Parameter adaptive LS-SVR based on multi-stages division and cluster sampling for remaining useful life prediction

Jian-wen Yan1,2, Xiao-hu Zhong1,2, San-min Guo1,2, Yu Fan2* and Qiang Zhang1

1School of Management, Hefei University of Technology, Hefei 230009, China
2Anhui Province Key Lab of Aerospace Structural Parts Forming Technology and Equipment, Institute of Industry & Equipment Technology, Hefei University of Technology, Hefei 230009, China

*Corresponding author’s e-mail: yufan@hfut.edu.cn

Abstract. Remaining useful life (RUL) prediction is an advanced technology to manage life cycles of equipment and to reduce maintenance cost. Support vector regressive (SVR) is one of the frequently-used data driven method in the field of prognostics, and is much suitable for analyzing small samples and multi-dimensional data. However, there is a limitation of traditional SVR model that the linearity will increase as the training data increase. In order to overcome it, this paper proposes the parameter adaptive least square support vector regressive (LS-SVR) based on multi-stages division and cluster sampling. It divides the degradation process of the equipment into several stages according to the value of degradation data and selects the model parameters stage by stage. The optimal model parameters of each stage are calculated through the novel testing data, which are obtained through appending the cluster sampling data to the original testing data. Therefore, the optimal model parameters can be selected adaptively online as the degradation evolves. Finally, the proposed method is verified systematically by a simulated dataset of fatigue crack growth and a real-world degradation dataset of GaAs-based lasers. It is shown that the proposed method is of better effectiveness, reliability and robustness.

1. Introduction

Equipment degradation and unexpected failures affect the safety, cost and quality in production. With the increasing automation and complexity, performance and health of the modern equipment will inevitably follow degradation before unplanned shutdown or even catastrophic failures. When the degradation reaches a certain extent, the equipment will not be able to complete normal tasks and functions. Prognostics and health management (PHM) as a key technology to ensure equipment safety and reliability, can effectively monitor equipment operation state and predict the remaining useful life (RUL), i.e., the time remaining before it can no longer perform its function once a potential fault occurs [1,2].

Generally, prognostics can be implemented using either physics-based or data-driven methods [3,4]. Physics-based methods need to have a broad understanding of physical mechanism, however there usually involves many physical parameters, and it is difficult to obtain an accurate physical model [5,6]. In addition, the identification of model parameters requires a large amount of experimental and failure history data, which is much too expensive and time consuming. In contrast, data-driven methods don’t require to work out a physics model to represent the physical failure mechanism, but do mine the hidden degradation information from the observation data or sensor measured data via some
data analysis tools and algorithms [3,7-9]. There are some classical data-driven methods for RUL prediction, such as auto-regressive (AR) models [10,11], fuzzy logic methods [12], hidden Markov models [13], Bayesian networks [8], artificial neural networks (ANN) [14-16], support vector regressive (SVR) methods [17,18], similarity-based RUL prediction (SbRP) methods [19-21] and stochastic process models [22-24], etc. SVR is the common application form of support vector machine (SVM) in the field of prognostics, and is much suitable for analyzing small samples and multi-dimensional data. The core idea is to use condition monitoring data to train a SVR model, and identify the model parameters (non-sensitive coefficients, penalty factors and kernel functions, etc.), accordingly the future degradation trend of equipment can be predicted based on the trained SVR model, therefore the RUL is obtained by comparing the predicted degradation level with the preset failure threshold. In comparison to other data-driven methods, SVR is of better preferable generalization performance in engineering practice due to that the real-world failure historical data are always limitation. For example, Nuhic et al. (2013) [25] employed SVR method to predict the RUL of lithium-ion batteries, the validation shows that it can reliably learn the degradation behavior of Li-ion cells. Loutas et al. (2013) [17] proposed a methodology for rolling bearing online RUL predictions based on probabilistic SVR, the results show adequate agreement with the corresponding actual RUL. Benkedjouh et al. (2015) [26] estimated and predicted successfully the amount of wear and calculate the RUL of the cutting tool by SVR method. Furthermore, Researchers usually combine the SVR model with other methods to predict the RUL of the equipment, in order to achieve better prediction results. Chen et al. (2018) [27] used SVR method to predict the RUL of the reference samples with right censored data, these processed samples are satisfactorily used for the similarity-based method to predict the RUL of the operating equipment, their combination has been successfully used for the RUL prediction of aircraft engines. Ordonez et al. (2019) [28] proposed a hybrid model incorporated of auto-regressive integrated moving average (ARIMA) model and SVR model to predict the RUL of aircraft engines. Wei et al. (2018) [29] employed a particle filter algorithm to suppress the measurement noises of current and voltage, the RUL of lithium-ion batteries was predicted based on the SVR model. In addition, the least square support vector regressive (LS-SVR) as the most widely used SVR model performs well in the field of RUL prediction. It transforms the problem of solving quadratic programming into the counterpart of solving linear least squares in high-dimensional space, and then simplifies the calculation process and improves the training speed.

