IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 16, 2020, accepted October 29, 2020, date of publication November 3, 2020, date of current version November 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3035622

A Degradation Fault Prognostic Method of Radar Transmitter Combining Multivariate Long Short-Term Memory Network and Multivariate Gaussian Distribution

YUTING ZHAI[1,2], AND SHAOJUN FANG[1], (Member, IEEE)
[1]Department of Information Science and Technology, Dalian Maritime University, Dalian 116026, China
[2]Department of Information Systems, Dalian Naval Academy, Dalian 116018, China

Corresponding author: Shaojun Fang (fangshj@dlmu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51809030.

ABSTRACT In the prognosis of radar transmitter degradation fault, there are some problems, such as the total sample size and fault sample size of sensor monitoring data are small, and the monitoring data can not reach the fault threshold. To solve these problems, a prediction model combining the multivariate long short-term memory networks with multivariate Gaussian distribution is proposed, in which the long short-term memory networks predict the subsequent time step of multi-sensor monitoring data, and the multivariate Gaussian distribution model constructed by few fault samples is used to realize the prediction of degraded faults. The key parameters of the model are determined by the data processing experiments of double monitoring points, and the validity of the model is verified by the data processing experiments of multiple monitoring points. The experimental results show that the degradation fault can be predicted effectively within 10% of the total time series of monitoring data. Compared with the traditional radar warning after failure, the model can effectively predict the degradation fault of the transmitter when the fault sample size is low and the fault threshold is not reached.

INDEX TERMS Multivariate long short-term memory network, multivariate Gaussian distribution, prognostic method, degradation fault of radar transmitter.

I. INTRODUCTION

Marine radar equipment is usually in a long-term working state. The transmitter is the core part of the radar. Although the transmitter has met the service life requirements of the device when it leaves the factory, it is prone to premature degradation under adverse conditions such as complex marine environments and frequent use. The maintenance after degradation of the transmitter has failed to meet the actual demand. Therefore, the prognosis of the degradation of the radar transmitter becomes very important. With the increasing complexity of marine radar systems, it becomes more and more difficult to predict the degradation fault of radar transmitters. Relying on professionals to monitor the characteristics of each monitoring point is not only inefficient but also difficult to predict the degradation fault. In this case, the concept of prognostic and health management (PHM) emerged [1]. It refers to using sensors to collect various data information of the system, with the help of various inference algorithms and intelligent models to monitor, predict and manage the state of the system, estimate the health of the system itself, and make effective predictions before the system fails.

The method of degradation fault prognosis for radar transmitter needs to meet the following requirements:

(1) The degradation fault prognosis for radar transmitter can be completed by using the small fault samples. For the radar transmitter, it has met certain service life requirements when it leaves the factory, and may not have degradation faults for a long period of monitoring time, that is, the number of fault samples is small.

(2) The degradation fault prognosis for radar transmitter can be realized without setting fault threshold artificially. Sometimes, monitoring indicators of the radar transmitter are

The associate editor coordinating the review of this manuscript and approving it for publication was Anton Kos.

within the normal operating range, but the radar transmitter is close to its useful life. Therefore, it does not make much sense to set a fault threshold on the data of the monitoring points.

(3) The degradation fault prognosis for radar transmitter can be achieved by using the small historical data. Due to the influence of ship navigation, the number of historical data samples is often insufficient, and the sensor can not provide equal interval continuous time series.

The prognosis research was originated from medicine [2] and is now widely used in the fields of machinery [3]–[6], electric energy [7]–[9], finance [10], [11], transportation [12]–[14], etc. Common fault prognosis methods include traditional physical fault model prognosis methods [15], [16] and the more popular data-driven prognosis methods in recent years [17]–[24]. For the radar transmitter, the physical fault model prognosis method is difficult to deal with the complexity and randomness of the system [25], so it will be subject to certain restrictions. In the data-driven prognosis method, although [17] and [18] have higher fault prognosis performance, they need to use thousands of historical data as the training sets. In [19] and [20], it is necessary to set the fault threshold to judge the predicted data. The historical data in [21] needs to provide time series with equal time intervals. In the experiment, the number of fault samples adopted by [22] and [23] is almost equal to that normal samples. Reference [24] verifies that the supervised learning method is not suitable when the number of fault samples is small. Unfortunately, none of the above studies can meet the requirements for the prognosis of radar transmitter degradation fault.

Research results on fault prognosis of radar emerge constantly in recent years. For instance, in [20], the fault prediction model based on analytic hierarchy process with BP neural network is used to predict the fault of the specific model digital chip in a certain radar analog-to-digital conversion module. Reference [26] analyzed the nonparametric regression kernel function prediction method to obtain the function relation between the predicted data series and historical data series, and calculate remaining useful life before radar failure. Reference [27] proposed an extraction method of key lifetime parameters for the transmitter/receiver module based on association rule. To research the health management of radar transmitting power supply, [28] proposed a new method to find suitable monitoring points, which focused on how to choose the optimal monitoring point. In [29], a diagnosis approach based on deep learning is proposed to predict the meteorological radar fault by a dynamic threshold. The above research results have played an important role in promoting the development of fault prognosis of radar. However, these techniques do not satisfy the requirements of radar transmitter degradation fault prognosis mentioned before.

Therefore, a new data-driven prognosis method is proposed to predict the transmitter degradation fault. The main contributions of this article can be summarized as follows:

(1) The degradation prognosis method of radar transmitter combining multivariate long short-term memory networks and multivariate Gaussian distribution is proposed.

(2) Solve the problem of radar transmitter degradation fault prognosis when the total sample size and fault sample size of sensor monitoring data are small, and the monitoring data below the fault threshold.

The paper is organized as follows: Section 2 analyzes the proposed prognostic algorithm theoretically. The experiment in Section 3 verifies the feasibility and portability of the algorithm and discusses the selection of key parameters. Section 4 gives the conclusions.

II. PROGNOSTIC METHOD ANALYSIS

In this article, a prediction model combining the multivariate long short-term memory networks with multivariate Gaussian distribution is proposed for evolving the tasks of prognosis of radar transmitter degradation fault. The motivations and solutions of this article can be concluded as follows:

(1) The data collected by each monitoring point of the radar transmitter contains sensor noise, so it needs to denoise. Here a classical denoising method is selected, namely wavelet denoising.

(2) There are multiple factors that cause degradation of the radar transmitter. Therefore, multiple monitoring points are set up in the radar transmitter, which can monitor the working data of the transmitter over time. The time sampling data of each monitoring point will form a time-series. In the data-driven prognosis method, the long short-term memory (LSTM) neural network has the characteristics of memorizing long-term information. The LSTM neural network has achieved excellent results in the prognosis of time-series data. Therefore, the proposed algorithm in this article uses LSTM to learn and predict the time-series data of each monitoring point.

