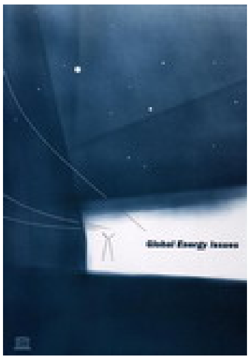


International Journal of Global Energy Issues

ISSN online: 1741-5128 - ISSN print: 0954-7118
https://www.inderscience.com/ijgei

A digital technology for energy-saving ventilation control in underground infrastructures: integrating coupled simulation and BP algorithm

Jing Li, Bo Wang, Tao Song

DOI: 10.1504/IJGEI.2026.10076331

Article History:

Received:	30 October 2025
Last revised:	16 January 2026
Accepted:	19 January 2026
Published online:	09 March 2026

A digital technology for energy-saving ventilation control in underground infrastructures: integrating coupled simulation and BP algorithm

Jing Li*, Bo Wang and Tao Song

Rocket Force University of Engineering,
Xi'an City, China
Email: 79062289@qq.com
Email: 53754924@qq.com
Email: songtriver@163.com
*Corresponding author

Abstract: To address the challenges of energy inefficiency and delayed response in ventilation management of energy-intensive underground infrastructures, this study proposes a novel digital technology named Spatial Temporal Attention-based Back Propagation (STABP), which integrates coupled simulation and BP algorithm. It aims to achieve accurate prediction and dynamic regulation of ventilation demand in underground spaces, thereby optimising energy use. This study constructs a cross-disciplinary technological framework that integrates physical process simulation with data-driven algorithms, representing a novel approach to managing technological complexity in building energy systems. Firstly, based on Kriging interpolation method, spatial reconstruction is performed on limited sensor data to generate high-resolution gridded pollutant concentration fields, which serve as input boundary conditions for coupled simulation. Then, the BP algorithm is used for rapid dimensionality reduction and error compensation of high-dimensional spatiotemporal fields. Time series data is processed using Long Short-Term Memory (LSTM).

Keywords: energy-intensive underground infrastructures; energy-saving ventilation control; coupled simulation; BP algorithm; spatio-temporal attention; LSTM; technology progress; digital energy; entrepreneurial strategy.

Reference to this paper should be made as follows: Li, J., Wang, B. and Song, T. (2026) 'A digital technology for energy-saving ventilation control in underground infrastructures: integrating coupled simulation and BP algorithm', *Int. J. Global Energy Issues*, Vol. 48, No. 7, pp.21–40.

Biographical notes: Jing Li received the master's degree from Xi'an University of Architecture and Technology, P.R. China. Now, she works in Rocket Force University of Engineering, and her research interests include heat and moisture transfer laws, indoor air quality control technology in underground engineering.

Bo Wang received the doctor's degree from PLA University of Science and Technology, P.R. China. Now, he works in Rocket Force University of Engineering, and his research interests include heat and moisture transfer laws, indoor air quality control technology in underground engineering.

Tao Song received his PhD from Harbin Institute of Technology, China, in 2021. He is currently working at the Rocket Force University of Engineering. His research interests include nanoparticle filtration, hydrocyclone separation, and heat pump technology.

1 Introduction

Subways, underground commercial centres, underground parking lots and underground industrial workshops are just a few examples of the deep underground infrastructure that has become an essential component of contemporary cities. They not only provide convenience for urban transportation and commercial activities, but also alleviate the tension of ground space to a certain extent (Demirkan et al., 2022; Li et al., 2023; Zhang et al., 2022). However, the enclosed nature of deep underground engineering makes it an energy-intensive environment, where ventilation systems account for a significant portion of total energy consumption, presenting an urgent issue for energy management. According to statistics from the International Energy Agency (IEA), the utilisation rate of global urban underground space has increased by approximately 35% over the past decade. It is projected that by 2030, the usable area of urban underground space will exceed 5 billion square meters. However, the enclosed nature and limited ventilation conditions of underground space pose dual challenges to air quality and energy management. On one hand, due to the low efficiency of ventilation systems, pollutants in underground space tend to accumulate, threatening the health of users. During peak hours in subway stations, the CO_2 concentration can rapidly rise from 400 ppm to 2000 ppm within 10–15 minutes due to dense passenger flow. Traditional ventilation systems often take hours to respond, resulting in passenger discomfort and energy waste. On the other hand, as the main source of energy consumption in underground space, ventilation systems account for 30–40% of total energy consumption. Traditional management methods lead to severe energy waste. Studies have shown that the energy waste rate of ventilation systems in underground space is as high as 25–35%, which increases operating costs and conflicts with the dual carbon goals.

Deep underground engineering has less air circulation than the outdoors, making it more likely for contaminants such carbon dioxide, volatile organic compounds and particulate matter to build up (Du et al., 2023). Elevated concentrations of these pollutants can have adverse effects on human health, easily causing respiratory diseases, dizziness and other symptoms and may even be life-threatening in severe cases. Wang et al. (2023) pointed out that this policy had significantly promoted the green development of the digital industry, especially in smart city construction, where digital technologies had improved resource utilisation efficiency and achieved precise control of pollution emissions. The management of these energy-intensive systems also faces problems such as operational inefficiency and strong regulatory lag, leading to unnecessary energy waste. Traditional air quality monitoring methods rely on limited fixed monitoring points, making it difficult to fully and real-time reflect the spatiotemporal distribution characteristics of pollutants in underground spaces. Moreover, experience-based regulation strategies often fail to respond promptly to changes in pollutant concentrations, resulting in poor regulation effects (Arroyo et al., 2022). The Back Propagation (BP) Neural Network is a classic feedforward neural network. It optimises weight parameters through the back propagation algorithm and can effectively handle nonlinear mapping relationships.

Traditional ventilation control methods rely on fixed sensor network layouts and preset fixed-threshold control strategies, which result in significant response delays. Specifically, fixed sensor networks cannot dynamically adapt to changes in the indoor environment and fail to promptly capture sudden changes in pollutant concentrations. When a high concentration of pollutants suddenly emerges from a new pollution source in a certain area, the fixed sensor network needs to wait for data to be transmitted from the monitoring point to the control centre, which then issues instructions. This process usually takes 3–5 minutes, during which pollutants may have spread to a wider range. In addition, traditional methods lack an adaptive mechanism for spatiotemporal dynamic changes and cannot adjust ventilation strategies in real time based on the physical characteristics of pollutant diffusion. In underground space environments, air flow is restricted by structures and pollutant diffusion paths are complex. Traditional methods cannot accurately predict pollutant propagation trends, leading to the failure of ventilation systems to respond to pollution incidents in a timely manner. In recent years, the Spatiotemporal Attention (STA) mechanism has been a prominent area of study in deep learning. When processing data with spatio-temporal characteristics, traditional neural networks often struggle to effectively capture important information in both spatial and temporal dimensions simultaneously. By introducing attention weights, the STA mechanism can dynamically allocate the model's degree of focus on different spatial locations and time steps, thereby better extracting key features from the data (Liu et al., 2024; Livingston et al., 2023). This study develops a STA method to further improve the model's capacity to concentrate on important spatiotemporal aspects. Through dynamic weight allocation, this mechanism enables the model to adaptively focus on regions with complex pollution source distributions or time periods with significant fluctuations in emission intensity.

