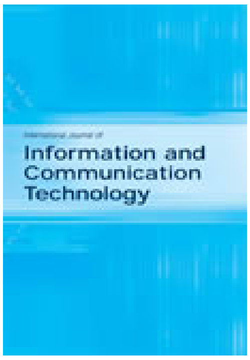


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

A time series neural network-based early warning system for thermal power station

Chen Yang, Hengyao Pan

Article History:

Received:	31 December 2024
Last revised:	13 January 2025
Accepted:	14 January 2025
Published online:	25 February 2025

A time series neural network-based early warning system for thermal power station

Chen Yang and Hengyao Pan*

School of Energy and Power Engineering,
Chongqing University,
Chongqing, 400044, China
Email: yxtyc@cqu.edu.cn
Email: panhengyao92@163.com
*Corresponding author

Abstract: Real-time monitoring and early warning are critical for power system stability and security as thermal power units become more crucial in power generation. This paper reports a time series neural network-based thermal power unit condition warning mechanism. First, time series analysis simulates the dynamic changes of the units to capture their long-term trend and time-varying characteristics over operation using multi-dimensional thermal power unit operation data. Time-series data is processed by long short-term memory (LSTM), which also learns features for highly precise defect prediction. Furthermore, designed in this work is a multi-dimensional performance assessment system based on fault detection rate (FDR), early warning lead time (EWLT), false alarm rate (FAR), and diagnostic accuracy (DA) to entirely assess the proposed method. Tests reveal that the time series neural network-based thermal power unit status warning system can let thermal power units run reliably.

Keywords: thermal power unit; condition monitoring; fault warning; time series analysis; long short-term memory; LSTM.

Reference to this paper should be made as follows: Yang, C. and Pan, H. (2025) 'A time series neural network-based early warning system for thermal power station', *Int. J. Information and Communication Technology*, Vol. 26, No. 4, pp.106–122.

Biographical notes: Chen Yang received his PhD in Engineering from Chongqing University in 1999. He is currently a Professor at the School of Energy and Power Engineering of Chongqing University. His research interests include advanced energy system simulation optimisation and intelligent control, as well as digital twin technology in energy systems.

Hengyao Pan received his MEng at Chongqing University in 2017. He is currently a Doctoral student at the School of Energy and Power Engineering at Chongqing University. His research interests include the integration of artificial intelligence, big data, and cloud computing technologies into energy systems.

1 Introduction

Thermal power units, one of the key tools in the power sector, hold a significant role in energy generation as industrialisation develops (Ma et al., 2017). Apart from their important contribution to the supply of energy, the steady running of thermal power plants is essential for the safety, economics, and environmental preservation of the power network. Nevertheless, long-term operation presents several difficulties for thermal power units, including aging equipment and frequent fault occurrence that makes condition monitoring and fault warning of thermal power units especially crucial. Usually depending on manual inspection and periodic testing, traditional thermal power unit monitoring systems have low efficiency and lagging response times, which makes it challenging to satisfy the need of the current power sector for intelligent and real-time equipment management (Stern, 2011).

The condition monitoring of thermal power units is progressively moving in the direction of intelligence and automation as information technology and intelligent technology – especially the broad application of sensor technology – especially the internet of things – big data analysis and other technologies – develop rapidly. Real-time monitoring and data analysis help the operating status of the unit to be dynamically assessed, anomalies to be found in a timely manner, and failure warnings to be executed, so preventing possible major equipment damage, lowering maintenance costs, extending the life of the equipment, and so improving the safety and dependability of power production (Lu et al., 2020). In the field of thermal power unit condition monitoring, data-driven intelligent monitoring and fault detection techniques – especially those based on time series analysis and machine learning – have lately become a major focus of research direction.

Based on conventional approaches and data-driven methodologies, fault monitoring and early warning systems for thermal power units mostly concentrate in two categories (Jieyang et al., 2023). Although to some extent they help in fault detection of thermal power units, traditional methods including empirical-based judgement, expert systems and signal processing techniques are difficult to adapt to the complex and dynamically changing operating environment due their over-reliance on human factors and fixed models.

Data-driven fault detection and early warning systems based on data have progressively become a hot topic for research in recent years as artificial intelligence technologies – especially machine learning and deep learning – have grown increasingly important. Real-time capture of several operating data (e.g., temperature, pressure, flow, current, etc.) from the unit and analysis of them in combination with sophisticated algorithms helps to efficiently identify possible fault risks (Ahmad et al., 2022). For instance, fault detection of thermal power units has extensively applied conventional machine learning techniques such as support vector machine (SVM), decision tree, and random forest (RF) (Raczko and Zagajewski, 2017). By understanding the features of the previous data of the equipment, these techniques can forecast its condition more precisely and alert of possible faults in advance.

