IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 13, 2017, accepted February 3, 2018, date of publication February 13, 2018, date of current version March 16, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2805280

Prognosis of Bearing Degradation Using Gradient Variable Forgetting Factor RLS Combined With Time Series Model

YANFEI LU[1], (Member, IEEE), QING LI[2], (Member, IEEE), ZHIPENG PAN[1], AND STEVEN Y. LIANG[1,2]

[1]George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA
[2]College of Mechanical Engineering, Donghua University, Shanghai 201620, China

Corresponding author: Yanfei Lu (ylu318@gatech.edu)

ABSTRACT Rolling element bearing is a critical component in many mechanical systems in view of its critical functionality. One of the major issues industries face today is the failure of bearings, which results in catastrophic consequences. Although various prognostic approaches were proposed for the degradation of bearings, the incapability of adaptation of those models yields inaccurate predictions under different running conditions of the bearings. To address this issue, this paper proposes a prognostic algorithm using the variable forgetting factor recursive least-square (VFF-RLS) combined with an auto-regressive and moving-average (ARMA) model. The structure and parameters of ARMA model were initially determined using the vibrational data of the bearing without significant defect presented. During the bearing degradation process, the ARMA model makes predictions of the future degradation trend. Once the future acquired signal becomes available, the error between the acquired and predicted vibrational signal is calculated. The VFF-RLS algorithm uses the calculated error, correlation matrix and other parameters to update the coefficients of the ARMA model. In addition, the VFF-RLS algorithm updates the forgetting factor during each iteration to achieve faster convergence and reduced error. The updated ARMA model makes new predictions and the adaptive process continues. To demonstrate the applicability of adaptive prognosis methodology, the accuracy of the prediction of the proposed model is tested using experimental and simulated data in comparison with an auto-regressive integrated moving average (ARIMA) model without adaptation. Results show accurate predictions of the vibrational signal and degradation trend of the bearings over the ARIMA model.

INDEX TERMS Adaptive algorithms, ball bearings, fault diagnosis, prognostics and health management, time-series analysis.

I. INTRODUCTION

Various mechanical systems such as gearboxes, helicopter rotors, and spindle assemblies of CNC machines rely on the running condition of the bearings. Even though the components of the bearings are made of materials with superior mechanical strength and fatigue life, limitations of the capabilities of the manufacturing process along with the defects of material lead to the formation of microcracks on the surfaces and sub-surfaces of the elements of bearings. Over time, the cracks will propagate because of periodic loadings and subsurface plastic flow. In addition, lack of lubrication, contamination, misalignment, corrosion, and improper loads cause premature failure of bearings. Because of its highly unpredictable failure time and mode, bearing failure is a major issue in rotating machineries. Therefore, companies are devoting significant effort to estimating the severity of bearing damage and predicting the failure time and degradation trend of bearings by implementing bearing diagnostic and prognostic schematics. Traditionally, the severity of bearing damage, such as the size of the defect, is measured by placing bearings under certain running conditions on a testing rig, disassembling the bearings after a period of time and measuring with an optical device [1]. However, the measurement of defect size is time-consuming. Thus, unobtrusive signal acquisition methods that use accelerometers or acoustic emission sensors are preferred. Because the time-domain acceleration data clearly reflect the degradation trend and are easily obtainable, the signal acquired by accelerometers is used in this paper to track the degradation process.

Bearing diagnosis can be classified into three major groups: time-domain, frequency-domain, and time-frequency-domain analysis. Time-domain methods use an acquired signal from sensors placed on a component of the machine to extract critical features for the diagnosis of the bearing condition. Some desired features of the diagnosis in the time-domain analysis are shock pulse counting, root mean square (RMS), peak values, crest factors, kurtosis, short-time energy, and short-time zero-crossing rate. Early research has demonstrated that time-domain data indicate certain patterns of the types of defects [1]–[5]. Using Fourier transform, researchers find that the frequency-domain data reveal the critical frequencies of vibration signals. Some widely used methods include bi-coherence analysis, cepstrum analysis, and the high-frequency response technique (HFRT) [1], which has been widely implemented. Using the HFRT, one study found that the impact generated by the defect site in the bearing elements normally excites resonance in other components in the system [6]. The high-frequency components directly reflect the damage levels of bearing during its service life.

The fundamental characteristic frequencies of the rolling elements of bearings are often buried during the signal monitoring because of the existence of noise and resonance. Therefore, researchers implemented different signal processing techniques to denoise the undesirable signals and locate important information in the system. For instance, Shiroishi *et al*. [7] utilized an adaptive line enhancer that increases the signal-to-noise ratio to detect small defects; Zhou *et al*. [8] implemented the Wiener filter to extract the bearing fault signature. Tian implemented spectral kurtosis and cross-correlation to detect incipient faults and location of faults without reference data [9]. Tian also implemented simulated annealing to optimize the spectral kurtosis to locate optimum frequency band for diagnosis [10]. The purpose of bearing prognostics is to predict the fault of the bearing before its catastrophic failure [11]. The most critical factors for optimizing the bearing maintenance schedule and further reducing cost are the accuracy and reliability of the prognostic methodology. Early researchers developed deterministic models from the physical understanding of the defect-propagation processes based on fracture mechanics [12], [13]. However, the propagation of defect is stochastic because most manufacturing processes induce phase transformation and grain structure change [14].

One of the most common prognosis is the prediction of the remaining useful life (RUL) of bearings. For instance, Qiu *et al*. [15] compared the performance of wavelet decomposition-based de-noising and wavelet filter-based de-noising methods in bearing prognosis. Caesarendra *et al*. [16] applied relevance vector machine and auto-regressive and moving-average (ARMA) model in simulated bearing data. All these methods are relatively complex and require a significant amount of computation, which results in issues for online application. To define the bearing failure point, researchers choose a specific size of defect to quantify damage of bearings [1], [7]. Shiroishi *et al*. [7] proposed the failure size as 0.01 in^2. However, this type of defect size should not be generalized to all types of bearings because the size of the defect is proportional to the size of the bearing elements. Therefore, defining a damage severity level for a specific type of bearings is superior to using the pre-defined defect size as the criteria of failure.