Air quality management in underground spaces faces two key challenges: First, traditional monitoring methods rely on limited fixed monitoring points, failing to reflect the spatiotemporal distribution characteristics of pollutants in underground spaces in real time and comprehensively. Second, empirical-based ventilation adjustment strategies cannot respond promptly to changes in pollutant concentrations, resulting in energy waste. Traditional air quality monitoring methods depend on fixed monitoring stations and intermittent sampling, which cannot capture sudden changes in pollutant concentrations in a timely manner. In pollution incidents caused by sudden industrial emissions or traffic accidents, the monitoring data from fixed stations often lags behind the actual occurrence time of pollution, leading to the failure of emergency response measures to be activated promptly. In addition, for pollution monitoring in underground spaces, traditional methods struggle to accurately reflect the complex distribution of pollution sources. Owing to the structural constraints on air flow in underground spaces, pollutant diffusion paths are complex, resulting in significant deviations in evaluations. This study aims to propose a Spatial Temporal Attention based BP (STABP) air quality prediction method and apply it to the health diagnosis of air quality in deep-buried underground projects. This technology integrates coupled simulation and BP algorithm, aiming to achieve accurate prediction and dynamic adjustment of ventilation demand in underground spaces, thereby optimising energy use. Firstly, multiple sensor nodes are deployed inside deep underground engineering to collect Air Quality Index (AQI) data and related environmental parameters. Then, pre-processing is performed on the collected data, including data cleaning and normalisation, to eliminate noise and outliers in the data and improve data quality. Next, the STABP model is constructed. In this model, time series data is processed using the Long Short-Term Memory (LSTM) network, spatial characteristics are extracted using the BP model, and attention weights are dynamically allocated using the STA mechanism. By learning from historical data, the model can establish complex relationships between air quality indices and environmental factors, and predict air quality at future moments.

2 Method

2.1 Foundation of air quality prediction for deep engineering

Compared with the outdoor environment, deep underground engineering usually has limited air exchange with the outside world, making it easy for pollutants to accumulate inside and difficult to diffuse quickly. This enclosed nature leads to significant differences in the dynamic changes of air pollution compared with the external environment. In subway stations, the carbon dioxide concentration in crowded areas may rise rapidly in a short time, and the adjustment of ventilation systems can effectively improve air quality only based on accurate predictions (Ali et al., 2022; Dong et al., 2023). The ventilation system is an important means to regulate the air quality of deep underground engineering, and its operating status directly affects the distribution and diffusion of air pollution. Air quality evaluation standards are important criteria to measure whether the air quality of deep underground engineering meets the standards. At present, widely adopted international air quality evaluation standards include the ‘WHO Guidelines for Indoor Air Quality’ and the standards of the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE). In China, the air quality of deep underground engineering mainly refers to ‘Standard for Air Quality of Underground Buildings’ (GB/T 18204.2) and ‘Code for Design of Metro’ (GB 50157). The pollutant concentration limits corresponding to the AQI are shown in Table 1. Among them, PM2.5 is one of the core pollutants in AQI calculation. Its concentration directly reflects the pollution degree of fine particulate matter (with a diameter of $\leq$2.5 μm) and is closely related to respiratory diseases. PM10 refers to inhalable particulate matter (with a diameter of $\leq$10 μm), which has potential hazards to the respiratory and cardiovascular systems.

Table 1 Pollutant concentration limit corresponding to AQI

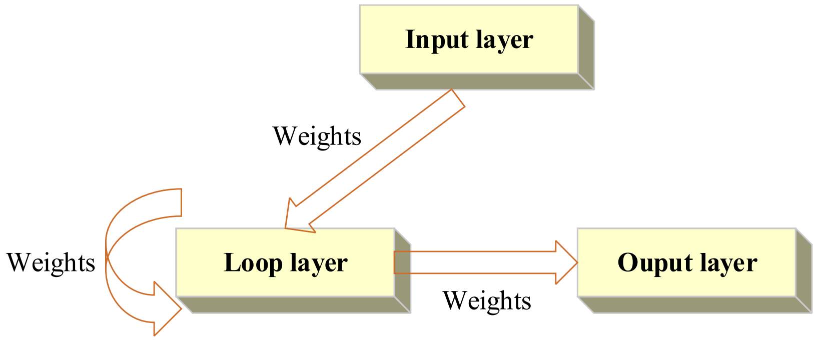
AQI	SO_2	NO_2	*CO*	O_3	*PM2.5*	*PM10*
0	0	0	0	0	0	0
50	50	40	2	100	35	50
100	150	80	4	160	75	150
150	475	180	15	245	115	250
200	800	280	24	265	150	350
300	1600	565	36	800	250	420

The Air Pollution Index (API) is a comprehensive indicator reflecting air quality conditions, typically used for outdoor air quality assessment. However, its basic principles can also be applied to the evaluation of air quality in deep underground engineering (Ge et al., 2024; Thanh et al., 2024; Wang et al., 2024b). A complete index is created through weighted summation, and the API calculation technique is based on the ratio of the concentration of key pollutants to their respective limitations. According to the API value, air quality can be split into multiple grades: Excellent air quality with no discernible effects on human health is achieved when the API is less than 50. Good air quality is achieved when the API is between 50 to 100, which may have minor effects on sensitive populations. The air quality is mildly contaminated when the API is between 100 and 150, which clearly affects sensitive populations. The population's health is significantly impacted by moderately polluted air when the API is between 150 and 200. The population's health is seriously impacted by the highly contaminated air quality when the API is 200 or above. In deep underground engineering, due to the enclosed space and complex pollution sources, air quality can easily reach the level of light pollution or even moderate pollution.

2.2 *Processing of spatio-temporal sequence information of air pollutants*

The topological topology of the Recurrent Neural Network (RNN) is depicted in Figure 1. Its ability to retain sequence information in memory is based on its recurrent structure. The RNN changes the hidden state by learning weights at each time step after receiving the current input and the hidden state from the previous time step. Then, it moves on to the subsequent time step. The current time step's components are often included in the RNN's input. The network can receive and process sequential data, including time series, natural language texts and other data with temporal correlation, thanks to the recurrent layer. The recurrent layer creates a new hidden state at each time step by receiving the hidden state from the previous time step as well as the current input. The network can identify temporal dependencies in the sequence thanks to this information transfer and memory mechanism.