But conventional machine learning techniques have some limits given the growing complexity of thermal power unit running circumstances. First, these techniques typically involve human feature extraction and perform poorly for nonlinear and time-varying system responses; second, it remains a difficulty to provide efficient and accurate failure warning in complex and changing surroundings. For this reason, the use of deep learning techniques has grown to be a major research trend in the field of thermal power unit monitoring in recent years. With their strong automatic feature extraction and modelling capacity, deep learning models including deep neural network (DNN), convolutional neural networks (CNN), and recurrent neural network (RNN) have produced amazing results in the fields of image recognition, speech recognition, natural language processing, etc.; they have also been progressively introduced into fault detection and early warning of thermal power units (Nassif et al., 2019).

Particularly the modelling approach based on time series data as the operation data of thermal power units is basically time series data, many researchers have started to investigate the integration of time series analysis and deep learning approaches to build more accurate fault warning models. Time series data illustrates the long-term trend and seasonal changes of the equipment in addition to the historical state of equipment operation. By means of thorough investigation of these time series data, the dynamic variations in equipment performance may be efficiently recorded and possible indicators of failure can be found.

For condition monitoring and defect warning of thermal power units, some pressing issues still need to be addressed even if current research has made great progress. Although most of the methods still depend on manually selecting features, current research concentrates on the application of a single model and lacks a comprehensive cross-model and cross-technology solution; moreover, the robustness and real-time performance of the models have yet to be improved in the face of complex and variable operating conditions. Therefore, a major focus of present research is still how to mix several new technologies to increase the accuracy, efficiency and adaptability of monitoring and early warning.

This work so suggests an accurate and effective thermal power unit status monitoring and fault warning system by means of clever algorithms combined with time series data analysis.

This work has original points of interest as follows:

1 A multi-model fusion method combining time series analysis and neural network is proposed. In this work, we create a thermal power unit status monitoring and defect warning model integrating several algorithms by creatively combining time series analysis approaches with deep learning methods. Accurate modelling of time series helps to capture the long-term trend and dynamic changes of equipment operation; incorporating the strong feature learning capability of neural networks enhances the accuracy and dependability of failure early warning.

2 Automatic feature extraction based on DNNs to reduce manual intervention. There are certain restrictions in conventional defect warning techniques, which usually depend on hand selection of features and rules of thumb. This work uses DNNs – especially long short-term memory (LSTM) – through which implicit characteristics in the timing data are automatically identified and learnt, therefore overcoming the constraints of manual feature extraction and increasing the adaptability and accuracy of the model.

3 Combining multi-source data for comprehensive analysis to enhance the robustness of the model. We realise all-round monitoring of unit status in this work by merging multi-dimensional operation data (e.g., temperature, pressure, current, etc.) of thermal power units with multi-source data fusion technology. Comprehensive

analysis of this data helps to improve the model's resilience, therefore facilitating effective early warning in challenging and dynamic settings.

4 An index method for early warning assessments suitable for thermal power plants is developed. This work constructs a set of multi-dimensional evaluation index system comprising fault detection rate (FDR), early warning lead time (EWLT), false alarm rate (FAR), and diagnostic accuracy (DA) in order to completely evaluate the performance of the model. These evaluation indexes more holistically assess the performance of the failure early warning system and thermal power unit status monitoring, therefore offering a basis for later optimisation and development.

2 Relevant technologies

2.1 Time series analysis

A statistical technique for modelling and forecasting data structured chronologically is time series analysis (Parmezan et al., 2019). Its main goal is to expose fundamental data patterns like trends, seasonality, and random fluctuations so enabling predictions about future behaviour. Dealing with time series data often calls for addressing data's smoothness. Smoothness is the property wherein the mean, variance, and self-covariance of a time series do not change with time. Often eliminated by differencing techniques, which generates a smooth time series, the trend component helps to make non-smooth time series fit for modelling.

The autoregressive moving average (ARMA) model is a fundamental method of time series analysis that models the dependent of a time series by autoregression (AR) on past data and moving average (MA) of the error term (Khan and Gupta, 2020). The ARMA model resembles this:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \tag{1}$$

where ϵ_t is the white noise error term; ϕ_i and θ_i are the model's parameters; p and q respectively indicate the AR and MA component orders. Y_t is the observed value at time point t.

Usually used for non-smooth time series is the autoregressive integral sliding average model (ARIMA) (Zheng et al., 2023). The ARIMA model combines the concepts of AR and MA with the expression by means of difference operations, hence smoothing the data.

$$\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)(1 - B)^d Y_t = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)\epsilon_t \tag{2}$$

where B is the lag operator; d is the number of differences; the other symbols match those of the ARMA model. By means of the AR and MA components, the ARIMA model is able to capture the dependence in the smooth series and efficiently eradicate the trend in the time series.

The seasonal ARIMA (SARIMA) model can be also applied in cases of seasonal data changes. Able to handle cyclically changing time series, the SARIMA model adds seasonal variations and seasonal components to ARIMA. Mathematically, it is:

$$\left(1 - \phi_1 B^s - \phi_2 B^{2s} - \cdots - \phi_p B^{ps}\right)(1 - B)^d Y_t = \left(1 + \theta_1 B^s + \theta_2 B^{2s} + \cdots + \theta_q B^{qs}\right)\epsilon_t \tag{3}$$

where B_s is the seasonal lag operator and s is the seasonal period. SARIMA can properly record seasonal variations using the seasonal term and seasonal difference.