This study uses an innovative combination of the ARMA model with the variable forgetting factor recursive least-square (VFF-RLS) algorithm to achieve prognosis of the vibrational signal of the bearings. In addition, the model overcomes the computational complexity and adapts to the stochastic nature of the fatigue behavior of bearings. The ARMA model is implemented instead of the ARIMA model because of the simplicity and commonality of the model in online application. The gradient based variable forgetting factor improves the convergence speed of the RLS algorithm and yields smaller error with fewer iterations of training. To prepare the signal for prognosis, a Butterworth band-pass filter and local regression smoothing filter is used along with the RMS of the signal to track the behavior of the bearing in the time-domain. The prognostic model implements the gradient based VFF-RLS algorithm to update the coefficients of the ARMA model, which predicts the degradation trend during the online monitoring process. Adaptation is needed during the online monitoring because the bearing degradation process does not exhibit a consistent increasing trend, and the process itself is highly stochastic. Even though vibrational signal in the time-domain is non-stationary, the in-process adaptation of the VFF-RLS algorithm tracks the changes of vibrational signal and reduces errors between the predicted and experimental values accordingly.

The rest of the paper is organized as follows: Section 2 describes the models used in bearing prognosis and introduces the ARMA+VFF-RLS model used in this paper. Section 3 includes testing the applicability of the ARMA+VFF-RLS model in online monitoring of the vibration of the bearings. To further validate the model's applicability in various conditions, the ARMA+VFF-RLS model was tested using simulated data in Section 4. Section 5 shows the improvement of the multi-iteration training on individual data points and compares the traditional RLS with the VFF-RLS. In addition, the ARIMA model is used as a benchmark to prove the improvement of the ARMA+VFF-RLS algorithm. Section 6 concludes the paper with comments and suggestions for future work.

II. ARMA+VFF-RLS MODEL

A. DETERMINISTIC CRACK PROPAGATION MODEL

A deterministic bearing diagnostic model is developed based on fracture mechanics and failure analysis for high-cycle fatigue. Various experimental data show that crack growth of bearing element under running conditions depends on a

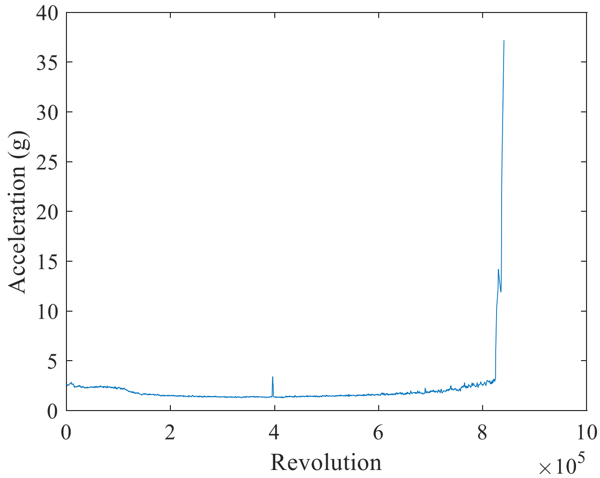


FIGURE 1. RMS value of vibrational signal.

variety of factors such as material properties, loading profile, manufacturing processes, size of bearing elements and other environmental effects. The most popular empirical model is Paris' law [17] shown as (1),

$$\frac{da}{dN} = C_0(\Delta K)^n \tag{1}$$

where $\frac{da}{dN}$ is the rate of the crack extension, instantaneous half crack length is denoted by a, the number of running cycles is represented by N, C_0 and n are constants related to the material properties and running conditions, and ΔK is the stress intensity factor range. Over the last decade, different researchers modified (1) to capture the failure mechanisms of bearings. For instance, Li *et al.* [18] applied a log-normal random variable to the original Paris' law to capture the stochastic nature of the defect propagations of bearings. This modified equation follows the general trend of the failure processes of bearings under stationary running conditions. However, other experimental data indicate that the failure of the bearing does not always follow Paris' law closely [19]. The measured vibration signals generally remain constant for a long period of time before sudden failure occurs. In addition, under certain circumstances, such as sharp edge rounding of the defects and reduction of the rotation speed, the measured vibrational signal decreases as shown in Fig.1. Because of the decreased vibration signal, the initiation of defect is extremely difficult to distinguish from the vibration signal. The deterministic models cannot resolve the two issues described. Therefore, implementing an adaptive model to capture the changes of the vibrational signal of bearings is beneficial.

B. ADAPTIVE RGNOSTIC MODEL

The prognostic model implemented reduces the complexity of computation compared with other models, such as the ARIMA model, in the online monitoring application. The proposed model is shown in Fig.2.

The model first takes in the signal from $x(0)$ to $x(t)$ in terms of output voltage of the accelerometer and passes the signal

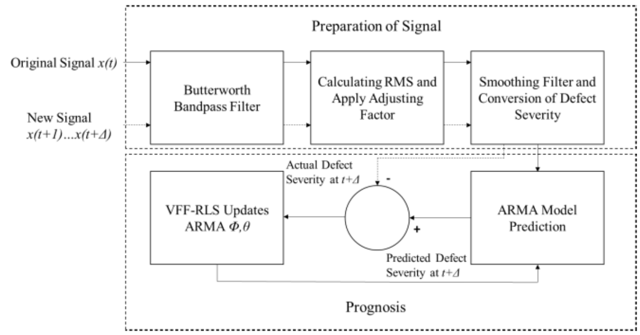


FIGURE 2. Adaptive bearing diagnostic and prognostic model.

to the bandpass filter to eliminate high-frequency noise. The future signals $x(t+1)\ldots x(t+\Delta)$ pass the same filter and are used to generate error terms for the RLS algorithm. After the signal is filtered, the RMS value of the vibration signal is calculated based on (2),

$$x_{rms} = \sqrt{\frac{1}{n}\left(x_1^2 + x_2^2 + \cdots + x_n^2\right)} \tag{2}$$

where n is the number of data points used in the calculation, and $x_1, x_2 \ldots x_n$ are the vibrational signals. The selection of n will affect the RMS of the vibrational signals. After the RMS value is calculated, a robust local regression smoothing filter is implemented to smooth out the leftover narrow-band oscillation in the RMS values. The regression smoothing filter computes regression weights for individual data points and uses linear least-squares regression to obtain the smoothed data. The magnitude of the RMS decreases after the filters are applied. Therefore, after completion of the filtering, a correction factor is applied to the filtered RMS to compensate for the decreased magnitude. The last step of the signal preparation is to convert the RMS value to either defect size or defect severity. Li [1] has identified that the value of RMS, kurtosis, and peak force of the vibration signal within a certain frequency range is related to the defect size. The relationship between the RMS and the defect size shows either a linear relationship with fitness value of $R^2 = 0.850$ or an exponential relationship with $R^2 = 0.843$. Similarly, a power law and a second order polynomial regression can be implemented between the kurtosis and the defect size using the measurement from Li [1] and Shiroishi *et al.* [7]. Because of the similarity in the experimental setup and bearing type with [1], the relationship between the RMS value and defect size in this paper can be described using (3),

$$D = a \times RMS + b \tag{3}$$

where D stands for the defect size (area) or defect severity, a is a coefficient related to a specific bearing running condition, and b is a constant with respect to the bearing type and the running condition. In the application of prognosis, an experiment which determines the values of coefficient a and constant $b$is necessary to establish the relationship between the severity of defect and RMS of the signal.

