} \omega |u_k(\omega)|^2}{\int_0^{\infty} |u_k(\omega)|^2} \tag{4}$$

$$\lambda^{n+1}(\omega) = \lambda^n(\omega) + \tau \left(f(\omega) - \sum\nolimits_k u_k^{n+1}(\omega) \right) \tag{5}$$

The mode number K and penalty factor a constitute critical parameters for VMD, necessitating their optimization during decomposition. To address this, MATLAB intelligent algorithms are used for parameter optimization, enabling adaptive decomposition of time series while mitigating the impact of subjective choices on decomposition robustness. This paper introduces envelope entropy as an evaluation criterion to guide VMD parameter optimization, calculated as:

$$\begin{cases} E_p = -\sum\limits_{j=1}^{N} p_j \lg p_j \\ p_j = \frac{a(j)}{\sum_{j=1}^{N} a(j)} \end{cases} \tag{6}$$

where, $j = 1,2,\ldots,N$, N is the number of sampling points, $a(j)$ is the Hilbert-demodulated envelope of $x(j)$, and $p(j)$ is the normalized form of $a(j)$. The average value of the envelope entropy for each decomposed modal component serves as the fitness function, calculated as:

$$Fitness(VMD(K,\alpha)) = \frac{1}{K} \sum\nolimits_{i=1}^{K} E_p(i) \tag{7}$$

2.2 Double-layer attention mechanism based on sample entropy weighting

Sample entropy is a statistical measure proposed by Richman and Moorman[18] that improves upon the approximate entropy algorithm by not counting self-matches. Sample entropy is currently one of the most widely used methods for calculating entropy feature values. Sample entropy is also a measure of time series complexity, the smaller the sample entropy, the lower the time complexity and the higher the self-similarity. The pseudo-code shown in below:

Step 1: The time series is defined as X.

Step 2: Iterate through all combinations of $X(i)$ and $X(j)$, where i ranges from 1 to $N-m$, excluding cases where i and j

are equal. There are a total of $N-m-1$ combinations. In this case, m is a window and N is the sequence length.

Step 3: Let the ratio of the approximate number to the total number be denoted by $B_i^m(r)$. Note that the total number at this point is $N-m-1$.

Step 4: Find $B^m(r)=\frac{1}{N-m}\sum_{i=1}^{N-m}B_i^m(r)$.

Step 5: Increase the dimension by m to $m+1$, repeat the above steps, and obtain $A^m(r)$.

Step 6: The sample entropy of the signal sequence is obtained as $SampEn(m,r)=\ln(B^m(r))-\ln(A^m(r))$.

The higher entropy values calculated for subsequences in the aforementioned process indicate greater complexity and potentially richer information content. We use sample entropy as prior knowledge, modulating the weighting coefficients for subsequence decomposition through elementwise multiplication. Compared to existing attention mechanisms in models such as Transformers, our introduction of an unsupervised, information-theoretic metric as prior knowledge effectively complements data-driven methods. Also, this achieves enhanced model performance at a reduced computational cost. Weighted processing of $SampEn^X$ obtained for time series X.

$$X^{SE}=X\cdot SampEn^X \tag{8}$$

Feature map X^{SE} is transformed to obtain feature map U^{SE}, defined as:

$$U^{SE}=V*X^{SE} \tag{9}$$

where, V is the learned filter kernel, and $*$ is the convolution operation.

Then, to solve the dependency problem, global spatial information is compressed into a channel descriptor, which allows the generated data to contain contextual information, as:

$$z=\frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}u(i,j) \tag{10}$$

Among them, H and W are the two dimension values of the feature map.

Then, a gate mechanism consisting of two fully connected layers is used to obtain the weights S, and the attention weights are weighted to the feature entropy in front to obtain the final output $\widetilde{X}$.

$$\widetilde{X}=S\cdot U^{SE} \tag{11}$$

Next, the feature map of the first layer of attention is used to generate two-dimensional vectors through maximum pooling and average pooling, respectively. The obtained feature maps are superimposed to form a new feature map, which is then converted through a convolution layer, and the spatial weights are output by the activation function.

$$T=\sigma\left(f^{n\times n}\left(\left[AvgPool\left(\widetilde{X}\right);MaxPool\left(\widetilde{X}\right)\right]\right)\right) \tag{12}$$

$$Y=G\left(\widetilde{X}\right)\cdot T \tag{13}$$

where, σ is the sigmoid function, $f^{n\times n}$ is the convolution operation with a convolution kernel size of $n\times n$, Y is the output feature map and $G\left(\widetilde{X}\right)$ is the division of the input feature map into a grid of points, each of which has a feature value. The overall structure is shown in Fig. 1.

2.3 Long short-term memory neural network

Long short-term memory neural networks[19,20] are special recurrent networks that are particularly well-suited for processing sequential data. LSTM was explicitly designed to address the issue of long-term dependencies, successfully overcoming the limitations of the original recurrent neural networks and has become the most widely used recurrent neural network today, with successful applications in many fields such as speech recognition, time series prediction, and natural language processing. In addition, compared to methods such as BiLSTM[21,22], LSTM can perform prediction tasks while avoiding leakage of future information, and exhibit lower computational costs and latency.

LSTM introduces input gates, forget gates, output gates and a cell state, which enable it to better handle long-term dependencies in sequences. The LSTM structure is shown in Fig. 2.

Forget gate: Through the operations of x and h_t, and after passing through the sigmoid function, a vector of 0 and 1 is obtained, where 0 represents that a certain part of the previous memory should be forgotten and 1 represents that the previous memory should be retained.

