

**ARTICLE**

# Cloud Based Monitoring and Diagnosis of Gas Turbine Generator Based on Unsupervised Learning

Xian Ma<sup>1</sup>, Tingyan Lv<sup>2,\*</sup>, Yingqiang Jin<sup>2</sup>, Rongmin Chen<sup>2</sup>, Dengxian Dong<sup>2</sup> and Yingtao Jia<sup>2</sup>

<sup>1</sup>China Academy of Industrial Internet, Beijing, 100041, China

<sup>2</sup>China Datang Co. Ltd., Beijing, 100032, China

\*Corresponding Author: Tingyan Lv. Email: maxiand@126.com

Received: 09 July 2020 Accepted: 09 November 2020

**ABSTRACT**

The large number of gas turbines in large power companies is difficult to manage. A large amount of the data from the generating units is not mined and utilized for fault analysis. This study focuses on F-class (9F.05) gas turbine generators and uses unsupervised learning and cloud computing technologies to analyse the faults for the gas turbines. Remote monitoring of the operational status are conducted. The study proposes a cloud computing service architecture for large gas turbine objects, which uses unsupervised learning models to monitor the operational state of the gas turbine. Faults such as chamber seal failure, load abnormality and temperature anomalies in the gas turbine system can be identified by using the method, which has an accuracy of 60%–80%.

**KEYWORDS**

Gas turbine generation; machine learning; cloud computing; monitoring and diagnostics

## 1 Introduction

At the present, the contribution of information and communication technology to the power generation industry is at the key point of quantitative change to qualitative change. In the new round of global electric power technology revolution and industrial transformation, the integration and development of the Internet and the power generation industry has broad prospects and unlimited potential. It is a priority to promote supply-side structural reform, further promote “Internet + advanced manufacturing industry”, speed up the construction and development of industrial Internet, and promote the development of the Internet and power generation industry. The deep integration of the Internet, big data, artificial intelligence and the real economy has become an unstoppable trend of current times [1,2].

The power generation industry continues to promote the upgrading of information technology, generating massive amounts of data every day. However, the existing technology and management mode in the power generation industry have become more important [3–5]. Data integration, analysis and management technology are backwards, due to the poor data sharing between various power plants, causing a strong information silo effect. There are serious data barriers in each power plant, and even when the volume of a single business or type of data is large, the lack of interconnection and sharing makes it less valuable. In addition, the data control ability is weak and lacks a unified and centralized



management of the crew [6]. In the era of big data, the level of data integration and the strength of data control ability affect the accuracy and real-time nature of data analysis. Moreover, the information fails to be utilized and in-depth exploitation of the data is lacking. The current database system is unable to discover the relationships existing in the data and predict the future trends of unit performance, especially lacking in mining data that hide knowledge behind the scenes. Further, the ability to resist information security risks needs to be further improved. The power industry is related to national strategic security and is a key area of the national economy. Network security protection needs to be further strengthened, from the current passive defence to a multi-level and active defence.

However, gas-steam combined cycle units, as an advanced power generation technology, have encountered some problems in the application of power production [7,8]. The main manifestations are highlighted below. First, data integration of the combustion engine system of the group company is low, and the scheduling level is weak. The combustion engine system has a high degree of integration and a wealth of sensor measurement points compared to a coal engine system, but these data have not formed a centralized use. An effective comparison between each other cannot be made. Second, to meet the needs of peak demand, gas units start and stop, and the need for efficient diagnostic technology is urgent. China's installed power capacity is high, and gas units, as a more flexible unit, are used in peak power production process, which leads to frequent start and stop of the unit. Frequent starting and stopping will cause the unit to malfunction, and at the same time, the efficiency of the unit is a higher requirement for calculating the ageing. Third, there is a lack of a unified management system for monitoring and diagnosis of the combustion engine system. Most of the gas turbines are equipped with control systems provided by the manufacturer, but the data structure and interfaces between the systems are different, often unable to have centralized and unified management [9].

The Siemens Fault Diagnostic System (D3000) combines more than 20 years of Siemens' experience in diagnostic calculations with an excellent combination of early warning and troubleshooting capabilities. The system is capable of modelling not only typical equipment but also systems and even processes. Trans-Alta Power of Canada commissioned Siemens to build an Operations Diagnostic Center, which was opened in March 2009. It was commissioned to monitor 11 units in 4 power plants with a monitoring capacity of 5000 MW and 40,000 monitoring points. The content includes vibration data, thermal performance, early intelligent alarms, KPIs and operating logs. Texas (TXU) Power Company in the USA commissioned Siemens to build a Power Optimization Center for a total of 24 power plants. Twenty-two units with a capacity of 18300 MW, including nuclear, coal, gas/oil and wind turbines, are monitored and controlled, with 50,000 monitoring points. Based on the concept of Internet of Things, big data and decentralized cloud platform, Huaneng Group built a big, unified standard, open interface and complete data management platform. The development of a real-time data unified platform for all the group's power generation equipment allowed online early warning and large-scale data analysis applications, such as energy saving analysis, operation analysis, condition analysis and evaluation, remote diagnostics, etc., in order to improve the production of power generation equipment on a management level. At the present, the data access work of the demonstration unit has been completed, with conditions for promotion.

At the present stage, power generation enterprises still generally adopt traditional after-treatment of equipment fault in equipment management, in which the operation personnel or inspection personnel judge and process the abnormal alarm by monitoring the parameters obtained from the DCS system of the unit, and then notify the maintenance personnel for processing.