The adaptive part of the model uses the ARMA model to predict the signal at time $t + \Delta$. Once the newly acquired signal is available, the error between the experimental and predicted signal is obtained. The error is passed to the VFF-RLS algorithm to correct the coefficients of the ARMA model for future prediction. The adaptive part of the model is implemented recursively so that the coefficients of the ARMA model are updated constantly during the online monitoring. The model overcomes the restrictions and limitations of a deterministic model and represents a robust and innovative method for bearing prognosis. The following section describes the ARMA model and VFF-RLS algorithm used in the prognostic model.

The ARMA model includes two parts: auto-regressive (AR) and moving-average (MA). An ARMA(p, q) model can be described using (4),

$$x_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i x_{t-i} - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{4}$$

where c is a constant and ε_t is an independent identically distributed random variable sampled from a normal distribution with zero mean value, $\varepsilon_t \sim N(0, \sigma^2)$. The variable σ^2 is the variance, ε_t is normally treated as a white noise with a variance of σ^2, x_t stands for the vibrational data in the time domain, and φ_i and θ_i are the coefficients of the AR and MA model respectively [20]. A general way of determining the ARMA model parameters p and q is the Akaike Information Criterion (AIC). The ARMA (p, q) model with the lowest AIC value is the most efficient model to describe a process.

The RLS algorithm can be initialized by setting $\boldsymbol{P}(0) = \delta^{-1}\boldsymbol{I}$, where $\boldsymbol{P}$ is the inverse correlation matrix, δ is a small number generally used as a scaling factor, and $\boldsymbol{I}$ is an identity matrix [21]. The size of $\boldsymbol{I}$ corresponds to the number of coefficients that the RLS algorithm updates. For instance, an ARMA (2,2) model described by (4) has five coefficients: one constant, two coefficients from the AR model, and two coefficients from the MA model. In this case, $\boldsymbol{I}$ represents a five by five identity matrix. The tap-weight vector, $\hat{\boldsymbol{w}}(0)$, was set to zero because the initial value does not affect the convergence of the RLS algorithm [1]. The length of the tap-weight vector equals the number of coefficients of the ARMA model. For an ARMA (2,2) model with five coefficients, the tap-weight vector $\hat{\boldsymbol{w}}(0)$ is a five by one vector. Then, for each instant of time, $n = 1, 2, 3, \ldots$, the algorithm calculates the following entities,

$$\mathbf{k}(\boldsymbol{n}) = \frac{\lambda^{-1}\mathbf{P}(\boldsymbol{n}-1)\mathbf{u}(\boldsymbol{n})}{1 + \lambda^{-1}\mathbf{u}^H(\boldsymbol{n})\mathbf{P}(\boldsymbol{n}-1)\mathbf{u}(\boldsymbol{n})} \tag{5}$$

where $\mathbf{k}$ is the gain vector, and λ is the forgetting factor. When λ is close to one, the RLS algorithm converges to a steady state error in a slow manner while yielding a small error, and for λ close to zero, the algorithm converges to a steady state error in a relatively fast manner while yielding a large error. The value of λ normally varies between 0.8 and 1. The input vector $\mathbf{u}$ has a dimensional size equal to the number of coefficients that need to be updated. Using the ARMA (2,2) example, $\mathbf{u}$ is a one by five vector. The error ξ between the experimental and predicted data is shown in (6),

$$\xi(n) = d(n) - \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n) \tag{6}$$

where $\boldsymbol{d}$ is the experimental value, $\hat{\mathbf{w}}^H(n-1)\mathbf{u}(n)$ represents the predicted value at time n, and ξ^* is the transpose of ξ. The error term ξ is supposed to decrease to zero as the algorithm updates the coefficients through the tap-weights

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{k}(n)\xi^*(n) \tag{7}$$

$$\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{u}^H(n)\mathbf{P}(n-1) \tag{8}$$

vector of the model. One advantage of this algorithm, mentioned by Li [1], is that the initial values of the coefficients that need to be updated do not affect the error convergence. In comparison with other algorithms, such as least-mean-square (LMS), the rate of convergence of RLS is an order of magnitude higher since the RLS algorithm does not require as many iterations to update the tap-weight vectors as does the LMS. This feature greatly improves its applicability in online bearing monitoring processes [21].

The gradient based variable forgetting factor algorithm improves the RLS algorithm convergence speed by changing the forgetting factor λ in (5). As demonstrated by So *et al.*, this algorithm overcomes the deficiency of the traditional RLS algorithm in time-varying models [22]. The gradient of the forgetting factor is obtained by deriving the dynamic equation of the mean square error. The algorithm adjusts the forgetting factor to minimize the mean square error and updates the forgetting factor recursively as shown in (9) [22],

$$\lambda(n+1) = [\lambda(n) - \mu\nabla\lambda(n)]_{\lambda-}^{\lambda+} \tag{9}$$

where $\nabla\lambda(n)$ is the gradient of the forgetting factor with respect to the mean square error, and μ is the step size in the gradient-descent algorithm. The recursive equation has an upper limit $\lambda+$ and a lower limit $\lambda-$ to avoid the divergence of this algorithm. The gradient of the mean square error σ_e^2 with respect to the forgetting factor is shown in (10) [22],

$$\frac{\partial\sigma_e^2(n+1)}{\partial\lambda} = \lambda_0\frac{\partial\sigma_e^2(n)}{\partial\lambda} + \frac{\partial C_1}{\partial\lambda}\sigma_e^2(n) + \frac{\partial C_2}{\partial\lambda}\sigma_\eta^2 \tag{10}$$

where λ_0 is the initial value of the forgetting factor, and σ_η^2 is the variance of the error between the experimental value and predicted value. With assumptions and simplifications, $\frac{\partial C_1}{\partial\lambda}$ and $\frac{\partial C_2}{\partial\lambda}$ can be simplified as (11) and (12) [22],

$$\frac{\partial C_1}{\partial\lambda} = 2 + \frac{2(N+2)(1-\lambda)[N\lambda^2 - (N+3)\lambda - 1]}{[N(1-\lambda)+2]^2} \tag{11}$$

$$\frac{\partial C_2}{\partial\lambda} = -2 - \frac{4(1-\lambda)[N\lambda^2 - (N+3)\lambda - 1]}{[N(1-\lambda)+2]^2} \tag{12}$$

where N is the length of the RLS filter. The choice of N depends on the signal to noise ratio. No specific guidance exists to select the optimized N number. For this ARMA(2,2) model, the filter length is set to five by trial and error. The