From the above, the effectiveness, reliability and robustness of SVR methods are well verified in engineering practice. Even so, there is an inevitable limitation that with the increase of sample set, linearity will increase which will lead to the increase of overfitting [30]. From a practical point of view, however, the degradation processes usually show two or even more distinct stages of different degradation rates during the whole life period [31], thus a single SVR model often place too much emphasis on the generalization capability which then lead to that the prediction accuracy isn’t satisfactory in some cases. To overcome this limitation, this paper proposes a parameter adaptive LS-SVR based on multi-stages division and cluster sampling for RUL prediction. The novelty is that it divides the degradation process into several stages, and through cluster sampling of the test data of each stage, as well as model construction stage by stage. Therefore, the optimal model parameters can be selected adaptively according to the degradation state, so that each stage of the degradation process can be fitted at the best extent, and the trained SVR model linearization is avoided.

The rest of this paper is organized as below. Section 2 introduces the SVR model and proposes the parameter adaptive LS-SVR based on multi-stages cluster sampling. Section 3 analyzes systematically the proposed method by a simulated dataset of fatigue crack growth and a real-world degradation dataset of GaAs-based semiconductor lasers. Finally, Section 4 concludes the whole paper.

2. Methodology

2.1. Introduction to SVR model
SVR is mainly used for prediction, which aims to find an optimal classification surface in order to minimize the error of the missed training samples set from the optimal classification surface. It’s the most common application of support vector machine (SVM). Given a training set $\{x_i, y_i\}_{i=1,2,...,n}$ with input data $x_i \in R^N$ and output data $y_i \in R$. Finding a function $f(x)$ on $R^N$ to infer the output $y$ corresponding to any $x$, is a regression problem.

Assume that the linear regression function for $f(x)$ constructed in the high-dimensional feature space is shown in (1):

$$f(x) = w \ast \phi(x) + b$$  \hspace{1cm} (1)

where $\phi(x)$ denotes a nonlinear mapping function.

Defining $\varepsilon$ linear insensitive loss function:

$$L(f(x), y, \varepsilon) = \begin{cases} 0, & |y - f(x)| \leq \varepsilon \\ |y - f(x) - \varepsilon|, & |y - f(x)| > \varepsilon \end{cases}$$  \hspace{1cm} (2)

The slack variables $\xi_i$ and $\xi_i^*$ are introduced, and the problem of calculating $w, b$ can be expressed as:

$$\begin{aligned} & \min \left[ \frac{1}{2} ||w||^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) \right] \\ & \text{s.t.} \begin{cases} -y_i - w \ast \phi(x_i) - b \leq \varepsilon + \xi_i \\ y_i + w \ast \phi(x_i) + b \leq \varepsilon + \xi_i^* \end{cases} \quad i = 1,2,...,n \\ & \xi_i \geq 0, \xi_i^* \geq 0 \end{aligned}$$  \hspace{1cm} (3)

According to eq. (3), if the penalty factor $C$ is larger, it is proved that the training error is large and the sample penalty of $\varepsilon$ is larger, otherwise the opposite is true; if $\varepsilon$ is smaller, the error of the regression function is smaller.