However, this method can only be implemented when the equipment has already failed, and then, the plant maintenance personnel can only stop the equipment and switch to an alternate equipment. Repairs are performed on the malfunctioning equipment, and if there is no spare equipment or if the spare equipment fails to start on time, it will affect the entire plant. This can have a devastating effect on the

production process and can even lead to abnormal shutdown of the plant. Current threshold-based alarms do not alert the operation supervisors of the onset of a deterioration trend. There is a need for a means of monitoring that can reflect the operating status of equipment in real time.

Real-time monitoring of equipment status and rapid diagnosis of faults are an important part of a modern power warning system. The definition of condition monitoring is as follows: Through data collection and data analysis of a certain equipment, the prediction of some important trends of the entire equipment or some components of the equipment allows for maximum utilization of the equipment based on the life characteristics of each component before maintenance is required. The use of condition monitoring technology in electrical equipment facilitates early detection of serious malfunctions in the event of minor abnormalities. Maintenance is performed only when the equipment needs maintenance, extending the maintenance interval, thus avoiding accidents caused by equipment failure downtime and resulting in lower maintenance costs. Fault diagnosis helps operators to locate the point of failure, thus reducing the probability of a major accident and reducing the risk of forced downtime. This can avoid production losses, improve the reliability and availability of equipment and to protect the stability of the power production process [10].

Neural networks, regression analysis and other computer modelling methods have developed significantly, allowing for the combination of advanced data analysis methods and traditional power equipment state monitoring [11–13]. These methods, when utilized in solving the problem in traditional power industry production equipment, cannot grasp the real-time operating state, real-time monitoring of equipment status, early warning of potential accidents and rapid fault location, and cannot reduce maintenance costs to the maximum extent [14,15].

This study provides an in-depth analysis of issues in power information technology and enterprise production and operation and to promotes the Internet, big data, artificial intelligence and traditional deep integration of the power generation industry, in an effort to achieve the acquisition, monitoring, analysis and diagnosis of gas-fired generating cluster of operation data. This study also promotes “Internet +”, cloud computing and big data hybrid application. To realize the whole life cycle management of the energy group gas turbine, regional, industrial and inter-enterprise barriers are cut, which contribute to national energy industry production. The research provides new explorations as well as lessons learned [16].

## **2 Cloud Computing Network Architecture**

### **2.1 IoT Architecture**

Regarding the architecture of the Internet of Things, there are many different views on the number and division of layers in the academic world, and this project adopts a 4-layer architecture model with a perception layer, a network layer, an application layer and a user layer.

The perception layer consists of various sensors and sensor gateways, including flow, vibration, temperature and pressure sensors, as well as gas concentration, bearing temperature and vibration sensors, etc. The monitoring equipment includes the main equipment of the power plant above 6 KV, including combustion engine, condenser, waste heat boiler, auxiliary engine, steam turbine, generators, cooling towers and critical electrical equipment. The real-time operating status of these devices will be transmitted to the network layer via sensors in the form of switches, analogue or pulse quantities.

The network layer is to provide support for the convergence, integration, delivery and control of information at the sensing layer, while providing support for the data centre IoT man-machine communication exchange and providing an information platform [17]. To improve the level of network security, the network data extraction of the monitoring and diagnostic centre uses a one-way shutter, so that data can only be transmitted from the production site to the data centre; the external network cannot

send information to the production area. Additionally, a fortress machine and secure access system are deployed to improve the stability of the network system.

The application layer is the structure of IoT and user layer, which is combined with the need for power data monitoring and diagnostics to enable intelligent IoT applications. The main functions are: 1. Data acquisition and integration; 2. Equipment performance deterioration and fault prediction and warning; 3. Equipment thermal efficiency analysis; 4. Power plant key performance indicator (KPI) analysis; 5. Unit control optimization decision recommendations; 6. Equipment condition maintenance sequencing.

The user layer is the information output target of the data centre, including each power plant and the head office. For each power plant, the data centre focuses on equipment failure, performance degradation and thermal efficiency analysis, etc. For the group company, the service focuses more on the information. For group companies, the service focuses more on information collection and reporting of key performance indicators.

## **2.2 Cloud Computing Model**

The traditional hardware architecture and server computing power can hardly meet the data processing and management needs generated by the Internet of Things of the gas turbine fleet. To achieve centralized and efficient processing of fleet data, it is necessary to establish a large-scale computing platform using cloud computing technology as a low-level data computing support. Cloud computing architecture is a massive data processing computing platform that utilizes central service resources, as shown in Fig. 1. According to the computing resources provided, cloud computing can be divided into a collection of several types of services according to the presentation. It is divided into Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). There are three typical service models of Software as a Service (SaaS). IaaS provides users with physical or virtual computing, storage, and networking resources. To use it, the user needs to provide the IaaS layer service provider with configuration information for the infrastructure, code to run on the underlying setup, a set of services to run on the underlying infrastructure and associated user data. PaaS is the environment in which cloud applications run on, providing application deployment and management services. With the PaaS layer of software tools and development languages, application developers can focus on the code and data without having to think about the underlying management of resources, such as servers, operating systems, networks and storage. SaaS is developed on a cloud-based platform. Applications, targeted towards the end business users, can address power generation enterprise computing functions by renting SaaS layer services.

As power producers are becoming a more critical infrastructure, their management services and network information security are particularly important for cloud computing [18]. The cybersecurity and management regulatory level should be built with network access and cyber-attack protection services.