Input gate: By adding the information that needed to be

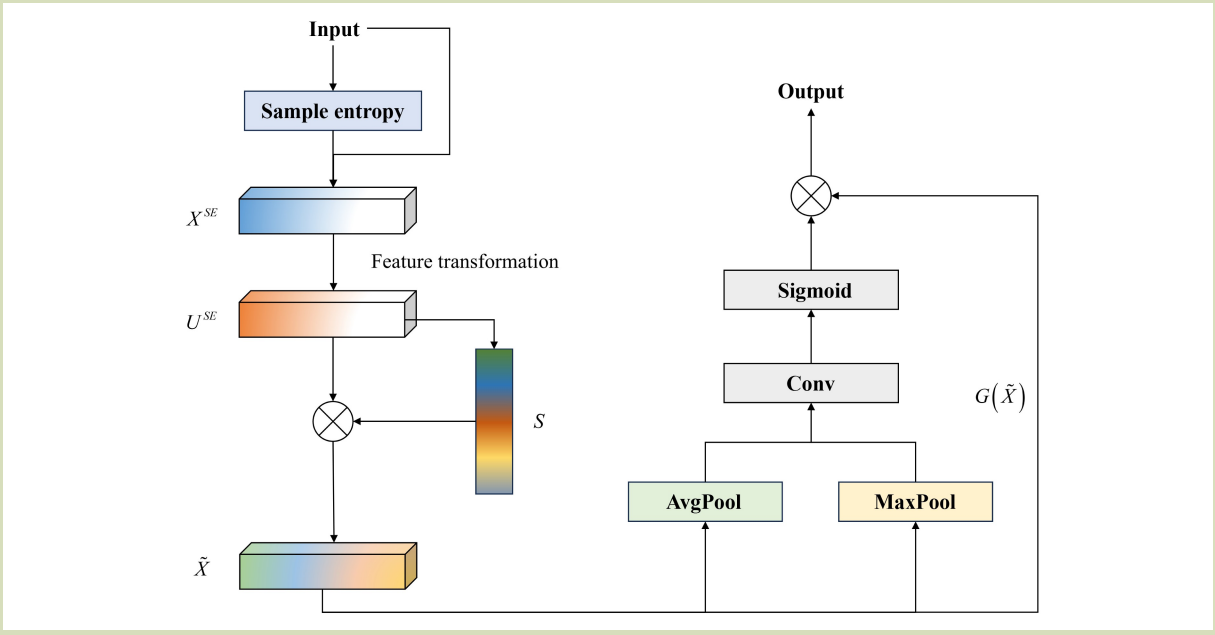


Fig. 1 Structure of the double-layer attention mechanism based on sample entropy weighting. AvgPool is average pooling, MaxPool is maximum pooling, Conv is convolution operation.

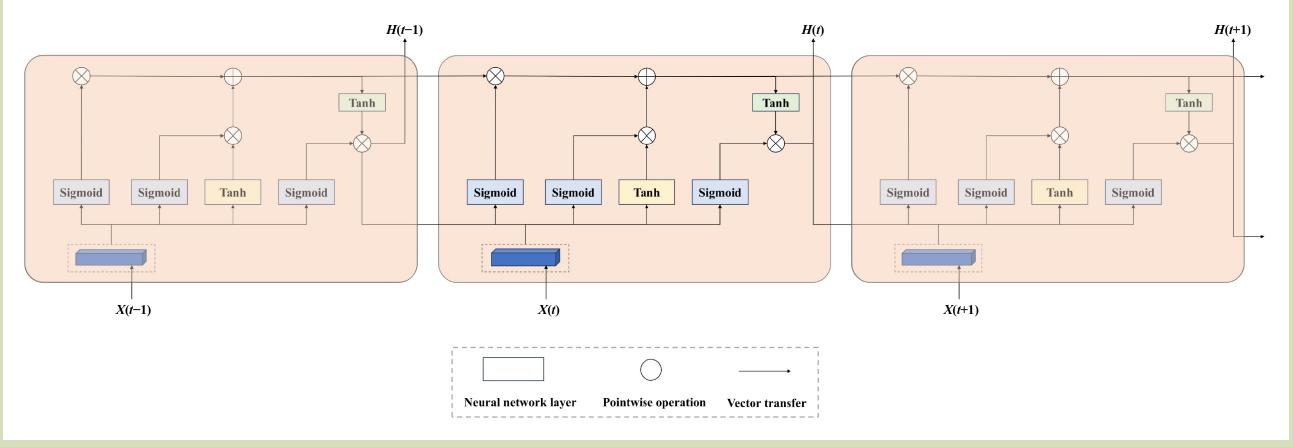


Fig. 2 Structure of long short-term memory neural network.

retained previously to the information that needs to be remembered now, a new memory state is obtained.

Output gate: Integrate C_t to obtain an output.

2.4 Hybrid model framework

The structure and hyperparameter information of the proposed VMD-SE-DLAM-LSTM hybrid model are given in Fig. 3 and Fig. 4, and Table 1.

3 Experiments and results

3.1 Dataset

This study used experimental time series data on CO, non-

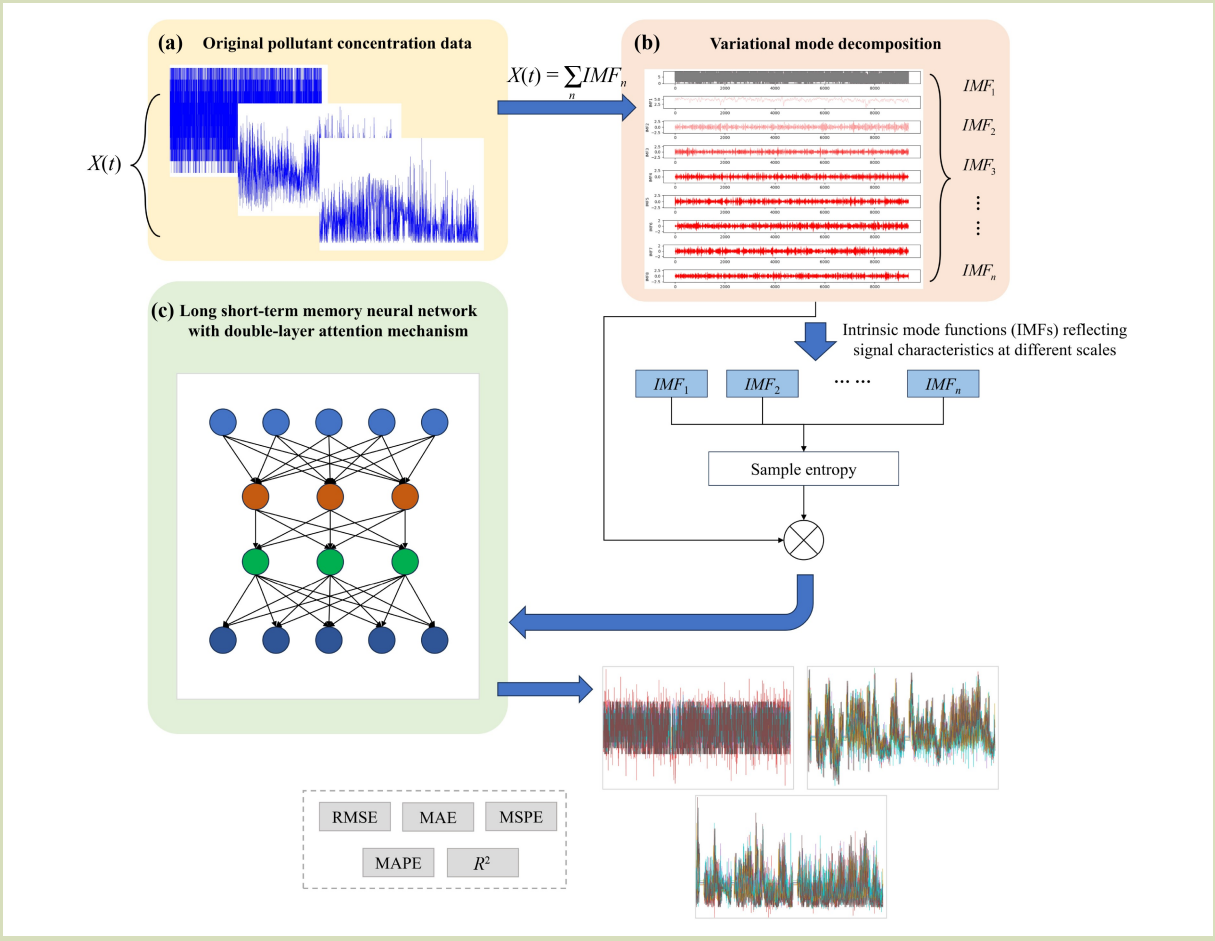


Fig. 3 VMD-SE-DLAM-LSTM structure. (a) Input of original pollutant concentration data; (b) generation of subsequences via variational mode decomposition, followed by calculation of sample entropy for these subsequences and weighting operations; (c) input into the long short-term memory neural network with double-layer attention mechanism to generate output prediction results.