Figure 1 RNN topology (see online version for colours)



RNN takes an input vector sequence $(x_1, x_2, ..., x_t)$ and extracts another vector sequence $(s_1, s_2, ..., s_t)$ to represent the hidden layer state. The hidden layer state can be expressed as:

$$net_s^t = Ux_1 + Ws_{t-1} + b_s \tag{1}$$

$$s_t = f\left(net_s^t\right) \tag{2}$$

The output sequence is represented by the hidden layer state, which can be expressed as:

$$net_y^t = Vs_t + b_y \tag{3}$$

$$y_t = g\left(net_y^t\right) \tag{4}$$

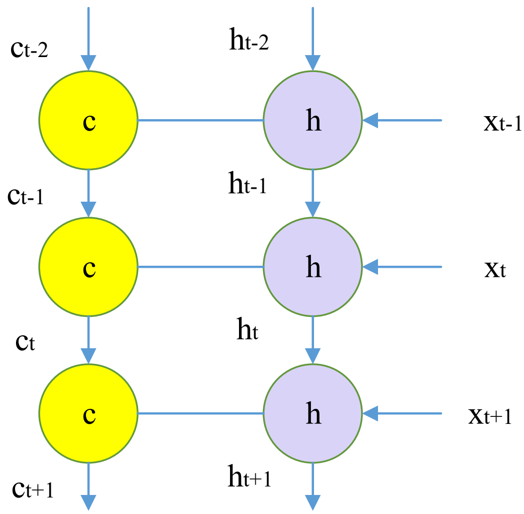
net_s^t and net_y^t represent weighted sums corresponding to pre-activation. $f(\bullet)$ and $g(\bullet)$ represent activation functions.

The current input and the hidden state from the previous time step are fed into the RNN at each time step of the training process. It uses the current input and the hidden state to determine the output of the current time step. The subsequent time step's input is derived from the current time step's output. Until the complete sequence is digested, this process is repeated.

RNN calculates the hidden state h_{t-1} at the current moment by combining the hidden state x_t at the previous moment with the input h_t at the current moment. This structure enables RNN to process time-dependent sequence data, such as time series and language series.

RNN faces issues of gradient vanishing and gradient explosion when processing long-sequence data. Gradient vanishing refers to the gradual reduction of gradient values during backpropagation, making it difficult for the network to learn long-distance dependencies. In contrast, gradient explosion refers to the gradual increase of gradient values, leading to unstable network training. These problems limit the effectiveness of RNN in practical applications (Guo et al., 2023; Radhakrishnan et al., 2022; Ramadan et al., 2024). LSTM solves the long-distance dependency problem effectively by introducing a gating mechanism to control the flow of information. The hidden layer of RNN, which is the LSTM neural network's module, primarily gains a control layer from the LSTM neural network. The conventional hidden layer can function using short-term memory if it contains a single state (the h layer). To preserve long-term states, the LSTM neural network has a control layer or c layer. A schematic diagram of the LSTM neural network's expansion over time is displayed in Figure 2.

Figure 2 Schematic diagram of time expansion of LSTM neural network (see online version for colours)



In LSTM, the role of forgetting gate is to decide which information to discard from the cell state. The calculation formula of forget gate is:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{5}$$

f_t is the output of forget gate, indicating the proportion of information to be forgotten. σ is the sigmoid activation function, and the output value is between 0 and 1. W_f is the weight matrix of forget gate, and b_f is the bias term.

The function of the input gate is to decide which new information needs to be written into the cell state. The input gate consists of two parts: the activation value i_t of the input gate and the candidate value $\tilde{C}_t$. The calculation equation is:

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{6}$$

i_t is the activation value of the input gate, indicating the proportion of information to be written. $\tilde{C}_t$ is a candidate value, indicating new information to be written. W_i and W_C are the weight matrices of input gates and candidate values. b_i and b_C are offset terms.

Cell state is the core of LSTM, which updates information through the control of forget gate and input gate. The updated equation of cell state is:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{7}$$

C_t is the cell state at the current moment. C_{t-1} is the cell state at the previous moment.

The function of the output gate is to decide what information to output from the cell state. The calculation equation of the output gate is:

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{8}$$

$$h_t = o_t \cdot \tanh\left(C_t\right) \tag{9}$$

o_t is the activation value of the output gate, indicating the proportion of information to be output. h_t is the hidden state of the current moment. W_o is the weight matrix of the output gate, and b_o is the bias term.

In the air quality management of deep underground engineering, the spatiotemporal evolution of pollutant concentrations exhibits high nonlinearity and uncertainty. Examples include the local concentration of exhaust emissions from construction machinery, spatial gradient changes caused by the operation of ventilation systems, and sudden pollution events triggered by fluctuations in the intensity of construction activities. Traditional models usually adopt fixed weights or global averaging strategies, making it difficult to effectively capture these dynamic features. The spatiotemporal attention mechanism greatly improves prediction accuracy and regulatory sensitivity by allowing the model to dynamically adapt the intensity of attention to various spatial regions and time periods depending on real-time data through the introduction of learnable weight parameters.

The goal of spatial attention mechanism is to extract the contribution weight of key areas from high-dimensional spatial features. Assume that the input spatial feature matrix is $X \in R^{N \times D}$, where N is the number of grid points and D is the feature dimension, such as pollutant concentration, temperature and humidity. The calculation equation of spatial attention weight $A_s \in R^{N \times N}$ is as follows:

$$A_s = softmax\left(\frac{XW_q XW_k^T}{\sqrt{d_k}}\right) \tag{10}$$

W_q and W_k are the spatial query matrix and bond matrix respectively, and d_k is the scaling factor, which is used to calculate the stable gradient. Through this equation, the model can automatically identify the hot spots of pollutant concentration distribution and give them higher weight.

The goal of temporal attention mechanism is to extract the dependence of key time steps from historical time series. Assume that the time feature sequence output by the LSTM module is $H = \left[h_1, h_2, \ldots, h_T\right] \in R^{T \times D}$, where T is the number of time steps. The calculation equation of time attention weight A_t is as follows:

$$A_t = softmax\left(\frac{HW_q HW_k^T}{\sqrt{d_k}}\right) \tag{11}$$

W_q and W_k are time query matrix and key matrix, respectively. Through this equation, the model can capture the key events of pollutant concentration changing with time and dynamically adjust its influence on the prediction results.