## **3 Monitoring and Diagnostic Methods**

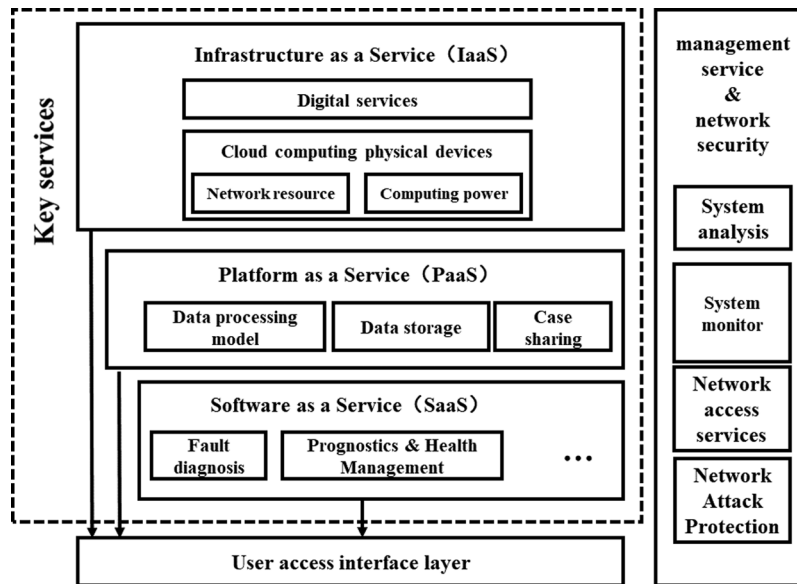
### **3.1 Data Processing**

The data monitoring and diagnostic models for IoT and cloud computing technologies are constructed on the basis of real time series of real-time transmitted data. The data quality has a high impact on the validity and accuracy of the model operation, as sensor failures, measurement noise or network transmissions can cause degradation of data quality. Therefore, when using the data in the model, one should treat the data further and to ensure that the data quality is improved. The following are the data governance methods used in this study [19].

(1) Data correlation filtering method: This uses the correlation between data to cut invalid data. This method is applied in the use of sensor redundancy measurements to calibrate the measurements of each

sensor; the similarity of the values of many redundant sensors should be high. The method is also effective in eliminating anomalous data from more correlated measurement points.

(2) Anomalous data screening: Anomalous values in the database and erroneous measurements may reduce the consistency and correlation of the data. These anomalous data are usually caused by faults in the sensors themselves, errors in the data transmission process, and changes in equipment maintenance. In this paper, a hypothesis testing approach is used to set a 95% data confidence interval to find the normal working data boundary condition and to filter abnormal data.



**Figure 1:** Cloud computing architecture for gas turbine monitoring and diagnostics

### 3.2 State Space Model

State space modelling is an effective way to monitor the state of the equipment. The first step in modelling is to select a sample from the reference data (X) and compose a state matrix (D), i.e., a process or equipment has n associated measurement points; set at some point a to sample it and collect from the n measurement points selected as a model [13].

$$X(i) = [x_1, x_2, \dots, x_n]^T \tag{1}$$

A large overhaul cycle is taken as the change node of the equipment condition, m modes are selected, and a state matrix (D) is formed.

$$D = [X(1) X(2) \dots X(m)] = \begin{bmatrix} x_1(1) & x_1(2) & \dots & x_1(m) \\ x_2(1) & x_2(2) & \dots & x_2(m) \\ \vdots & \vdots & \vdots & \vdots \\ x_n(1) & x_n(2) & \dots & x_n(m) \end{bmatrix}_{n \times m} \tag{2}$$

Each column of vectors in the state matrix represents a normal operating condition of the device. The subspace (D) formed by the historical patterns is capable of representing the entire dynamic process of the normal operation of the device. The composition of the entire state matrix is the learning of the operating characteristics of the device, and using a combination of these modes, produces an estimated model.

At some point, an input pattern  $X_{in}$  consists of a single reading for each sensor in the model.

$$X_{in} = [x_{1in}, x_{2in}, \dots, x_{nin}]^T \quad (3)$$

Comparing the similarity of the input pattern  $X_{in}$  with each pattern in the state matrix (D) yields the similarity vector (a), which contains the same number of elements as the training matrix (pattern) stored in the state matrix.

$$a = D^T \otimes x_{in} \quad (4)$$

Converting a similar vector representing the degree of similarity into a weight vector (W)

$$w_0 = G^{-1} \cdot a \quad (5)$$

$$G^{-1} = (D^T \otimes D)^{-1} \quad (6)$$

$$w = \frac{w_0}{\sum w_0} \quad (7)$$

where  $\otimes$  is a non-linear operator, chosen as the (Euclidean distance) between two vectors.

$$\otimes(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (8)$$

Estimates are generated by a linear combination of samples and weights.

$$x_{out} = D \cdot w \quad (9)$$

$$x_{out} = \frac{D \cdot (D^T \otimes D)^{-1} \cdot (D^T \otimes x_{in})}{\sum (D^T \otimes D)^{-1} (D^T \otimes x_{in})} \quad (10)$$

The performance and status of the equipment are monitored by calculating the residuals between the estimated value and the actual operating data. If the residuals exceed the threshold limit, an alert for monitoring and diagnostics is triggered.