(3) The correlation between the multivariate time-series data monitored by each monitoring point is not clear, the data fluctuates within the normal range, and the number of degradation fault samples is small, so this article use the multivariate Gaussian distribution model to establish correlation for the time-series. The modelling process only needs a few fault samples as cross-validation sets to learn multivariate Gaussian distribution parameters, then the multivariate Gaussian distribution model can be used to determine whether the predicted value of LSTM is a degradation fault.

The algorithm mainly considers three parts (Figure 1). They are wavelet threshold denoising, long short-term memory network and multivariate Gaussian distribution,which are described in detail below.

A. WAVELET THRESHOLD DENOISING

Since most of the time-series data monitored by the sensor contain noise, it must be denoised and preprocessed. Here the classic wavelet threshold denoising method is chosen. The basic idea of wavelet denoising [30] is that the original sequence after wavelet transformation will generate a

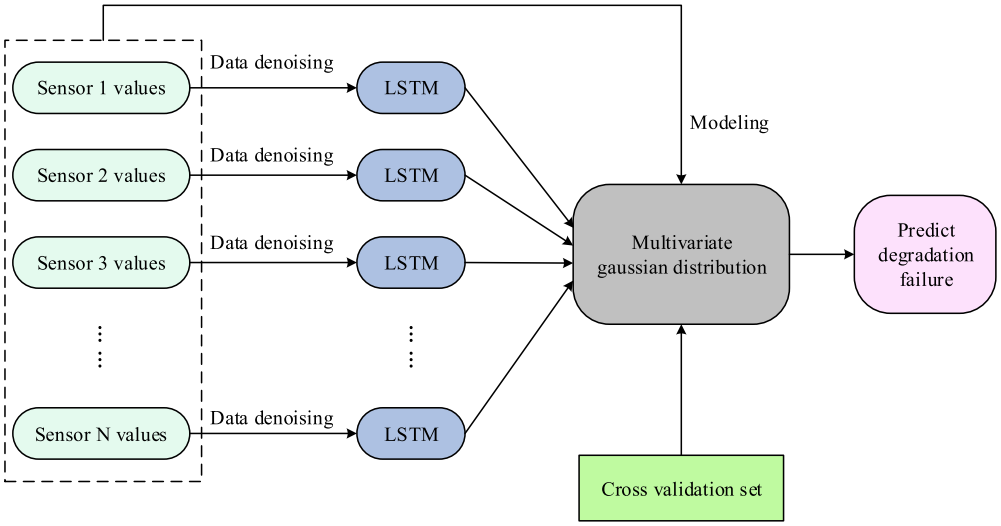


FIGURE 1. Model structure and algorithm flow chart.

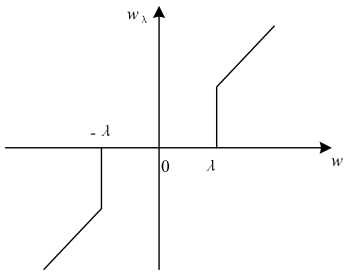
wavelet coefficient w, the wavelet coefficient of real data is larger and the wavelet coefficient of noise is smaller. Set a suitable threshold λ, the wavelet coefficients larger than λ are considered to be generated by real data, keep them, while those coefficients smaller than the threshold are considered to be generated by noise, set to zero to achieve the purpose of denoising. The selection of threshold function is the core of threshold denoising. Classic threshold functions include hard threshold function (Figure 2a) and soft threshold function (Figure 2b). The expressions are Equation (1) and Equation (2) respectively. The hard threshold function will produce additional oscillation and jump points after denoising, which does not have the smoothness of the original data. The soft threshold [31] has a good overall continuity of wavelet coefficients, and no additional oscillation will be generated.

$$w_\lambda = \begin{cases} w & |w| \geq \lambda \\ 0 & |w| < \lambda \end{cases} \tag{1}$$

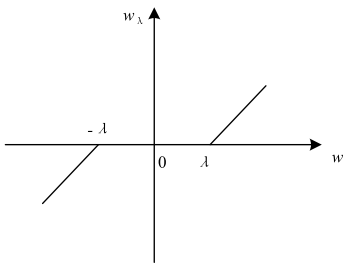
$$w_\lambda = \begin{cases} \text{sgn}(w)(|w| - \lambda) & |w| \geq \lambda \\ 0 & |w| < \lambda \end{cases} \tag{2}$$

B. LONG SHORT-TERM MEMORY NETWORK

LSTM is a special Recurrent Neural Network (RNN), which can learn long-term dependent information, originally proposed by Hochreiter and Schmidhuber [32]. It has been widely used in various time-series prognosis problems and has achieved great success. To know what LSTM is, it is necessary to start with RNN. The traditional deep neural network contains a cyclic structure, as shown in Figure 3a. This cycle seems very mysterious, but RNN can think of this cycle as multiple copies of the same neural network, each neural network module will pass the message to the next, as shown



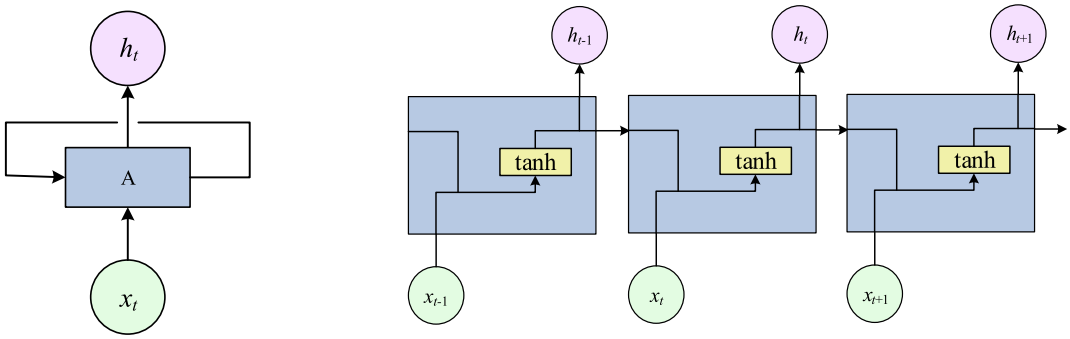
(a) Hard threshold function.



(b) Soft threshold function.

FIGURE 2. Threshold function.

in Figure 3b. RNN can connect the previous information to the current task, but unfortunately, when the distance between the previous information and the current task point increase, RNN will lose the ability to connect the long-distance information to the current task point. The main reason is that RNN uses back-propagation training in long-term sequences, and the growth or contraction of back-propagation gradients accumulates at each time step. This accumulation will cause the gradient to explode or disappear in many time steps [33] and cause long-term dependence of RNN. The repeating module in LSTM can avoid the problem of long-term dependence by the design shown in Figure 4.



(a) Loop structure of traditional deep neural network.

(b) Loop structure of RNN.

FIGURE 3. Network loop structure.