methane hydrocarbons (NMHC) and NO_2 from the UCI air quality dataset[23] to validate the effectiveness of the proposed model. The dataset contained 9,358 hourly average responses from five metal oxide chemical sensor arrays embedded in air quality chemical multi-sensor devices. The device was located in a field within a severely polluted area of an Italian city. The recorded data spanned from March 2004 to February 2005, representing the longest recorded response of the air quality chemical sensor device deployed at the site. Sensor data exhibits missing values, which were reconstructed for this study using a time-series-based linear interpolation method. Then, we normalized the data by applying min-max normalization to map it to the range [0, 1], thereby eliminating the effects of differing scales and units of measurement[24]. The hourly average concentrations of CO, NMHC, benzene, total nitrogen oxides and NO_2 were provided by a reference certified analyzer located at the same site. The data used in this paper was presented in Table 2 and Figs. 5–7. The dataset was partitioned into a 70% training set and a 30% testing set. The framework code is run on a Windows 11 operating system, Intel Core i5-12600KF CPU, Nvidia GeForce RTX 3060 and Pycharm 2021.

3.2 Performance evaluation metrics

The four primary evaluation metrics used in this study to assess the predictions obtained include the mean absolute error (MAE), mean absolute percentage error (MAPE), mean square percentage error (MSPE), root mean square error (RMSE) and R^2, with the calculation formula being[25,26]:

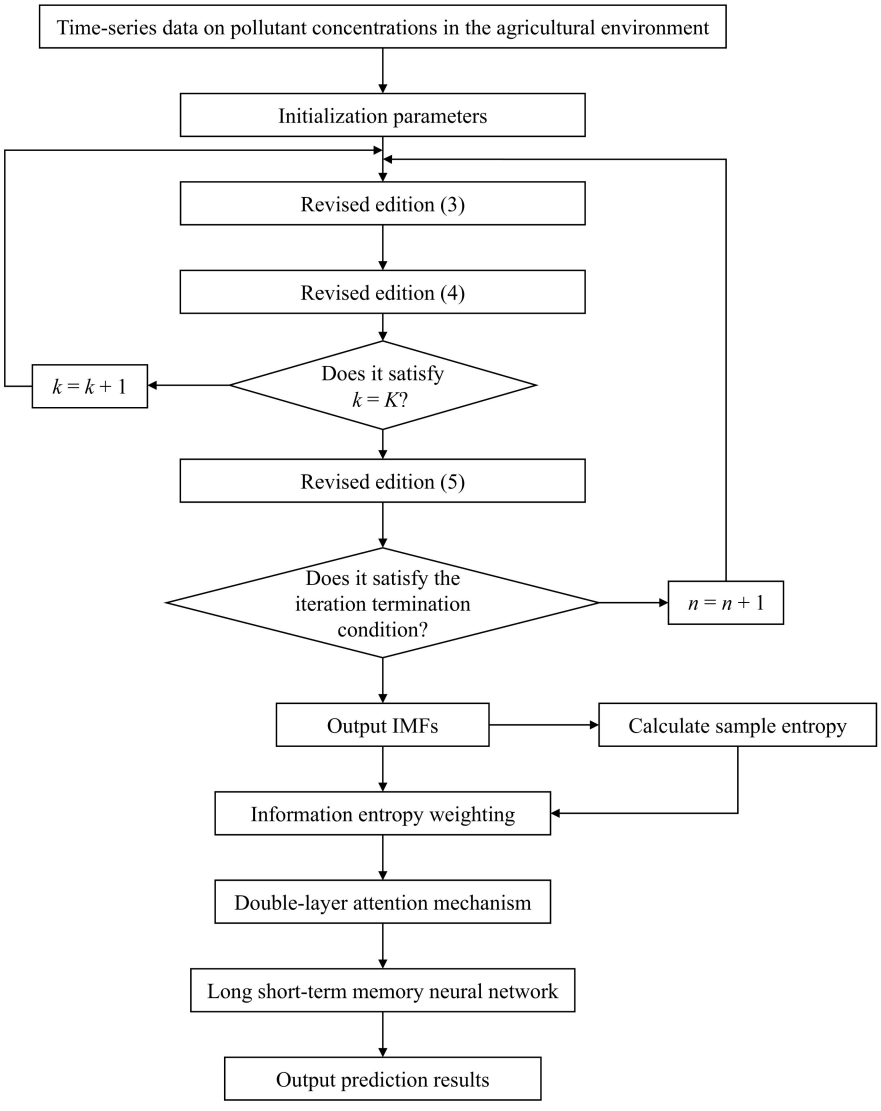


Fig. 4 VMD-SE-DLAM-LSTM workflow chart. The time-series data first undergoes variational mode decomposition to determine if it satisfies the iteration termination condition. If satisfied, IMFs are output and sample entropy is calculated. Subsequently, information entropy weighting is applied, and prediction results are generated through the double-layer attention mechanism and long short-term memory neural network.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(X(t)-\widehat{X}(t)\right)^2} \quad (14)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|X(t)-\widehat{X}(t)\right| \quad (15)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{X(t)-\widehat{X}(t)}{X(t)}\right|\times 100\% \quad (16)$$

$$MSPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{X(t)-\widehat{X}(t)}{X(t)}\right)^2\times 100\% \quad (17)$$

$$R^2 = \frac{\sum_{i=1}^{N}\left(X(t)-\overline{X}(t)\right)\left(Y(t)-\overline{Y}(t)\right)}{\sqrt{\sum_{i=1}^{N}\left(X(t)-\overline{X}(t)\right)^2}\sqrt{\sum_{i=1}^{N}\left(Y(t)-\overline{Y}(t)\right)^2}} \quad (18)$$

where, N is the sample size, $X(t)$ is the actual value, and $\widehat{X}(t)$ is the predicted value. RMSE, MAE, MSPE, and MAPE are

Table 1 Information pertaining to VMD-SE-DLAM-LSTM hyperparameters

Hyperparameter	Value
Optimizer	Adam
Initial learning rate	1e-3
Batch size	32
Weight decay	1e-4
Dropout rate	0.2
Transformer layers	4
Epochs	400

inversely proportional to prediction accuracy, and R^2 is directly proportional to prediction accuracy.

3.3 Model evaluation

To demonstrate the capability of the proposed model in predicting agricultural environment pollutants, we selected standard models, including ARIMA[27], recurrent neural network[28], one-dimensional convolutional neural network[29], support vector machine[30], deep belief network[31], state-of-the-art models including gated recurrent unit[32], Maity et al.[33] designed the state-of-the-art machine learning method MLSFDD and Das et al.[34] proposed the AQMS model for comparison. Additionally, to validate the effectiveness of the modules designed in this study, ablation experiments were conducted on the decomposition algorithm and attention mechanism and compared with the proposed model.