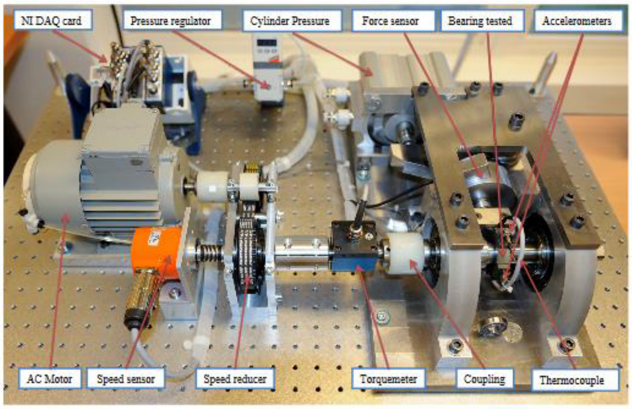


FIGURE 3. Experimental setup for ibration analysis of bearings [19].

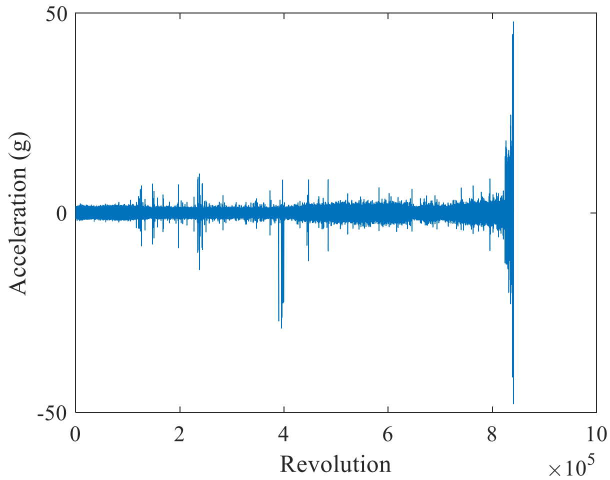


FIGURE 4. Unprocessed vibrational signal.

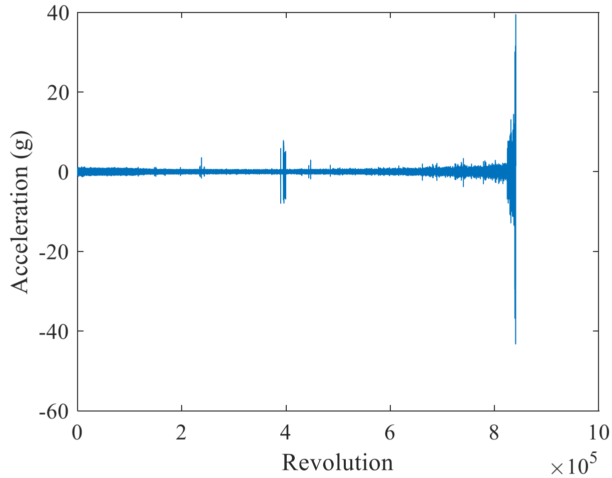


FIGURE 5. Vibrational signal bandpassed from 1 kHz to 5 kHz.

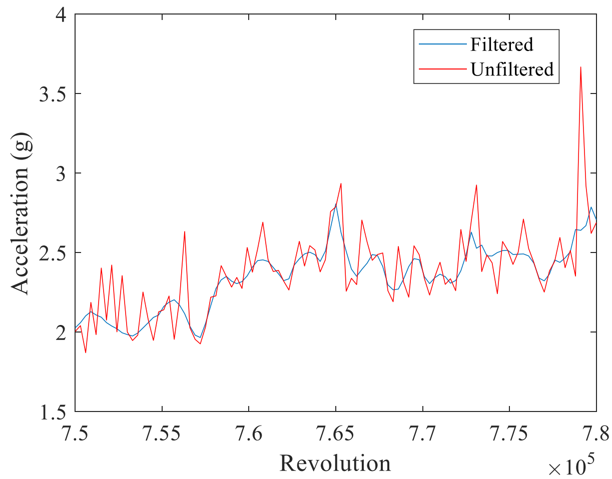


FIGURE 6. Comparison of RMS value with and without filtering.

VFF-RLS algorithm significantly reduces the number of iterations to minimize the mean square error as described in [22]. In the next section, the diagnostic and prognostic model is tested using an experimental bearing vibration data.

III. EXPERIMENTAL DATA

To investigate the feasibility of the proposed approach for online monitoring of bearings, we used the bearing data measured by Nectoux *et al.* [19].

The experimental setup is shown in Fig.3. The vibration signal was collected by the DYTRAN 3035B accelerometer with a 25.6 kHz sampling frequency to avoid aliasing. The expected resonance frequency of the bearing is within 5 kHz as mentioned in [1]. Therefore, the 25.6 kHz sampling frequency is adequately larger than the Nyquist frequency. The bearing was running at 1800 rpm with a 4000 N load, which is the dynamic load rating of the bearing. The failure criteria is defined as when the measured acceleration exceeds 20 g. Fig.4 shows the raw signal in the time domain. The data show that the raw experimental signal is comparatively noisy.

Because only using the VFF-RLS algorithm for noisy signals generates undesirable oscillations in predictions, bandpass and smoothing filters are implemented to exclude the noise within the data and facilitate the convergence speed of the VFF-RLS algorithm. The filtered signal using a fifth order Butterworth bandpass filter (1 kHz to 5 kHz) is shown in Fig.5. We picked the frequency range of the filter based on previous experiment conducted by Li [1], which indicates that the high frequency components reflect the defect signal clearer than low frequency components in the time domain.

Since the magnitude of the signal is decreased by using the Butterworth filter, the signal requires rectification before it is used to calculate the RMS. Each value of the RMS is obtained using 2560 sample points. The resulting RMS after applying a correction factor is shown in Fig.6.

