A Data-driven Based Decomposition-integration Method for Remanufacturing Cost Prediction of End-of-Life Products

Zhigang Jiang[a *], Zhouyang Ding[b], Ying Liu[c] , Yan Wang[d], Xiaoli Hu[e], Yihua Yang[f]

[a]Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science & Technology, Wuhan 430081, China

[b]Key Laboratory of Metallurgical Equipment and Control Technology, Wuhan University of Science & Technology, Wuhan 430081, China

[c] Institute of mechanical and manufacturing engineering, School of Engineering, Cardiff University, Cardiff CF24 3AA

[d]Department of Computing, Engineering and Mathematics, University of Brighton, Brighton, BN2 4GJ, United Kingdom

[e] School of Mechanical Engineering, Shanghai Dianji University, Shanghai 201306, China

[f]Sevalo Construction Machinery Remanufacturing Co., Ltd , Wuhan 430040, China

Abstract: Remanufacturing cost prediction is conducive to visually judging the remanufacturability of End-of-Life (EOL) products from economic perspective. However, due to the randomness, non-linearity of remanufacturing cost and the lack of sufficient data samples. The general method for predicting the remanufacturing cost of EOL products is very low precision. To this end, a data-driven based decomposition-integration method is proposed to predict remanufacturing cost of EOL products. The approach is based on historical remanufacturing cost data to build a model for prediction. First of all, the remanufacturing cost of individual EOL product is arranged as a time series in reprocessing order. The Improved Local Mean Decomposition (ILMD) is employed to decompose remanufacturing cost time series data into several components with smooth, periodic fluctuation and use this as input. BP neural network based on Particle Swarm Optimization (PSO-BP) algorithm is utilized to predict the cost of each component. Finally, the predicted components are added to obtain the final prediction result. To illustrate and verify the feasibility of the proposed method, the remanufacturing cost of DH220 excavator is applied as the sample data, and empirical results show that the proposed model is statistically superior to other benchmark models owing to its high prediction accuracy and less computation time. And proposed method can be utilized as an effective tool to analyze and predict remanufacturing cost of EOL products.

Keywords: Remanufacturing cost prediction; data-driven; decomposition-integration method; End-of-Life products

1 Introduction

Traditional manufacturing industry is under increasing pressures to be resource efficiency, thus to consume less resources and to reduce pollution on the environment [1]. Remanufacturing is a significant value-recovery approach industrial process that can return End-of-Life (EOL) products to as-good-as-new ones [2~4], thus releases the huge hidden economic value of EOL products. It is becoming a strategic emerging industry in many countries, e.g. China and USA [5]. However, at present, there is an urgent need for an efficient and effective decision-making system to decide whether an EOL product should be remanufactured. The assessment and decision-making of remanufacturability for EOL products should be, concentrated on both technical and economic feasibility. To this end, some evaluation models and reliability theories for remanufacturability have been presented. Zhang et al. [6] established a decision model for the remanufacturing of used parts from the perspective of predicting the remaining service life. A remanufacturing evaluation model in which technical feasibility, economic and environmental benefit [7, 8] were considered in a product remanufacturing analysis system based on Analytic Hierarchy Process (AHP) and Case-Based Reasoning (CBR) [9]. On the basis of quantifying the cost, environmental impact and dismantling time of the remanufacturing process, Lee et al. [10] proposed a multi-objective methodology for determining the appropriate EOL options for components. Similarly, Liu et al. [11] also used cost, environmental impact and labor time as indicators to measure the sustainability of EOL products.

Although the above methods and models have provided valuable guidelines for the remanufacturing decision-making in enterprises, there still remain some notable limitations. Since most of these methods can only evaluate the remanufacturability of a product qualitatively rather than quantitatively, these may lead to inefficient and inaccurate results for decision-makings, and the misinterpretation results may even result in a financial loss. Therefore, there are limited real-world applications of these approaches. Our anecdotal evidences suggested that if the overall remanufacturing cost of EOL products is less than 50%-80% of that of the manufacture of new products, then the remanufacturing is profitable [12]. Predicting the remanufacturing cost of EOL products accurately will paramount for decision making regarding whether or not an EOL product should be remanufactured.

Nevertheless, remanufacturing cost prediction is faced with many difficulties. Not only is the remanufacturing cost affected by uncertain market factors, but also as a special manufacturing process, remanufacturing faces many uncertainties, e.g., the remanufacturing rate of recycled products, the arrival time of the EOL products, reprocessing route, the quantity of recycled products, and the purchase quantity of new parts, etc. These factors of remanufacturing cost make it difficult to create a model to predict remanufacturing cost that possesses, good nonlinear approximation ability. To this end, this paper takes the prediction of remanufacturing cost for EOL products as the research objective, focuses on the decomposition of less sample data and adopts a well-fitted algorithm, and the prediction accuracy of the model is further improved. In the current researches, the remanufacturing cost prediction models can be divided into two categories: traditional linear prediction methods (such as vector auto-regression) and intelligent algorithm model with high fitting accuracy. As for traditional linear models, linear regression models, Grey Theory (GT) methods, etc. have popularly been used for remanufacturing cost prediction. For instance, according to the statistical characteristics of the random distribution of product remanufacturing cost, Liu et al. [13] established a linear regression model to predict the remanufacturing cost for EOL products. In view of the difference of remanufacturing cost due to the failure severity of the machine tools, Zhang et al. [14] established a cost analysis and prediction model of remanufactured machine tools based on the method of job cost and CBR. Goodall et al. [15] also used CBR to estimate the cost of remanufacturing activities under uncertain circumstances. Sabharwal and Garg [16] developed a model to evaluate the economic viability of remanufacturing using graph theory and calculated the cost-benefit indices of five common products. Based on the remanufacturing process of heavy-duty engines, and analyzing the influencing factors of remanufacturing cost, Sang et al. [17] used GT to establish a state-based remanufacturing cost prediction model.

In terms of intelligent algorithm model, Support Vector Machine (SVM), Artificial Neural Network (ANN) and Least Squares Support Vector Regression (LSSVR) have become the most popular methods for remanufacturing cost prediction, and they are actually superior to traditional linear methods. For example, Qin et al. [18] proposed a neural network based cost model to estimate the remanufacturing cost of an engineering machinery hydraulic cylinder. Considering the randomness uncertainty and limited available data samples of remanufacturing cost, a new method based on SVM was proposed by Xiang et al. [19] to predict the remanufacturing cost. Zhang et al. [20] combined the semi-supervised learning algorithm with the LSSVR to predict the remanufacturing cost of EOL components and parts in the case of a small number of samples. Furthermore, Gao Ting [21] developed a remanufacturing cost prediction model using BP neural network to estimate the cost of remanufacturing of EOL hydraulic cylinders. But these intelligent models also have some shortcomings, such as parameter sensitivity, potential over-fitting, input is difficult to define, low prediction accuracy due to small amount of data, etc.