Introduce Lagrange function, convert to dual form, as follow:

$$\begin{aligned} & \max \left[ -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) K(x_i, x_j) - \sum_{i=1}^{n} (\alpha_i + \alpha_i^*) \varepsilon + \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) y_i \right] \\ & \text{s.t.} \begin{cases} \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i \leq C \\ 0 \leq \alpha_i^* \leq C \end{cases} \end{aligned}$$  \hspace{1cm} (4)

where $K(x_i, x_j) = \phi(x_i) \phi(x_j)$ is the kernel function.

Assume that the optimal solution of eq. (4) is $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n), (\alpha_i^*, \alpha_2^*, ..., \alpha_n^*)$, and then:

$$\begin{aligned} w^* &= \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \phi(x) \\ b^* &= \frac{1}{N_{SSV}} \sum_{0 < \alpha_i < \varepsilon} [y_i - \sum_{j} (\alpha_j - \alpha_j^*) K(x_i, x_j) - \varepsilon] + \sum_{0 < \alpha_i^* < \varepsilon} [y_i - \sum_{j} (\alpha_j - \alpha_j^*) K(x_i, x_j) + \varepsilon] \end{aligned}$$  \hspace{1cm} (5)

where $N_{SSV}$ is the number of standard support vectors and $m$ is the number of support vectors. Then the approximate regression function is expressed as:

$$f(x) = \omega^* \ast \phi(x) + b^* = \sum_{i} (\alpha_i - \alpha_i^*) K(x_i, x) + b^*$$  \hspace{1cm} (6)

In addition, the key of SVR is choosing a proper kernel function, which directly affects the nonlinear mapping of samples. In order to obtain better generalization and nonlinear regression ability, the RBF kernel function is selected for SVR model construction in this paper.

2.2. The proposed method for RUL prediction

In engineering practice, even the same batch of products operate under the same operating conditions, their degradation rates are often different with each other, even these differences become more and more obvious over time. Such issue leads that the real-world failure historical data present varied degradation curves. Therefore, with the increase of training sample set the linearity will increase in the trained SVR model, thus the prediction error of the trained SVR model for an operating product will increase. However, the degradation process of a product can usually be divided into two or more stages, each stage present different degradation patterns. For example, Hu et al. (2016) [32] divided the degradation processes of generator bearings into four stages based on the change points of confidence levels. Kimtho et al. (2013) [33] divided the degradation processes of bearings into five
stages by analyzing the changes of frequency amplitudes in the power spectral density. Their studies show that subdivide the degradation process is beneficial to improve the prediction accuracy. Therefore, based on this view point, this paper proposes a parameter adaptive LS-SVR method to overcome the limitation of traditional SVR model, it’s based on multi-stages division and cluster sampling, and can be effectively used for RUL prediction. The flowchart of the proposed method is shown in Fig. 1.

In sampling theory, the coherent sample set composed of several connected basic units is called a cluster. For cluster sampling, the population is first divided into different sub-clusters according to some criteria, and then each sub-cluster is taken as the sampling unit, finally several sub-clusters are extracted from them by random principle for analysis. The shortcoming of cluster sampling is that the sampling error is often larger than that of simple random sampling, due to that the difference between different sub-clusters is often remarkable. However, the aim of employing cluster sampling in this paper is to increase the weight of corresponding sub-clusters in SVR model construction, so that such shortcoming can be ignored. The RUL prediction model is constructed at offline modeling part. Reference to Fig. 1, when a suitable health indicator (HI) is obtained to represent the whole degradation process, it can be divided subjectively into stages according to the its value, with the difference between each stage is as large as possible. Suppose that, a process has K stages, of which 0 stage is the initial state; 1, ⋯, K-1 stages are the degradation state of acceptable degradation level; K stage is the failure state. Then through M times cluster sampling of the test data of each stage, as well as model construction stage by stage. Therefore, the optimal model parameters can be selected adaptively according to the degradation state, so that each stage of the degradation process can be fitted at the best extent. Additionally, the test error of the SVR model decreases as the value of M increases, and finally converges at some level, which is verified in the later sections. Finally, based on the optimal number of cluster sampling, the multi-stages prediction model is constructed with the optimal parameters of each stage. Therefore, for an online operating product, the prediction accuracy of the proposed method is better than the traditional SVR model.