The smoothing filter eliminates most of the narrow band noises, which reduces fluctuation in prediction. After the smoothing filter has been successfully implemented, the conversion equation between the RMS value and the defect severity is applied based on the user's requirement. The conversion between the RMS value and defect severity uses a similar structure as shown in (3). A defect severity level from one

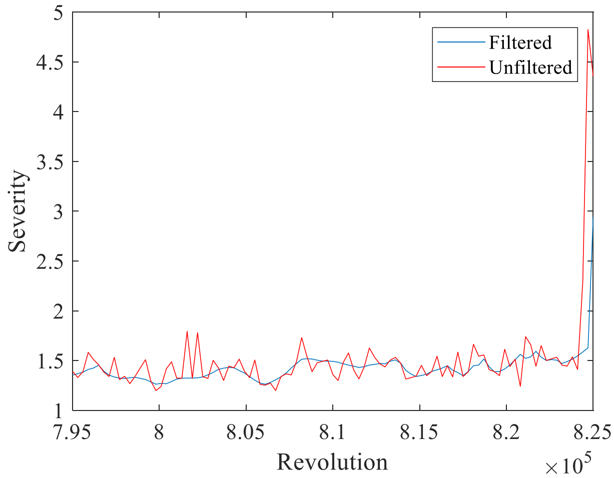


FIGURE 7. Smoothed experimental defect severity.

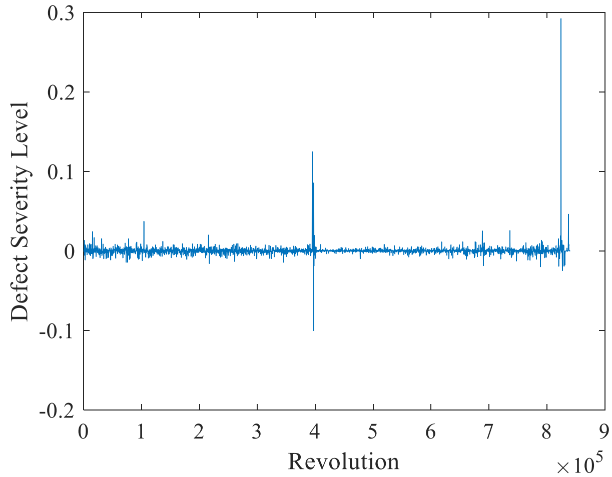


FIGURE 9. ARMA+VFF-RLS error with $\delta = 0.01$ and $\lambda = 0.9$.

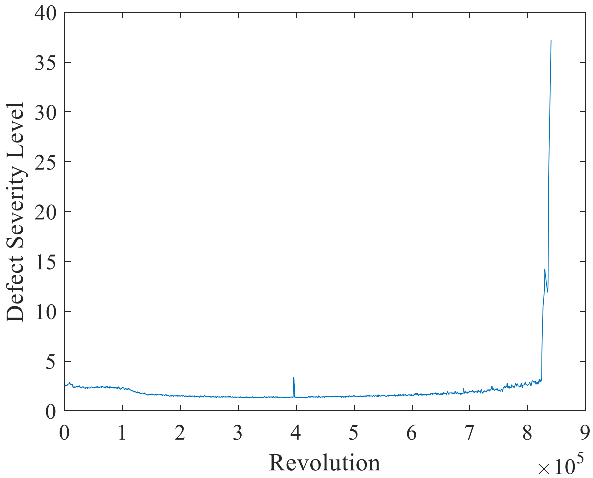


FIGURE 8. ARMA+VFF-RLS prediction with $\delta = 0.01$ and $\lambda = 0.9$.

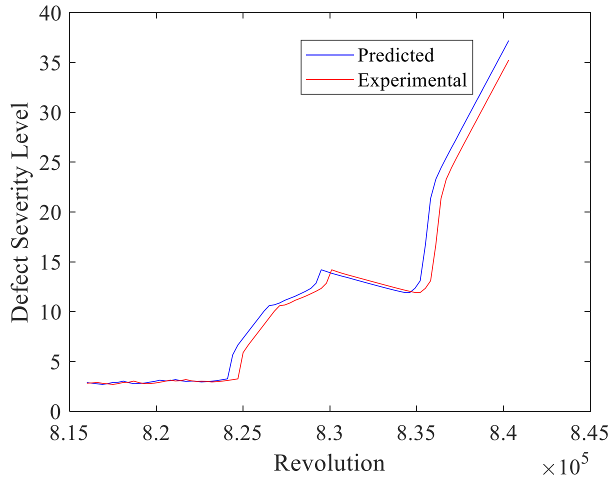


FIGURE 10. Zoomed interval of prediction.

to ten is assigned based on the failure requirement in the experiment performed by Nectoux *et al.* [19], with a severity of ten indicating a bearing failure. The experimental defect severity is shown in Fig.7.

To find the most efficient model to present the signal, the AIC is implemented to find the ARMA model structure. The AIC test result suggests an ARMA (2,2) model structure for this specific experiment. The VFF-RLS correlation matrix parameter δ is selected to be 0.01 and the forgetting factor λ is initially set to 0.9. The influence of the convergence of error by λ and δ is demonstrated by Zhang and Zhang [23], which mentioned that a large forgetting factor reduces the steady state error while sacrificing the speed of convergence. The benefit of the VFF-RLS is that it overcomes the deficiency of the fixed value forgetting factor. Fig.8 and 9 show the predicted trend and modeling error during the prognostic process with $\delta = 0.01$. The spikes from Fig.9 of the prediction error indicate a sudden change in the experimental signal, which possibly indicates damage.

By comparing Fig.5 and 9, we observed that the sudden increase of error around four hundred thousand cycles in Fig.9 is related to an unusual increase of the signal in Fig.5. Because the duration of the increased error is short, we attribute this change into possible operational error while conducting the experiment. Fig.10 and 11 are an enlarged portion of Fig.8 and 9 from eight hundred twenty thousand cycles to eight hundred forty thousand cycles. Although large fluctuations are observed during the updates of the coefficients of the ARMA model, the relatively fast convergence feature of the VFF-RLS algorithm efficiently reduces the error to zero as shown in Fig.11 around eight hundred twenty-five thousand cycles.

IV. SIMULATED DATA

The model was also tested on simulated bearing data created using an exponential function with sinusoidal functions of various frequencies. The simulated bearing signal can be seen in Fig.12.