To do so, a data-driven based decomposition-integration approach is proposed to predict remanufacturing cost of EOL products. In general, the decomposition-integration prediction model consists of three steps, including decomposition of original data, individual prediction of each decomposition mode, and integrated prediction [22, 23]. Significantly, because of their good functions in internal factors identification, prediction accuracy, etc., similar decomposition-integration models have achieved good results in prediction. Zhang et al. [24] introduced Empirical Mode Decomposition (EMD) to decompose wind power sequence, and experiments showed that this method can effectively reduce the errors and improve the prediction accuracy. Qiu et al. [25] presented a decomposition-integration method for short-term electricity price forecasting by utilizing the EMD, Kernel Ridge Regression (KRR) and Support Vector Regression (SVR), and results show that the data decomposition method can effectively improve the prediction accuracy. Lu et al. [26] used the Ensemble Empirical Mode Decomposition (EEMD) and Auto Regressive and Moving Average model (ARMA) to establish a decomposition-integration learning paradigm, which effectively predicted the grid price of power grids. In addition, similar decomposition-integration models have achieved good results in the prediction of power load [27-29], air passenger traffic [30] and so on. The above researches show that the decomposition-ensemble strategy can significantly improve the performance of the prediction model. However, adapting decomposition-integration approach in remanufacturing cost prediction has not been addressed before. This model can effectively avoid the disadvantages of low prediction accuracy caused by the small amount and large randomness of data, and is especially suitable for remanufacturing cost prediction.

This paper proposes a decomposition-integration approach based on data-driven to predict the remanufacturing cost of EOL products, which can support remanufacturing decision and implementation. The approach is based on historical remanufacturing cost data to build a model for prediction. The method involves three main steps, i.e., data decomposition, individual prediction and integrated prediction. At the first stage, the remanufacturing cost of individual EOL product is arranged as a time series in reprocessing order. Then the Improved Local Mean Decomposition (ILMD) is employed to decompose these time series into a series of components with smooth, periodic fluctuation. Multiple different data sequences can be obtained by data decomposition by ILMD method to compensate for the low prediction accuracy caused by the small amount of data. The second stage is the individual prediction for each component of decomposition. Thus, one specific Artificial Intelligence (AI) technique, i.e., BP neural network based on Particle Swarm Optimization (PSO-BP), is utilized to model and predict all modes. The final stage is integrated prediction, which adds the prediction results of the modes at the same time point to obtain the final results. To illustrate and verify the feasibility of the proposed method, the remanufacturing cost prediction of an excavator is applied as case study, and seven benchmark models (including two single AI models and five similar decomposition-integration models) are introduced for comparison purpose. Specifically, the innovations in this paper focus on the following three aspects: (1) The data decomposition by ILMD is to obtain the smooth and periodic data components of remanufacturing cost time series, and transform the small sample data into a large amount of data. Decomposed data can improve the accuracy of prediction results. (2) A decomposition-integrated prediction model combining ILMD and BP algorithm is proposed to predict the remanufacturing cost of EOL products. The BP neural network is improved by Particle Swarm Optimization algorithm, which plays an important role in improving the prediction accuracy. (3) Through the direct data input model, the data-driven method makes the prediction results more objective and accurate. Empirical research shows that the proposed method is feasible.

2 Model Development

There are two main obstacles to remanufacturing cost prediction of EOL products. One is that the uncertainty of the remanufacturing process leads to the cost instability, the traditional linear prediction is difficult. The other is that the scale of the remanufacturing industry is limited, and the amount of available data accumulated is less, which results in the low accuracy of the prediction. For this purpose, a data-driven method for predicting the remanufacturing cost of EOL products is proposed in this section, which is the cost prediction model based data decomposition-integration. The remanufacturing cost of EOL products is affected not only by processing process, etc., but also by the relationship between supply and demand in the market. Therefore, the remanufacturing cost change with time, so these data can also be regarded as time series data. Arrange the cost of the product into time series data according to the time sequence of product remanufacture, time series prediction model can be established to predict the future remanufacturing cost of EOL products.

In this model, time series data are decomposed by ILMD, which makes the data capacity increase, and the decomposed components are periodic and stable. Cost prediction is performed by using the PSO-BP algorithm to avoid nonlinear and non-stationary of data. The model has the following three advantages: (1) The amount of data is expanded by data decomposition, which is very suitable for remanufacturing cost prediction with a small amount of sample data. For example, the original data has 100 groups, which are decomposed into 7 components, and each component also has 100 sets of data, so after decomposing, there are a total of 700 groups. (2) Data decomposition reduces the order of magnitude of data, making each component periodic and easy to calculate. (3) By decomposing the data, the original non-stationary and irregular cost sequence can be decomposed into some steady and regular components, and the prediction accuracy is improved finally.

Section 2.1 generally explains the proposed model, and Sections 2.2-2.4 describe in detail the three main stages of the model and the techniques/methods used in each stage.

2.1 Overall structure of the proposed model

In order to improve the prediction accuracy, the data-driven based decomposition-integration approach is proposed for predicting remanufacturing cost of EOL products, as shown in Fig. 1. The proposed method consists of three main stages, i.e., the historical cost data is decomposed into several data sequences with different characteristics by ILMD, and a powerful AI intelligent algorithm (i.e., PSO-BP) is employed to predict each mode separately, then the integrated prediction obtains the final results by adding the prediction results of the respective decomposition modes at the same time point.

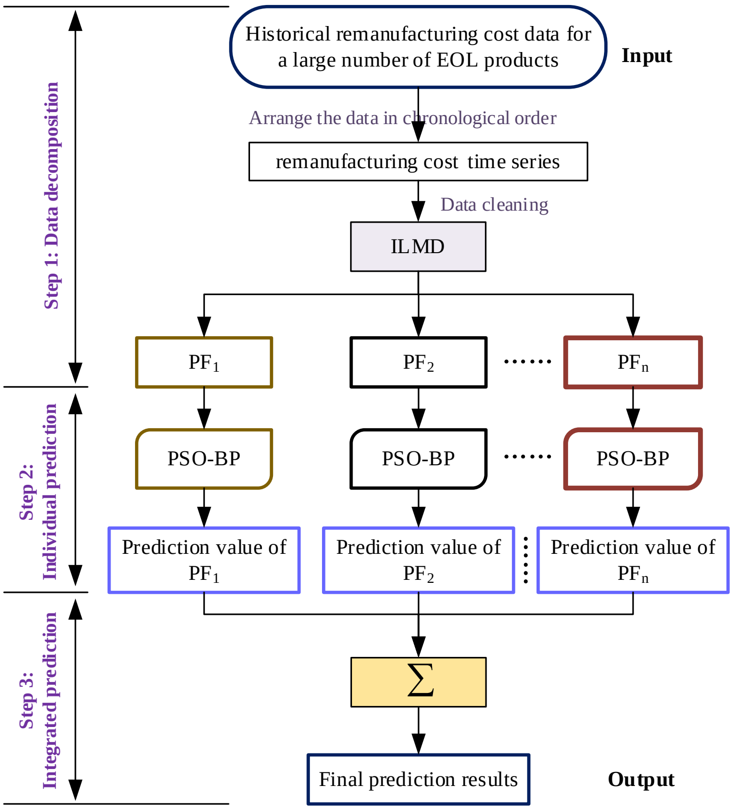


Fig. 1 Framework of data-driven based decomposition-integration model for remanufacturing cost prediction