Time series data in many actual issues may include complicated nonlinear interactions, which the conventional ARIMA model finds challenging to sufficiently model. At this point, one can include a support vector regression (SVR) model, appropriate for time series forecasting and able to manage challenging nonlinear issues. The SVM regression formula is:

$$y_t = w \cdot \phi(x_t) + b \tag{4}$$

where b is a bias term; w is the weight of the model; $\phi(x_t)$ is a nonlinear mapping function. Especially appropriate for handling high-dimensional data and nonlinear patterns, SVR fits time series data by building a hyperplane.

Furthermore extensively applied for time series forecasting are neural network models, particularly RNN and its variations LSTM and gated recurrent unit (GRU). Appropriate for handling large time series and complex nonlinear patterns, these models can automatically learn temporal dependencies in time series (Natarajan et al., 2023).

Time series analysis may reasonably predict and forecast different kinds of time series data using these several approaches. The time series analysis approach offers a trustworthy theoretical framework and technological support for the accurate prediction of equipment failure in the application of thermal power unit condition warning.

2.2 Neural networks

Deep learning technology is developing quickly, so neural network techniques have become a major instrument for time series data processing. Neural networks can automatically extract characteristics from unprocessed data and adjust to complicated nonlinear relationships in the data unlike conventional statistical approaches. This reveals considerable promise for neural networks in time series prediction, anomaly detection, and trend analysis. Particularly for complicated dynamic systems and extended time series, neural networks exhibit more adaptability and accuracy than conventional techniques.

Constructing a mapping between inputs and outputs by means of one or more layers of neurons, feedforward neural networks (FNN) are the most fundamental neural network model (Ojha et al., 2017). By mapping previous data to future values, FNN is able to reasonably capture nonlinear correlations in data for time series prediction. The fundamental FNN formula is:

$$y = f(Wx + b) \tag{5}$$

where x is the input, W is the weight; b is the bias term; f is the activation function; y is the output. FNN creates expected values and fits the training data via parameter adjustment. FNN cannot, however, capture dynamic changes in time series as later recurrent networks can or manage temporal dependencies.

RNNs were developed to help to tackle this challenge. By using the output of the past instant as the input of the current moment, RNNs let the network remember prior data and detect temporal dependencies in a time series. The fundamental RNN formula is:

$$h_t = \sigma\left(W_{xh} x_t + W_{hh} h_{t-1} + b\right) \tag{6}$$

where W_{xh} and W_{hh} are the weight matrices; σ is the activation function; h_t is the hidden state at the current moment; x_t is the input at the current moment. The RNN can efficiently record short-term dependencies in the data using this recursive structure. Standard RNNs may, however, run into training challenges while handling extended sequences by experiencing gradient vanishing and gradient explosion.

LSTM is intended to address this issue. By including gating mechanism, particularly in cases of long-term dependencies, LSTM greatly enhances the performance of conventional RNN (Kong et al., 2017). LSTM's secret memory units – which govern information flow through input gates, output gates, and forgetting gates – so avoiding the gradient problem – are its key. LSTM has as its basic formula:

$$h_t = o_t \cdot \tanh(c_t) \tag{7}$$

where o_t is the output gate; c_t is the memory cell state right now; h_t is the hidden state right now. LSTM can thus manage very long time-series data and efficiently capture long-term dependencies in time series using this framework (Song et al., 2020).

Like LSTM, GRU is also a better RNN that simplifies the gating mechanism thereby lowering the computing cost of the model. In many applications, GRU shows comparable results to LSTM; it is particularly appropriate in cases of restricted processing capability. GRU's formula is as follows:

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \tag{8}$$

where z_t is the update gate; $\tilde{h}_t$ is the candidate hidden state. Using this framework, GRU avoids some difficult computations in LSTM while effectively doing time series modelling.

Apart from RNN and its variations, CNN has been extensively applied in time series analysis (Canizo et al., 2019). CNNs were first utilised for image processing, but their strong local feature extraction powers help them to also effectively process time series data. CNN uses local characteristics extracted by a convolutional layer in time series forecasting applications subsequently using a fully connected layer. CNN's convolutional operation can be shown as:

$$y_t = \sum_{k=0}^{K-1} w_k \cdot x_{t-k} + b \tag{9}$$

where y_t as the output, x_{t-k} as the input data, w_k as the convolution kernel's weight, b as the bias term, and K as the convolution kernel's size. CNNs may effectively extract locally dependent characteristics in the time series using this convolutional structure, hence enhancing the prediction accuracy.

Particularly LSTM, GRU, CNN, neural network techniques are able to manage intricate nonlinear time series data and fit for applications requiring local or long time-dependent feature collecting. Neural networks can effectively early warning and decision support, predict equipment failure in advance, and automatically learn the laws from a lot of historical data for the thermal power unit status warning problem.