$$Residual = x_{in} - x_{out} \quad (11)$$

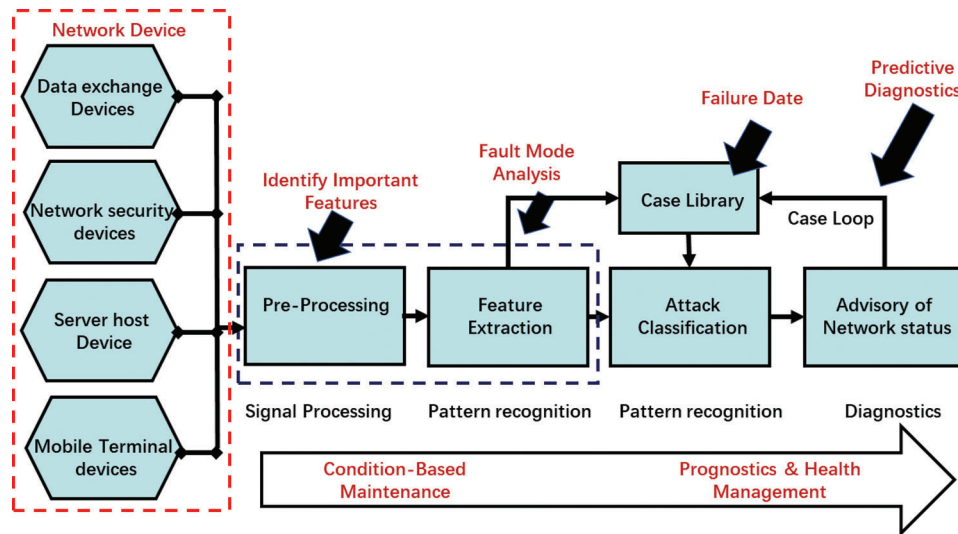
By subtracting the input mode from the estimated mode to produce a residual, the variable with the lower remaining value is treated as normal.

### 3.3 Construction of Monitoring and Diagnostic Models

The construction of a monitoring and diagnostic model can divide into two stages. One is the offline model construction stage, and the other is the online maintenance and correction stage after the model has been constructed offline and is in use. In the offline model construction phase, historical data are pre-processed. The relevant physical attributes of the monitoring device are selected as the input measurement points. The typical conditions under which the device operates are selected, and the conditions should cover as much of the device's operation as possible [20]. Common operating conditions include shutdowns, load lifts, and rapid drops. To improve the quality and stability of the model, the historical data require data pre-processing.

With reference to the model principle of state space, equipment-related physical quantities are used to build the state equations, and model fitting training is conducted to adjust the input data and residual threshold boundaries. The model is tested using available actual data to see if the residual alarm will be triggered.

In actual use, the state of the equipment is often not stable, and the monitoring and diagnostic model construction process is shown in Fig. 2. The change of equipment, the replacement of parts, and the switching of operating modes will change the normal operating conditions of the equipment.



**Figure 2:** Monitoring and diagnostic model construction method for gas turbine generator

Thus, online model construction and maintenance should be performed on the changed equipment to avoid false alarms. At this stage, the first real-time data are used for data cleaning, the failure condition is removed, and the model is re-trained.

A dynamic state space model for fusion and comparison of many sensor parameters are proposed for the high coupling between parameters of combined cycle units to solve the problem of spatial state monitoring and fault prediction of complex systems. The model based on the similarity principle, that is, the equipment is stable under normal conditions, and the measured data under similar conditions are similar to each other. The state model is populated with measurements from normal operation and the similarity of the real-time measurements is compared to the state in the state matrix. Then, the predicted value is deduced, according to the measured value and the predicted value of the residual changes, combined with a predetermined threshold, to give a fault warning.

However, with the maintenance of equipment, transformation, etc., the characteristics of the equipment are also more likely to change, when the original state model cannot adapt to the new devices. This model can add new states into the state space, replacing vectors that are no longer applicable, thus enabling dynamic tracking of the state model and updating.

The dynamic state spatial model is different from the traditional spatial model in that it can adjust the model volume in real time to follow the actual operating state of the equipment and adjust the model, i.e., adding or removing modelling data that do not meet the normal operating conditions of the equipment at any time, or adding associated measurement points to the model according to the operating state of the equipment to make the model more suitable for the actual operating state of the equipment.

The early warning of the dynamic state spatial model is different from that of the traditional DCS system alarm in power plants, which can send out an early warning at the early stage of equipment deterioration. At the same time, and different from the single-point alarm mechanism of the DCS system, the dynamic state space model can achieve many measurement points, and many evidence of equipment failure and

performance degradation can be warned. Additionally, thresholds are used to determine the dynamic boundaries of equipment failure.

The dynamic state space model can monitor the deviation between the actual operating state of the unit and the standard operating state in real time. The model is based on the transmission of real-time data from sensors in the field, which can help data analysts to determine the various parameters of the plant under different operating conditions. The analyst can detect when the actual operating state of the equipment deviates from the normal operating state, and whether the changes are in the normal range or anomalies. An eradication model can determine the scope and severity of the fault, from unplanned, passive inspection to planned, active inspection, thus providing time for power plant operation decisions.

### 3.4 Training Algorithm

The network is trained with algorithm, during the training phase wavelet parameters  $a$ ,  $b$ , etc. [21] are adjusted to minimize the error, given the  $d_i^p$  as the  $i$ th target output of  $p$ th input pattern, the evaluation function can be written as:

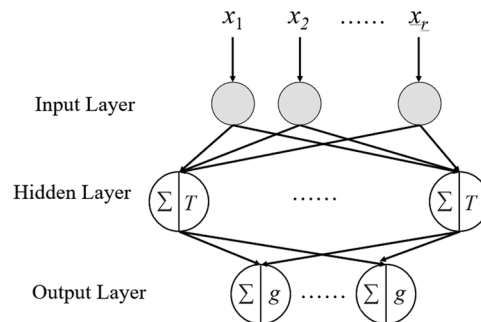
$$E = \frac{1}{2} \sum_{P=1}^P \sum_{I=1}^N (d_i^p - y_i^p)^2 \quad (12)$$

The selection of the mother wavelet is very important and depends on particular application. For this wavelet neural network, wavelet has been chosen to serve as a basis function, which has been the preferred choice in most work dealing with wavelet neural network, due to its simple explicit expression.