Step 1: Data decomposition

The remanufacturing cost of individual EOL product is arranged as a time series according to the order of repair. And through the data cleaning, that is, the data that deviates far from the average value is eliminated to ensure the accuracy of the prediction. The time series $x_t, (t = 1,2,...,T)$ is decomposed into n product function components （PF）, $PF_i(t), (i = 1,2,...,n)$ and one residue $u_k(t)$ by using ILMD technology.

Step 2: Individual prediction

In the second step, the PSO-BP algorithm, is used as a predictive tool for each mode. Thus, the prediction results $r_i(t)$ of the corresponding component can be obtained.

Step 3: Integrated prediction

The predicted values of all PF components and margins are added together to obtain the final predicted value, i.e., $\vec{C}_t = \sum_{i=1}^{n} \vec{PF}_{it} + \vec{u}_t$, where $\vec{PF}_{it}$ is the predicted value of PF_i at time t, and $\vec{u}_t$ is the predicted value of the margin at time t. In particular, the direct addition method is simple, but is more reliable and can reduce computational complexity.

The above three main steps and related techniques of the proposed method are detailed in Sections 2.2 to 2.4, respectively.

2.2 Data decomposition

At the stage of data decomposition, an effective data decomposition technique, i.e., the ILMD, is used. To improve the accuracy of the cost modelling, some researchers tried to decompose the time series by using the data decomposition method, mine the components of the data information at different time scales. An appropriate prediction model is established for different components, and the prediction effect is improved to a certain extent. The more widely used decomposition methods are EMD [31], EEMD [32], local mean decomposition (LMD) [33], and each decomposition method has its own advantages and deficiencies. The classic LMD uses the moving average method for calculating the local mean function and the envelope estimation function. For low frequency signals, this produces a significant error [34]. Therefore, the classic LMD method will not achieve a better decomposition effect. The improved LMD (ILMD) has been applied in this paper. The cubic mean square interpolation method is used to replace the original moving average method to obtain its local mean function and envelope estimation function. The ILMD does not need to be recalculated, and its error is small, so it has great advantages in computational efficiency and accuracy. The decomposition process of ILMD is as follows:

(1) Find all the local extreme points n_i of the original signal $x(t)$, and then find the average of all the adjacent extreme points. The formula is as follows:

$$m_i = \frac{n_i + n_{i+1}}{2} \tag{1}$$

Connect all adjacent average values m_i with a straight line, and then use cubic spline interpolation to obtain its local mean function $m_{11}(t)$.

(2) Calculate the envelope estimate:

$$a_i = \frac{n_i - n_{i+1}}{2} \tag{2}$$

All adjacent envelope estimates a_i a are connected by a straight line, and the envelope estimation function $a_{11}(t)$ is obtained by cubic spline interpolation.

(3) Separate $m_{11}(t)$ from the original signal to get $h_{11}(t)$, which is calculated as follows:

$$h_{11}(t) = x(t) - m_{11}(t) \tag{3}$$

(4) Divide $h_{11}(t)$ obtained by the above formula by $a_{11}(t)$ to obtain $s_{11}(t)$. The calculation formula is as follows:

$$s_{11}(t) = h_{11}(t) / a_{11}(t) \tag{4}$$

(5) Repeat the above four steps for $s_{11}(t)$, if the envelope evaluation function $a_{12}(t)$ of $s_{11}(t)$ is not equal to 1, it means that $s_{11}(t)$ is not a pure FM signal, then repeat the above four steps again, until $s_{11}(t)$ is a pure FM signal, that is its envelope evaluation function $a_{1(n+1)}(t) = 1$ and its termination condition is $\lim_{n\to\infty} a_{1n}(t) = 1$. Without affecting the decomposition effect, a variation range Δ can also be set in order to reduce the number of iterations, and the iteration is stopped when $1\text{-}\Delta \le a_{1n}(t) \le 1 + \Delta$ is satisfied.

(6) Many envelope estimation functions are generated during the iteration process. Multiply all of them to obtain an envelope signal.

$$a_1(t) = \prod_{q=1}^{n} a_{1q}(t) \tag{5}$$

(7) The first PF component PF1 is the product of the pure FM signal $s_{1n}(t)$ and the envelope signal

$a_1(t)$, i.e., $PF1 = a_1(t)s_{1n}(t)$.

(8) Separate the PF1 from the original signal $x(t)$ to get the signal $u_1(t)$, repeat the above process for $u_1(t)$, and the loop is repeated m times until $u_m(t)$ is a monotonic function. As follows:

$$\begin{gathered} u_1(t) = x(t) - PF_1(t) \\ u_2(t) = u_1(t) - PF_2(t) \\ \cdots \\ u_m(t) = u_{m-1}(t) - PF_m(t) \end{gathered} \tag{6}$$

Finally, the original signal $x(t)$ can be expressed in the following form:

$$x(t) = \sum_{p=1}^{m} PF_p(t) + u_m(t) \tag{7}$$

2.3 Individual prediction

Intelligent algorithms are the best tools for applying in the individual prediction stage due to its good adaptability and fitting ability. For this reason, researchers in remanufacturing cost prediction apply AI prediction and computational intelligence such as ANN, SVM, and LSSVR to such highly random, non-linear, non-stationary complex systems. Among many AI models, ANN have excellent features such as large-scale distributed parallel processing, nonlinear, self-organizing, self-learning, and associative memory, which can be used as an advanced means of cost prediction. The neural network (NN) can be used for time series prediction, which refers to using a neural network to approximate a time series, and the first k values $x_n, x_{n-1}, x_{n-2}, \cdots, x_{n-k} + 1$ of the time series can be used to predict the next m values $x_{n+1}, x_{n+2}, \cdots, x_{n+m}$. Described as follows:

$$x_{n+m} = F(x_n, x_{n-1}, x_{n-2}, \cdots, x_{n-k} + 1) \tag{8}$$

That is, the neural network is used to fit the function F and use it to derive future values. In the prediction process, the obtained prediction value can be used as an input of the next prediction to calculate further prediction values for iterative multi-step prediction.