3. Case studies
In this section, we first use simulation study to demonstrate the reliability and robustness of the proposed model. Then, it is applied to a real case study of GaAs-based semiconductor lasers.
3.1. Simulation study
In this subsection, the prediction is illustrated and the performance is evaluated through numerical simulations. The degradation trajectories dataset of 30 fatigue-crack-growth units is obtained by the following polynomial-type Wiener process model [23]:

$$s(t) = s_0 + a \int_0^t \tau b^\tau - 1 d\tau + \int_0^t \sigma B(\tau)$$  \hspace{1cm} (8)

The initial crack length $s_0$ is 0 inch, and the failure threshold is set to be 50 inches. The parameter $a$ follows a Gamma distribution $G(\alpha, \beta)$. The preset values of the model parameters are given in Table 1. And the degradation trajectories of 30 units are shown in Fig. 2, which present the growth of the fatigue crack along with the stress cycles. The numerical simulations are all terminated at 25 million of cycles. These degradation trajectories are full of time variability and nonlinearity, can be considered as the real data for analysis. We choose 20 units randomly as the training units, and the remaining 10 units as the testing ones.

| Model parameters | $\alpha$ | $\beta$ | b | $\sigma_B^2$ |
|------------------|---------|--------|---|------------|
| Preset values    | 10      | 0.2    | 1.5| 4          |

Table 1 Preset parameters of the Wiener process model

In this experiment, we divide the degradation process into stages by every 10 inches:
• Stage 0 – initial state, HI=0;
• Stage 1 – degradation state, 0<HI<10;
• Stage 2 – degradation state, 10\leq HI<20;
• Stage 3 – degradation state, 20\leq HI<30;
• Stage 4 – degradation state, 30\leq HI<40;
• Stage 5 – degradation state, 40\leq HI<50;
• Stage 6 – failure state, HI\geq 50.

Clearly, the RUL prediction is performed at stage 1-5. In order to demonstrate the performance of the proposed method, we employ mean absolute error (MAE) as the performance metric, and the lower the MAE, the better the performance. Fig. 3 illustrates the performance changes with the increasing number of cluster sampling for one of the testing unit, while others are much the same as it and don’t be provided here. As the number of cluster sampling increases, the performance of the
The proposed method converges gradually at a constant level which is much better than that of the traditional SVR model without cluster sampling (i.e. the number of cluster sampling M=0). Therefore, in compare with the traditional SVR model, the proposed method can effectively improve the accuracy of online prediction. Although the time consumption of model construction increases as the number of cluster sampling increases, these works are carried out offline.

Fig. 3 The performance changes with the increasing number of cluster sampling for one of the testing unit

Due to the limited space, the test results of the 10 testing units are not provided here. As a replacement, we provide the performance metrics of them with the RUL prediction starts at 4 million of cycles, summarized in Fig. 4. The performance of the proposed method is compared with the traditional SVR method and the SbRP method. The MAE of the predicted results are shown in Table 2. Comparing these three methods, it shows clearly that the performance of the proposed method is the best one, and its reliability and robustness are all better than the other two methods.

Fig. 4 The comparison of the predictive metrics between the proposed method and the traditional SVR method as well as the SbRP method
Table 2 The comparison of the predicted results for the simulated data

| Testing unit | MAE of the proposed method | MAE of the traditional SVR method | MAE of the SbRP method |
|--------------|-----------------------------|-----------------------------------|------------------------|
| 1            | 0.407                       | 0.423                             | 0.400                  |
| 2            | 0.497                       | 0.574                             | 0.694                  |
| 3            | 0.553                       | 0.679                             | 0.560                  |
| 4            | 0.684                       | 0.87                              | 0.525                  |
| 5            | 0.903                       | 1.022                             | 1.720                  |
| 6            | 1.05                        | 1.434                             | 1.211                  |
| 7            | 1.275                       | 1.363                             | 2.183                  |
| 8            | 1.709                       | 2.641                             | 2.54                   |
| 9            | 2.115                       | 2.302                             | 2.761                  |
| 10           | 2.222                       | 2.411                             | 3.078                  |
| Mean         | 1.142                       | 1.372                             | 1.567                  |