2.3 An air pollution prediction model based on integrated coupled STABP

The STABP model is an integrated framework that combines physical process simulation with data-driven algorithms to achieve accurate prediction and dynamic adjustment of ventilation demand in underground spaces. The model consists of five core components: spatial data reconstruction based on Kriging interpolation, high-dimensional spatiotemporal field processing using BP Neural Network (BPNN), time series data processing via Long-Short-Term Memory (LSTM), dynamic weight assignment based on spatiotemporal attention mechanism, and pollutant concentration prediction. It follows a sequential processing workflow: first, spatial data reconstruction generates a high-resolution pollutant concentration field. Second, the BPNN performs rapid dimensionality reduction and error compensation for the high-dimensional spatiotemporal field. Third, LSTM processes time series data; fourth, the spatiotemporal attention mechanism dynamically allocates weights to focus on critical spatiotemporal information. Finally, the model outputs the predicted pollutant concentration.

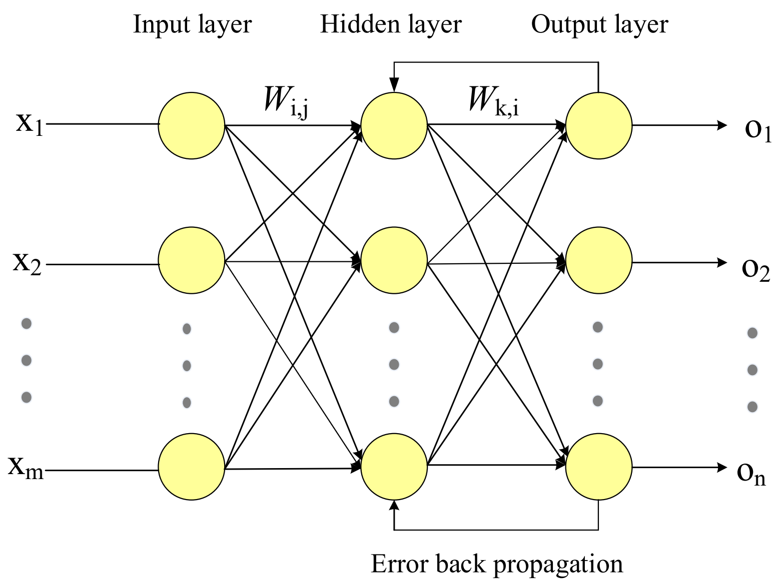
In deep underground engineering, the deployment of sensors is usually limited, which leads to insufficient spatial coverage of data (Wang et al., 2024a). To expand the spatial characteristics of the data, this study first adopts the Kriging interpolation method for spatial interpolation in unknown research areas. Kriging interpolation is a geostatistical method based on spatial autocorrelation, which can estimate the values of unknown points based on data from known points. Through Kriging interpolation, a denser spatial data grid can be generated, thereby providing more comprehensive spatial information for subsequent models. The core of Kriging interpolation lies in measuring the autocorrelation of spatial data through the semi variogram. The semi variogram describes the spatial dependence between data points; the smaller its value, the stronger the correlation between data points. The specific expression of the Kriging interpolation method can be stated as follows:

$$Z^*(x_0) = \mu + \sum_{i=1}^{n} \lambda_i \left(Z(x_i) - \mu \right) \tag{12}$$

$Z^*(x_0)$ is the predicted value of position x_0. μ is the average value of the whole area, $Z(x_i)$ is the observed value of known position x_i, and λ_i is the weight coefficient.

As a core component, the BP Neural Network (BPNN) undertakes the key task of extracting critical features from high-dimensional spatial data. BPNN architecture is displayed in Figure 3. Its design goal is to convert the high-resolution gridded pollutant concentration fields generated by Kriging interpolation into low-dimensional and discriminative spatial feature representations through the nonlinear mapping capability of multi-layer perceptron, thereby providing basic support for the generation of subsequent dynamic regulation strategies. The design of BPNN is mainly based on the following considerations: (1) The input feature dimension is 12. According to the empirical rule, the number of hidden layer nodes is usually set to 1.5-2 times that of input nodes, so the number of nodes in the first hidden layer is set to 24. (2) The number of nodes in the second hidden layer is set to 2/3 of that in the first layer, i.e., 16 nodes, which aims to gradually reduce the feature dimension and extract higher-level feature representations. (3) The rationality of this structure is verified through ablation experiments.

Figure 3 BPNN architecture (see online version for colours)



In this study, the input layer receives the pollutant concentration field data generated by Kriging interpolation, with a dimension of $N \times D$, where N represents the number of spatial grid points and D represents the feature dimension of each grid point. The design of the input layer must fully consider the sparsity and non-uniformity of sensor data in deep underground engineering scenarios. Therefore, normalisation processing is adopted to eliminate the influence of different feature dimensions, and missing data is supplemented through filling strategies to ensure the integrity of the input data. The goal of the output layer is to generate a dimensionally reduced spatial feature matrix $H_s \in R^{N \times d}$, where d is the dimension after feature dimensionality reduction. This matrix realises the abstract representation of the spatial distribution pattern of pollutant concentrations by linearly combining the outputs of the hidden layers and incorporating activation functions such as Softmax or Tanh.

The forward transmission of input and the backward transmission of errors are the two fundamental operations that make up the computation of the BPNN algorithm. The existing weights and thresholds are used to determine how input data is sent. Equation (13) displays the output of the *k*-th node in the hidden layer.

$$H_k = \phi\left(\sum_{j=1}^{m} \omega_{ij} x_j + \theta_i\right) \tag{13}$$

ϕ represents the transfer function of the hidden layer. ω_{ij} represents the weight, and θ_i represents the node threshold of the hidden layer.

The output of the *s*-th node in the output layer is shown in equation (14).

$$o_s = \psi\left[\sum_{i=1}^{q} \omega_{sj} \phi\left(\sum_{j=1}^{m} \omega_{ij} x_j + \theta_i\right) + a_s\right] \tag{14}$$

ψ represents the transfer function of the output layer, usually using a linear function. ω_{ij} represents the weight between the hidden layer and the input layer. a_s represents the node threshold of the output layer.