As a kind of ANN, BP neural network is the most widely used due to its good approximation ability and mature training methods. BP neural network is a multilayer feedforward neural network (MFNN), which can approximate any nonlinear continuous function with arbitrary precision, has strong fault tolerance and fast processing speed, and is suitable for prediction. Therefore, Particle Swarm Optimization BP neural network is introduced as an individual predictive tool for components in this research.

2.3.1 BP neural network

BP neural network, one of the most widely used neural network models, is a multilayer feedforward network which propagates in the opposite direction of error [35]. The topological structure of the BP neural network model comprises input layer, hidden layer and output layer. The learning process of BP network is divided into two stages:

(1) Input a known learning sample, and calculate the output of each neuron from the first layer of the network backward through the set network structure and the weight and threshold of the previous iteration.

(2) Modify the weight and threshold, and calculate the influence (gradient) of each weight and threshold on the total error from the last layer, and modify the weights and thresholds accordingly.

The above two processes are repeated alternately, and the error is transmitted back to the layer by layer to correct the weight and threshold between the layers until the convergence is reached. The PSO fitness function adopted in this paper is determined according to this rule.

In neural network training, the so-called "over-fitting" problem may occur. The sample error for the training set can be small, but the error is large for new samples other than the training set. That is the network remembers the trained samples, but lacks the ability to generalize the new samples. This paper improves the generalization ability of the network by correcting the network error performance function and the error performance adjustment rate.

2.3.2 PSO algorithm

PSO is an intelligent optimization algorithm that can optimize the internal structure of BP neural network, improve the convergence speed, effectively prevent falling into local extremum, and quickly search for global optimal solution.

Suppose the velocity and position vector of particle i are $v_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$ and $x_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$, where D represents the dimension of space, then

$$v_{id}(t+1) = \omega v_{id}(t) + cr_1(p_{id}(t) - X_{id}(t)) + c_2 r_2(p_{gd}(t) - X_{id}(t)) \qquad (9)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \qquad (10)$$

Where c_1 and c_2 are acceleration coefficients; $p_{id}(t)$ and $p_{gd}(t)$ are the optimal positions of particle i and particle group at time t; r_1 and r_2 are random numbers between 0 and 1; ω is inertia factor.

2.3.3 PSO-BP model

In the process of cost prediction, since the initial parameters of the network are randomly set, it may cause a large error or non-convergence in the process of network training. Therefore, PSO is used to improve the initial parameters of BP neural network. In the process of optimization, each particle individual is a set of weights and thresholds. After the particle finds the optimal position, the information carried by the particle is the optimal weight value and threshold value. The information is assigned to the network and then trained and predicted. The algorithm flow of the PSO-BP model is shown in Fig. 2. The prediction steps of PSO-BP model are as follows:

Step 1: The BP neural network is initialized, and the data is normalized.

Step 2: Initialization parameters. Determine the velocity $v_{\max}$, $v_{\min}$, position $x_{\max}$, $x_{\min}$, c_1, c_2, r_1, r_2, $w_{\max}$ and $w_{\min}$, population number N, maximum iteration number T, objective function E, minimum error e in the PSO algorithm.

Step 3: Randomly generate N particles pop, each of which is a set of potential solutions, the N group weight threshold.

Step 4: Set the e as the index, the fitness values of each particle are calculated, and P_{best} and G_{best} are obtained by comparison.

Step 5: When the number of iterations reaches T or the fitness value $E < e$, the particle pop at this time is output. Otherwise, let the number of iterations $t = t + 1$ and return to *step 3*.

Step 6: The optimal solution of the PSO algorithm is employed as the initial weight and threshold of the BP neural network for cost prediction.

Step 7: The BP neural network is trained and tested with the normalized cost data, and the error is compared with the actual cost. If the error is met, the prediction result is output, otherwise go to *step 2*.

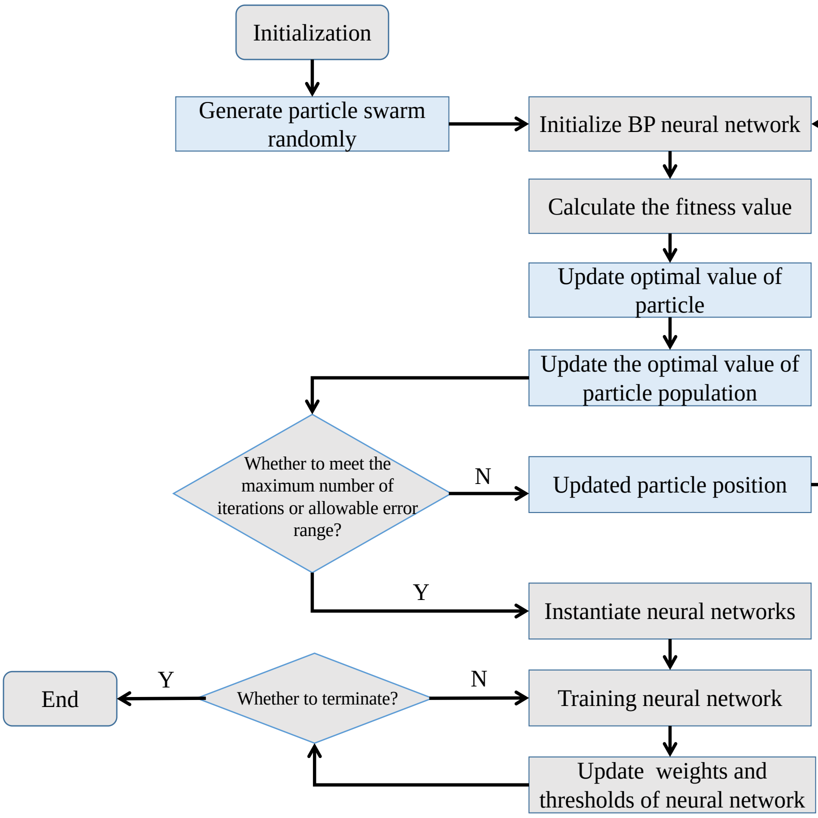


Fig. 2 Flow of PSO neural network algorithm

2.4 Integrated prediction

In the stage of integrated prediction, BP neural network or linear regression etc., can be employed to obtain the final result. However, the integration of results by direct addition has an irreplaceable advantage in terms of calculation speed and accuracy.

$$\vec{c}_t = f\left(\vec{d}_t(1), \vec{d}_t(2), \ldots, \vec{d}_t(k)\right) \tag{11}$$

where $\vec{c}_t$ denotes the final cost prediction result at time t ,and $\vec{d}_t(j)$ is the prediction value for the *j*-th mode. Owing to the original sequence data is decomposed into a plurality of modes by ILDM, the prediction values of the components at each moment are further obtained by separately predicting. The predicted values of each component at the same time point are added to obtain the final prediction result at that time point, i.e., $\vec{c}_t = \sum_{j=1}^{k} d_t(j)$. Therefore, a simple but effective integration method is adopted, i.e., the ADD.