Using the same diagnostic model as shown in Fig.2, the smoothed RMS is shown in Fig.13. The discontinuity of the RMS value in Fig.13 is because of the window size of

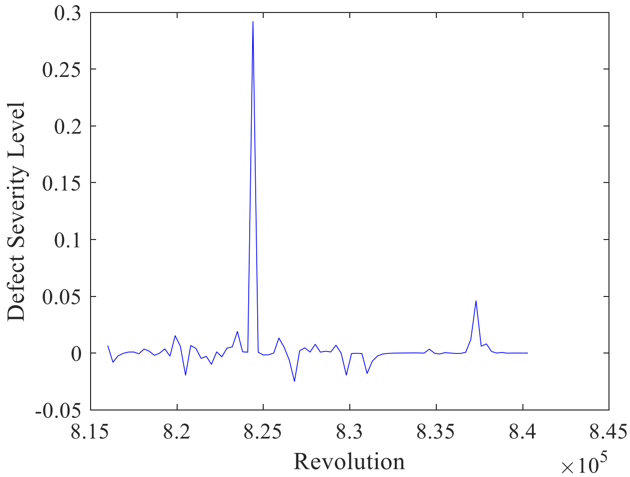


FIGURE 11. Zoomed interval of prediction error.

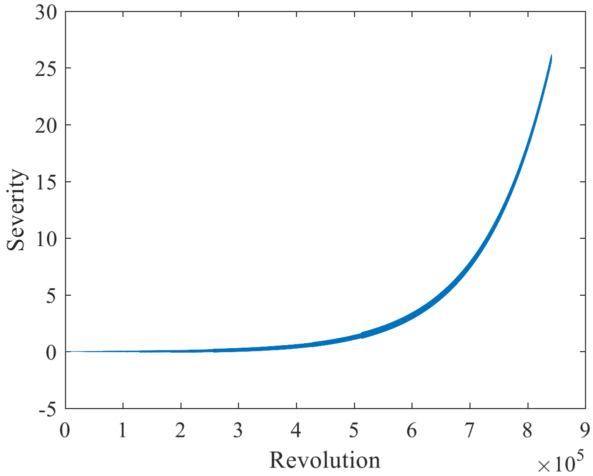


FIGURE 12. Simulated data.

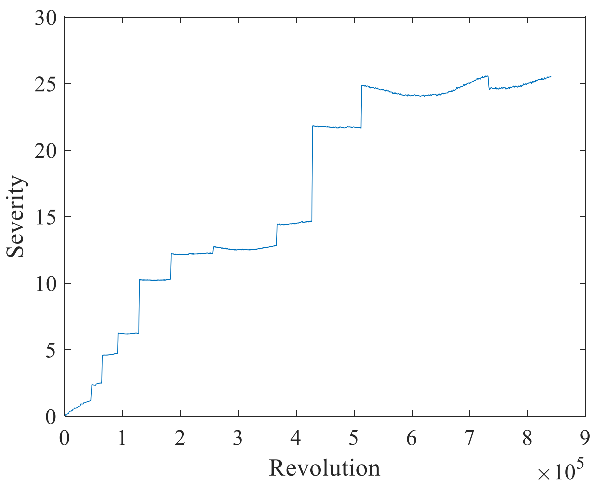


FIGURE 13. RMS value for simulated data.

calculating the RMS. A smaller window could be used to reduce the discontinuity shown in Fig.13.

In addition, the prediction error of the ARMA+VFF-RLS is shown in Fig.14. Both the simulated and experimental data

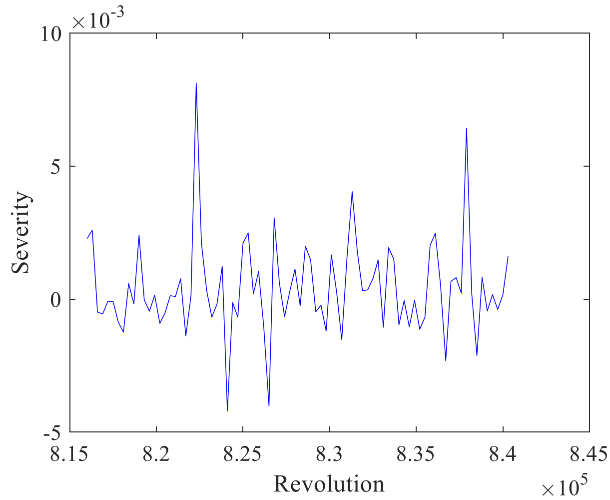


FIGURE 14. Prediction error of simulated data.

TABLE 1. Error comparison of different iteration number.

	Mean Square Error	Mean Absolute Percentage Error	Maximum Error	Percentage of Change
Iteration n=1	0.0164	1.26%	3.18	N/A
Iteration n=5	0.0063	0.96%	2.64	17.0%
Iteration n=15	0.0020	0.64%	1.35	48.9%
Iteration n=50	0.0001	0.17%	0.43	68.1%
Iteration n=100	5.47×10^{-6}	0.02%	0.09	79.1%

demonstrate the successful prediction of the bearing failure using the ARMA+VFF-RLS model in different patterns of degradation.

V. MODEL COMPARISON

An adaptive ARMA+RLS model is used on both experimental and simulated data as a comparison. In this section, we compared the multi-iteration training of the RLS algorithm with the proposed VFF-RLS algorithm. For the multi-iteration RLS, each individual data point was used multiple times in the RLS algorithm to update the coefficients of the ARMA model. The results show that the error of prediction can be reduced significantly through multi-iteration data training by 79.1%. The result of the different iterations of the training is shown in Table 1. It can be seen that the increment in the number of iterations significantly reduces the error in terms of mean square error (MSE), mean absolute percentage error (MAPE), and maximum error. The error converges to a satisfactory value defined by the user after fifteen iterations. However, the added computation time may result in hardware challenge and extra memory required in practical applications.

The gradient based VFF-RLS algorithm overcomes the disadvantage of the multi-iteration RLS algorithm by

TABLE 2. Error comparison of VFF-RLS and RLS.

	VFF-RLS	RLS	VFF-RLS (iteration=5)	RLS (iteration=5)
MSE	0.0850	0.1478	5.47×10^{-5}	1.85×10^{-4}
MAPE	1.4775	1.5509	0.1228	0.1927
Max Error	10.0383	10.4087	0.2920	0.5034

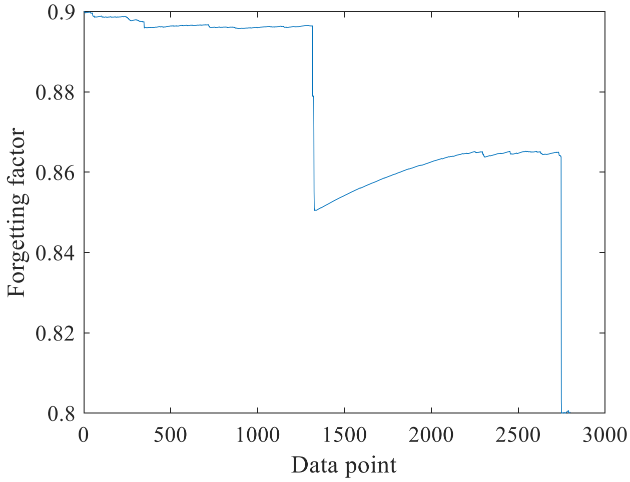


FIGURE 15. Forgetting factor change during adaptation.