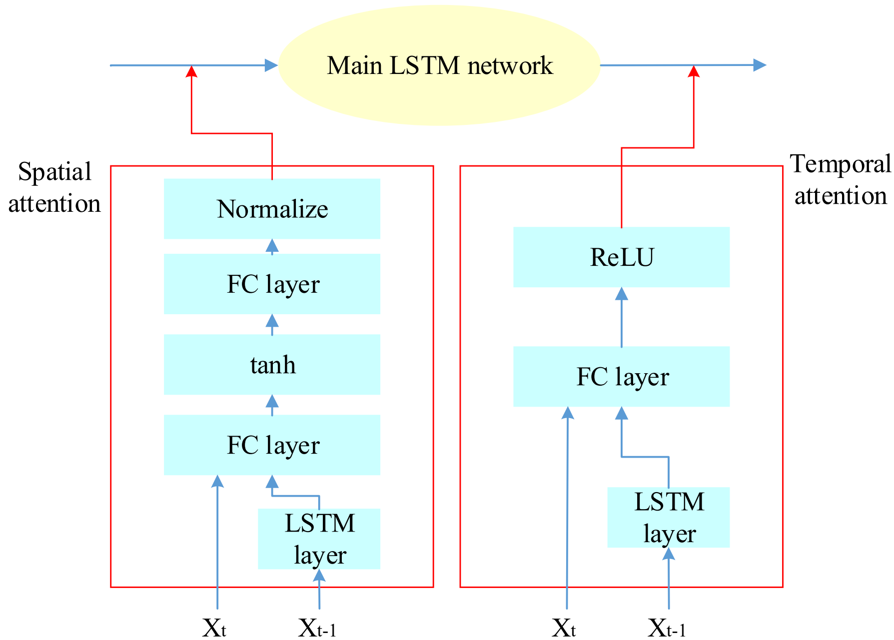
The BPNN is trained by calculating the output error based on the expected and actual outcomes, and then transmitting the result backward. The gradient descent method involves continuously adjusting the network's weights and thresholds to bring the output closer to the reference value over time. In equation (15), the error function is displayed.

$$E_p = \frac{1}{2}\sum_{k=1}^{n}\left(T_k - o_k\right)^2 \tag{15}$$

T_k represents the output reference value of the *k*-th node.

Each convolutional layer outputs a set of feature maps, which represent the features of the input data at different spatial scales (Fahad et al., 2023; Mutovina et al., 2025; Pérez-Beltrán et al., 2024). The model's middle layer uses a stacked Spatio-Temporal Attention-Long-Short-Term Memory network (STA-LSTM) after spatial features have been extracted. Through the attention mechanism, STA-LSTM dynamically distributes the degree of attention to various spatial locations and time steps at each time step, in addition to taking into account the input at that instant and the hidden state at that moment. The model can more easily capture the spatiotemporal dependencies in the data thanks to this mechanism. The STA mechanism is the core part of STA-LSTM. It dynamically distributes the model's attention to different spatial positions and time steps by calculating attention weights.

Figure 4 STA-LSTM architecture (see online version for colours)



In the STABP air pollution prediction model, the Spatial Attention Module (SAM) and Temporal Attention Module (TAM) are illustrated in Figures 5 and 6, respectively. In areas close to pollution sources, changes in pollutant concentrations may be more significant, whereas in areas far from pollution sources, such changes may be relatively minor (Rad et al., 2022; Tripathi et al., 2024; Zhao et al., 2024). In the STABP model, the spatiotemporal attention mechanism is integrated in the feature extraction stage, located after the BPNN extracts spatial features and before the LSTM processes time series data. The SAM calculation is based on the spatial features extracted by BPNN. Its computation process includes a series of convolutional layers, followed by a softmax function for weight normalisation (Sharifi et al., 2023). The weights are multiplied by the spatial features to emphasise critical spatial regions. SAM is applied to the output of BPNN, which is then fed into the LSTM layer. TAM calculation is based on the time series features extracted by LSTM. Its computation process also includes a series of convolutional layers, followed by a softmax function for weight normalisation. The weights are multiplied by the time series features to emphasise critical time steps. TAM is applied to the output of LSTM, which is then fed into the final prediction layer. During training, the attention weights are optimised through backpropagation, similar to other parameters of the model. The attention mechanism is trained end-to-end with the rest of the network, enabling the model to adaptively focus on the most relevant spatial locations and time points based on input data. This dynamic allocation of attention weights allows the model to better capture the complex spatiotemporal patterns of pollutant concentrations, especially during sudden changes in pollutant concentrations. Through dynamic allocation of temporal attention weights, TAM allows the model to focus more on time steps that are more crucial for prediction.

Figure 5 SAM module (see online version for colours)

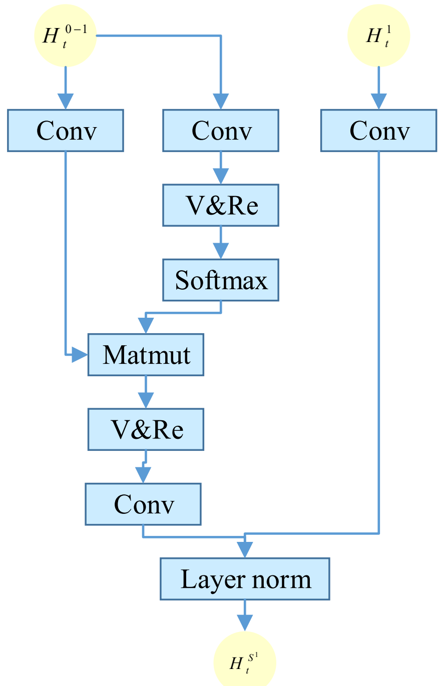
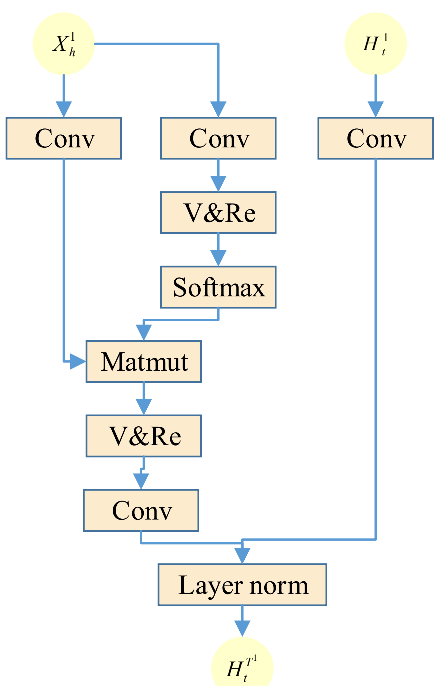


Figure 6 TAM module (see online version for colours)



The design of the spatiotemporal attention mechanism fully considers the physical properties of different pollutants. For CO, which spreads rapidly and is highly affected by air flow, the attention mechanism assigns higher weights to the temporal dimension, enabling the model to pay more attention to recent concentration changes. Specifically, the temporal attention weight coefficient for CO is set to 0.75, and the spatial attention weight coefficient is 0.25. For $PM_{2.5}$, which spreads slowly and is easily affected by meteorological conditions, the attention mechanism assigns higher weights to the spatial dimension, with the temporal attention weight coefficient set to 0.4 and the spatial attention weight coefficient to 0.6. In model implementation, these weight coefficients are determined through physical property analysis and automatically optimised through backpropagation during training.