3 Case study

In order to verify the effectiveness of the proposed method, the remanufacturing cost prediction of the excavators DH220 in a construction machinery company is taken as the case study. The company specializes in the remanufacturing of construction machinery, the main products are various types of excavators (e.g., model DH220). As shown in Fig. 3, it is the DH220 excavator and the remanufacturing process of excavator. An empirical study is designed in Section 3.1 and the results of the experiment are further discussed in Section 3.2.

Fig. 3 Excavator of model DH220 and remanufacturing process of excavator

3.1 Experiment design

3.1.1 Data descriptions and prediction process

Hereon, the historical remanufacturing cost of DH220 excavator from June to December 2017 in a remanufacturing company is used as the original sample, for a total of 85 sets of data, as shown in Table 1. After data cleaning to eliminate outliers, a total of 80 sets of data are valid and can be used as input to the predictive model, as shown in Fig. 4. The sample data are divided into two subsets, i.e., the preceding 2/3 of total observations are selected as the training data input model, and the last 25 sets of data are used as test samples to detect the accuracy of the prediction model. The data-driven based decomposition-integration model proposed in this paper is adapted to predict the remanufacturing cost of the excavator. The prediction processes are as follows:

Table 1 Historical data of remanufacturing cost of DH220 excavator

Product NO. / Repair time Ranking T	Cost (¥) /yuan	Product NO. / Repair time Ranking T	Cost (¥) /yuan	Product NO. / Repair time Ranking T	Cost (¥) /yuan
1	393656	8	351145	15	339612
2	406136	9	392699	16	277098
3	400338	10	426110	……	……
4	445091	11	446508	……	……
5	384368	12	441273	83	297124
6	422581	13	290964	84	307812
7	367647	14	486811	85	352145

Step 1: Preprocess the historical data, make the data into a time series in the order in which they are remanufactured. Time series data of a single product is expressed as $P(t,c)$, where t indicates time, which is the number of the *t*-th product, and c represents the cost.

Step 2: ILMD decomposition. The time data sequence obtained in the Step 1 is decomposed according to the method of improving the local mean decomposition. After the decomposition, the simulation result finally obtains 4 components ($PFi,(i=1,2,3,4)$ and 1 margin (represented by $R5$). The decomposition of the latter 25 sets of data is mainly employed for subsequent predictions.

Step 3: Train the BP neural network. For BP neural network, it must be trained to get the suitable network that is for prediction. Hereon, BP neural network is established for the 4 components and 1 margin obtained in the second step, and they are trained separately. The input layer of BP neural network has 6 nodes and the output layer has 1 node. The number of hidden layer nodes is determined during the experiment. The preceding six values of time series are enlisted as input to predict the following values, and so on. The training method and rules are as follows: the first group of data (each group of 6 values) is only used as the input of the second group of data, and the other groups of data are used as both the output of the previous set of data and the input of the next set of data. The parameters are optimized by PSO algorithm to improve its convergence speed and prediction accuracy.

Step 4: BP neural network prediction. The prediction process proposed model is shown in Fig. 5. The last 25 sets of data of each component decomposed by ILMD in step 2 is applied as the test set. When predicting the decomposition value of the *n*-th time of the *i*-th component, for the BP neural network, the input layer inputs the data of the six time points of $PFi_{n-6}, PFi_{n-5}, \cdots, PFi_{n-1}$. For example, to get the value at the 11th time point, the PF1 value of the nearest 6 points ($PF1_5, PF1_6, \cdots, PF1_{10}$) is employed as the input to obtain the predicted value.

Step 5: Find the predicted value. Finally, all the predicted values of the PF component and the margin are added together to obtain the final predicted value.

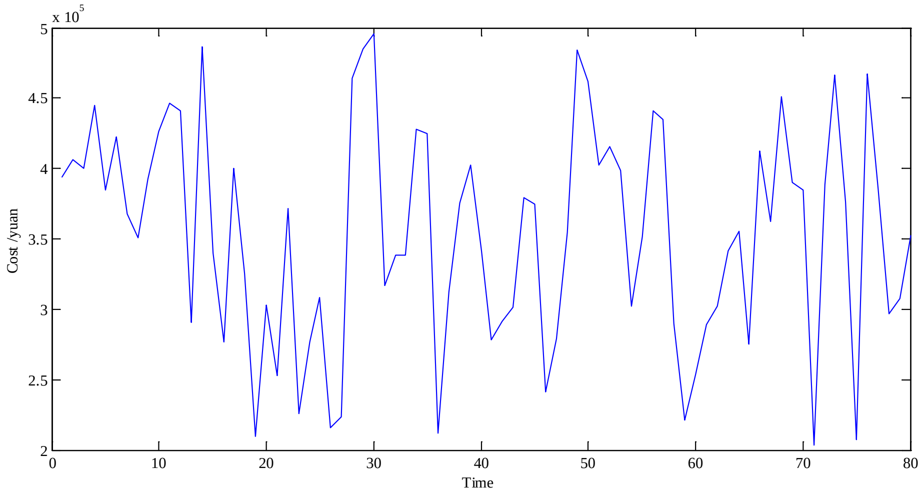


Fig. 4 Original data of remanufacturing cost of excavators

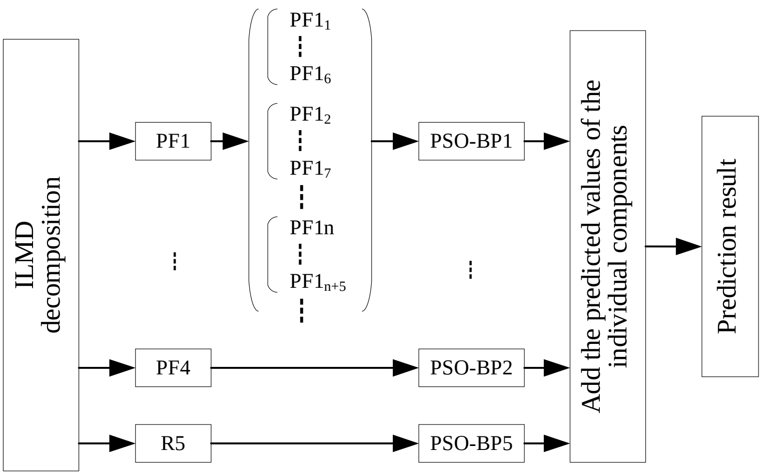


Fig. 5 Prediction process of model

3.1.2 Evaluation criteria

In order to more accurately and reasonably describe the predictive performance of various prediction methods, two indicators are selected here as evaluation criterion, i.e., mean square error (MSE) and mean absolute percent error (MAPE). The expressions of these two error indicators are:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(c_i - \vec{c}_i)^2 \tag{12}$$

$$MAPE = \frac{1}{N}\left(\sum_{i=1}^{N}\left|\frac{c_i - \vec{c}_i}{c_i}\right| \times 100\%\right) \tag{13}$$

where c_i is the actual cost value, $\vec{c}_i$ is the predicted cost value, N is the total number of predicted values.