Then we can call  $T(t)$  basic wavelet

$$T_{a,b}(t) = \cos(1.75t_z) e^{-t_z^2/2\theta^2} \quad (13)$$

The neural network in this paper is designed as three layers with input, hidden and output layer (Fig. 3). To enhance the precision of wavelet neural network and accelerate the convergence speed, this paper adopts momentum algorithm. In this algorithm, there is a value proportional former changed value of weight adding to current weight during the error back-propagation period.  $y_j$  the  $j$ th component of the output vector.  $x_k$  the  $k$ th component of the input vector. The training algorithm is shown as follows:



**Figure 3:** Structure of neural network

$$w_{jk}(t+1) = w_{jk}(t) + \Delta w_{jk}(t+1) \quad (14)$$

$$v_{jk}(t+1) = v_{jk}(t) + \Delta v_{jk}(t+1) \quad (15)$$



$$a_{jk}(t+1) = a_{jk}(t) + \Delta a_{jk}(t+1) \quad (16)$$

$$b_{jk}(t+1) = b_{jk}(t) + \Delta b_{jk}(t+1) \quad (17)$$

$$\theta_{jk}(t+1) = \theta_{jk}(t) + \Delta \theta_{jk}(t+1) \quad (18)$$

The modifier value is shown as follows:

$$w_{jk}(t+1) = -\varphi \frac{dE}{dw_{jk}} + \mu \Delta w_{jk}(t) \quad (19)$$

$$v_{jk}(t+1) = -\varphi \frac{dE}{dv_{jk}} + \mu \Delta v_{jk}(t) \quad (20)$$

$$a_{jk}(t+1) = -\varphi \frac{dE}{da_{jk}} + \mu \Delta a_{jk}(t) \quad (21)$$

$$b_{jk}(t+1) = -\varphi \frac{dE}{db_{jk}} + \mu \Delta b_{jk}(t) \quad (22)$$

$$\theta_{jk}(t+1) = -\varphi \frac{dE}{d\theta_{jk}} + \mu \Delta \theta_{jk}(t) \quad (23)$$

The expression:

$$net_j = \sum_{k=0}^L w_{jk} x_{jk} \quad (24)$$

as a result

$$T_{a,b}(net_j) = T\left(\frac{net_j - b_j}{a_j}\right) \quad (25)$$

The training algorithm of the network used to minimize the error, and make the output of network match the target output.

## 4 Application of Monitoring and Diagnostic Technology

### 4.1 Application for Power Plants

For better validation of the model application, in this study, the F-class gas turbine (9F.05) will be used to fit the operating parameters of the equipment using the model. The gas turbine information is detailed in [Tab. 1](#).

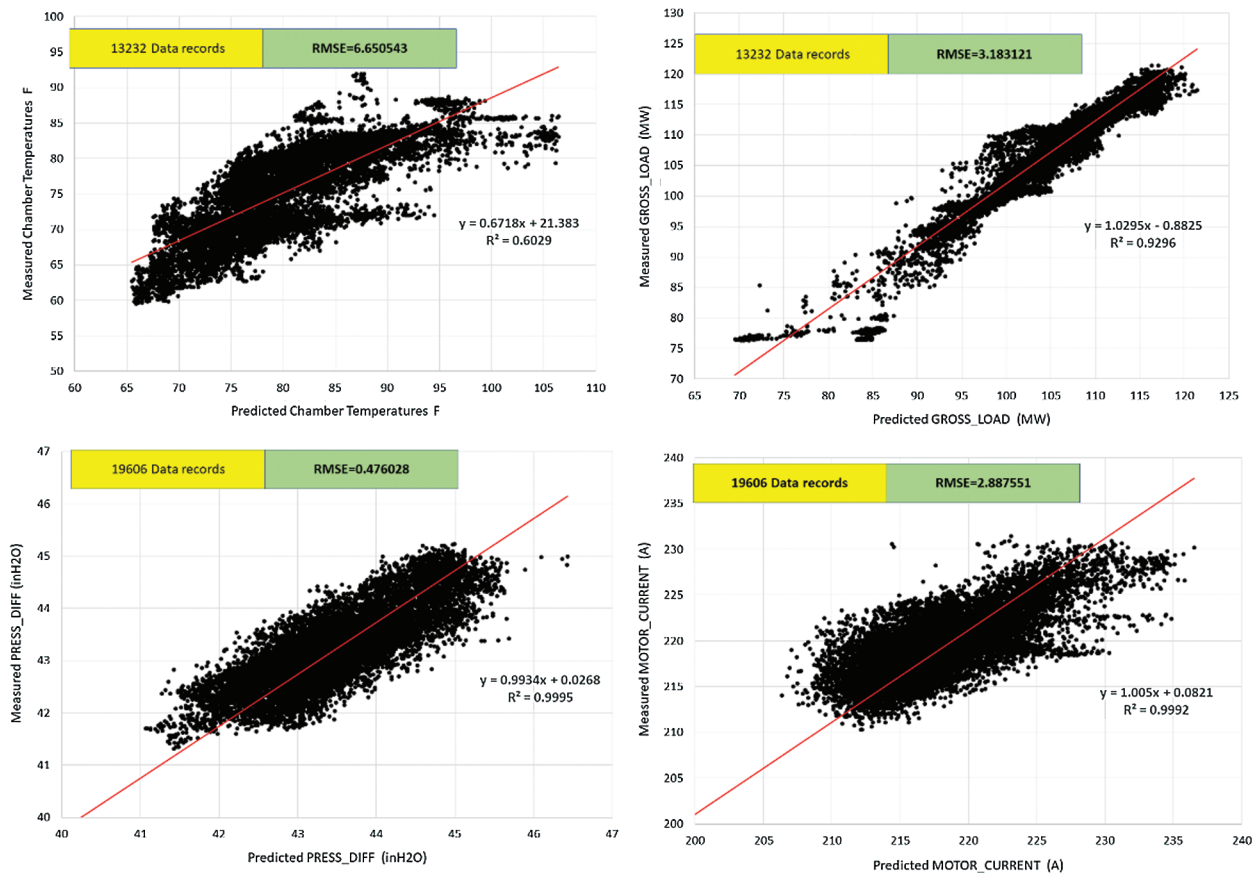
[Fig. 4](#) shows the parameters of gas turbine chamber temperature, pressure differential, load, and motor current calculated by the state-space model. The results show that the model has a good fit to the key parameters of the gas turbine operation.