3.3.1 Model performance comparison

ARIMA is a model based on autoregressive and moving average time series analysis, which uses a linear combination of past observations to predict future values. The recurrent neural network is a deep learning model used to learn temporal features, where units are connected in a long chain and evolve recursively in a sequential manner. CNN uses the capabilities

Table 2 Descriptive statistics of the UCI air quality dataset

Pollutant	Maximum	Minimum	Mean	Standard deviation
CO	9.0	0.0	4.7	2.8
Non-methane hydrocarbons	2010	430	1020	341
NO_2	63.0	1.0	12.0	9.1

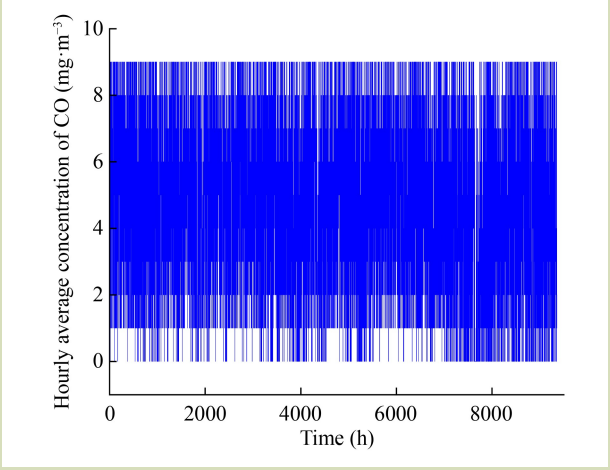


Fig. 5 Time series of CO concentration. The hourly average concentration of CO varies with the sampling time.

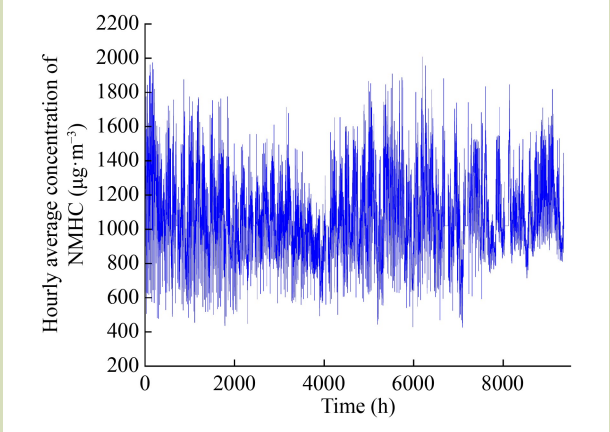


Fig. 6 Time series of non-methane hydrocarbon concentration (NMHC). The hourly average concentration of NMHC varies with the sampling time.

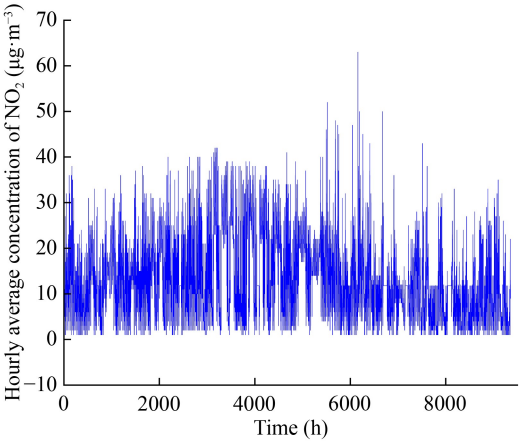


Fig. 7 Time series of NO_2 concentration. The hourly average concentration of NO_2 varies with the sampling time.

of convolutional kernels to sense changes in historical data over a period of time and make predictions. A support vector machine is a standard supervised learning algorithm. A deep belief network is a generative model composed of stacked layers of restricted Boltzmann machines, capable of capturing high-level abstract features in data and demonstrating strong representational capabilities for complex data structures. The gated recurrent unit model introduces a gating mechanism to address the gradient vanishing and explosion issues in existing recurrent neural networks and provides an effective means of capturing long-term dependencies.

Through experimental comparison and analysis, as shown in Figs. 8–10, it was found that our proposed method achieved the best performance in predicting the concentrations of three pollutants: CO, NMHC and NO_2. Compared with standard models, the prediction correlation R^2 of the proposed model reached 0.62, 0.76 and 0.65, respectively. The RMSE, MAE, MAPE and MSPE evaluation metrics are generally lower than those models, proving that the VMD-SE-DLAM-LSTM model has better stability and more accurate prediction capabilities. As shown in Table 3, the reduction in MAE was particularly pronounced, with MAE values of 5.92, 6.85 and 2.38, representing reductions of 4.8% to 59.5%. When compared with the state-of-the-art models, our proposed model outperforms the best-performing gated recurrent unit model by 6.9%, 10.1%, and 4.8% in the R^2 metric. Also, the computational complexity of the VMD algorithm is $O(KN \cdot \log N)$ per iteration, while that of SE-DLAM is $O(N^2 \cdot d)$, where K is the number of modes, N is the sequence length, and d is the feature dimension. It is evident that the computational complexity of both modules falls within an acceptable range, rendering them feasible for practical applicability. Our proposed model has a significant advantage in pollutant concentration prediction and can be integrated into mobile robots for agricultural environment monitoring and early warning.

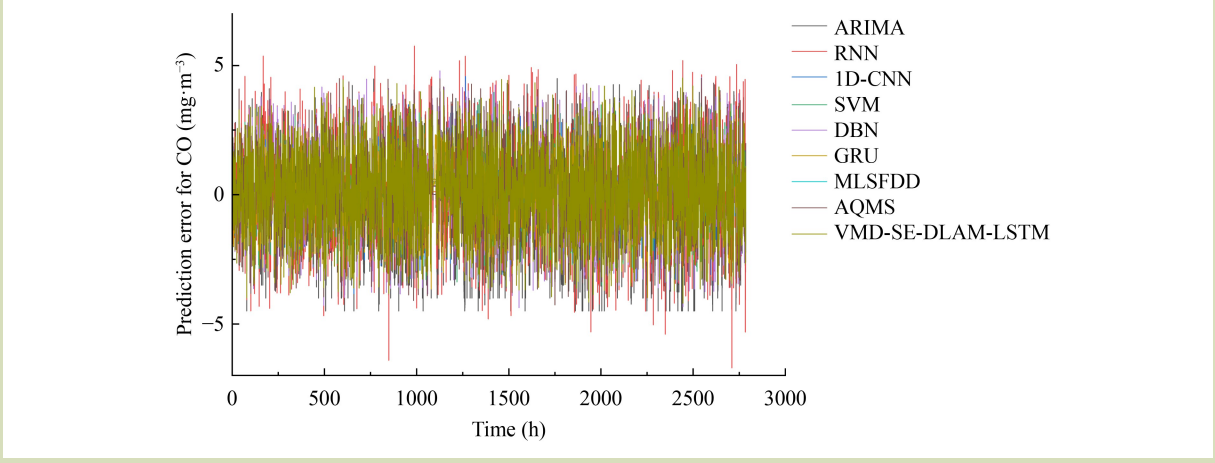


Fig. 8 Predicted results of CO concentration test set. Comparison of predictive performance between different models. RNN is recurrent neural network, 1D-CNN is one-dimensional convolutional neural network, SVM is support vector machine, DBN is deep belief network, GRU is gated recurrent unit.