3.1.3 Benchmark models

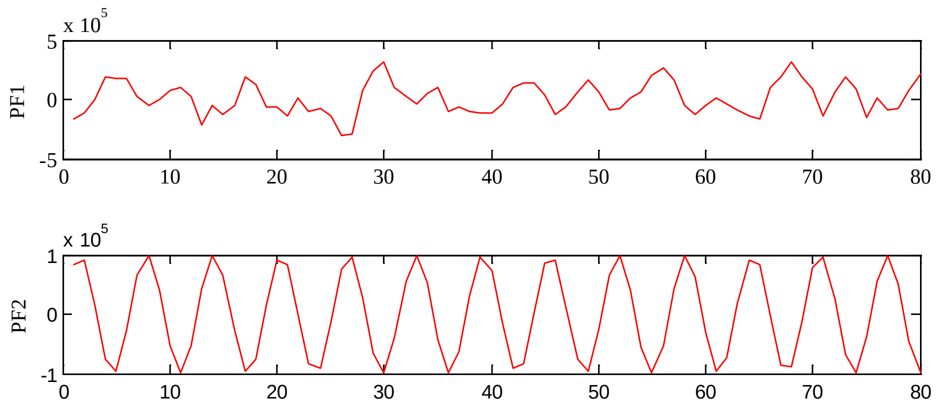
In this paper, some other prediction models are introduced, and some of the mainstream prediction techniques are utilized as benchmark models, which are compared with the proposed models, so as to verify the superiority of the proposed model. Accordingly, the most popular AI single model (such as SVM) and a similar decomposition-integration model are formulated.

The existing mainstream prediction techniques are roughly divided into two categories, i.e., traditional metrological models and artificial intelligence models. Therefore, in the field of artificial intelligence, a general BP neural network and a BP neural network based on Particle Swarm Optimization are selected as the benchmark model. For the decomposition-integration model, some typical decomposition methods such as EMD, EEMD, LMD, etc., combined with the same individual prediction and integrated prediction methods as the benchmark model.

Ultimately, there are seven benchmark models, including two popular AI single models (BP and PSO-BP), and five similar decomposition-integration models (i.e., EMD-PSO-BP, EEMD-PSO-BP, LMD-PSO-BP, ILMD-SVM, ILMD-BP), which are compared with proposed model (i.e., ILMD-PSO-BP). In these abbreviations, the first part and second part (the last two or one acronym) correspond to data decomposition and individual prediction techniques, respectively.

3.2 Prediction results

3.2.1 Data decomposition results



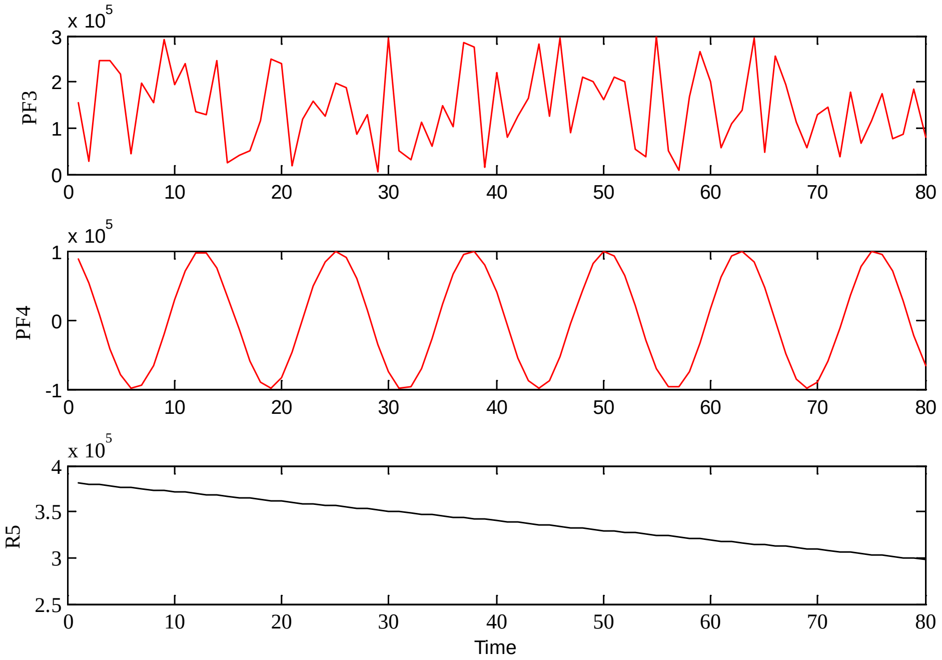


Fig. 6 Data decomposition results of remanufacturing cost via ILMD

The results of the original cost data decomposition are shown in Fig. 6. It can be clearly seen that sequence data of historical remanufacturing cost for the excavator is decomposed into five modes, i.e., four kinds of PFs and one margin (R5). As can be seen from Fig. 6, the original sequence is decomposed into several simple waveforms. Each waveform shows obvious regularity, and all kinds of prediction techniques are better for the prediction of time series with single or few composite trends. At the same time, appropriate prediction models can be selected according to the characteristics of each component. In this way, the accuracy of the whole prediction can be improved by the individual prediction of each component. Therefore, the PSO-BP model is used to predict the components, and then the final prediction results are obtained by adding.

3.2.2 Prediction results and comparative analysis

The results of the remanufacturing cost prediction based on the ILMD and PSO-BP neural network models are shown in Fig. 7. From Fig. 7, it is known that the predicted value of the remanufacturing cost time series is quite close to the actual value. To further illustrate effects of proposed modeling and prediction for remanufacturing cost, the prediction effects of the seven benchmark models presented in Section 3.1.3 are compared, as shown in Fig. 7.

As for prediction accuracy, Fig. 8 reports the MAPE and MSE criteria for the prediction of different models using the same data. Three main significant conclusions can be obtained from the figure. First of all, the performance of the proposed method is significantly better all benchmark models approach, including the corresponding single AI model or a similar decomposition-integration prediction method, with minimal MAPE and MSE values in all cases. Further results show that the proposed decomposition-integration prediction method can significantly improve the performance of the model.

Second, when the proposed model (i.e., ILMD-PSO-BP) is compared with tow single AI models (i.e., BP and PSO-BP), the performance of the former is significantly better than that of the single model. Especially in all cases, the MAPE and MSE values of the ILMD-PSO-BP model are much smaller than the MAPE and MSE values of BP (or PSO-BP). It can be considered that the decomposition-integration model can help to explore the inherent hidden relationship of the remanufacturing cost, thereby further improving the prediction accuracy.