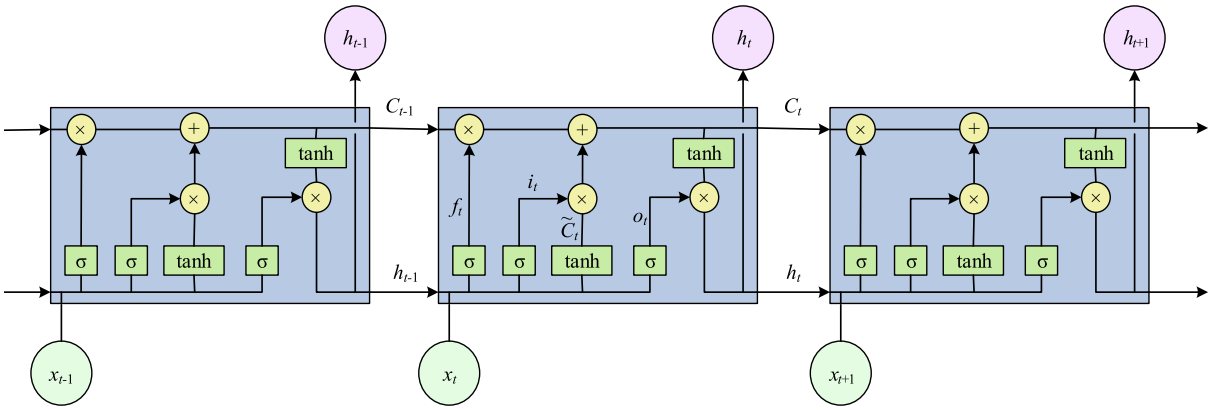


FIGURE 4. Loop structure of LSTM.

The working process of each LSTM cell can be described by the following equation:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{3}$$
$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{4}$$
$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{5}$$
$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{6}$$
$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{7}$$
$$h_t = o_t \times \tanh\left(C_t\right) \tag{8}$$

Among them, W_f, W_i, W_C, W_o are the weights corresponding to the forget gate f_t, input gate i_t, modulation gate $\tilde{C}_t$ and output gate o_t. b_f, b_i, b_C, b_o are the bias parameters corresponding to each function gate.

When using LSTM to predict the data, it can be divided into a single-step prognosis with updates and a multi-step prognosis without updates. Single-step prognosis with updates means that you can access the real value of the previous time step to update the network before predicting the next time step, while the multi-step prognosis without updates can only use the predicted value of the previous time step to predict the next time step. Usually, the prediction result of the single-step prognosis with updates is better than the multi-step prognosis without updates. The quality of the LSTM prediction result can be measured by the root mean square error (RMSE), as shown in Equation (9), where $h(x_i)$ is the forecasted value and y_i is the observed value. The smaller the RMSE value, the better the LSTM prediction result. The above structural characteristics of LSTM make it have outstanding performance in time-series data prediction, and the radar transmitter degradation fault prognosis is to predict the data (time-series) of each monitoring point of the transmitter, so LSTM can meet our needs.

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left[h\left(x_i\right) - y_i\right]^2} \tag{9}$$

C. MULTIVARIATE GAUSSIAN DISTRIBUTION

Multivariate Gaussian distribution is a high-dimensional generalization of univariate normal distribution. It can automatically capture the correlation between multiple features and establish a model, so the multivariate Gaussian distribution is suitable for multi-feature anomaly prediction [34].

For multivariate feature variable set $x_1 = \left\{x_1^1, x_1^2, \ldots, x_1^m\right\}$, $x_2 = \left\{x_2^1, x_2^2, \ldots, x_2^m\right\}$, $\ldots$, $x_n = \left\{x_n^1, x_n^2, \ldots, x_n^m\right\}$, each feature variable follows Gaussian distribution, it can be written as Formula (10), and its probability density is Equation (11). The expectation and variance of each feature variable set are shown in Equation (12) and (13). At this time, to determine whether the new sample $x^{\text{test}} = \left[x_1^{\text{test}}, x_2^{\text{test}}, \ldots, x_n^{\text{test}}\right]$ is abnormal, you need to calculate $P\left(x^{\text{test}}\right)$ and compare it with the threshold parameter ε. However, in practice, there is often the high value of $P\left(x^{\text{test}}\right)$, and the new sample is indeed an abnormal sample, which cannot effectively predict the abnormality of multiple features. This is caused by the traditional Gaussian distribution model ignoring the correlation between multiple characteristic variables. The multivariate Gaussian distribution can solve the above problems by establishing a correlation model between multivariate feature variables.

$$x_1 \sim N\left(\mu_1, \sigma_1^2\right), x_2 \sim N\left(\mu_2, \sigma_2^2\right), \ldots, x_n \sim N\left(\mu_n, \sigma_n^2\right) \tag{10}$$

$$P(x) = \prod_{j=1}^{n} P\left(x_j; \mu_j, \sigma_j^2\right) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{\left(x_j - \mu_j\right)^2}{2\sigma_j^2}\right] \tag{11}$$

$$\mu_j = \frac{1}{m}\sum_{i=1}^{m} x_j^i, \quad j = 1, 2, \ldots, n \tag{12}$$

$$\sigma_j^2 = \frac{1}{m}\sum_{i=1}^{m}\left(x_j^i - \mu_j\right)^2 \tag{13}$$

The multivariate Gaussian distribution model can be expressed as Equation (14), the expected vector is shown in Equation (15), and Equation (16) is a covariance matrix, which contains the correlation information between each feature variable. When using multivariate Gaussian distribution to predict fault data, the method use training data to obtain a multivariate Gaussian model, and then use Equation (14) to calculate the probability density of new samples, finally, compare the obtained probability value with the adaptive threshold ε to predict whether it is a fault data.

$$P\left(x; \mu, \sigma^2\right) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\left(x^i - \hat{\mu}\right)^{\mathrm{T}} \Sigma^{-1}\left(x^i - \hat{\mu}\right)\right] \tag{14}$$

$$\hat{\mu} = [\mu_1, \mu_2, \ldots, \mu_n], \quad \hat{\mu} \in \mathbb{R}^n \tag{15}$$

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m}\left(x^i - \hat{\mu}\right)\left(x^i - \hat{\mu}\right)^{\mathrm{T}}, \quad \Sigma \in \mathbb{R}^{n \times n} \tag{16}$$

The evaluation standard of a multivariate Gaussian distribution model can be measured by F_1 value, its calculation is shown in Equation (17)-(19), the closer the F_1 value to 1, the better the model prediction effect. where P is the precision metric, R is the recall metric. n_{tp} is the number of true positives: the ground truth label says it's an anomaly and our algorithm correctly predicted it as an anomaly. n_{fp} is the number of false positives: the ground truth label says it's not an anomaly, but our algorithm incorrectly predicted it as an anomaly. n_{fn} is the number of false negatives: the ground truth label says it's an anomaly, but our algorithm incorrectly predicted it as not being anomalous. The selection of the threshold parameter ε can be determined by setting the cross-validation set. In order to obtain the maximum F_1 value on the cross-validation set, different thresholds should be put into test. At this time, the corresponding threshold is ε. In the process of training the model, the training data can be all normal samples, and only a few fault samples (or even only one fault sample) are needed on the cross-validation set to determine the threshold. The entire modeling process takes into account the correlation between the characteristic variables and does not require a large number of fault samples, which can also meet our requirements for radar transmitter degradation fault prognosis.