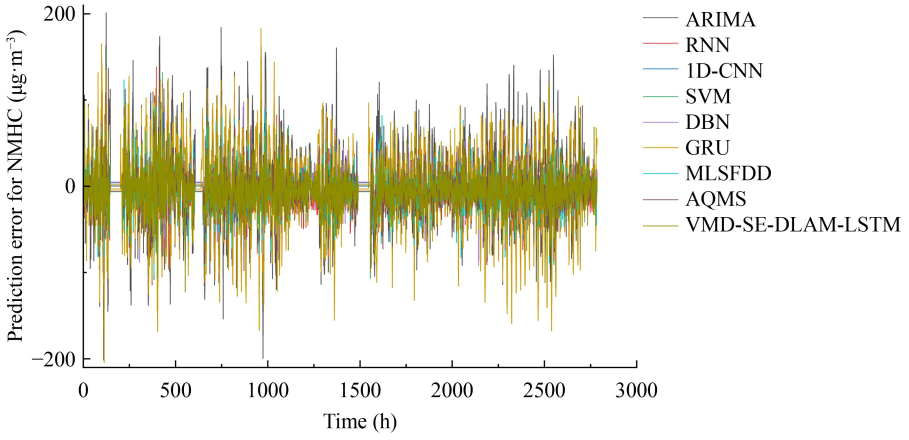


Fig. 9 Predicted results of non-methane hydrocarbon concentration test set. Comparison of predictive performance between different models. RNN is recurrent neural network, 1D-CNN is one-dimensional convolutional neural network, SVM is support vector machine, DBN is deep belief network, GRU is gated recurrent unit.

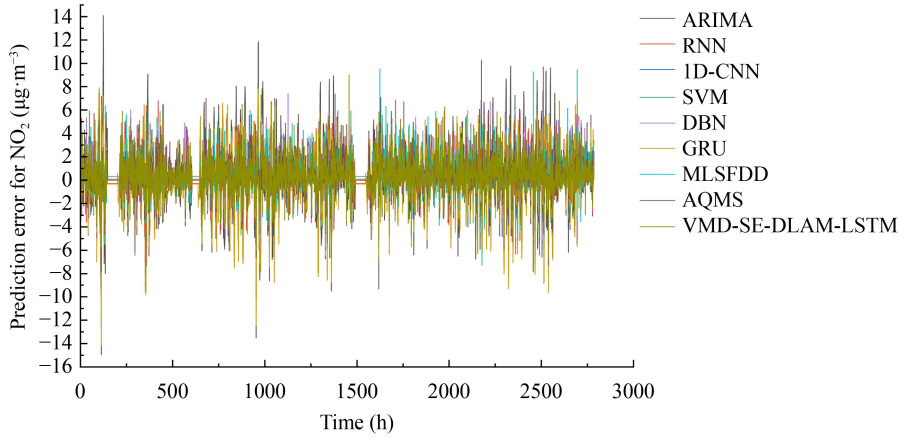


Fig. 10 Predicted results of NO_2 concentration test set. Comparison of predictive performance between different models. RNN is recurrent neural network, 1D-CNN is one-dimensional convolutional neural network, SVM is support vector machine, DBN is deep belief network, GRU is gated recurrent unit.

3.3.2 Comparison of decomposition algorithms

To validate the contribution of the variational mode decomposition method selected in this study to the predictive performance of the model, we designed multiple ablation experiments, using different decomposition algorithms to process the original data, to highlight the importance of the variational mode decomposition method. This paper selected empirical mode decomposition[35], wavelet packet decomposition[36], and local mean decomposition[37] for experimental comparison and analysis, with the results shown in the Table 4 and Figs. 11–13. It can be seen that the VMD method used has the most significant effect on improving the prediction performance of the model. For the CO pollutant, the prediction error is significantly reduced, with the R^2 metric improving by 12.7% to 29.2%. Table 4 indicates that the VMD-SE-DLAM-LSTM method achieved error metrics of 3.02, 5.92, 0.50, and 1.73, respectively, with maximum error reductions of 24.9%, 2.0%, 32.4% and 19.9%. For the NMHC pollutant, the

Table 3 Comparison of predictive performance between different models. RNN is recurrent neural network, 1D-CNN is one-dimensional convolutional neural network, SVM is support vector machine, DBN is deep belief network, GRU is gated recurrent unit

Pollutant	Metric	Model								
		ARIMA	RNN	1D-CNN	SVM	DBN	GRU	MLSFDD	AQMS	VMD-SE-DLAM-LSTM
CO	RMSE	3.45	3.82	3.04	3.05	3.43	2.85	3.42	3.11	3.02
	MAE	11.91	14.61	9.21	12.10	11.77	12.28	10.54	7.54	5.92
	MAPE	0.80	1.02	1.65	2.23	0.91	0.73	1.58	1.00	0.50
	MSPE	2.07	2.63	2.94	3.53	2.74	2.15	2.22	1.95	1.73
	R^2	0.30	0.54	0.45	0.28	0.50	0.58	0.60	0.61	0.62
NMHC	RMSE	12.34	11.75	12.95	10.49	11.71	10.33	11.92	12.28	10.51
	MAE	15.24	13.81	11.75	7.51	13.71	7.52	12.63	9.36	6.85
	MAPE	0.83	0.81	0.88	0.92	0.81	0.90	0.89	0.80	0.62
	MSPE	0.21	0.12	0.22	0.36	0.34	0.30	0.46	0.23	0.09
	R^2	0.67	0.70	0.75	0.71	0.70	0.69	0.70	0.69	0.76
NO_2	RMSE	5.88	5.17	4.28	4.83	5.00	4.59	4.52	4.91	4.17
	MAE	3.62	2.68	2.83	4.04	2.50	2.65	2.60	2.44	2.38
	MAPE	0.74	0.97	1.01	1.46	0.86	0.93	0.80	0.83	0.46
	MSPE	2.46	3.31	3.61	4.75	2.60	2.22	2.26	2.40	1.39
	R^2	0.45	0.51	0.54	0.40	0.46	0.62	0.59	0.63	0.65

Table 4 Comparison of predictive performance between different decomposition algorithms

Pollutant	Decomposition algorithm	Metric				
		RMSE	MAE	MAPE	MSPE	R^2
CO	EMD-SE-DLAM-LSTM	3.33	5.87	0.62	2.01	0.55
	WPD-SE-DLAM-LSTM	3.98	6.04	0.63	1.82	0.51
	LMD-SE-DLAM-LSTM	4.02	5.97	0.74	2.16	0.48
	VMD-SE-DLAM-LSTM	3.02	5.92	0.50	1.73	0.62
NMHC	EMD-SE-DLAM-LSTM	10.38	7.00	0.59	0.13	0.71
	WPD-SE-DLAM-LSTM	12.54	8.17	1.43	0.42	0.59
	LMD-SE-DLAM-LSTM	10.76	8.31	1.25	0.38	0.64
	VMD-SE-DLAM-LSTM	10.51	6.85	0.62	0.09	0.76
NO_2	EMD-SE-DLAM-LSTM	4.08	2.52	0.66	1.92	0.58
	WPD-SE-DLAM-LSTM	4.20	2.55	0.57	1.77	0.61
	LMD-SE-DLAM-LSTM	6.24	3.08	1.03	2.94	0.50
	VMD-SE-DLAM-LSTM	4.17	2.38	0.46	1.39	0.65

model using the VMD algorithm achieved the best performance, with RMSE, MAE, MAPE, MSPE, and R^2 reaching 10.51, 6.85, 0.62, 0.09 and 0.76, respectively. For the NO_2 pollutant, the local mean decomposition method performed the worst, while empirical mode decomposition and wavelet packet decomposition had similar effects, and the