3.2. Application to GaAs-based semiconductor lasers

In this subsection, the proposed method is applied to a degradation dataset of GaAs laser devices, which is taken from Meeker and Escobar (1998) [34]. Degradation of the laser device is measured as the operation current, which increases over time. Degradation data from 15 testing units are collected over time with measure frequency 250 h and terminal time 4000 h, they are shown in Table 3. This dataset has been extensively studied in degradation modeling [35-37] and RUL prediction [38]. In this paper, a laser is assumed to have failed if the percentage of its operating current exceeds a predefined threshold level FT=6%. Note that the observation time points are recorded at every 250 h, this leads to the inability to know the true life, i.e. the time arriving the failure threshold. Due to the reasonable degradation process of laser is an approximate linear degradation [35-37], the most possible failure time considered as the real one can be calculated by local linear regression method.

Table 3 GaAs-based laser performance degradation data under thermal stress of 80°C (percentage increase of working current)

| t/h | No. | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 250 | 0.47| 0.71| 0.71| 0.36| 0.27| 0.36| 0.36| 0.46| 0.51| 0.41| 0.44| 0.39| 0.3  | 0.44| 0.51|
| 500 | 0.93| 1.22| 1.17| 0.62| 0.61| 1.39| 0.92| 1.07| 0.93| 1.49| 1   | 0.8 | 0.74| 0.7  | 0.83|
| 750 | 2.11| 1.9 | 1.73| 1.36| 1.11| 1.95| 1.21| 1.42| 1.57| 2.38| 1.57| 1.35| 1.52 | 1.05| 1.29|
| 1000| 2.72| 2.3 | 1.99| 1.95| 1.77| 2.86| 1.46| 1.77| 1.96| 3   | 1.96| 1.74| 1.85 | 1.35| 1.52|
| 1250| 3.51| 2.87| 2.53| 2.3 | 2.06| 3.46| 1.93| 2.1 | 2.59| 3.84| 2.51| 2.98| 2.39 | 1.8  | 1.91|
| 1500| 4.34| 3.75| 2.97| 2.95| 2.58| 3.81| 2.39| 2.4 | 3.29| 4.5 | 2.84| 3.59| 2.95 | 2.55 | 2.27|
| 1750| 4.91| 4.42| 3.3 | 3.39| 2.99| 4.53| 2.68| 2.78| 3.61| 5.25| 3.47| 4.03| 3.51 | 2.83| 2.78|
| 2000| 5.48| 4.99| 3.94| 3.79| 3.38| 5.35| 2.94| 3.02| 4.11| 6.26| 4.01| 4.44| 3.92 | 3.39 | 3.42|
| 2250| 5.99| 5.51| 4.16| 4.11| 4.05| 5.92| 3.42| 3.29| 4.6 | 7.05| 4.51| 4.79| 5.03 | 3.72 | 3.78|
| 2500| 6.72| 6.07| 4.45| 4.5 | 4.63| 6.71| 4.09| 3.75| 4.91| 7.8 | 4.8 | 5.22| 5.47 | 4.09 | 4.1 |
| 2750| 7.13| 6.64| 4.89| 4.72| 5.24| 7.7 | 4.58| 4.16| 5.34| 8.32| 5.2 | 5.48| 5.84 | 4.83 | 4.38|
| 3000| 8   | 7.16| 5.27| 4.98| 5.62| 8.61| 4.84| 4.76| 5.84| 8.93| 5.66| 5.96| 6.5  | 5.41 | 4.63|
| 3250| 8.92| 7.78| 5.69| 5.28| 6.04| 9.15| 5.1 | 5.16| 6.4 | 9.55| 6.2 | 6.23| 6.94 | 5.76 | 5.38|
| 3500| 9.49| 8.42| 6.02| 5.61| 6.32| 9.95| 5.57| 5.46| 6.84| 10.45| 6.54| 6.99| 7.39 | 6.14 | 5.84|
| 3750| 9.87| 8.91| 6.45| 5.95| 7.1 | 10.49| 6.1 | 5.81| 7.2 | 11.28| 6.96| 7.37| 7.85 | 6.51 | 6.16|
| 4000| 10.94| 9.28| 6.88| 6.14| 7.59| 11.01| 7.17| 6.24| 7.88| 12.21| 7.42| 7.88| 8.09 | 6.88 | 6.62|