3 Theory of condition monitoring and early warning of thermal power units

Ensuring the safe and stable functioning of thermal power units depends much on condition monitoring and early warning of thermal power units. Its main goal is to examine the unit's real-time data in order to find aberrant trends and timely project possible problems. Usually comprising many fundamental processes of data collecting, feature extraction, model building, and early warning issuing, traditional condition monitoring and early warning systems

Monitoring starts first with data collecting. Through the sensor network, including temperature, pressure, vibration, flow, etc., thermal power units gather, in real-time, the operating parameters of every important component of the unit. Usually time series data, these records have a definite dependent relationship between each time point. Thus, sensible feature selection and processing are rather important.

Second, the process of analysis depends absolutely on feature extraction. Effective features have to be retrieved from the raw data after data collecting for additional study. These characteristics comprise, among others statistical numbers (e.g., mean, variance, etc.), frequency domain features (e.g., frequency features following Fourier transform), and trends, fluctuations, etc. of the time series. Feature extraction aims to convert redundant, high-dimensional raw data into low-dimensional features reflecting the unit's health condition, therefore enabling early warning and fault diagnostics later on.

Establishing a defect diagnostic model comes next from feature extraction. Conventional models consist in discriminant analysis models, statistically based regression models, etc. By means of historical data, these techniques are trained to generate a predictive model capable of spotting equipment flaws or deviations. This model allows the system to evaluate the unit's real-time health condition during operation and ascertain whether failure risk exists.

At last, warning signals are set out when the monitoring system identifies anomalies in the running condition of the unit or forecasts a future breakdown. Usually depending on defined criteria or model output forecasts, early warning signals are issued. In this process, a crucial phase is threshold setting. Usually, analyses of past data define a safe range for some important parameters (e.g., temperature, pressure, etc.). The system generates an alert when the real-time data crosses the range. More and more monitoring systems are using data-driven techniques including machine-learning-based models to automatically modify these thresholds and forecasts for smarter warnings, hence improving the accuracy of predictions.

Although conventional techniques can somewhat provide condition monitoring and early warning of thermal power units, their accuracy and adaptability are more limited. Traditional approaches might have more restrictions for handling complicated and nonlinear system failures since they mostly depend on rule setting and empirical judgement. Monitoring and warning systems based on deep learning and neural networks have progressively taken front stage in research as technology develops since they exhibit more flexibility and prediction accuracy when addressing dynamic and complex systems.

4 Methodological framework for condition monitoring and early warning of thermal power units

This methodological framework based on ARIMA and LSTM proposes a thermal power unit condition monitoring and warning approach by combining time series analysis with neural networks. It can effectively extract the temporal features in the data and achieve the accurate monitoring of the unit status and fault warning by employing the ARIMA model for the preliminary analysis of time series data together with the deep learning capacity of LSTM network. The exact framework's structure and the unique linkages are shown here.

4.1 Data acquisition and pre-processing

Through sensors, the thermal power unit condition monitoring system gathers in real-time important operating parameters of the unit, including temperature, pressure, flow rate, etc. Usually showing temporal order, this data may have noise and missing values. Data preparation is a necessary first step to guarantee the stability and correctness of the model. Operations in normalisation, missing value filling, and noise reduction define data preparation mostly. Particularly for time-series data, sliding averages or filters help to remove noise and fill missing values using suitable interpolation techniques.

$$X_t = X_{t-1} + \frac{X_{t+1} - X_{t-1}}{2} \tag{10}$$

where X_t is the estimation of the missing value; X_{t-1} and X_{t+1} are respectively the data before and after the missing value.

4.2 Time series modelling (ARIMA)

ARIMA model is a classical approach extensively applied for forecasting and modelling that can help to model trends in time series data. By capturing the AR, I, and MA components of the data, ARIMA model can extract patterns and seasonality in time series data, therefore enabling projections of future states. Appropriate for smooth time series data, ARIMA model can manage trend and seasonality elements.

ARIMA model has a general form as:

$$Y_t = \alpha + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{11}$$

where the observed time series value is Y_t.

The ARIMA model's fundamental idea is to estimate the future trend of the unit state by use of AR and sliding averages thereby modelling the dependencies in the time series.

4.3 LSTM network modelling

While LSTM network displays a strong modelling capability for complicated nonlinear interactions, ARIMA model can efficiently capture linear features in time series. Particularly appropriate for analysing and forecasting time series data, LSTM is a unique form of RNN. By use of its internal gating mechanism – that of forgetting gate, input gate, and output gate – LSTM can efficiently preserve the information in the long time series. By efficiently preserving information in extended time series, LSTM can handle the long-term reliance issue.

LSTM has as its primary structural equations these:

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right) \tag{12}$$

$$i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right) \tag{13}$$

$$\tilde{C}_t = \tanh\left(W_C\left[h_{t-1}, x_t\right] + b_C\right) \tag{14}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{15}$$

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{16}$$

$$h_t = o_t \cdot \tanh\left(C_t\right) \tag{17}$$

By progressively optimising the parameters over the back-propagation technique during the training process, the LSTM model is able to automatically capture the complicated temporal correlations in the data and hence generate effective prediction of the thermal power unit state.