The temperature value refers to the temperature of the entire compartment contained in the turbine frame of the combustion engine. Pressure refers to the pressure difference between the inside and outside of the exhaust frame, which indicates the tightness of the chamber. The data in our study are based on real data of the unit. The selected condition is the speed of the unit up to 3000 rpm. All other parameters are at the normal parameters of the current operation of the unit. The early warning threshold is usually set at  $\pm 10\%$  of this value, and is also adjusted according to the field operation. The power load is between 70 MW–200 MW (normal load), the ambient temperature is: humidity, temperature and pressure, the combustion

is pre-mixed and DLN2.6+ is available. The source data for the fault prediction model should contain the complete operating conditions of the unit, including the unit load adjustment, ambient temperature changes, etc.

**Table 1:** The gas turbine parameters (9F.05)

Plant Performance	9F.05
Net Output (MW)	314
Net Heat Rate (Btu/kWh, LHV)	8,846
Net Heat Rate (kJ/kWh, LHV)	9,333
Net Efficiency (% , LHV)	38.6%
The Ambient Conditions (ISO Conditions)	59°F/15°C, 60% RH
The Turbine Pressure Ratio (Expansion Ratio)	17:1
Inlet Air Flow (Kg/S)	847



**Figure 4:** Fitting results for operating parameters of gas turbine generator

The dataset of long-term gas turbine performance data includes a number of outlier data records. These few outliers limit the prediction accuracy that can be achieved by optimized data matching. In Fig. 4, the

predicted chamber temperatures fitted using neural networks have good similarity with an R squared of 0.6. And The results of the combustion engine load fit have good similarity with an RMSE of 3.18. The analysis of the differential pressure fit revealed an R squared of 0.99. These show that the main operating parameters of gas turbine can be well fitted by unsupervised learning algorithms.

The burner fleet of IoT user-oriented applications monitor and record equipment operating history data and receive the operating fault alarm signals of main and auxiliary engine equipment, followed by timely processing. For example, when the unit load remains stable, but the chamber pressure and temperature drop at the same time, the model indicates that the chamber seal may have failed. The use of such remote monitoring and diagnostic technology can monitor the status of equipment, reducing the workload of operators and inspection personnel. Using the model residual analysis method, one can analyse the small slow changes in equipment failure state, to ensure that combustion equipment is under safe and economical operation of the company.

#### 4.2 Effectiveness Evaluation

There are currently three statistical methods to examine the accuracies of early warnings, and the principles of the three calculation methods are explained below [22]. The valid case early warning statistics method calculates the accurate rate of early warnings through the ratio of valid cases to the total number of early warnings. This is shown by the following formula.

The shortcomings of this approach are as follows:

$$\text{Alerts Accuracy} = \left[ \frac{\text{Customer Confirmed Cases}}{\text{Total alerts}} \right] \quad (26)$$

Many alerts sometimes produce only one case. Counting the number of cases and the total number of alerts are two different statistics and are not comparable. Cases and alerts have different processing cycles. Alerts that do not produce cases will be processed much faster than those that do. The method will reduce effective alerts in disguise. When equipment malfunctions occur, many alerts are often generated, and many alerts can only be combined to generate one case.

Transient early warning statistics:

$$\text{Alerts Falserate} = 1 - \frac{[\text{Dismissed Transient alerts}]}{[\text{Total Active alerts}]} \quad (27)$$

This method defines the ratio of the number of discharged transient warnings to the total allowed warnings in a statistical period as the false alarm rate. The warning accuracy is obtained by converting the false alarm rate.

The total allowable alerts for a training-qualified state space are generally considered to be a fixed value, which is the total number of data centre monitoring equipment models, i.e., 60%. Each model can trigger an early warning, and no more than 60% of the total number of models should warn of the activity allowed in the system. See the following formula for details.

The shortcomings of this approach:

$$\text{Alerts Falserate} = 1 - \frac{[\text{Dismissed Transient alerts}]}{[\text{Dismissed alerts}]} \quad (28)$$

This statistical method suffers from the problem that the numerator and denominator are not completely exclusive. Any early warning is closed after it has been generated and processed. In any case, processing alerts will end up existing in the system in three states: model maintenance, closed loop resolution, and

transient alert. However, the total number of alerts in this statistical method is the number of transient alerts (molecules) that are still active, and the number of transient alerts that have been released (molecules) are not present in the total number of early warnings.