$$F_1 = \frac{2PR}{P+R}, \quad F_1 \in [0, 1] \tag{17}$$

$$P = \frac{n_{\text{tp}}}{n_{\text{tp}} + n_{\text{fp}}} \tag{18}$$

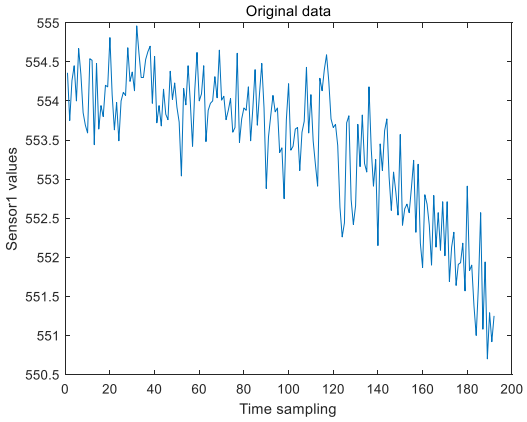
$$R = \frac{n_{\text{tp}}}{n_{\text{tp}} + n_{\text{fn}}} \tag{19}$$

III. EXPERIMENT AND ANALYSIS

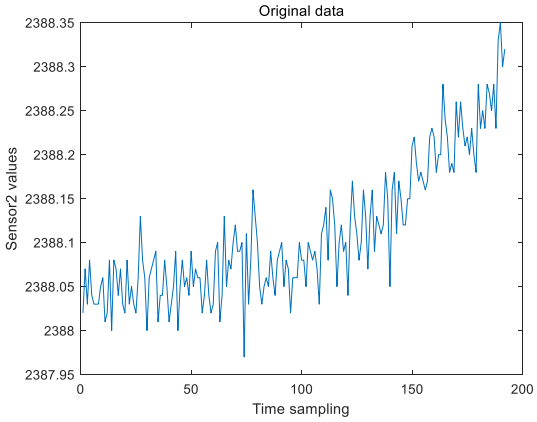
To verify the effectiveness of our proposed method, two experiments are set up in this section. Experiment 1: it is the experiment to perform degradation fault prognosis on the data of radar transmitter monitoring points. For the convenience of visualization, the data of two monitoring points are selected in the first experiment. Experiment 2: it is the degradation fault prognosis experiment on the data of multiple monitoring points of the radar transmitter. It should be noted that Experiment 2 increases the number of monitoring points based on Experiment 1 and verifies the feasibility of the method under different degradation fault conditions. The purpose of the first experiment is to prove the rationality of our proposed method in the form of visualization. The second experiment is to verify the feasibility of the method's technology implementation. The experiments are implemented on the MATLAB 2019b platform.

A. RATIONALITY EXPERIMENT OF RADAR TRANSMITTER DEGRADATION FAULT PROGNOSIS

This experiment select the time-series historical data monitored by the two monitoring points before the degradation fault of the radar transmitter. Each monitoring point has 192 historical data points, as shown in Figure 5. The horizontal axis represents the time step of the sampling point, and the vertical axis represents the data value of the monitoring sensor. Because the sensor has data noise, the two time-series are first denoised by wavelet. Specific parameter settings include: the threshold function is the soft threshold function,

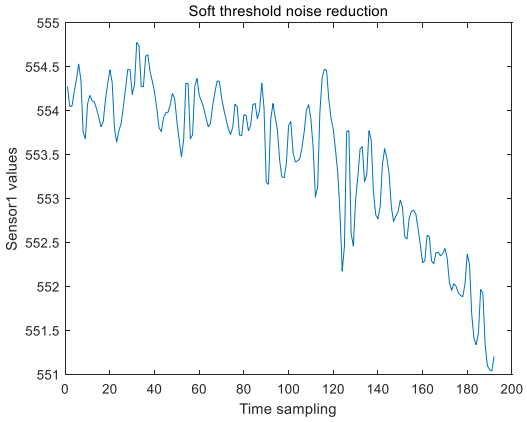


(a) Sensor 1 values.

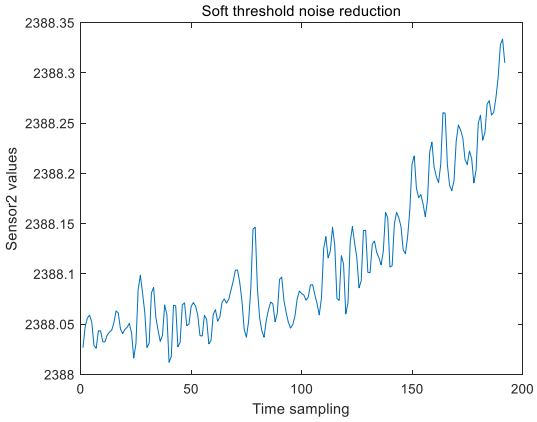


(b) Sensor 2 values.

FIGURE 5. Radar transmitter degradation fault data.



(a) Sensor 1 values after denoising.



(b) Sensor 2 values after denoising.

FIGURE 6. Radar transmitter degradation fault data after denoising.

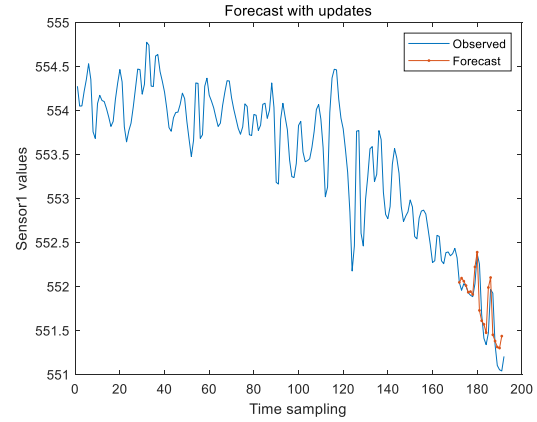
the threshold is the heuristic threshold, the wavelet decomposition level is 1, the wavelet base is Daubechies (db4). The result obtained after denoising is shown in Figure 6.

Second, the experiment use LSTM to train the first 90% samples (first 173 samples) of the denoised time-series, while the last 10% samples (last 19 samples) to predict and compare with the real observed data. Different LSTM parameter settings will get different prediction results. The key parameters are determined by the experimental data. The detailed parameter settings and prediction results, see Table 1.