4.4 Fault warning mechanism

Based on LSTM model prediction, the design of the early warning mechanism seeks to ascertain whether a thermal power unit is in a probable fault state. More specifically, the system sets an alert to let the operator know when the LSTM model's prediction value surpasses a specified preset threshold therefore enabling maintenance or inspection.

The early warning system operates with a simple formula:

$$\hat{y}_t = f\left(x_t; \theta\right) \tag{18}$$

$$\text{if } \hat{y}_t > \text{Threshold, Trigger Alarm} \tag{19}$$

where Threshold is the defect warning threshold established depending on historical data and expert knowledge; $\hat{y}_t$ is the condition of the unit expected by the LSTM model.

Should the system's prediction value above this level, a defect could develop and the system will send an early warning signal asking the pertinent personnel to act.

4.5 Integrated optimisation

LSTM paired with ARIMA for integrated learning helps to raise the prediction accuracy and resilience. Combining the strengths of several models helps the system to better handle challenging operating conditions and lower the bias resulting from one model.

One often used type of integrated optimisation technique is weighted averaging, that is,

$$y_{final} = \sum_{i=1}^{n} w_i y_i \tag{20}$$

where y_i is the prediction result of every base model; y_{final} is the ultimate prediction result; W_i is the weight of every base model.

This weighting system may dynamically change the weights of every model based on performance, therefore enhancing the general prediction performance.

Algorithm 1 exhibits the framework's workflow.

Algorithm 1 Pseudo-code for fire power plant state monitoring and early warning

```
Input: Historical time series data of power plant parameters (e.g., temperature, pressure, flow),
       Model parameters (e.g., ARIMA, LSTM), Predefined threshold for early warning,
       Number of iterations for training LSTM, Learning rate, Sliding window size.
Output: Optimised ARIMA model, Trained LSTM model, Early warning output.
begin
    Load historical time series data;
    Preprocess the data (handle missing values, noise filtering, and normalisation);
    Divide data into training and test sets;
    Initialise ARIMA model parameters (p, d, q);
    Fit ARIMA model on training data;
    Predict time series trend using ARIMA (ARIMA prediction);
    Initialise LSTM model with random weights;
    Define LSTM architecture (input layer, LSTM layers, dense layer);
    Train LSTM model using the training data with backpropagation;
    for each epoch to max_epochs do
        Input time series data to LSTM model;
        Forward pass through LSTM network to obtain predictions;
        Calculate loss (e.g., Mean Squared Error between prediction and actual values);
        Compute gradients of loss function with respect to LSTM weights;
        Update LSTM weights using the optimiser (e.g., Adam or SGD);
    end for
    Evaluate LSTM performance on test data;
    if prediction_error > threshold then
        Trigger early warning signal (e.g., fire alarm or maintenance request);
    end if
    Optionally, use ARIMA and LSTM predictions together (ensemble learning);
    if ensemble prediction error > threshold then
        Trigger early warning signal;
    end if
    return Optimised ARIMA model, Trained LSTM model, Early warning output;
end
```

Based on a combination ARIMA and LSTM model, this methodological framework can efficiently extract important information from the time series data of thermal power units for fault prediction and health evaluation. The overall system provides effective unit condition monitoring and warning with great robustness and adaptive capability by means of the phases of data collecting, preprocessing, time series modelling, LSTM modelling, warning mechanism and integration and optimisation. This framework is expected to offer a suitable technical solution for maintenance of thermal power plants and early warning of faults.

5 Experimental results and analyses

5.1 Datasets

Derived from the online monitoring system of a thermal power plant of Huaneng Group, the dataset used in this study spans a broad spectrum of states and environmental conditions in the long-term operation of the unit. Key operational metrics like temperature, pressure, flow rate, vibration, etc. of the boiler, turbine, generator, and other major equipment in the dataset are strongly linked to the operational health state of the unit and offer significant information for this study.

The dataset logs time-series data gathered by several sensors, noting the running settings of every main piece of equipment in the unit –boiler, turbine, generator, etc.? With an acquisition frequency of one minute and a data time range of three years, every record shows the condition of the unit at a given instant in time. Table 1 lists the salient characteristics of the dataset:

Table 1 Dataset statistical information

Dataset subset	*Feature variables*	*Description*
Boiler system data	Boiler pressure, boiler temperature, combustion efficiency	These data describe the operating status and efficiency of the boiler.
Turbine system data	Inlet temperature, outlet temperature, vibration frequency	These data describe the operational status and stability of the turbine.
Generator system data	Generator load, current, voltage	These data reflect the load capacity and electrical parameters of the generator.
Environmental data	Ambient temperature, humidity	These data provide background information on the external environment's impact on the unit.

To guarantee the quality of the data, data preprocessing including standardisation, noise reduction and missing value filling was done. Linear interpolation filled missing values; sliding average approach reduced noise; data were normalised to harmonise the scale of several features.

Three categories – normal state, abnormal state and fault state – separate the unit's operational state and fault circumstances identified by the labels of the dataset. Labelling generation depends on expert system diagnosis findings and past error data. The dataset guarantees the variety of model training since it comprises more than 100,000 records spanning the states of the units under various operating settings.