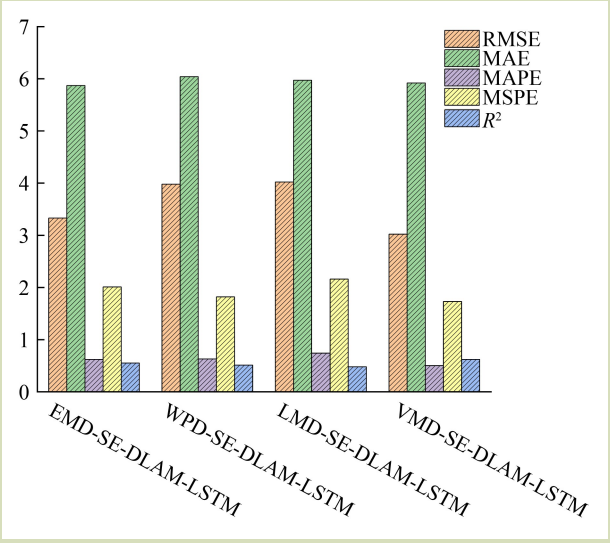


Fig. 11 Comparison of CO concentration prediction evaluation metrics for different decomposition algorithms. EMD is empirical mode decomposition, WPD is wavelet packet decomposition, LMD is local mean decomposition. The vertical axis represents the numerical values of the evaluation metrics.

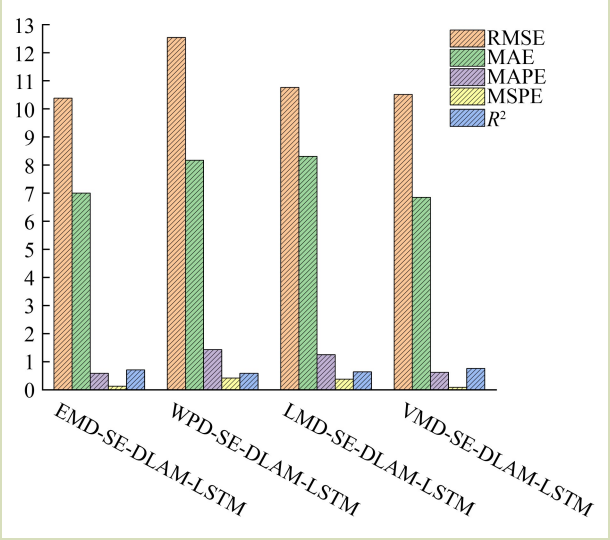


Fig. 12 Comparison of NMHC concentration prediction evaluation metrics for different decomposition algorithms. EMD is empirical mode decomposition, WPD is wavelet packet decomposition, LMD is local mean decomposition. The vertical axis represents the numerical values of the evaluation metrics.

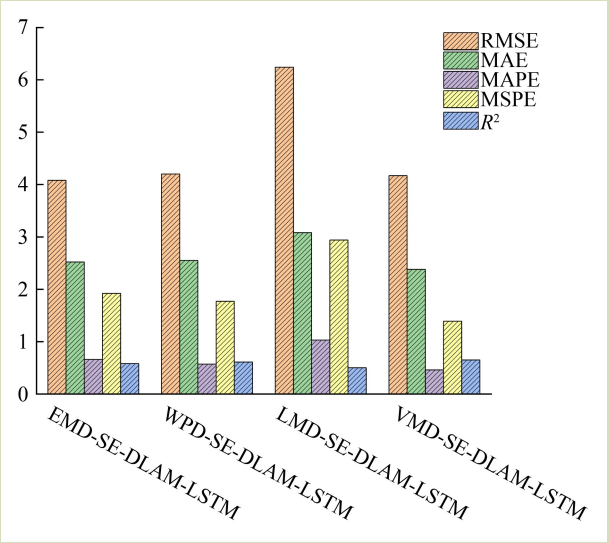


Fig. 13 Comparison of NO_2 concentration prediction evaluation metrics for different decomposition algorithms. EMD is empirical mode decomposition, WPD is wavelet packet decomposition, LMD is local mean decomposition. The vertical axis represents the numerical values of the evaluation metrics.

VMD method achieved the best prediction correlation at 0.65. In summary, the time series prediction experiments for the three pollutants demonstrated that the VMD algorithm used in this study has a unique benefit in enhancing pollutant prediction capabilities.

3.3.3 Comparison of attention mechanisms

To validate the contribution of the proposed attention mechanism with sample entropy weighting in predicting agricultural environment pollutants, this was compared with other attention mechanisms, including the channel attention mechanism[38] and the spatial attention mechanism[39]. The results are given in Table 5 and Figs. 14–16.

Ablation experiments with different decomposition algorithms demonstrated that the VMD algorithm is crucial in the proposed framework. Therefore, we conducted ablation experiments on different attention mechanisms based on VMD-LSTM. Channel attention mechanism, spatial attention mechanism and the attention mechanism designed in this study can all improve the prediction performance of the model to a certain extent. Of these, the spatial attention mechanism outperformed the channel attention mechanism, improving the

Table 5 Comparison of predictive performance between different attention mechanisms

Pollutant	Attention mechanism	Metric				
		RMSE	MAE	MAPE	MSPE	R^2
CO	VMD-CAM-LSTM	3.69	12.35	0.78	2.45	0.54
	VMD-SAM-LSTM	3.31	7.59	0.64	1.80	0.60
	VMD-SE-DLAM-LSTM	3.02	5.92	0.50	1.73	0.62
NMHC	VMD-CAM-LSTM	11.34	10.87	0.94	0.17	0.70
	VMD-SAM-LSTM	10.68	8.95	0.86	0.20	0.74
	VMD-SE-DLAM-LSTM	10.51	6.85	0.62	0.09	0.76
NO_2	VMD-CAM-LSTM	4.72	3.06	0.50	1.37	0.53
	VMD-SAM-LSTM	5.11	2.88	0.47	1.56	0.59
	VMD-SE-DLAM-LSTM	4.17	2.38	0.46	1.39	0.65

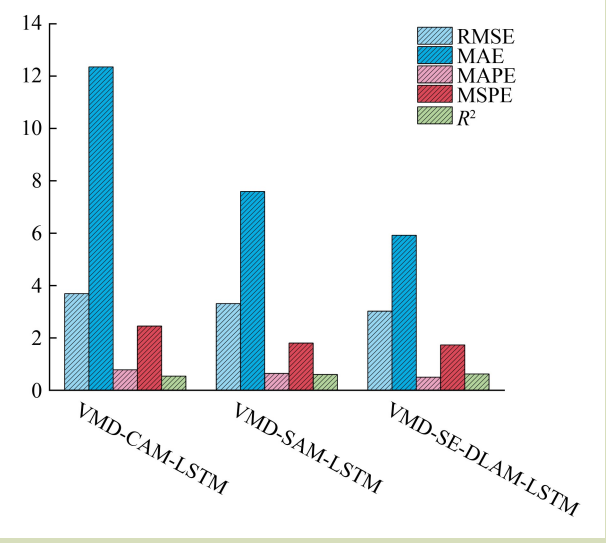


Fig. 14 Comparison of CO concentration prediction evaluation metrics for different attention mechanisms. CAM is channel attention mechanism, SAM is spatial attention mechanism. The vertical axis represents the numerical values of the evaluation metrics.