Third, the decomposition-integration model using ILMD as a decomposition technique (i.e., ILMD-PSO-BP) is more accurate and effective than the other methods (i.e., EMD-PSO-BP, EEMD-PSO-BP, LMD-PSO-BP, ILMD-SVM, ILMD-BP) in most cases. This shows that in data analysis, ILMD is a more efficient decomposition method than its original form LMD and the common two technologies EMD/EEMD, because it successfully solves the problem of mode mixing.

Generally speaking, an important conclusion is that the proposed prediction model can statistically prove superior to all considered benchmark models in terms of remanufacturing cost prediction EOL products. Especially in various models, it can be seen from the comparison of results of MSE and MAPE that the proposed model achieves the highest prediction accuracy. In addition, the computational time of the proposed method is relatively short, further illustrating the efficiency of the model. The above results testify the feasibility of the proposed method in predicting performance, so corresponding decomposition-integration method can be utilized as an effective tool for remanufacturing cost analysis and prediction.

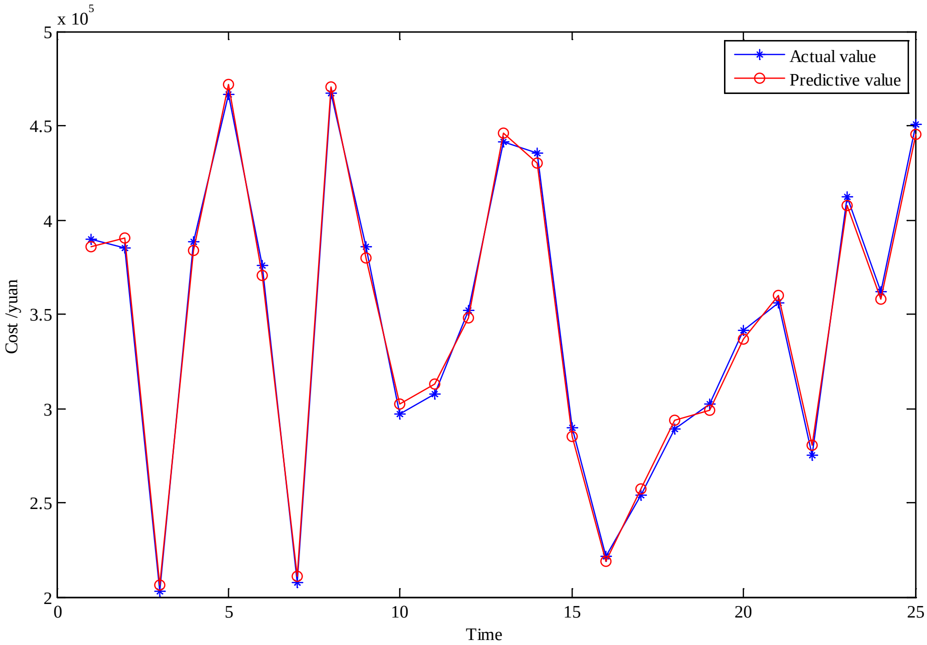


Fig. 7 The prediction results of the proposed decomposition-integration method

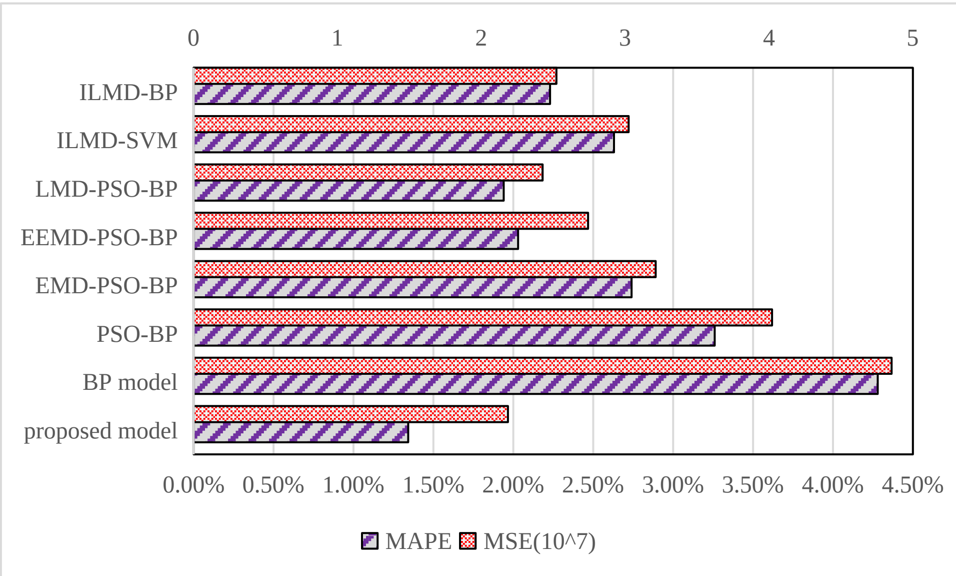


Fig. 8 Performance comparison of eight models for remanufacturing cost prediction

4 Conclusions

A data-driven based decomposition-integration method has been presented for predicting remanufacturing cost of EOL products. Actually, three main steps are involved in the model, i.e., data decomposition by ILMD to generate several new components, individual prediction for each decomposed component by an AI tool ((i.e., PSO-BP)), and integrated prediction via adding the predicted values of all components for obtaining final prediction results. In the proposed data-driven prediction model, all the decomposed components are deeply analyzed to reduce the computational complexity and improve the prediction accuracy.

In the case study, the proposed method is applied to the predict the remanufacturing cost of EOL excavators. Above all, the accuracy of the method is verified by comparing the predicted value with the actual value of samples, and furthermore, it is compared with seven benchmark models. The results show that the proposed data-driven based decomposition integration model is more advantageous than all the considered benchmark models in terms of improving prediction accuracy and reducing the computational cost. And it can be utilized as an effective tool to analyze and predict remanufacturing cost of EOL products.

Besides remanufacturing cost data, this method can be applied to other prediction tasks, such as power load forecasting, wind power forecasting, etc., particularly for data with small data volume, strong randomness, and unclear internal relationships. In addition, based on the interesting "data-driven modeling" idea, by formulating data decomposition, individual prediction and integrated prediction technology, other more powerful decomposition-integrated models can be developed accordingly. These issues will be resolved in subsequent studies.

Acknowledgments

The work described in this paper was supported by the National Natural Science Foundation of China (51675388); the Plateau Disciplines in Shanghai; 2018 Green manufacturing system integration project of Ministry of information industry; College Students Science and Technology Innovation Fund of Wuhan University of Science and Technology (18ZRB074); Wuhan University of Science and Technology Graduate Innovation and Entrepreneurship Fund (JCX201866). These financial contributions are gratefully acknowledged.