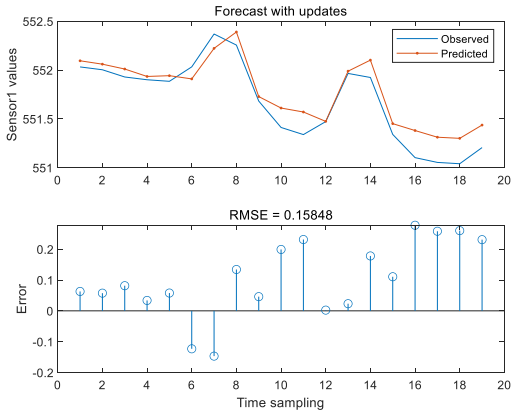
It can be seen from the table that the RMSE of a single-step LSMT with updates is much lower than the RMSE of a multi-step LSMT without updates. On this basis, it compared the impact of different iteration times on the prediction results. When the number of iterations is 300, the RMSE values of the time-series monitored by both sensors are low, the RMSE of sensor 1 is 0.1585 and the RMSE of sensor 2 is 0.0283. The number of iterations 300 is determined first. LSTM can obtain a better prediction effect when the number of iterations is 300 and the initial learning rate in the table is 0.005. In the same way, the number of network layers is determined to be 200. From this, the selection of key parameters of LSTM is completed. Subsequent experiments have selected the above key parameters. At this time, the key

TABLE 1. Effect of different parameters of LSTM on RMSE.

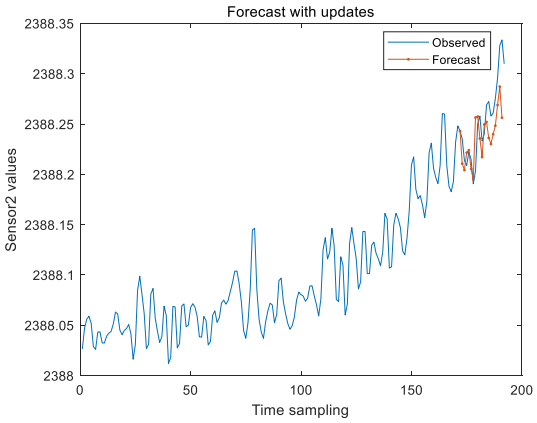
Parameter	Values	RMSE of Sensor1	RMSE of Sensor2
LSTM with updates Number of iterations	200	0.2644	0.0328
	300	0.1585	0.0283
	400	0.2502	0.0291
LSTM without updates Number of iterations	200	0.6107	0.0560
	300	0.4705	0.0518
	400	0.5485	0.0526
Initial learn rate	0.005	0.1585	0.0283
	0.01	0.2955	0.0357
	0.02	0.3731	0.0312
Network layer	100	0.2602	0.0342
	200	0.1585	0.0283
	300	0.2120	0.0298
	400	0.3534	0.0316



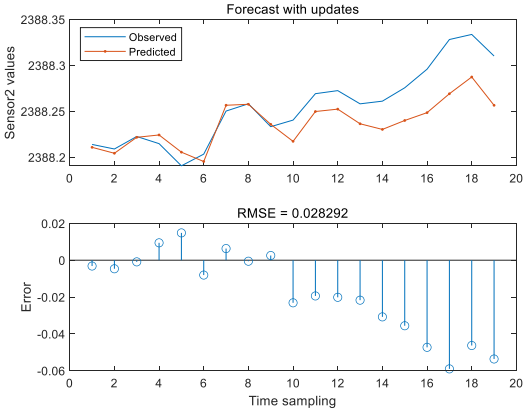
(a) Sensor 1 prognosis result.



(b) Sensor 1 predicted data partial enlargement and RMSE value.



(c) Sensor 2 prognosis result.



(d) Sensor 2 predicted data partial enlargement and RMSE value.

FIGURE 7. LSTM prognosis result.

parameters can be used to make a prognosis for the last 10% data of the two sensors. The predicted result is shown in Figure 7.

Finally, the experiment use the 90% of the data at each monitoring point to build a multivariate Gaussian distribution model and use the data predicted by LSTM for testing. The cross-validation set is needed to determine the threshold parameter ε and the F_1 value in the model. The cross-validation set is another two sets of time-series data monitored by the same monitoring point under the same degradation fault condition in another period. Although the degradation fault data samples are not included in the time-series of the cross-validation set, the degradation fault will occur at the next time step of the last time point of the data. Therefore, the data of the last few time steps in the time-series of the cross-validation set can be regarded as the failure sample. The number of fault samples in the cross-validation set is different, the prognostic results are also different. See Table 2 for details.

It can be seen from the table that when the number of failure samples is 2 and 5, the F_1 value of the model is higher, which is 0.667. When the number of fault samples is 2,

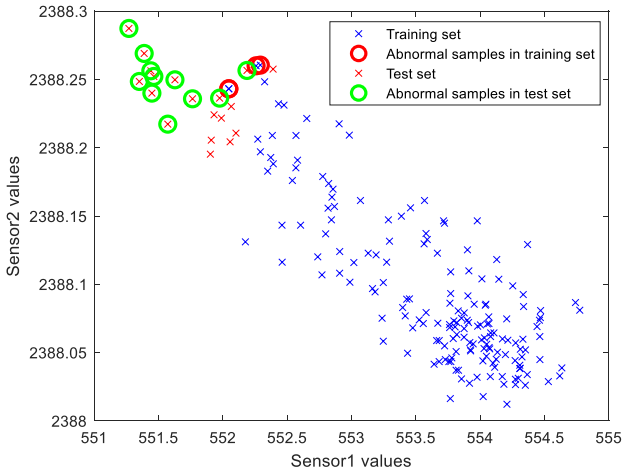
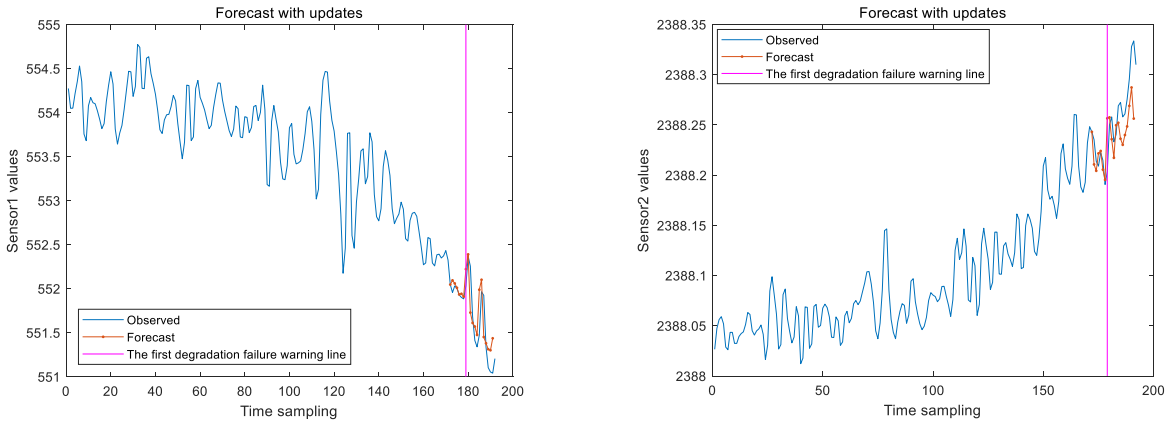


FIGURE 8. Visualization of experimental sample points.

the first degradation fault prognosis point is the 7th time step of the prediction sample (13 time steps in advance to predict the degradation fault). When the number of fault samples is 5, the first degradation fault prognosis point is the third

TABLE 2. Influence of the different fault samples number in the cross-validation set on the prognostic results.