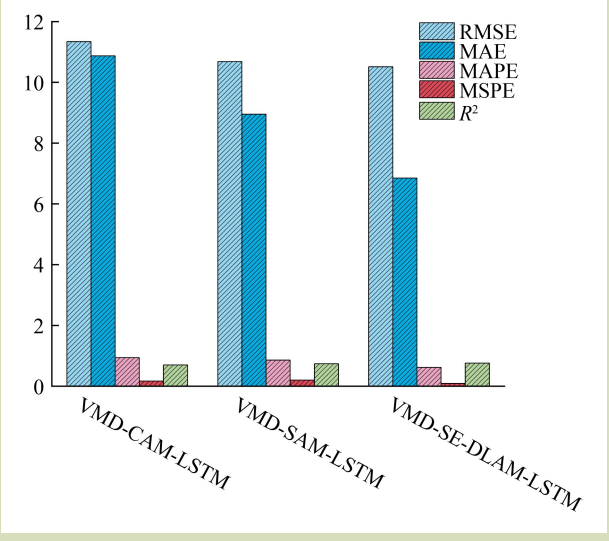


Fig. 15 Comparison of non-methane hydrocarbon concentration prediction evaluation metrics for different attention mechanisms. CAM is channel attention mechanism, SAM is spatial attention mechanism. The vertical axis represents the numerical values of the evaluation metrics.

prediction correlation of CO, NMHC and NO_2 by 11.1%, 5.7%, and 11.3%, respectively. The SE-DLAM gave optimal performance, achieving RMSE, MAE, MAPE, MSPE and R^2 values of 10.51, 6.85, 0.62, 0.09 and 0.76, respectively, for NMHC. In summary, the double-layer attention mechanism based on sample entropy weighting maximizes model performance and can be applied within the framework.

4 Conclusions

This paper proposes a novel hybrid deep learning model for predicting the concentration of agricultural environment pollutants and has demonstrated its feasibility for implementation on mobile robots for automatic monitoring and early warning.

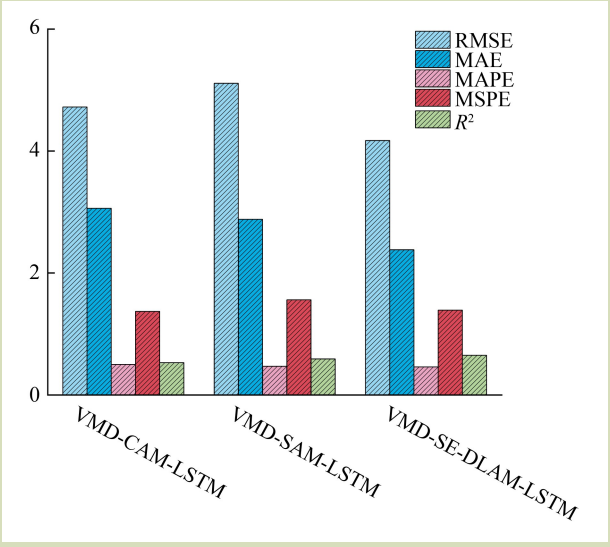


Fig. 16 Comparison of NO_2 concentration prediction evaluation metrics for different attention mechanisms. CAM is channel attention mechanism, SAM is spatial attention mechanism. The vertical axis represents the numerical values of the evaluation metrics.

Comparison of experimental results showed that the proposed VMD-SE-DLAM-LSTM model performed the best, achieving competitive performance compared to standard and state-of-the-art models. The VMD algorithm effectively enhances the predictive performance of the model, achieving lower nonlinearity and more distinct features in the decomposed subsequences compared to other decomposition algorithms. The sample entropy-weighted double-layer attention mechanism proposed significantly enhances pollutant prediction performance by evaluating the feature value of decomposed subsequences and identifying key regions, enabling the model to more easily learn the latent features in the data. The VMD-SE-DLAM-LSTM model achieved high-quality results in time series prediction experiments for three pollutants, demonstrating its stability, accuracy and versatility. In the future, it can be deployed in mobile robots for automatic monitoring and early warning research in agricultural environments. However, this study has its limitation and there is a need for careful evaluation of some practical challenges, including as sensor noise, missing data and communication delays. Subsequent work will use model distillation, edge computing development, the introduction of varying degrees of noise and delay compensation techniques to further enhance the reliability and robustness of agricultural environmental pollutant concentration predictions.

Acknowledgements

The study is fully supported by the National Natural Science Foundation of China (52072412).

Compliance with ethics guidelines

Jiangxun Liu, Zijun Liu, and Hui Liu (member of FASE declare that they have no conflicts of interest or financial conflicts to disclose. This article does not contain any studies with human or animal subjects performed by any of the authors.

REFERENCES

1. Lu Q, Han Q L. Mobile robot networks for environmental monitoring: a cooperative receding horizon temporal logic control approach. *IEEE Transactions on Cybernetics*, 2019, **49**(2): 698–711
2. Notomista G, Pacchierotti C, Giordano P R. Online robot trajectory optimization for persistent environmental monitoring. *IEEE Control Systems Letters*, 2022, **6**: 1472–1477
3. Wu X Y, Wang S, Wang X N, He G G. A wireless real-time monitoring system on intelligent underwater pollution cleaning robot. *Journal of Computational Methods in Sciences and Engineering*, 2021, **21**(4): 1051–1060
4. Chen C, Qiu A G, Chen H Y, Chen Y J, Liu X, Li D. Prediction of pollutant concentration based on spatial-temporal attention, ResNet and ConvLSTM. *Sensors*, 2023, **23**(21): 8863
5. Liu B C, Guo X L, Lai M Z, Wang Q S. Air pollutant concentration forecasting using long short-term memory based on wavelet transform and information gain: a case study of Beijing. *Computational Intelligence and Neuroscience*, 2020, **2020**(1): 8834699
6. Zhang B, Zou G J, Qin D M, Ni Q, Mao H W, Li M Z. RCL-Learning: ResNet and convolutional long short-term memory-based spatiotemporal air pollutant concentration prediction