An early warning is usually indicated by a malfunction in the equipment. Alerts can also be issued when the equipment performance deteriorates or when a new condition (e.g., test, switchover) occurs. There are different attributes for active and discharged alerts. Active alerts are still being processed, while discharged alerts have outcomes. The early warning accuracy should be evaluated according to the different statistical methods.

As mentioned earlier, early warnings can be divided into active warnings and discharged warnings. Active alerts are actually still being processed and their validity cannot be defined. However, lifted alerts are those that have been analysed and concluded by the data centre, and the analysis of lifted alerts is important for evaluating the performance of the data centre. The method defines the early warning false alarm rate as the ratio of transient warnings to the total number of discharged warnings.

This method of early warning statistics, as proposed in this study, has the following advantages.

1. This method does not count all early warnings, but rather the lifting of early warnings that have already been concluded as a result of the analysis process. Since early warnings have different states at different stages, active warnings and lifted warnings cannot be mixed together for calculation and comparison. This method is used to evaluate the accuracy of the released alerts.

2. The numerator and denominator of the statistics is itself a relationship of inclusion, that is, the lifting of the transient warning is in the total of the lifting of early warning, with some statistical significance.

3. Model maintenance tuning is considered as an effective early warning. In the model sense, the warning generation itself is not necessarily an equipment failure; it actually represents a kind of equipment that does not match the model. Anomalies, or new operating conditions that do not exist in the model, occur. The identification of the new condition is also done through early warning to prompt the new condition data into the model for training. Thus, early warning does not only indicate equipment anomalies but is also an important process and tool for model optimization and machine learning.

By utilizing this method for our models, the accuracy is between 60–80%. It should also be noted that the false alarm rate of the model is at odds with the prediction time of the model. The sooner you want to predict failures, the more false alarms tend to result, which need to be thought of as de-screening.

#### ***4.3 Application for Energy Group Companies***

Through online monitoring of the gas turbine fleet, the group can obtain the basic data and information of the whole group's combustion engine fleet, providing a data basis for industrial management and the formulation and implementation of energy savings and emission reduction policies. Potential safety hazards can be discovered in a timely manner to facilitate safe supervision and operation. The level of information industrialization integration of the power production sector can be improved as a whole. The integration of the control of production and operation of the combustion engine fleet can be better realized. Production costs and management costs can be reduced, and energy utilization efficiency can be improved, thereby achieving energy savings and emission reduction.

Through the implementation and in-depth application described in this study, obvious social benefits can be brought to large power companies. This study adopts the concept of industrial Internet, combines advanced information technology with traditional power generation industry, and uses information technological means to change the traditional industrial management mode of operation and management. This reduces the cost of operation and management, enhances the level of enterprise management, shortens the maintenance time, and reduces the cost of maintenance. Longer maintenance intervals,

improved energy efficiency, fewer failed starts, more efficient engineering analysis, increased plant capacity, lower production cost and improved system reliability can be achieved.

The Internet of Things, big data, cloud computing and other advanced technologies can provide early warning, fault diagnosis and performance optimization for the whole life cycle of gas turbine equipment and value-added services, improve the reliability of gas-fired generating units, optimize their operating economics, reduce operation and maintenance costs, and promote energy conservation. For example, the diagnostic technology detected a deviation on 9F.05 gas turbine at a power plant. Specifically, the exhaust temperature spread registered a step increase from 33°C to 69°C, while the model estimated a value of 33°C. Corresponding deviations were noted on multiple exhaust thermocouples around the fuel nozzles. After receiving the alert, the operator investigated the issue and decided to increase the load. This decision was to rule out any relation of high exhaust spread with load changes. In parallel, the customer prepared for maintenance by arranging the fuel nozzles and associated materials for execution during next planned outage. Later, the spread further increased from ~69°C to ~151°C during a brief unit run as the load increased to ~22 MW on gas fuel operation.

Moreover, it can provide environmental protection, cut regional, industry, and corporate barriers to promote clean energy generation based on achieving full life management of gas turbine power generation, expand and extend technical services in the region, explore a new model of clean power production and operation, and take this opportunity to digest and assimilate. The experiences and realization of the application towards the entire power generation industry are achieved.

On the basis of realizing the whole life management of a large number of combustion engines, we will explore a new mode of clean power generation, production and operation, and take this opportunity to digest and absorb the experience and realize the application towards the entire power generation enterprise, which will become a development direction of the energy industry and power systems, promoting the traditional industry and producing far-reaching effects.

## 5 Conclusion

A monitoring and diagnostic technique that integrates the use of mathematical and physical models is developed. A dynamic state space model including multi-sensor parameter comparison is proposed to address the high coupling between parameters of the combustion engine system. The complex system states and the problem of low stability of fault monitoring results are addressed. Building typical failure modes from combustion engine operating history data improves the accuracy of fault triggering alarms through physical models. The technology can identify faults in the early stages of equipment failure based on potential problems with the equipment.

The research builds a cloud platform for monitoring and diagnosis of a combustion engine fleet to achieve whole-system, whole-process knowledge innovation management. The research and application of the combustion engine fleet monitoring and diagnostic cloud platform technology has led to the realization of full life cycle management of large power group of combustion engines, eliminating regional, industry and inter-firm barriers.