2.4 Experimental design

This study selects Jiangsu Province as the research area to verify the effectiveness of the STABP model. The data set covers air pollution data and meteorological data recorded every 3 hours by monitoring stations in various cities of Jiangsu Province from 1 January 2022 to 31 December 2024. These data are freely available through public channels: air quality data are obtained from the China National Environmental Monitoring Centre (http://www.cnemc.cn/), and meteorological data are from the European Centre for Medium-Range Weather Forecasts (https://www.ecmwf.int). The STABP model is trained using air quality data and meteorological data recorded every 3 hours at monitoring stations in various cities of Jiangsu Province from 1 January 2022 to 31 December 2024. The data preprocessing steps are as follows: (1) *Data cleaning*: The 3σ criterion is used to remove outliers and missing values – any data point outside three standard deviations from the mean is considered an outlier and eliminated. (2) *Data normalisation*: Min-max normalisation is applied to scale all features to the range [0, 1]. (3) *Data splitting*: The data set is divided into a training set (70%), a validation set (15%) and a test set (15%) in chronological order to avoid data leakage.

The model is trained using the Adam optimiser with a learning rate of 0.001, a batch size of 64 and 500 iterations. The training process is monitored using a validation set to prevent overfitting. An early stopping strategy is implemented with a patience of 50 iterations, meaning training stops if the validation loss does not improve for 50 consecutive iterations. To evaluate model performance, 5-fold cross-validation is used to ensure the robustness of results. In each fold, 80% of the training data is used for training and 20% for validation. The final model is selected based on the average performance across all folds. Model parameters are adjusted via a grid search method, testing combinations of different learning rates (0.0001, 0.0005, 0.001, 0.005), batch sizes (32, 64, 128) and iteration counts (200, 400, 600). The optimal combination is determined as a learning rate of 0.001, a batch size of 64 and 500 iterations.

To guarantee data quality and applicability, the original data undergoes the required preparation steps prior to model training. First, outliers and missing values are eliminated by data cleaning. Outliers can be brought on by monitoring errors or severe weather, whereas missing results can be the consequence of malfunctioning monitoring equipment or problems with data transmission. By guaranteeing the data's correctness and integrity, data cleansing offers a solid data base for further analysis. Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are often used evaluation metrics in air pollution prediction. These three metrics are used in this study to compute prediction errors, which are then compared to actual values to assess the STABP model's prediction accuracy. Let $\hat{n}_i$ denote the predicted value for a certain future time period, n_i denote the actual observed value and m denote the number of samples. The calculation methods of the three metrics are as follows:

$$MAE\left(x,\hat{x}\right)=\frac{1}{m}\sum_{i=m}^{m}\left|x_i-\hat{x}_i\right| \tag{16}$$

$$MAPE\left(x_i,\hat{x}_i\right)=\frac{1}{m}\sum_{i=m}^{m}\frac{\left|x_i-\hat{x}_i\right|}{\hat{f}_i} \tag{17}$$

$$RMSE\left(x_i,\hat{x}_i\right)=\sqrt{\frac{1}{m}\sum_{i=m}^{m}\left|n_i-\hat{n}_i\right|^2} \tag{18}$$

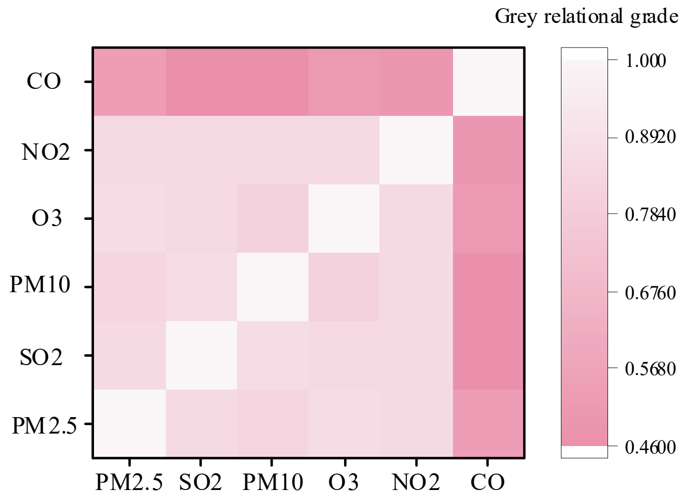
3 Results and discussion

3.1 The correlation analysis of air pollutants

In this study, air pollution concentrations are taken as the reference sequence, while meteorological data and concentrations of other pollutants serve as comparison sequences. By analysing the grey relational degree indices among multiple sets of data sequences, a set of quantitative correlation coefficients can be obtained. The value of these coefficients intuitively reflects the closeness of the relationship between the compared sequence and the reference sequence: a higher calculated value indicates a stronger correlation between the two sequences. This quantitative analysis method based on grey relational theory provides a reliable mathematical basis for evaluating the dynamic relationships between different variables. Figure 7 shows that there are strong correlations between certain characteristic variables, and these correlations may have important indicative significance in air pollution prediction models. Feature combinations with high correlation degrees may indicate common pollution sources or similar physicochemical processes, and such information is of great reference value for understanding the complex mechanisms of air pollution and optimising prediction models. The analysis reveals that CO has a moderate negative correlation with NO_2 and O_3, and a strong positive correlation with PM10. PM2.5 has strong positive correlations with other pollutants such as PM_{10}, SO_2 and NO_2, indicating that they share common pollution sources or occur under similar environmental conditions.

Analysis of the correlations between major pollutants (including CO, $PM_{2.5}$ and NO_2) reveals that the correlation coefficient is 0.65 between CO and $PM_{2.5}$, 0.72 between CO and NO_2 and 0.81 between PM2.5 and NO_2. Based on these correlation analyses, a multi-pollutant coupled input strategy is designed: pollutant pairs with a correlation coefficient greater than 0.7 are combined into feature pairs and fed into the input layer of the model, while pollutants with a correlation coefficient less than 0.7 are used as individual input features. This feature engineering strategy enables the model to more effectively capture the interactive relationships between pollutants. In specific implementation, pollutant pairs with a correlation coefficient greater than 0.7 are treated as input feature pairs, and each feature pair contains two pollutant concentration values to form new input features.

Figure 7 Correlation analysis of pollutant data (see online version for colours)



3.2 *The prediction performance analysis of STABP model*

Figure 8 compares the PM2.5 concentrations predicted by the STABP model with the actual values. It shows that the STABP model exhibits good performance in capturing the fluctuation trends of PM2.5 concentrations. In most cases, the fluctuation trends of predicted values are consistent with those of actual values. Particularly during periods when PM2.5 concentrations show significant fluctuations, the model can predict these changes relatively accurately. This shows that the STABP model can efficiently learn the spatiotemporal features of changes in PM2.5 concentrations, resulting in a high prediction accuracy.