model. *Expert Systems with Applications*, 2022, **207**: 118017

7. Bai L, Liu Z, Wang J Z. Novel hybrid extreme learning machine and multi-objective optimization algorithm for air pollution prediction. *Applied Mathematical Modelling*, 2022, **106**: 177–198
8. Li D, Wang J, Tian D W, Chen C, Xiao X X, Wang L, Wen Z, Yang M Z, Zou G J. Residual neural network with spatiotemporal attention integrated with temporal self-attention based on long short-term memory network for air pollutant concentration prediction. *Atmospheric Environment*, 2024, **329**: 120531
9. Wan A P, Yang J, Chen T, Yang J X, Li K, Zhou Q L. Dynamic pollution emission prediction method of a combined heat and power system based on the hybrid CNN-LSTM model and attention mechanism. *Environmental Science and Pollution Research*, 2022, **29**(46): 69918–69931
10. An F R, Yang D, Sun X Y, Wei H B, Chen F L. A machine learning model integrating spatiotemporal attention and residual learning for predicting periodic air pollutant concentrations. *Environmental Modelling & Software*, 2025, **188**: 106438
11. Wang W J, Ma T Y, Wang L R. Air pollutant concentration prediction based on a new hybrid model, feature selection, and secondary decomposition. *Air Quality, Atmosphere & Health*, 2023, **16**(10): 2019–2033
12. Xiao Y J, Wang X K, Wang J Q, Zhang H Y. An adaptive decomposition and ensemble model for short-term air pollutant concentration forecast using ICEEMDAN-ICA. *Technological Forecasting and Social Change*, 2021, **166**: 120655
13. Wang X L, Tao L Z, Fu M L, Wang Q. Air pollutant concentration forecasting with WTMP: wavelet transform-based multilayer perceptron. *Atmosphere*, 2024, **15**(11): 1296
14. Wang J J, Xu W J, Dong J, Zhang Y. Two-stage deep learning hybrid framework based on multi-factor multi-scale and intelligent optimization for air pollutant prediction and early warning. *Stochastic Environmental Research and Risk Assessment*, 2022, **36**(10): 3417–3437
15. Dragomiretskiy K, Zosso D. Variational mode decomposition. *IEEE Transactions on Signal Processing*, 2014, **62**(3): 531–544
16. Li F H, Li R, Tian L L, Chen L, Liu J. Data-driven time-frequency analysis method based on variational mode decomposition and its application to gear fault diagnosis in variable working conditions. *Mechanical Systems and Signal Processing*, 2019, **116**: 462–479
17. Wang D M, Sun Y, Xiao J L, Lu J Y. Research on pipeline leakage signal denoising using variational mode decomposition and energy value. *Petroleum Science and Technology*, 2025, **43**(2): 202–218
18. Richman J S, Moorman J R. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology Heart and Circulatory Physiology*, 2000, **278**(6): H2039–H2049
19. Liang J Y, Lu Y X, Su M M. Hga-lstm: LSTM architecture and hyperparameter search by hybrid GA for air pollution prediction. *Genetic Programming and Evolvable Machines*, 2024, **25**(2): 20
20. Sang S, Li L. A novel variant of LSTM stock prediction method incorporating attention mechanism. *Mathematics*, 2024, **12**(7): 945
21. Gidon J S, Borah J, Sahoo S, Majumdar S, Fujita M. Bidirectional LSTM model for accurate and real-time landslide detection: a case study in Mawiongrim, Meghalaya, India. *IEEE Internet of Things Journal*, 2024, **11**(3): 3792–3800
22. Borah J, Nadzir M S M, Cayetano M G, Ghayvat H, Majumdar S, Srivastava G. Timezone-aware auto-regressive long short-term memory model for multipollutant prediction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2025, **55**(1): 344–352
23. De Vito S, Massera E, Piga M, Martinotto L, Di Francia G. On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario. *Sensors and Actuators B: Chemical*, 2008, **129**(2): 750–757
24. Borah J, Chakraborty T, Nadzir M S M, Cayetano M G, Benedetto F, Majumdar S. A novel hybrid approach for efficiently forecasting air quality data. *IEEE Sensors Letters*, 2025, **9**(1): 6001204
25. Chen D Y, Liu H, Li Y F, Duan Z. An improved deep learning model for soybean future price prediction with hybrid data preprocessing strategy. *Frontiers of Agricultural Science and Engineering*, 2025, **12**(2): 208–230
26. Su M S, Liu H, Yu C Q, Duan Z. A novel AQI forecasting method based on fusing temporal correlation forecasting with spatial correlation forecasting. *Atmospheric Pollution Research*, 2023, **14**(4): 101717
27. Rezaiy R, Shabri A. An innovative hybrid W-EEMD-ARIMA model for drought forecasting using the standardized precipitation index. *Natural Hazards*, 2024, **120**(14): 13513–13542
28. Lee H J, Kim Y S, Lee W S, Choi I H, Lee C K. RNN-based sequence-aware recommenders for tourist attractions. *CAAI Transactions on Intelligence Technology*, 2025, **10**(4): 1077–1088
29. Yang B F, Zhu X L, Liu Y T, Liu H X. A single-channel EEG based automatic sleep stage classification method leveraging deep one-dimensional convolutional neural network and hidden Markov model. *Biomedical Signal Processing and Control*, 2021, **68**: 102581
30. Acharya S, Kar T, Smal U C, Patra P K. Performance comparison between SVM and LS-SVM for rice leaf disease detection. *EAI Endorsed Transactions on Scalable Information Systems*, 2023, **10**(6): 1–7
31. Wang W Z, Wang Z W, Liu X, Li W J, Li Q F, Zhang Y G,

Chen Q C, Guo S Y, Xu Z. Frequency response mode prediction of power system after large disturbances based on deep belief neural network. *IEEE Access*, 2023, **11**: 113653–113666

32. Flores A, Tito-Chura H, Yana-Mamani V. An ensemble GRU approach for wind speed forecasting with data augmentation. *International Journal of Advanced Computer Science and Applications*, 2021, **12**(6): 569–574
33. Maity T, Bhawani A N, Samanta J, Saha P, Majumdar S, Srivastava G. MLSFDD: machine-learning-based smart fire detection device for precision agriculture. *IEEE Sensors Journal*, 2025, **25**(5): 8921–8928
34. Das R, Acharyya A, Majumdar S. Influences of human presence on the indoor air quality of educational institutions: concurrent multipollutant sensing approach. *IEEE Open Journal of Instrumentation and Measurement*, 2025, **4**: 9500508
35. Li H Y, Wang C J, Zhao D. Filter bank properties of envelope modified EMD methods. *IET Signal Processing*, 2018, **12**(7): 844–851
36. Liu H, Mi X W, Li Y F. Comparison of two new intelligent wind speed forecasting approaches based on Wavelet Packet Decomposition, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and Artificial Neural Networks. *Energy Conversion and Management*, 2018, **155**: 188–200
37. Xiong Z Q, Han C, Zhang G R. Fault diagnosis of anti-friction bearings based on Bi-dimensional ensemble local mean decomposition and optimized dynamic least square support vector machine. *Scientific Reports*, 2023, **13**(1): 17784
38. Chen X L, Chang H, Cao B N, Lu Y B, Lin D M. Prediction of continuous blood pressure using multiple gated recurrent unit embedded in SENet. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 2022, **26**(2): 256–263
39. Qi L, Qin X W, Gao F, Dong J Y, Gao X B. SAWU-Net: spatial attention weighted unmixing network for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 2023, **20**: 5505205