TABLE 3. Comparison of RLS and VFF-RLS.

	VFF-RLS	RLS
Iteration	10	100
Max Error	0.078	0.09

optimizing the RLS convergence speed with the changes of the forgetting factor. The step size μ is selected to be 0.4 by trial and error, and the upper limit and lower limit of the forgetting factor is set as 0.995 and 0.8. The comparison result can be seen in Table 2. The VFF-RLS, in comparison with RLS, has marginal improvement during single data iteration in a time-varying environment, as shown in Table 2. However, with multiple iterations, the VFF-RLS yields smaller error than the traditional RLS algorithm. The error is reduced more than 50% in terms of the MSE and 42% in terms of the maximum error.

The change of forgetting factor can be seen from Fig.15. The result of Fig.15 indicates that a large or constant forgetting factor does not guarantee the maximum speed of convergence and minimization of error. Instead, the forgetting factor varies all the time during the VFF-RLS algorithm. Table 3 demonstrates the improvement by using the VFF-RLS over the RLS. The VFF-RLS improves the convergence speed by ten times and yields a smaller maximum error for the bearing prognostic model.

The ARMA+VFF-RLS model is also compared with a widely used time-series forecasting model, the ARIMA

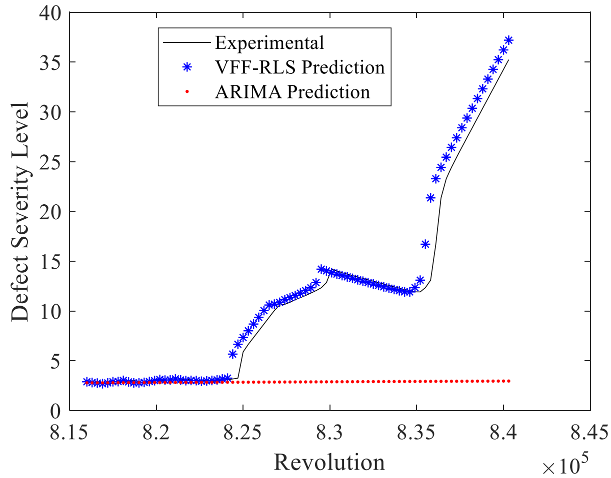


FIGURE 16. Comparison of prediction of different models.

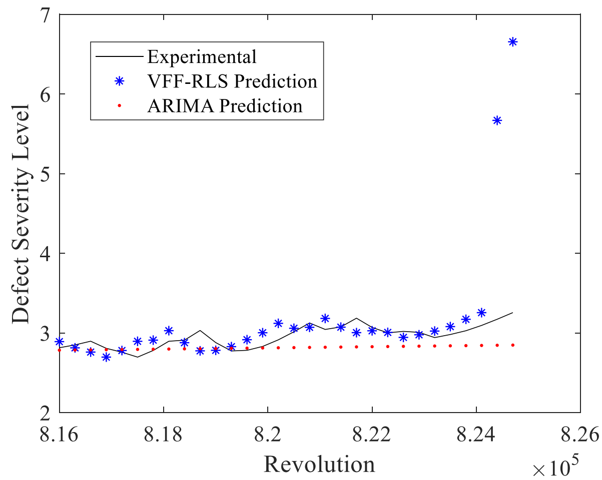


FIGURE 17. Comparison of prediction of different models for stationary signal.

model. The structure of the ARIMA model is selected to be (2,1,2) based on the AIC. The value of the prediction is shown in Fig.16 and 17 along with the prediction from the ARMA+VFF-RLS model.

It can be observed that the ARIMA model is unable to adapt to the non-stationary data. In addition, the error of the ARMA+VFF-RLS method is smaller in the stationary part of the signal. In conclusion, the ARMA+VFF-RLS model is superior than the time-series ARIMA model in predicting both stationary and non-stationary trends.

VI. CONCLUSION

This paper presents an adaptive prognostic model that is implemented in online monitoring processes of bearings to predict the degradation trend of bearings. The model first establishes a linear relationship between the vibrational signal and defect-severity level of the bearing. The prognostic model uses an ARMA (2,2) model to predict the defect-severity level of the bearing with the VFF-RLS adaptation.

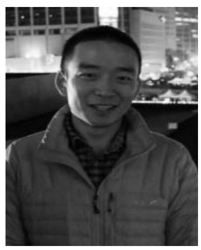
The adaptive model overcomes the inflexibility of the deterministic model by constantly varying the coefficients of the ARMA model. Previous documented models only account for the increase of the magnitude in the vibration signals of bearings. In comparison, this model can accommodate various types of signals and patterns of degradation during the online monitoring process. The VFF-RLS further improves the RLS by increasing the model convergence speed while reducing the error of prediction. The model also yields a better prediction over a widely used ARIMA model which is used as a benchmark model for comparison.

Both experimental and simulated data were tested to verify the applicability and robustness of the model. The adaptive model effectively predicts the defect propagation process by optimizing the coefficients of the ARMA model. The error between the experimental value and predicted value is reduced significantly by the adaptation. The VFF-RLS further optimizes the forgetting factor to minimize the MSE and increase rate of convergence. The iteration number of training data is closely related to the accuracy of the prognosis. A large iteration number generates small error; however, it requires more computation power and memory. The proposed VFF-RLS overcomes the disadvantage of the slow convergence, which resulted from a large forgetting factor.

Since the AIC was used initially in the early stage of bearing vibration signal, the ARMA model structure could change over time during the process of degradation. Further research can be performed to add structural adaptation mechanisms to modify the time-series model. In addition, other signal features such as energy ratio could be combined with the proposed model to predict the trend of degradation.