	Fault samples number in the cross-validation set				
	1	2	3	4	5
Threshold ε	2.78×10^{-2}	1.10×10^{-2}	2.78×10^{-2}	2.78×10^{-2}	2.78×10^{-2}
F_1	0.500	0.667	0.571	0.500	0.667
Fault sample's time sampling point in the test set	8–19	7,9–13,15–19	3–5,7–19	3–5,7–19	3–5,7–19



(a) The degradation failure prognosis warning line of sensor 1.

(b) The degradation failure prognosis warning line of sensor 2.

FIGURE 9. The degradation failure prognosis warning line.

time step of the prediction sample (17 time steps in advance to predict the degradation fault). The more fault samples in the cross-validation set, the better the forward-looking of the degradation fault prediction. Of course, if there are only two fault samples, the prognosis of the degradation fault can also be met. Therefore, in the subsequent experiments, the number of fault samples in the cross-validation set is two.

Through the above analysis, the experiment use an LSTM with 300 iterations, an initial learning rate of 0.005, and a network layer of 200 to predict the last 10% data of the two monitoring points. Then, the multivariate Gaussian distribution model is used to predict the degradation faults. The visualized sample diagram of the experimental results is shown in Figure 8. It can be seen from the figure that the samples predicted by LSTM are outliers relative to most normal samples, so they are regarded as fault samples. Figure 9 shows the warning line for the degradation fault prognosis. The experimental results show that the radar transmitter degradation fault prognosis method proposed in this article is reasonable and applicable.

B. FEASIBILITY EXPERIMENT OF RADAR TRANSMITTER DEGRADATION FAULT PROGNOSIS

To visualize the experimental samples and experimental results in Experiment 1, the sensor data from only two monitoring points is selected for the experiment. But usually, there are multiple monitoring points around the radar transmitter to monitor its performance. In Experiment 2, the number of monitoring points will increase to verify the feasibility of the method proposed in this article, and verify the portability of the method under different degradation fault. The degradation fault of this experiment is the same as Experiment 1 (degradation fault 1). The data of three monitoring points, four monitoring points, five monitoring points and eight monitoring points of the radar transmitter are selected for the experiment. The experimental data is shown in Figure 10. Each monitoring point has 192 historical data points. The first 90% of the data is used for training and the last 10% is used for testing. The experimental parameter settings are the same as the key parameters determined in Experiment 1. The specific experimental results are shown in Table 3.

It can be seen from the table that with the increase of the monitoring points number, the more advanced the time step for the fault samples predicted in the test set, and also the smaller the F_1 value of the model. The main reason for the above results is that the multivariate Gaussian distribution considers the correlation between multiple monitoring points, which increases the number of predicted to be fault samples, resulting in the decrease of the precision P and the F_1 value. If the experiment want to ensure the performance of the multivariate Gaussian distribution model, it can select 3 or 4 key sensors monitoring point data for modeling. If the experiment want to obtain an early degradation failure warning, it can select more sensors monitoring point data for modeling.

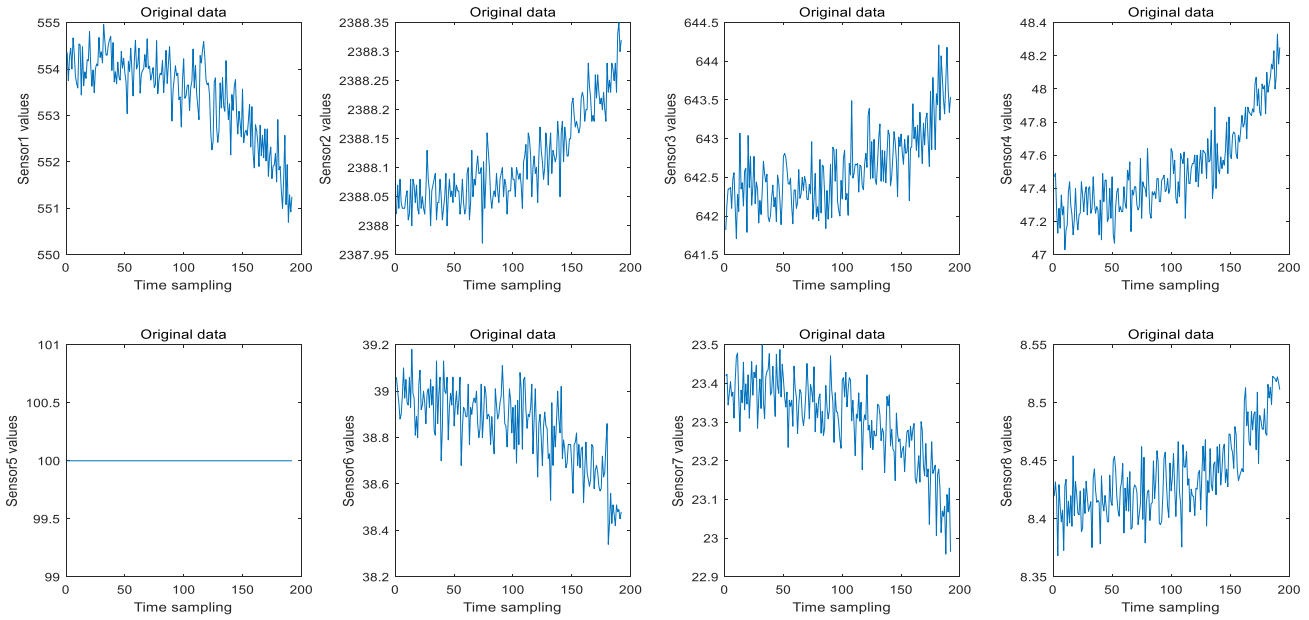


FIGURE 10. Monitoring data of 8 sensors under the condition of degradation fault 1.

TABLE 3. Experimental results of degradation fault prognosis of multiple monitoring points under the condition of degradation fault 1.