In order to demonstrate the effectiveness of the proposed method, we choose unit 6-15 as the training units while others as the testing ones. RUL prediction starts at 750 h, using separately the proposed method, the traditional SVR method and the SbRP method. For the proposed method, we divide the degradation process into stages by every 1% increase of working current:

- Stage 0 – initial state, HI=0;
- Stage 1 – degradation state, 0<HI<1%;
- Stage 2 – degradation state, 1%≤HI<2%;
- Stage 3 – degradation state, 2%≤HI<3%;
• Stage 4– degradation state, 3% ≤ HI < 4%;
• Stage 5– degradation state, 4% ≤ HI < 5%;
• Stage 6– degradation state, 5% ≤ HI < 6%;
• Stage 7– failure state, HI ≥ 6%.

Take the unit 3 for example, the predicted results are plotted in Fig. 5, the results by the proposed method is relatively closer to the actual value. In addition, Fig. 6 provides the predicted results of unit 1, unit 2, unit 4 and unit 5. For unit 1 and unit 2, we can see that the predicted results by traditional SVR method don't converge near the real value over time, even the gap expands when the degradation state closes to failure, which is unacceptable in engineering practice. This issue is mainly caused by the limitation of SVR model that the linearity will increase as the training data increase, so that the traditional SVR model may get poor performance for some nonlinear units. Moreover, the predicted results for unit 1 and unit 2 by SbRP method are also of large gap with the real value during the whole lifecycle. Therefore, the proposed method has better reliability and robustness than the other two methods.

![Fig. 5 The comparison of the predicted results of testing unit 3](image)
Fig. 6 The comparison of the predicted results of testing unit 1, unit 2, unit 4 and unit 5

Table 4 The comparison of the predicted results for the GaAs-based lasers data

| Testing unit | MAE of the proposed method | MAE of the traditional SVR method | MAE of the SbRP method |
|--------------|---------------------------|-----------------------------------|------------------------|
| 1            | 80.765                    | 411.257                           | 451.276                |
| 2            | 99.799                    | 323.853                           | 340.039                |
| 3            | 189.368                   | 222.83                            | 303.851                |
| 4            | 347.734                   | 400.756                           | 540.246                |
| 5            | 22.060                    | 68.630                            | 99.000                 |
| Mean         | **149.145**               | **285.465**                       | **346.882**            |

To demonstrate further the effectiveness of the proposed method, Table 4 provides the MAE of the predicted results for the testing units. The proposed method is of the best performance for all the five testing units. Therefore, the proposed method can be used effectively for RUL prediction, and has better reliability and robustness than the traditional SVR method and the SbRP method.

4. Conclusions

This paper proposes a novel LS-SVR method based on multi-stages division and cluster sampling. It divides the degradation process into several stages, and through cluster sampling of the test data of each stage, as well as model construction stage by stage. Therefore, the optimal model parameters can be selected adaptively as the degradation evolves, so that each stage of the degradation process can be fitted at the best extent. This method overcomes the limitation of traditional SVR model that the linearity will increase as the training data increase.

To demonstrate the advantages of the proposed method, a simulated dataset of fatigue crack growth and a real-world degradation dataset of GaAs-based lasers are applied to study. And the predicted results are compared with the counterparts by the traditional SVR method and the SbRP method.
Results show that the proposed method is of better reliability and robustness than the traditional SVR method and the SbRP method. The proposed method extends the studies of SVR methods for RUL prediction, and is of effectiveness in engineering practice.