Training, validation, and testing sets totalling an 80%, 10%, and 10% ratio temporally separate the dataset. This split enables the model to acquire strong generalisation capacity in state changes and variations in several time periods.

5.2 Assessment of indicators

5.2.1 Fault detection rate

This statistic especially gauges, in all circumstances when errors actually arise, the percentage of faults the model accurately detects. In condition monitoring of thermal power plants, where timeliness and accuracy of fault detection is crucial, FDR helps to quantify the fault prediction capacity of the model. It is computed applying the formula:

$$\text{FDR} = \frac{TP}{TP + FN} \tag{21}$$

where TP marks successfully diagnosed fault cases and FN marks missed fault instances. A high FDR indicates that the system can give early warning and rapidly and precisely identify faults.

5.2.2 Early warning lead time

This indicator gauges the model's ability to provide an early warning before a fault strikes, hence measuring its degree of advanceability. It is especially relevant to early warning systems for thermal power units and guarantees rapid reaction and corrective action. The computation model is:

$$\text{EWLT} = \text{Time of fault occurrence} - \text{Time of first prediction} \tag{22}$$

A higher EWLT shows that the early warning system is more able to detect trouble signals ahead and lower the frequency of unplanned mishaps.

5.2.3 False alarm rate

The FAR gauges, in the absence of defects, the fraction of cases the model falsely forecasts as faults. It can clarify the system's allergy – that is, the number of false alarms among the expected positive cases – that is, how bad the allergy of the system is. It is determined with the formula:

$$\text{FAR} = \frac{FP}{FP + FN} \tag{23}$$

where TN is true negative and FP is false positive – false positive. In condition monitoring of thermal power plants, high false positives can cause needless system downtime and therefore it is quite crucial.

5.2.4 Diagnostic accuracy

Diagnosis is accuracy gauges whether the model can identify the kind of equipment condition of fault or type of error. It takes into account the accurate identification of fault types, so unlike the traditional classification accuracy. In multi-equipment, multi-state thermal power units, this statistic is especially crucial since the system must not only identify the occurrence of a problem but also properly determine the kind of failure. It's computed using the formula:

$$\text{DA} = \frac{TP_{fault\ type}}{TP_{fault\ type} + FP_{fault\ type} + FN_{fault\ type}} \tag{24}$$

where TP_{fault} type is the correctly diagnosed fault type; FP_{fault} type is the misdiagnosed fault type; FN_{fault} type is the fault type overlooked. High DA indicates that the model not only points up flaws but also accurately detects fault types to guide next maintenance actions.

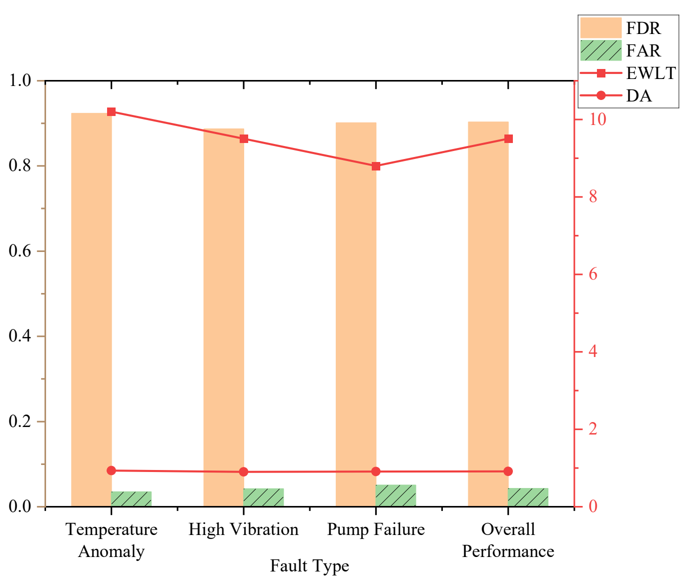
5.3 Fault detection and early warning performance evaluation experiment

The experiment's primary goal is to assess the model's real-time performance and accuracy in defect detection of thermal power units. Strictly differentiated in time to prevent data leaking, the test dataset and the training dataset help to guarantee the fairness and suitability of the experiment. Every test data point is tagged with the real fault occurrence time; so, the model must alert ahead of time about fault occurrence.

Data preparation, model training and test evaluation comprise the experimental process. The data were first split into a training set (70%) and a test set (30%), after being preprocessed and normalised to eliminate outliers and fill in missing data. The model is then trained using past data from thermal power plants and employs cross-valuation to prevent overfitting. At last, the trained model's capacity to identify unidentified fault samples is assessed by means of the test set data.

Figure 1 shows the experimental findings, which list the model's performance under several fault kinds.

Figure 1 Experimental results of fault detection and warning performance evaluation (see online version for colours)



The model boasts a FDR of 92.3%, a fault warning time of 10.2 minutes, a FAR of 3.5%, and a DA of 93.4% for the temperature anomaly fault. The FDR under an excessive vibration fault type is 88.7%; the fault warning time is 9.5 minutes; the FAR is 4.2%; and the DA is 90.1%. The FAR was 5.1%, the FDR was 90.1%, the fault warning duration was 8.8 minutes, and the DA was 91.2%. The pump defect was with an overall performance of 90.3% FDR, 9.5 minutes fault warning duration, 4.3% FAR, and 91.3% DA the model performs better on all fault kinds.