	Number of the monitoring points			
	3	4	5	8
Threshold ε	5.15×10^{-2}	9.99×10^{-3}	1.40×10^{-2}	2.73×10^{-3}
F_1	0.625	0.576	0.476	0.435
Fault sample's time sampling point in the test set	4,7–19	3,4,7–19	1–19	1–19

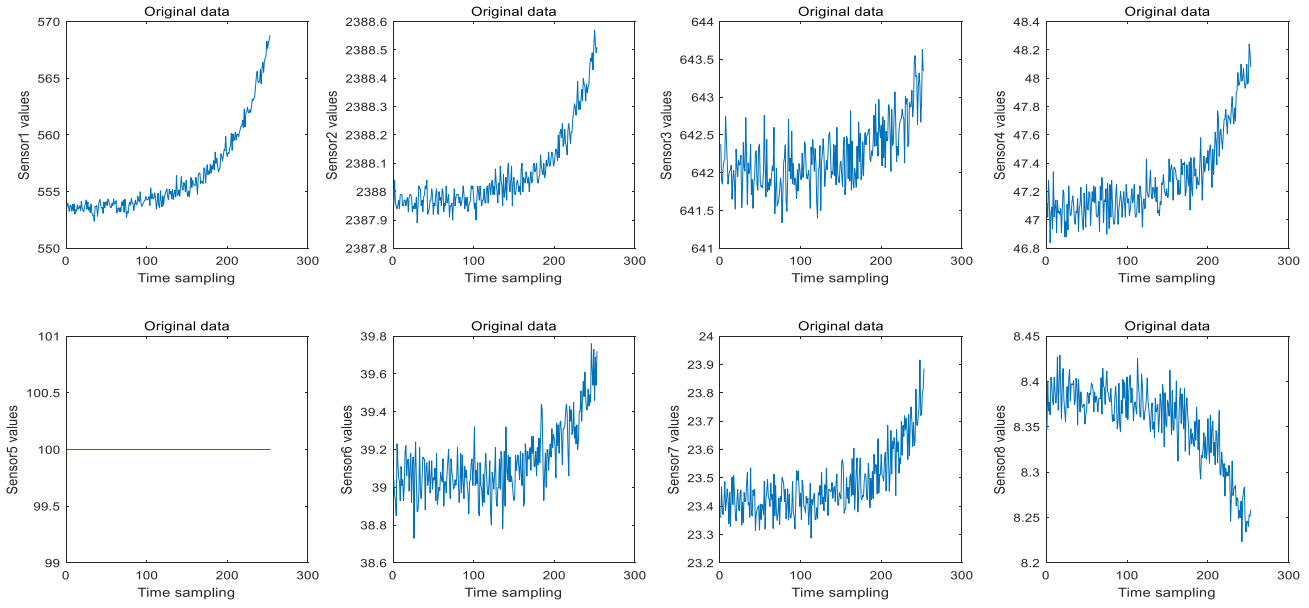


FIGURE 11. Monitoring data of 8 sensors under the condition of degradation fault 2.

To further verify the portability of the proposed method on the prognosis of different degradation fault, this task conduct experiments using degradation fault(degradation fault 2) data different from experiment 1, the experimental data is shown in Figure 11. Each monitoring point has 253 historical data points. The first 90%(first 228 samples) of the data is used

TABLE 4. Experimental results of degradation fault prognosis of multiple monitoring points under the condition of degradation fault 2.

	Number of the monitoring points			
	3	4	5	8
Threshold ε	9.09×10^{-3}	9.20×10^{-3}	3.23×10^{-3}	5.95×10^{-3}
F_1	0.659	0.552	0.442	0.430
Fault sample's time sampling point in the test set	3,5–25	1–25	1–25	1–25

TABLE 5. Comparison between the proposed method and the existing method.

Reference	Artificial threshod	The sample size	The hyper-parameter settings
[20]	Yes	Not mentioned	Input layer nodes=9;hidden layer nodes=9;output layer nodes=1
[26]	Yes	1120	Kernel function=probability density function of normal distribution
[29]	No	5500	Not mentioned
This work	No	<300	Iterations=300;learning rate=0.005;network layer=200

for training and the last 10%(last 25 samples) is used for testing. The experimental method and experimental parameter settings are the same as those in Experiment 1, and the experimental results are shown in Table 4. Compared with the experimental results in Table 3, it is possible to obtain similar results to the degradation fault 1 experiment, which can fully verify the feasibility and portability of the method proposed in this article.

Table 5 compares the proposed degradation fault prognostic method with previous method. In [20], the fault prediction model based on analytic hierarchy process and BP neural network is used to predict the fault of the model digital chip in radar analog-to-digital conversion module. The error of the prediction model is less than 5%. Non-parametric regression model is studied in reference [26] to predict remaining useful life of radar, the error of the model is 2.05%. LSTM model and dynamic threshold are used to detect the faults of meteorological radar in reference [29]. It can be seen from the table that compared with the reference [20], [26] and [29], the method proposed in this article has a smaller sample size and does not need to set the fault threshold artificially.

IV. CONCLUSION

This article proposed a radar transmitter fault prognosis method that combines multivariate long short-term memory network and multivariate Gaussian distribution. This method can be used for radar transmitter deterioration fault prediction when the total sample size and fault sample size of sensor monitoring data are small, and the monitoring data can not reach the fault threshold. This article discusses the prognostic model and workflow of radar transmitter degradation fault. We determine the key parameters of the model through experiments and set up multiple monitoring points and different degradation fault experiments to verify the feasibility and portability of the model. The prognostic method proposed in this article considers the correlation between the data of multiple monitoring points and reduces the use of training samples and fault samples when the data of each monitoring point does not reach the failure threshold. The model can effectively realize the prediction of different degradation failures of radar transmitters. The degradation fault can be predicted within 10% time step of the total time step of the monitoring data before the degradation fault occurs, while at the same time giving the warning time sampling point of the degradation fault. Experimental results show that this method is effective and feasible.

REFERENCES

[1] A. Ahmadi, T. Fransson, A. Crona, M. Klein, and P. Soderholm, "Integration of RCM and PHM for the next generation of aircraft," in *Proc. IEEE Aerosp. Conf.*, Mar. 2009, pp. 1–9.
[2] F. Paycha, "Diagnosis, therapeutics, prognosis, and computers," *IRE Trans. Med. Electron.*, vol. ME-7, no. 4, pp. 288–290, Oct. 1960.
[3] X. Jin, Y. Sun, Z. Que, Y. Wang, and T. W. S. Chow, "Anomaly detection and fault prognosis for bearings," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 9, pp. 2046–2054, Sep. 2016.
[4] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, "Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications," *Mech. Syst. Signal Process.*, vol. 42, nos. 1–2, pp. 314–334, 2014, doi: 10.1016/j.ymssp.2013.06.004.
[5] J. Long, S. Zhang, and C. Li, "Evolving deep echo state networks for intelligent fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4928–4937, Jul. 2020.
[6] J. Long, Z. Sun, C. Li, Y. Hong, Y. Bai, and S. Zhang, "A novel sparse echo autoencoder network for data-driven fault diagnosis of delta 3-D printers," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 3, pp. 683–692, Mar. 2020.
[7] A. El Mejdoubi, H. Chaoui, H. Gualous, P. Van Den Bossche, N. Omar, and J. Van Mierlo, "Lithium-ion batteries health prognosis considering aging conditions," *IEEE Trans. Power Electron.*, vol. 34, no. 7, pp. 6834–6844, Jul. 2019.
[8] W. Yan, B. Zhang, W. Dou, D. Liu, and Y. Peng, "Low-cost adaptive lebesgue sampling particle filtering approach for real-time li-ion battery diagnosis and prognosis," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 4, pp. 1601–1611, Oct. 2017.
[9] L. Shi, Y. He, B. Li, T. Cheng, Y. Huang, and Y. Sui, "Transmission tower tilt angle on-line prognosis by using solar-powered LoRa sensor node and sliding XGBoost predictor," *IEEE Access*, vol. 7, pp. 86168–86176, 2019.
[10] R. Ren, D. D. Wu, and T. Liu, "Forecasting stock market movement direction using sentiment analysis and support vector machine," *IEEE Syst. J.*, vol. 13, no. 1, pp. 760–770, Mar. 2019.