Figure 8 Prediction effect based on STABP model (see online version for colours)

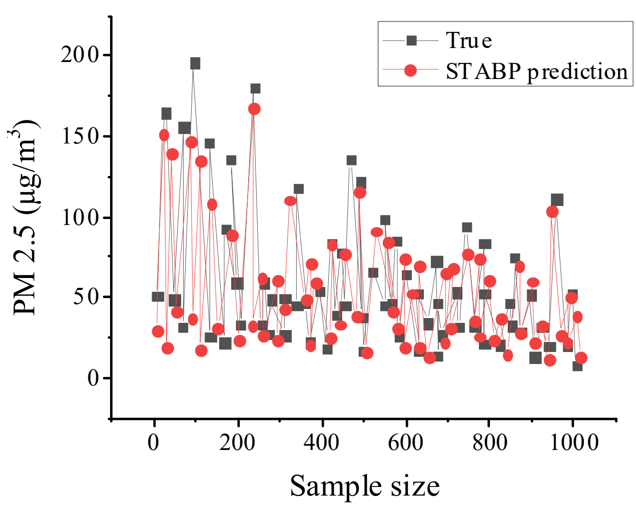
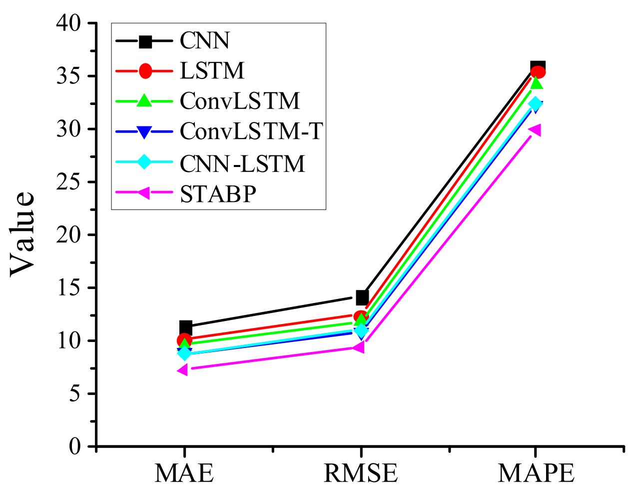


Figure 9 compares the prediction performance of different models. Among the comparative models, the CNN model consists of 3 convolutional layers (kernel size 3×3, output channels 32, 64, 128, respectively) and 2 fully connected layers (256 and 128 nodes, respectively). The LSTM model is configured with 2 LSTM layers (128 nodes per layer) and a dropout rate of 0.2, designed for time series prediction. The ConvLSTM model includes 2 ConvLSTM layers (64 nodes per layer, kernel size 3×3). The ConvLSTM-T model comprises 2 ConvLSTM layers (64 nodes per layer, kernel size 3×3) plus 1 temporal convolutional layer. The CNN-LSTM model follows the pipeline: 3 CNN layers (output channels 32, 64, 128) → 2 LSTM layers (128 nodes per layer). All comparative models use the Adam optimiser with a learning rate of 0.001. The proposed STABP hybrid model outperforms other models in terms of indicators such as MAE, RMSE and MAPE, with values of 7.31, 9.43 and 29.96, respectively. Although LSTM has the ability to process time series data, its prediction effect when used alone is not ideal. This result further confirms the importance of introducing temporal and spatial attention mechanisms for learning time series data. In addition to better extracting temporal and spatial information from time series data, these mechanisms also improve neural networks' capacity for learning, which raises the model's capacity for prediction.

Figure 9 Comparison of prediction performance of different models (see online version for colours)



4 Conclusions

This study provides a scalable technological solution for energy entrepreneurs and firms aiming to innovate in the smart building and urban energy management sector. This study addresses the problems of energy inefficiency and operational lag in managing urban underground infrastructures of deep underground engineering by proposing an STABP model that integrates coupled simulation and BP algorithm. Through the collaborative optimisation of physical process simulation and data-driven algorithms, this model realises accurate prediction and dynamic regulation of pollutant concentrations in underground spaces. The STABP model first performs spatial reconstruction on limited sensor data based on the Kriging interpolation method to generate a high-resolution gridded pollutant concentration field. Then, a BPNN is used for rapid dimensionality reduction and error compensation of the high-dimensional spatiotemporal field, combined with an LSTM module to process time series information. Lastly, to improve the model's capacity to concentrate on important spatiotemporal information, the STA mechanism dynamically distributes weights. Experiments using Jiangsu Province's air quality and meteorological data from 2022 to 2024 confirmed the STABP model's superior performance in PM2.5 prediction. Compared to standard models, its MAE, RMSE and MAPE indices are much improved at 7.31, 9.43 and 29.96, respectively. The results show that the STA mechanism effectively improves the model's adaptability to the complexity of pollution source distribution and fluctuations in emission intensity, and exhibits higher sensitivity especially in capturing sudden changes in pollutant concentrations. The STABP model exemplifies how digital transformation (AI integration) can drive business model innovation for clean energy services in urban contexts.

In practical applications, a high-density sensor network is deployed in underground projects, including spatially distributed pollutant sensors and meteorological sensors, to collect air quality data in real time. Next, the collected data undergoes preliminary processing via edge computing devices, including data cleaning, normalisation and feature extraction. The processed data is then input into the STABP model for real-time prediction. Finally, based on the prediction results, the operating status of ventilation equipment is automatically adjusted, and actual ventilation effect data is collected to support continuous optimisation and updating of the model. For specific underground engineering applications, it is recommended to first conduct small-scale pilots in key locations such as subway stations. 3–5 typical areas are selected as pilot sites, deploy sensor networks and ventilation control systems, collect 3 months of operational data and gradually expand the application scope after verifying the model's effectiveness.

Despite the promising performance of the STABP model, several aspects require future improvements. First, the model's robustness to sensor data missing or noise needs to be enhanced. In practical scenarios, sensor data may be affected by environmental factors or hardware issues, leading to data loss or anomalies. Future work can integrate advanced imputation techniques or robust loss functions to address such situations. For sensor data missing, nearest-neighbour sensor data can be used for filling – considering spatial correlation, the 3 sensors closest to the missing point are selected for weighted average filling, with weights inversely proportional to distance. In practical applications, a data missing processing module can be further designed to automatically detect missing values before data input into the model and select corresponding processing strategies based on the type of missing data. Additionally, subsequent work can explore coupling the STABP model with other more advanced algorithms or technologies to further improve model performance. Combining the STABP model with Transformer-based architectures may more effectively capture long-range dependencies.

Declarations

All authors declare that they have no conflicts of interest.

References

Ali, A., Aliyuda, K., Elmitwally, N. and Bello, A.M. (2022) 'Towards more accurate and explainable supervised learning-based prediction of deliverability for underground natural gas storage', *Applied Energy*, Vol. 327. Doi: 10.1016/j.apenergy.2022.120098.

Arroyo, P., Gómez-Suárez, J., Herrero, J.L. and Lozano, J. (2022) 'Electrochemical gas sensing module combined with unmanned aerial vehicles for air quality monitoring', *Sensors and Actuators B: Chemical*, Vol. 364. Doi: 10.1016/j.snb.2022.131815.