**Funding Statement:** This paper is supported by the China Datang Group Corporation of Science and Technology Project Plans (DT/KJ/2013-42).

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

1. Goel, P., Datta, A., Mannan, M. S. (2017). Industrial alarm systems: Challenges and opportunities. *Journal of Loss Prevention in the Process Industries*, 50, 23–36. DOI 10.1016/j.jlp.2017.09.001.
2. Sanin, C., Zhang, H., Shafiq, I., Waris, M. M. Oliveira, C. S. D. et al. (2019). Experience based knowledge representation for Internet of Things and Cyber Physical Systems with case studies. *Future Generation Computer Systems*, 92, 604–616. DOI 10.1016/j.future.2018.01.062.
3. Yang, Y., Li, X., Yang, Z., Wei, Q. Wang, N. et al. (2018). The application of cyber physical system for thermal power plants: Data-driven modeling. *Energies*, 11(4), 690. DOI 10.3390/en11040690.
4. Hou, G., Bai, X., Zhang, J., Zhao, Z. (2014). Multi-model predictive control based on neural network and its application in power plant. *Proceeding of the 11th World Congress on Intelligent Control and Automation*, 1, 4379–4383. DOI 10.1109/WCICA.2014.7053450.
5. Rostek, K., Morytko, U., Jankowska, A. J. E. (2015). Early detection and prediction of leaks in fluidized-bed boilers using artificial neural networks. *Energy*, 89, 914–923.
6. Gong, G. J., Sun, Y., Cai, M. M., Wu, R. Z., Tang, L. R. (2011). Research of network architecture and implementing scheme for the internet of things towards the smart grid. *Power System Protection & Control*, 39 (20), 52–58.
7. Ben Ali, J., Fnaiech, N., Saidi, L., Chebel-Morello, B., Fnaiech, F. J. A. A. (2015). Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Applied Acoustics*, 89(3), 16–27.
8. Chen, X., Wang, P., Hao, Y., Zhao, M. J. N. (2018). Evidential KNN-based condition monitoring and early warning method with applications in power plant. *Neurocomputing*, 315, 18–32.
9. Wei, C. Y., Xu, J., Liao, S. Y., Sun, Y. Z., Jiang, Y. B. (2018). A bi-level scheduling model for virtual power plants with aggregated thermostatically controlled loads and renewable energy. *Applied Energy*, 224(C), 659–670.
10. Li, Z., Xiong, X., Wang, W., Zhou, N., Wang, J. J. I. G. T. et al. (2016). Early warning method for transmission line galloping based on SVM and AdaBoost bi-level classifiers. *IET Generation Transmission & Distribution*, 10(14), 3499–3570.
11. Lu, H., Zhang, Y., Wu, C., Sun, W. J. E. P. (2012). Dynamic model identification of the main Steam temperature for supercritical once-through boiler. *Energy Procedia*, 17, 1704–1709.
12. Mohamad, W. M. W., Ahmad, T., Karim, N. A. A., Ashaari, A. (2018). Fuzzy arithmetical modeling of a steam turbine and a boiler system. *Mathematical Modelling and Analysis*, 23(1), 101–116.
13. Zhang, W., Liu, J., Gao, M., Pan, C., Huusom, J. K. (2019). A fault early warning method for auxiliary equipment based on multivariate state estimation technique and sliding window similarity. *Computers in Industry*, 107, 67–80. DOI 10.1016/j.compind.2019.01.003.
14. Caesarendra, W., Lee, J. M., Ha, J., Choi, B. K. (2015). Slew bearing early damage detection based on multivariate state estimation technique and sequential probability ratio test. *IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, 1161–1166.
15. Jiang, Z., Hu, M., Feng, K., Hao, W. J. S. (2018). A SVDD and K-Means based early warning method for dual-rotor equipment under time-varying operating conditions. *Shock and Vibration*, 2018(937), 1–14.
16. Li, J. Q., Yu, F. R., Deng, G., Luo, C., Ming, Z. et al. (2017). Industrial internet: A survey on the enabling technologies, applications, and challenges. *IEEE Communications Surveys and Tutorials*, 19(3), 1504–1526. DOI 10.1109/COMST.2017.2691349.
17. Di, Y., Jin, C., Bagheri, B., Shi, Z., Ardakani, H. D. et al. (2018). Fault prediction of power electronics modules and systems under complex working conditions. *Computers in Industry*, 97, 1–9.
18. Ciancamerla, E., Fresilli, B., Minichino, M., Patriarca, T., Iassinovski, S. (2014). An electrical grid and its SCADA under cyber attacks: Modelling versus a Hybrid Test Bed. *2014 IEEE Proc. International Carnahan Conference on Security Technology*, 1–6.
19. Goel, P., Datta, A., Mannan, M. S. (2017). Application of Big Data analytics in process safety and risk management. *Proceedings of the 2017 IEEE International Conference on Big Data (Big Data)*, 1143–1152.

20. Cui, C., Lin, W., Yang, Y., Kuang, X., Xiao, Y. J. M. (2018). A novel fault measure and early warning system for air compressor. *Measurement*, 135, 593–605.
21. Qingyang, X., Xianyao, M., Ning, W. (2008). Gas turbine fault diagnosis based on ART2 Neural Network. *Proceedings of the International Conference on Wavelet Analysis & Pattern Recognition*.
22. Zhou, Q., Xiong, T., Wang, M., Xiang, C., Xu, Q. J. E. (2017). Diagnosis and early warning of wind turbine faults based on cluster analysis theory and modified ANFIS. *Energies*, 10(7), 898.