[11] J.-S. Chou and T.-K. Nguyen, "Forward forecast of stock price using sliding-window Metaheuristic-optimized machine-learning regression," *IEEE Trans. Ind. Informat.*, vol. 14, no. 7, pp. 3132–3142, Jul. 2018.
[12] W. Zhao, Y. Gao, T. Ji, X. Wan, F. Ye, and G. Bai, "Deep temporal convolutional networks for short-term traffic flow forecasting," *IEEE Access*, vol. 7, pp. 114496–114507, 2019.
[13] Y. Hou, P. Edara, and C. Sun, "Traffic flow forecasting for urban work zones," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 1761–1770, Aug. 2015.
[14] P. Cai, Y. Wang, and G. Lu, "Tunable and transferable RBF model for short-term traffic forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 11, pp. 4134–4144, Nov. 2019.
[15] N. Sinnadurai, A. A. Shukla, and M. Pecht, "A critique of the reliability analysis center failure-rate-model for plastic encapsulated microcircuits," *IEEE Trans. Rel.*, vol. 47, no. 2, pp. 110–113, Jun. 1998.
[16] R. Martens and M. Pecht, "Effects and interactions of design parameters for gold-plated electric contacts," *J. Mater. Sci. Mater. Electron.*, vol. 11, no. 3, pp. 209–218, 2000, doi: 10.1023/A:1008992731436.
[17] T. Zimnickas, J. Vanagas, K. Dambrauskas, A. Kalvaitis, and M. Ažubalis, "Application of advanced vibration monitoring systems and long short-term memory networks for brushless DC motor stator fault monitoring and classification," *Energies*, vol. 13, no. 4, p. 820, Feb. 2020. [Online]. Available: https://www.mdpi.com/1996-1073/13/4/820
[18] A. Yin, Y. Yan, Z. Zhang, C. Li, and R.-V. Sánchez, "Fault diagnosis of wind turbine gearbox based on the optimized LSTM neural network with cosine loss," *Sensors*, vol. 20, no. 8, p. 2339, Apr. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/8/2339
[19] C. Wang, N. Lu, S. Wang, Y. Cheng, and B. Jiang, "Dynamic long short-term memory neural-network-based indirect remaining-useful-life prognosis for satellite lithium-ion battery," *Appl. Sci.*, vol. 8, no. 11, p. 2078, Oct. 2018. [Online]. Available: https://www.mdpi.com/2076-3417/8/11/2078
[20] Z. Dai, L. Yao, and J. Qin, "Research on fault prediction of radar electronic components based on analytic hierarchy process and BP neural network," in *Proc. 12th Int. Conf. Intell. Comput. Technol. Autom. (ICICTA)*, Oct. 2019, pp. 91–95.
[21] Y.-L. Kong, Q. Huang, C. Wang, J. Chen, J. Chen, and D. He, "Long short-term memory neural networks for online disturbance detection in satellite image time series," *Remote Sens.*, vol. 10, no. 3, p. 452, Mar. 2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/3/452
[22] Y. An and D. Liu, "Multivariate Gaussian-based false data detection against cyber-attacks," *IEEE Access*, vol. 7, pp. 119804–119812, 2019.
[23] C. Xiao, Z. Liu, T. Zhang, and L. Zhang, "On fault prediction for wind turbine pitch system using radar chart and support vector machine approach," *Energies*, vol. 12, no. 14, p. 2693, Jul. 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/14/2693
[24] F. Pittino, M. Puggl, T. Moldaschl, and C. Hirschl, "Automatic anomaly detection on in-production manufacturing machines using statistical learning methods," *Sensors*, vol. 20, no. 8, p. 2344, Apr. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/8/2344
[25] P. Visanadham and P. Singh, *Failure Modes and Mechanisms in Electronic Packages*. New York, NY, USA: CRC Press, 1998.
[26] L. Qianqian, Z. Jingyuan, and C. Bing, "Study on life prediction of radar based on non-parametric regression model," in *Proc. 13th IEEE Int. Conf. Electron. Meas. Instrum. (ICEMI)*, Oct. 2017, pp. 586–590.
[27] X. Hou, J. Yang, B. Deng, L. Xia, Y. Zhang, and Z. Zhang, "A key lifetime parameters extraction method for T/R module based on association rule," in *Proc. Int. Conf. Sens., Diagnostics, Prognostics, Control (SDPC)*, Aug. 2018, pp. 165–170.
[28] Q. Li and Y. Lv, "Radar transmitting power supply health monitoring based on circuit modeling and simulation technology," in *Proc. IEEE Int. Conf. Power, Intell. Comput. Syst. (ICPICS)*, Jul. 2019, pp. 584–588.
[29] W. Li, W. Zhou, Y. M. Wang, C. Shen, X. Zhang, and X. Li, "Meteorological radar fault diagnosis based on deep learning," in *Proc. Int. Conf. Meteorol. Observ. (ICMO)*, Dec. 2019, pp. 1–4.
[30] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 617–643, Mar. 1992, doi: 10.1109/18.119727.
[31] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995, doi: 10.1109/18.382009.
[32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[33] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.
[34] A. Ng, *Machine Learning*. Accessed: Dec. 1, 2019. [Online]. Available: https://www.coursera.org/

YUTING ZHAI was born in Yantai, Shandong, China in 1988. She received the B.Eng. degree from Yantai University, Yantai, China, in 2011, and the M.Eng. degree from Dalian Jiaotong University, China, in 2014. She is currently pursuing the Ph.D. degree with the Department of Information Science and Technology, Dalian Maritime University. She is also a Lecturer with the Department of Information Systems, Dalian Naval Academy. Her current research interests include degradation fault prognostic method of radar, deep learning, and radar health management.

SHAOJUN FANG (Member, IEEE) received the Ph.D. degree in communication and information systems from Dalian Maritime University (DLMU), Liaoning, China, in 2001.

Since 1982, he has been with DLMU, where he is currently the Head Professor with the Department of Information Science and Technology. His recent research interests include RF monitoring, patch antennas, and computational electromagnetics. He has authored or coauthored three books and over 100 journal articles and conference papers. He was a recipient of the Best Doctor's Dissertation Award of Liaoning Province, in 2002, and the Outstanding Teacher Award of the Ministry of Transport of China.