References

[1] Vogl, G. W., Weiss, B. A., Helu, M. (2019). A review of diagnostic and prognostic capabilities and best practices for manufacturing. *Journal of Intelligent Manufacturing*, 30(1), 79–95.

[2] Javed, K., Gouriveau, R., Zerhouni, N., Nectoux, P. (2015). Enabling health monitoring approach based on vibration data for accurate prognostics. *IEEE Transactions on Industrial Electronics*, 62(1), 647-656.

[3] Si, X. S., Wang, W., Hu, C. H., Zhou, D. H. (2011). Remaining useful life estimation – a review on the statistical data driven approaches. *European Journal of Operational Research*, 213, 1-14.

[4] Hu, Y. W., Liu, S. J., Lu, H. T., Zhang, H. C. (2019). Remaining useful life model and assessment of mechanical products: a brief review and a note on the state space model method. *Chinese Journal of Mechanical Engineering*, 32(1), 15.

[5] An, D., Kim, N. H., Choi, J. H. (2015). Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. *Reliability Engineering & System Safety*, 133, 223-236.

[6] Li, C. J., Lee, H. (2005). Gear fatigue crack prognosis using embedded model, gear dynamic model and fracture mechanics. *Mechanical Systems and Signal Processing*, 19(4), 836-846.

[7] Kontar, R., Zhou, S. Y., Sankavaram, C., Du, X. Y., Zhang, Y. L. (2018). Nonparametric-condition-based remaining useful life prediction incorporating external factors. *IEEE Transactions on Reliability*, 67(1), 41-52.

[8] Mosallam, A., Medjaher, K., Zerhouni, N. (2016). Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing*, 27(5), 1037-1048.

[9] Fan, Z. L., Liu, G. B., Si, X. S., Zhang, Q., Zhang, Q. H. (2013). Degradation data-driven approach for remaining useful life estimation. *Journal of Systems Engineering and Electronics*, 24(1), 173-182.

[10] Long, B., Xian, W. M., Jiang, L., Liu, Z. (2013). An improved autoregressive model by particle swarm optimization for prognostics of lithium-ion batteries. *Microelectronics Reliability*, 53(6), 821-831.

[11] Barraza-barraza, D., Tercero-Gomez, V. G., Beruvides, M. G., Limon-Robles, J. (2017). An adaptive ARX model to estimate the RUL of aluminum plates based on its crack growth. *Mechanical Systems & Signal Processing*, 82, 519-536.

[12] Fleischer, C., Waag, W., Bai, Z., Sauer, D. U. (2013). On-line self-learning time forward voltage prognosis for lithium-ion batteries using adaptive neuro-fuzzy inference system. *Journal of Power Sources*, 243, 728-749.

[13] Dong, M., He, D. (2007). A segmental hidden semi-Markov model (HSMM)-based diagnostics and prognostics framework and methodology. *Mechanical Systems and Signal Processing*, 21(5), 2248-2266.

[14] Gebraeel, N. Z., Lawley, M. A. (2008). A neural network degradation model for computing and updating residual life distributions. *IEEE Transactions on Automation Science and Engineering*, 5(1), 154-163.

[15] Tian, Z. G. (2012). An artificial neural network approach for remaining useful life prediction of equipment subject to condition monitoring. *Journal of Intelligent Manufacturing*, 23(2), 227-237.

[16] Mahamad, A. K., Saon, S., Hiyama, T. (2010). Predicting remaining useful life of rotating machinery based artificial neural network. *Computers & Mathematics with Applications*, 60(4), 1078-1087.
[17] Loutas, T. H., Roulias, D., Georgoulas, G. (2013). Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic e-support vectors regression. *IEEE Transactions on Reliability*, 62(4), 821-832.

[18] Khelif, R., Chebel-Morello, B., Malinowski, S., Laajili, E., Zerhouni, N. (2017). Direct remaining useful life estimation based on support vector regression. *IEEE Transactions on Industrial Electronics*, 64(3), 2276-2285.

[19] Zio, E., Maio, F. D. (2010). A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear system. *Reliability Engineering & System Safety*, 95(1), 49-57.