References

[1] Zhang. Y, Ren. S, Liu. Y, Si. S, A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products, Journal of Cleaner Production, 2016.

[2] Li. L, Li. C, Tang. Y, Du, Y, An integrated approach of reverse engineering aided remanufacturing process for worn components, Robotics & Computer Integrated Manufacturing. 48 (2017) 39-50.

[3] Ye Bowen, The research of application big Data analytics to the parts remanufacturing economic evaluation, Huazhong University of Science and Technology, 2014.

[4] Ijomah. W. L, Mcmahon. C. A, Hammond. G. P, Newman. S. T, Development of design for remanufacturing guidelines to support sustainable manufacturing, Robotics & Computer Integrated Manufacturing. 23 (6) (2007) 712-71916.

[5] Jiang. Z, Jiang. Y, Wang. Y, Zhang. H, Cao. H, Tian, G, A hybrid approach of rough set and case-based reasoning to remanufacturing process planning, Journal of Intelligent Manufacturing. 30(1)(2019) 19-32.

[6] Zhang X, Remanufacturing Scheme Decision Model and Application Based on Remaining Useful Life Estimation, Journal of Mechanical Engineering. 49 (7) (2013) 51.

[7] Du. Y, Cao. H, Chen. X. Wang. B, Reuse-oriented redesign method of used products based on axiomatic design theory and QFD, Journal of Cleaner Production. 39 (1) (2013) 79-86.

[8] Ghazalli. Z, Murata. A, Development of an AHP-CBR evaluation system for remanufacturing: end-of-life selection strategy, International Journal of Sustainable Engineering. 4 (1) (2011) 14.

[9] Lee. S. G, Lye. S. W, Khoo. M. K, A multi-objective methodology for evaluating product end-of-life options and disassembly, International Journal of Advanced Manufacturing Technology. 18 (2) (2001) 148-156.

[10] Ma. J, Gül E. Okudan Kremer, A sustainable modular product design approach with key components and uncertain end-of-life strategy consideration, International Journal of Advanced Manufacturing Technology. 85(1-4) (2016) 741-763.

[11] Liyan Shen, Reproducibility Evaluation of Mechanical and Electrical Products, Shan Dong University, 2008.

[12] LIU Zhi-feng, BI. Dao-kun, LIU Tao, Estimating for remanufacturing cost of mechanical and electrical products based on linear regression, Machinery Design＆Manufacture. 9(2011) 258-260.

[13] Xugang Zhang, Study on Remanufacturability Evaluation and Remanufacturing Process Decision Methods of Machine Tool, Wuhan University of Science and Technology, 2014.

[14] Goodall. P, Graham. I, Harding. J, Conway. P, Schleyer. S, West. A, Cost estimation for remanufacture with limited and uncertain information using case based reasoning, Journal of Remanufacturing. 5(1) (2015) 7.

[15] Sabharwal. S, Garg. S, Determining cost effectiveness index of remanufacturing: a graph theoretic approach, International Journal of Production Economics. 144 (2) (2013) 521-532.

[16] Sang Fan, Li Enzhong, Zheng Handong, Peijing Shi, Prediction of Remanufacturing Cost Combination of Heavy-Duty Engine Based on State, Surface Engineering and Remanufacturing. 3(2017) 18-22.

[17] Qin P P, Yun S, Shu J W, Estimation of the remanufacturing cost of engineering machinery hydraulic cylinder, 3d International Conference on Advanced Information and Communication Technology for Education. 2015.

[18] XIANG Hong, XIA Xu-hui, WANG Lei, JIANG Zhi-gang, A Prediction Method for Remanufacturing Cost of Mechanical and Electrical Products Based on Support Vector Machine, Machinery Design＆Manufacture, 4(2016) 268-272.

[19] ZHANG Xugang, AO Xiuyi, ZhANG Hua, Research on Remanufacturing Cost Prediction Model of Used Parts Considering Failure features, China Mechanical Engineering. 29(21) (2018) 2594-2599.

[20] Gao Ting, Feasibility analysis of hydraulic cylinder remanufacturing, Wuhan University of Science and Technology, 2013.

[21] Liu. H, Tian. H. Q, Pan. D. F, Li. Y. F, Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks, Applied Energy. 107(2013) 191-208.

[22] Yu. L, Wang. Z, Tang. L, A decomposition-ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting, Applied Energy. 156(2015) 251-267.

[23] Zhang. W, Liu. F, Zheng. X, Li. Y, A hybrid EMD-SVM based short-term wind power forecasting model, IEEE Pes Asia-pacific Power & Energy Engineering Conference, 2016.

[24] Qiu. X, Suganthan. P. N, Amaratunga. G. A. J, Short-term electricity price forecasting with empirical mode decomposition based ensemble kernel machines, Procedia Computer Science. 108(2017) 1308-1317.

[25] Yanchao. L. U, Weining. W, Biao. Z, Yan. Z, Wire price forecasting of grid project based on eemd and arma, Electric Power Construction. 34(11) (2013) 39-44.

[26] Huifang Qu, Guoqiang Tang, Qiying Lao, Oil Price Forecasting Based on EMD and BP_AdaBoost Neural Network, Open Journal of Statistics. 8(2018) 660-669.

[27] Yu. L, Wang. S, Lai. K. K, Forecasting crude oil price with an emd-based neural network ensemble learning paradigm, Energy Economics. 30(5) (2008) 2623-2635.

[28] Tang. L, Dai. W, Yu. L, Wang. S, A novel ceemd-based eelm ensemble learning paradigm for crude oil price forecasting, International Journal of Information Technology & Decision Making. 14(01) (2015) 141-169.

[29] Xiaozhen L, Han Q, Shouyang W, An integrated forecasting model for air passenger traffic in China based on singular spectrum analysis, Systems Engineering-Theory & Practice. 37(6) (2017) 1479-1488.

[30] Huang. N. E, Shen. Z, Long. S. R, Wu. M. C, Shih. H. H, Zheng. Q, The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings A. 454(1971) (1998) 903-995.

[31] Wu Z, Ensemble empirical mode decomposition : a noise assisted data analysis method, Adv Adapt Data Anal. 1(2009).

[32] Smith J S, The local mean decomposition and its application to EEG perception data, Journal of the Royal Society Interface. 2(5) (2005) 443-454.

[33] Song. H, Huang. C, Liu. H, Chen. T, Li. J, Luo. Y, A new power quality disturbance detection method based on the improved lmd, Proceedings of the Csee. 34(10) (2014) 1700-1708.

[34] Changguo Lv, Research on Grid Resource Scheduling Based On BP Algorithm, Harbin University of Science and Technology, 2007.

[35] Ying Yuan, Study on the Application of Particle Swarm Optimization in System Power Economic Dispatch, Xi'an University of Technology, 2010.