From the experimental findings, it is clear that the model has a high FDR and DA on various fault types, especially on the temperature abnormality and vibration too high fault types, the FDR reaches 92.3% and 88.7%, respectively. The somewhat short fault warning period indicates that the model can efficiently prevent too many false alarms and offer timely warnings before defects start to manifest themselves.

Particularly the high DA, which indicates that the model is able to effectively identify different kinds of faults, the model shows good performance in terms of accuracy and fault detection capability and is able to precisely predict and identify faults in thermal power units in a rather short period of time. Furthermore, demonstrating the ability of the model to offer efficient advance warning for the operation of thermal power units to guarantee their safety is the performance of the warning time.

5.4 Comparative experiments

We performed comparison experiments with several standard fault detection techniques in order to fully evaluate the efficiency of the suggested method in condition monitoring and early warning of thermal power units. The approaches of comparison consist in:

The ARIMA model is a classical time series forecasting tool used in modelling and forecasting linear time series. We evaluate its performance against the proposed technique as a benchmark model.

Widely employed in sequence prediction applications, LSTM can efficiently manage time series data including long-term dependencies. LSTM is often used in fault detection to replicate intricate time-series data.

Widely employed in classification and regression analysis, SVM is a supervised learning technique particularly suited for handling nonlinear issues and hence appropriate for abnormal state identification in fault detection activities.

RF can efficiently record intricate correlations between features and makes judgments by combining several decision trees. RF is fit for many nonlinear situations and can manage high dimensional data in fault detection.

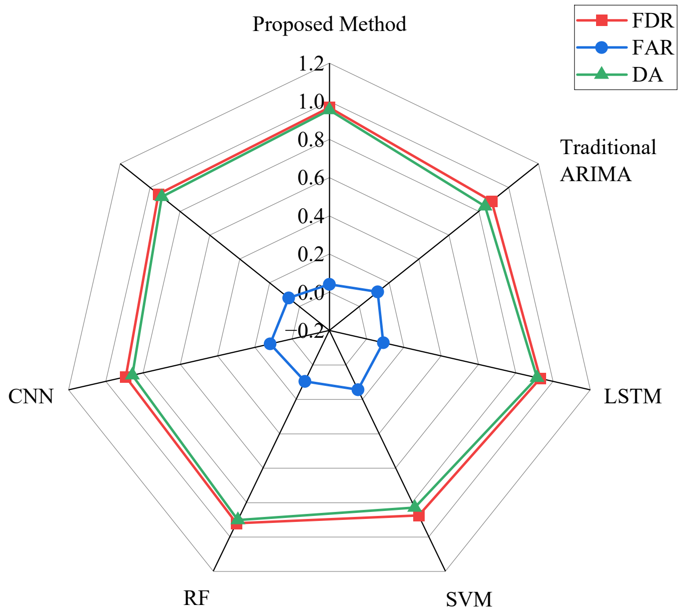
CNN: Generally used for image recognition tasks, CNN can also extract local features via convolutional layers when working with time-series data and can capture spatio-temporal characteristics in the data, so supporting condition monitoring of thermal power units.

Widely utilised in classification problems and especially fit for processing unbalanced data and high dimensional data, XGBoost is an optimisation approach based on gradient boosting trees.

All models are trained and tested on the same dataset comprising time-series data of normal and fault states of thermal power units. We tweaked all models with hyperparameters and employed 10-fold cross-valuation to guarantee the stability of the experimental results. LSTM models among deep learning models were trained with ideal hyperparameters.

Figure 2 displays the experimental findings.

Figure 2 Results of the comparison experiment (see online version for colours)



Particularly in FDR and DA, the suggested approach beats the other comparative models in all the assessment indices based on the experimental results: it reaches 96.8% and 95.4%, respectively. This shows that under normal circumstances the technique can lower false alarms and more precisely point up the flaws in thermal power plants.

Compared to existing techniques, the suggested method demonstrates superior fault identification capability on FDR and can better capture aberrant signals in unit operation, so lowering the occurrence of missed detection. With a low level of 4.1%, which is much lower than the other techniques, the suggested approach on FAR stays at a low level indicating less false alarms under typical settings. With an accuracy of 95.4%, the suggested approach on DA shows the great efficiency of the method in thermal power unit condition monitoring; it also surpasses all the other comparing models.

6 Conclusions

This work presents a thermal power unit condition monitoring and warning system based on ARIMA-LSTM. An intelligent system able of real-time monitoring and predicting the operating state of thermal power units is built by integrating the time series prediction ability of ARIMA model with the processing advantage of LSTM model for long time-dependent data. We first preprocessed the historical operational data of thermal power units and investigated them in time series employing the ARIMA model to capture the trends and cyclical fluctuations in them throughout the building of the model. The residuals then are modelled using an LSTM network, therefore enhancing the capacity of the model to forecast nonlinear and intricate time series data.