[20] You, M. Y., Meng, G. (2013). A framework of similarity-based residual life prediction approaches using degradation histories with failure, preventive maintenance, and suspension events. *IEEE Transactions on Reliability*, 62(1), 127-135.

[21] Gu M. Y., Chen Y. L. (2019). Two improvements of similarity-based residual life prediction methods. *Journal of Intelligent Manufacturing*, 30(1), 303-315.

[22] Noortwijk, J. M. V. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94(1), 2-21.

[23] Li, N. P., Lei, Y. G., Yan, T., Li, N. B., Han, T. (2019). A Wiener-process- model-based method for remaining useful life prediction considering unit-to-unit variability. *IEEE Transactions on Industrial Electronics*, 66(3), 2092-2101.

[24] Son, K. L., Fouladirad, M., Barros, A., Levrat, E., Lung, B. (2013). Remaining useful life estimation based on stochastic deterioration models: a comparative study. *Reliability Engineering and System Safety*, 112, 165-175.

[25] Nuhic, A., Terzimehic, T., Soczka-Guth, T., Buchholz, M., Dietmayer, K. (2013). Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. *Journal of Power Sources*, 239, 680–688.

[26] Benkedjouh, T., Medjaher, K., Zerhouni, N., Rechak, S. (2013). Health assessment and life prediction of cutting tools based on support vector regression. *Journal of Intelligent Manufacturing*, 26(2), 213-223.

[27] Chen, Z. Z., Cao, S. C., Mao Z. J. (2018). Remaining useful life estimation of aircraft engines using a modified similarity and supporting vector machine (SVM) approach. *Energies*, 11(1), 1-14.

[28] Ordonez, C., Lasheras, F. S., Roca-Pardinas, J., Juez, F. J. D. (2019). A hybrid ARIMA–SVM model for the study of the remaining useful life of aircraft engines. *Journal of Computational and Applied Mathematics*, 346, 184-191.

[29] Wei, J. W., Dong, G. Z., Chen, Z. H. (2018). Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression. *IEEE Transactions on Industrial Electronics*, 65(7), 5634-5643.

[30] Pei, H., Hu, C. H., Si, X. S., Zhang, J. X., Pang, Z. N., Zhang, P. (2019). Review of machine learning based remaining useful life prediction methods for equipment. *Journal of Mechanical Engineering*, 55(8), 1-13.

[31] Wen, Y. X., Wu, J. G., Das, D., Tseng, T. L. (2018). Degradation modeling and RUL prediction using wiener process subject to multiple change points and unit heterogeneity. *Reliability Engineering & System Safety*, 176, 113-124.

[32] Hu, Y. G., Li, H., Liao, X. L., Song, E. B., Liu, H. T., Chen, Z. (2016). A probability evaluation method of early deterioriation condition for the critical components of wind turbine generator systems. Mechanical Systems and Signal Processing, 76–77, 729–741.

[33] Kimotho, J. K., Sondermann-Woelke, C., Meyer, T., Sextro, W. (2013). Machinery prognostic method based on multi-class support vector machines and hybrid differential evolution–particle swarm optimization. *Chemical Engineering Transactions*, 33, 619–624.

[34] Meeker, W., Escobar, L. A. (1998). Statistical methods for reliability data. *Probability and Statistics*, New York: Wiley.
[35] Wang, X., Xu, D. H. (2010). An inverse gaussian process model for degradation data. *Technometrics*, 52(2), 188-197.

[36] Ye, Z. S., Chen, N. (2014). The inverse gaussian process as a degradation model. *Technometrics*, 56(3), 302-311.

[37] Pan, D. H., Liu, J. B., Cao, J. D. (2016). Remaining useful life estimation using an inverse gaussian degradation model. *Neurocomputing*, 185, 64-72.

[38] Liu, Z., Wang, Q., Song, C. L., Cheng, Y. H. (2017). Similarity-based difference analysis approach for remaining useful life prediction of GaAs-based semiconductor lasers. *IEEE Access*, 5, 21508-21523.