Demirkan, D.C., Duzgun, H.S., Juganda, A., Brune, J. and Bogin, G. (2022) 'Real-time methane prediction in underground longwall coal mining using AI', *Energies*, Vol. 15, No. 17. Doi: 10.3390/en15176486.

Dong, J., Goodman, N. and Rajagopalan, P. (2023) 'A review of artificial neural network models applied to predict indoor air quality in schools', *International Journal of Environmental Research and Public Health*, Vol. 20, No. 15. Doi: 10.3390/ijerph20156441.

Du, W., Chen, L., Wang, H., Shan, Z., Zhou, Z., Li, W. and Wang, Y. (2023) 'Deciphering urban traffic impacts on air quality by deep learning and emission inventory', *Journal of Environmental Sciences*, Vol. 124, pp.745–757.

Fahad, S., Su, F., Khan, S. U., Naeem, M. R. and Wei, K. (2023) 'Implementing a novel deep learning technique for rainfall forecasting via climatic variables: an approach via hierarchical clustering analysis', *Science of The Total Environment*, Vol. 854. Doi: 10.1016/j.scitotenv.2022.158760.

Ge, M., Feng, Z. and Meng, Q. (2024) 'Urban planning and green building technologies based on artificial intelligence: principles, applications, and global case study analysis', *Applied Science and Engineering Journal for Advanced Research*, Vol. 3, No. 5, pp.18–27.

Guo, Z., Yang, C., Wang, D. and Liu, H. (2023) 'A novel deep learning model integrating CNN and GRU to predict particulate matter concentrations', *Process Safety and Environmental Protection*, Vol. 173, pp.604–613.

Li, L., Wang, J., Franklin, M., Yin, Q., Wu, J., Camps-Valls, G. and Reichstein, M. (2023) 'Improving air quality assessment using physics-inspired deep graph learning', *NPJ Climate and Atmospheric Science*, Vol. 6, No. 1. Doi: 10.1038/s41612-023-00475-3.

Liu, Q., Cui, B. and Liu, Z. (2024) 'Air quality class prediction using machine learning methods based on monitoring data and secondary modeling', *Atmosphere*, Vol. 15, No. 5. Doi: 10.3390/atmos15050553.

Livingston, S.J., Kanmani, S.D., Ebenezer, A.S., Sam, D. and Joshi, A. (2023) 'An ensembled method for air quality monitoring and control using machine learning', *Measurement: Sensors*, Vol. 30. Doi: 10.1016/j.measen.2023.100914.

Mutovina, N., Nurtay, M., Kalinin, A., Tomilov, A. and Tomilova, N. (2025) 'Application of artificial intelligence and machine learning in expert systems for the mining industry: modern methods and technologies', *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 15, No. 3, pp.3291–3308.

Pérez-Beltrán, C.H., Robles, A.D., Rodríguez, N.A., Ortega-Gavilán, F. and Jiménez-Carvelo, A.M. (2024) 'Artificial intelligence and water quality: from drinking water to wastewater', *TrAC Trends in Analytical Chemistry*, Vol. 172. Doi: 10.1016/j.trac.2024.117597.

Rad, A.K., Shamshiri, R.R., Naghipour, A., Razmi, S.O., Shariati, M., Golkar, F. and Balasundram, S.K. (2022) 'Machine learning for determining interactions between air pollutants and environmental parameters in three cities of Iran', *Sustainability*, Vol. 14, No. 13. Doi: 10.3390/su14138027.

Radhakrishnan, K., Ramakrishnan, D., Khalaf, O. I., Uddin, M., Chen, C. L. and Wu, C. M. (2022) 'A novel deep learning-based cooperative communication channel model for wireless underground sensor networks', *Sensors*, Vol. 22, No. 12. Doi: 10.3390/s22124475.

Ramadan, M.N., Ali, M.A., Khoo, S.Y., Alkhedher, M. and Alherbawi, M. (2024) 'Real-time IoT-powered AI system for monitoring and forecasting of air pollution in industrial environment', *Ecotoxicology and Environmental Safety*, Vol. 283. Doi: 10.1016/j.ecoenv.2024.116856.

Sharifi, S.A., Zaeimdar, M., Jozi, S.A. and Hejazi, R. (2023) 'Effects of soil, water and air pollution with heavy metal ions around lead and zinc mining and processing factories', *Water, Air, & Soil Pollution*, Vol. 234, No. 12. Doi: 10.1007/s11270-023-06758-y.

Thanh, H.V., Dai, Z., Du, Z., Yin, H., Yan, B., Soltanian, M.R. and Abualigah, L. (2024) 'Artificial intelligence-based prediction of hydrogen adsorption in various kerogen types: implications for underground hydrogen storage and cleaner production', *International Journal of Hydrogen Energy*, Vol. 57, pp.1000–1009.

Tripathi, A.K., Aruna, M., Parida, S., Nandan, D., Elumalai, P.V., Prakash, E. and Rao, K.S. (2024) 'Integrated smart dust monitoring and prediction system for surface mine sites using IoT and machine learning techniques', *Scientific Reports*, Vol. 14, No. 1. Doi: 10.1038/s41598-024-58021-x.

Wang, C., He, Q., Li, Z., Yu, J., Bello, I.T., Zheng, K. and Ni, M. (2024a) 'A novel in-tube reformer for solid oxide fuel cell for performance improvement and efficient thermal management: a numerical study based on artificial neural network and genetic algorithm', *Applied Energy*, Vol. 357. Doi: 10.1016/j.apenergy.2023.122030.

Wang, J., Wang, D., Zhang, F., Yoo, C. and Liu, H. (2024b) 'Soft sensor for predicting indoor PM2.5 concentration in subway with adaptive boosting deep learning model', *Journal of Hazardous Materials*, Vol. 465. Doi: 10.1016/j.jhazmat.2023.133074.

Wang, Z., Liang, F., Li, C., Xiong, W., Chen, Y. and Xie, F. (2023) 'Does China's low-carbon city pilot policy promote green development? Evidence from the digital industry', *Journal of Innovation and Knowledge*, Vol. 8, No. 2. Doi: 10.1016/j.jik.2023.100339.

Zhang, Q., Han, Y., Li, V.O. and Lam, J.C. (2022) 'Deep-AIR: a hybrid CNN-LSTM framework for fine-grained air pollution estimation and forecast in metropolitan cities', *IEEE Access*, Vol. 10, pp.55818–55841.

Zhao, Q., Wang, H. and Chen, C. (2024) 'Underground hydrogen storage: a recovery prediction using pore network modeling and machine learning', *Fuel*, Vol. 357. Doi: 10.1016/j.fuel.2023.130051.