Though it has some restrictions, the ARIMA-LSTM-based thermal power unit condition monitoring and warning approach suggested in this paper has attained improved experimental results. First of all, the stability and accuracy of the model could suffer if the employed thermal power unit dataset in the study has some noise or missing. Second, the model applied in this work was trained on a particular thermal power unit dataset; so, the limits of the dataset could influence the generalisation capacity of the model. Furthermore, this work mostly depends on time-series data for fault monitoring of thermal power units and ignores additional multimodal data (e.g., environmental data, vibration data, etc.) that might influence the unit state.

Future research on the following elements can be directed depending on the outcomes of current one:

1 Multimodal data fusion. Future studies can incorporate more sensor data, ambient data and other multimodal information to increase the accuracy and robustness of thermal power unit condition monitoring by data fusion. Deep learning techniques can help the model to analyse many kinds of data, particularly picture and sound data, so improving its integrated capacity for decision-making.

2 Migration learning and incremental learning. Migration learning and incremental learning strategies might be included to solve the generalisation ability of the model. While incremental learning can help the model to continually learn and update and enhance its prediction capacity under the continuous input of new data, migration learning allows knowledge from other similar domains to be employed to improve the adaptability of the model in new contexts.

3 Optimisation of real-time monitoring and warning system. Future studies can enhance the current model to improve its response speed and processing capacity so attaining real-time monitoring and failure warning for thermal power plants. One can investigate lightweight neural network designs or apply distributed computing and edge computing technologies to get quick reaction of real-time data.

4 Further classification and diagnosis of fault types. While categorisation and diagnosis of fault types can be further improved in the future, present studies concentrate on fault detection and early warning. More accurate diagnosis of various kinds of problems is given by deep learning models, thereby enabling targeted maintenance and optimisation as well as increasing the running efficiency of the unit.

Declarations

All authors declare that they have no conflicts of interest.

References

Ahmad, T., Madonski, R., Zhang, D. et al. (2022) 'Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: key developments, challenges, and future research opportunities in the context of smart grid paradigm', *Renewable and Sustainable Energy Reviews*, Vol. 160, p.112128.

Canizo, M., Triguero, I., Conde, A. et al. (2019) 'Multi-head CNN-RNN for multi-time series anomaly detection: an industrial case study', *Neurocomputing*, Vol. 363, pp.246–260.

Jieyang, P., Kimmig, A., Dongkun, W. et al. (2023) 'A systematic review of data-driven approaches to fault diagnosis and early warning', *Journal of Intelligent Manufacturing*, Vol. 34, No. 8, pp. 3277–3304.

Khan, F.M. and Gupta, R. (2020) 'ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India', *Journal of Safety Science and Resilience*, Vol. 1, No. 1, pp.12–18.

Kong, W., Dong, Z.Y., Jia, Y. et al. (2017) 'Short-term residential load forecasting based on LSTM recurrent neural network', *IEEE Transactions on Smart Grid*, Vol. 10, No. 1, pp.841–851.

Lu, Q., Xie, X., Parlikad, A.K. et al. (2020) 'Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance', *Automation in Construction*, Vol. 118, p.103277.

Ma, X., Wang, Y. and Wang, C. (2017) 'Low-carbon development of China's thermal power industry based on an international comparison: review, analysis and forecast', *Renewable and Sustainable Energy Reviews*, Vol. 80, pp.942–970.

Nassif, A.B., Shahin, I., Attili, I. et al. (2019) 'Speech recognition using deep neural networks: a systematic review', *IEEE Access*, Vol. 7, pp.19143–19165.

Natarajan, S., Kumar, M., Gadde, S.K.K. et al. (2023) 'Outbreak prediction of COVID-19 using recurrent neural network with gated recurrent units', *Materials Today: Proceedings*, Vol. 80, pp.3433–3437.

Ojha, V.K., Abraham, A. and Snášel, V. (2017) 'Metaheuristic design of feedforward neural networks: a review of two decades of research', *Engineering Applications of Artificial Intelligence*, Vol. 60, pp.97–116.

Parmezan, A.R.S., Souza, V.M. and Batista, G.E. (2019) 'Evaluation of statistical and machine learning models for time series prediction: identifying the state-of-the-art and the best conditions for the use of each model', *Information Sciences*, Vol. 484, pp.302–337.

Raczko, E. and Zagajewski, B. (2017) 'Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images', *European Journal of Remote Sensing*, Vol. 50, No. 1, pp.144–154.

Song, X., Liu, Y., Xue, L. et al. (2020) 'Time-series well performance prediction based on long short-term memory (LSTM) neural network model', *Journal of Petroleum Science and Engineering*, Vol. 186, p.106682.

Stern, D.I. (2011) 'The role of energy in economic growth', *Annals of the New York Academy of Sciences*, Vol. 1219, No. 1, pp.26–51.

Zheng, W., Li, H., Yang, X. et al. (2023) 'Trends and prediction in the incidence rate of hepatitis C in Shandong Province in China from 2004 to 2030', *Preventive Medicine*, Vol. 177, p.107749.