REFERENCES

[1] Y. Li, "Dynamic prognostics of rolling element bearing condition," Ph.D. dissertation, School Mech. Eng., Georgia Inst. Technol., Atlanta, GA, USA, 1999.
[2] O. G. Gustafsson and T. Tallian, "Detection of damage in assembled rolling element bearings," *ASLE Trans.*, vol. 5, no. 1, pp. 197–209, 1962.
[3] D. Dyer and R. M. Stewart, "Detection of rolling element bearing damage by statistical vibration analysis," *ASME J. Mech. Des*, vol. 100, no. 2, pp. 229–235, 1978.
[4] R. J. Alfredson and J. Mathew, "Time domain methods for monitoring the condition of rolling element bearings," *Inst. Eng. (Australia), Trans. Mech. Eng.*, vol. 10, no. 2, pp. 102–107, 1985.
[5] H. R. Martin and F. Honarvar, "Application of statistical moments to bearing failure detection," *Appl. Acoust.*, vol. 44, no. 1, pp. 67–77, 1995.
[6] P. D. McFadden and J. D. Smith, "Vibration monitoring of rolling element bearings by the high-frequency resonance technique—A review," *Tribol. Int.*, vol. 17, no. 1, pp. 3–10, 1984.
[7] J. Shiroishi, Y. Li, S. Liang, T. Kurfess, and S. Danyluk, "Bearing condition diagnostics via vibration and acoustic emission measurements," *Mech. Syst. Signal Process.*, vol. 11, no. 5, pp. 693–705, Sep. 1997.
[8] W. Zhou, B. Lu, T. G. Habetler, and R. G. Harley, "Incipient bearing fault detection via motor stator current noise cancellation using Wiener filter," *IEEE Trans. Ind. Appl.*, vol. 45, no. 4, pp. 1309–1317, Jul. 2009.
[9] J. Tian, C. Morillo, M. H. Azarian, and M. Pecht, "Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis," *IEEE Trans. Ind. Electron.*, vol. 63, no. 3, pp. 1793–1803, Mar. 2016.
[10] J. Tian, C. Morillo, and M. G. Pecht, "Rolling element bearing fault diagnosis using simulated annealing optimized spectral kurtosis," in *Proc. IEEE Conf. Prognostics Health Manage. (PHM)*, Jun. 2013, pp. 1–5.
[11] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, 2006.
[12] Y. Murakami, M. Kaneta, and H. Yatsuzuka, "Analysis of surface crack propagation in lubricated rolling contact," *ASLE Trans.*, vol. 28, no. 1, pp. 60–68, 1985.
[13] L. M. Keer and M. Bryant, "A pitting model for rolling contact fatigue," *J. Lubrication Technol.*, vol. 105, no. 2, pp. 198–205, 1983.
[14] Z. Pan, S. Y. Liang, H. Garmestani, and D. S. Shih, "Prediction of machining-induced phase transformation and grain growth of Ti-6Al-4 V alloy," *Int. J. Adv. Manuf. Technol.*, vol. 87, nos. 1–4, pp. 859–866, 2016.
[15] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *J. Sound Vibrat.*, vol. 289, nos. 4–5, pp. 1066–1090, 2006.
[16] W. Caesarendra, A. Widodo, H. T. Pham, and B.-S. Yang, "Machine degradation prognostic based on RVM and ARMA/GARCH model for bearing fault simulated data," in *Proc. IEEE Prognostics Health Manage. Conf. (PHM)*, Jan. 2010, pp. 1–6.
[17] P. C. Paris, M. P. Gomez, and W. E. Anderson, "A rational analytic theory of fatigue," *Trend Eng.*, vol. 13, no. 1, pp. 9–14, 1961.
[18] Y. Li, T. R. Kurfess, and S. Y. Liang, "Stochastic prognostics for rolling element bearings," *Mech. Syst. Signal Process.*, vol. 14, no. 5, pp. 747–762, 2000.
[19] P. Nectoux *et al.*, "PRONOSTIA: An experimental platform for bearings accelerated degradation tests," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (PHM)*, Jun. 2012, pp. 1–8.
[20] P. Whittle, *Hypothesis Testing in Time Series Analysis*. Uppsala, Sweden: Almqvist & Wiksells, 1951.
[21] S. S. Haykin, "Recursive least square algorithm," in *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.
[22] C. F. So, S. C. Ng, and S. H. Leung, "Gradient based variable forgetting factor RLS algorithm," *Signal Process.*, vol. 83, no. 6, pp. 1163–1175, 2003.
[23] S. Zhang and J. Zhang, "An RLS algorithm with evolving forgetting factor," in *Proc. IEEE 7th Int. Workshop Signal Design Appl. Commun. (IWSDA)*, Sep. 2015, pp. 24–27.



YANFEI LU (M'17) received the B.S. degree in mechanical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2015, where he is currently pursuing the Ph.D. degree in mechanical engineering.

He was with MedShape Inc., Atlanta, as a Manufacturing and Design Engineer. His current research is in intelligent prognosis and process optimization.

QING LI (M'17) received the master's degree from the Shanghai University of Science and Engineering in 2015. He is currently pursuing the Ph.D. degree with the College of Mechanical Engineering, Donghua University, Shanghai, China. He is also a Visiting Researcher with the Georgia Institute of Technology, Atlanta, GA, USA. His technical interests lie in dynamic signal/image processing, fault diagnosis for rotating machinery, and machine learning. He has authored over 20 archival journal papers and professional conference articles in these areas. He is also an invited peer reviewer for many prestigious international journals, such as *Material & Design* and *Frontiers of Information Technology & Electronic Engineering*. He is a member of the American Society of Mechanical Engineers and the Vibration Engineering Society of China.

ZHIPENG PAN received the B.S. degree in mechanical engineering from HUST, Wuhan, China, in 2013. He is currently pursuing the Ph.D. degree in mechanical engineering with the Georgia Institute of Technology, Atlanta, GA, USA.

Since 2013, he has been a Research Assistant with the Precision Machining Research Center, Georgia Tech. He is currently with an Internship Program at the Schlumberger Research Center. He has authored 19 papers during his study at Georgia Tech.

STEVEN Y. LIANG received the Ph.D. degree in mechanical engineering from the University of California at Berkeley in 1987. He is currently the Morris M. Bryan, Jr. Professor in mechanical engineering for Advanced Manufacturing Systems (permanently appointed) at the Georgia Institute of Technology. He was the Institute's founding Director of Precision Machining Research Consortium, Director of Manufacturing Education Program, and an Associate Director of the Manufacturing Research Center from 1996 to 2008. His technical interests include advanced engineering technologies, precision manufacturing, and technology innovation. He is fellow of ASME and the Society of Manufacturing Engineers. He served as the President of the North American Manufacturing Research Institution and the Chair of the the Manufacturing Engineering Division, The American Society of Mechanical Engineers (ASME). He was a recipient of many prestigious awards, including the SME Robert B. Douglas Outstanding Young Manufacturing Engineer Award, the Society of Automotive Engineers Ralph R. Teetor Educational Award, and the SME Blackall Machine Tool and Gage Award. He has been elected the National Thousand-Elite Expert of China and Shanghai Thousand-Elite Expert.