International Journal of Fluid Machinery and Systems
Vol. 5, No. 4, October-December 2012
DOI: http://dx.doi.org/ 10.5293/IJFMS.2012.5.4.143
ISSN (Online): 1882-9554

Original Paper

Machine Condition Prognostics Based on Grey Model and Survival Probability

Stenly Tangkuman[1,2], Bo-Suk Yang[2] and Seon-Jin Kim[2]

[1]Department of Mechanical Engineering, Sam Ratulangi University,
Kampus Unsrat, Bahu, Manado 95115, Indonesia, stenlytangkuman@yahoo.com
[2]Department of Mechanical & Automotive Engineering, Pukyong National University,
San 100, Yongdang-dong, Nam-gu, Busan 608-739, Korea, bsyang@pknu.ac.kr, sjkim@pknu.ac.kr

Abstract

Predicting the future condition of machine and assessing the remaining useful life are the center of prognostics. This paper contributes a new prognostic method based on grey model and survival probability. The first step of the method is building a normal condition model then determining the error indicator. In the second step, the survival probability value is obtained based on the error indicator. Finally, grey model coupled with one-step-ahead forecasting technique are employed in the last step. This work has developed a modified grey model in order to improve the accuracy of prediction. For evaluating the proposed method, real trending data of low methane compressor acquired from condition monitoring routine were employed.

Keywords: Prognostics, Grey model, Survival probability, Condition monitoring, Maintenance.

1. Introduction

Nowadays, manufacturing industries are highly expected to continuously reduce maintenance costs, operation downtime, and safety hazards. More efficient maintenance strategies such as condition-based maintenance (CBM) are being implemented to handle the situation. The important factors in CBM are predicting the future condition of machine and assessing the remaining useful life (RUL) [1].

It is critically necessary to assess the RUL of a machine or an asset while in use since it has impacts on the planning of maintenance activities [2]. Condition monitoring (CM) can provide information on current working age and state of the system that may affect its future life. This information then can be used for prediction of future condition and remaining useful life of the system or components and planning of maintenance activities. Thus, RUL prediction deals with prognostics of machine health condition based on measured data from CM [3].

The existing prognostics methods can be classified into three categories, which are physics model-based prognostics, knowledge-based prognostics, and data driven prognostics [4]. Physic model-based implements a mathematic formulation, e.g. Paris law to predict defect propagation of machine components. This technique requires less data than data-driven technique and can be highly accurate if physic model remain consistent across the system. The demerit of physic model-based is difficult to approach the real-life system due to complex and stochastic. Knowledge-based prognostics solves problem by mimicking how human expert makes decisions. However, the disadvantages of this approach are hard to obtain domain knowledge and hard to convert domain knowledge to rules. Data driven prognostics is based upon statistical and learning techniques, most of which originated from the theory of pattern recognition.

Widely used data driven methodologies are Artificial Neural Network (ANN), Bayesian-related method, hidden Markov models (HMM) and hidden semi-Markov models (HSMM), hazard rate (HR), and proportional HR. Byington et al. have developed a neural network methodology for remaining life prediction of aircraft actuator components [5]. Gebraeel et al. developed Bayesian updating methods that use real-time condition monitoring information to develop a closed-form residual-life distribution for the monitored device [6]. Huang and Xi deal with a new scheme for the prediction of a ball bearing's remaining useful life based on self-organizing map (SOM) and back propagation neural network methods [7]. Gebraeel and Lawley used a neural network-based degradation model to compute and continuously update residual life distributions of partially degradated components [8]. Baruah and Chinnam employed HMMs for modeling sensor signals emanating from the machine, and in turn, identify the health state of the cutting tool as well as facilitate estimation of remaining useful life [9]. Liao et al. presented the proportional hazards model and logistic regression model, which relates the multiple degradation features of sensor signals to the specific reliability indices of the unit, and enable us to predict its RUL [10]. However, disadvantages of the above methodologies

Received March 26 2012; revised July 2 2012; accepted for publication August 26 2012: Review conducted by Prof. Hideaki Tamaki. (Paper number O12005J)
Corresponding author: Stenly Tangkuman, Dr. Eng, stenlytangkuman@yahoo.com, stenly@pknu.ac.kr

are hard to fit domain knowledge to ANN, need model retraining if operating conditions change, need a lot of historical state transition and fatal data, and the model is complicated.

Among the data driven approaches, the grey model which have been successfully used in other domains before, has been introduced to deal with prognostics. Ku and Huang explored the application of grey model for predicting and monitoring production processes [11]. Subsequently, Gu et al. successfully developed grey prediction model in the failure prognostics for electronics [12].

This new theory, grey model, was originally proposed by Deng [13]. The advantages from the use of grey model are it can be applied to circumstances with the minimum data down to some observations, and utilizes a first-order differential equation to characterize a system [14]. On the other hand, nowadays it is desirable to develop a RUL estimation model based on very few data situations [15]. Grey model can be a promising model to respond this current challenge. So far, rare application of grey prediction model has been reported in machine prognostics or RUL prediction.

This paper presents a new prognostics method based on grey model and survival probability. The proposed method consists of three steps. The first step of the method is building a normal condition model then determining the error indicator. In the second step, the survival probability value is obtained based on the error indicator. Finally, grey model coupled with one-step-ahead forecasting technique is employed in the last step.

The remaining parts of this paper are organized as follows. In Section 2, we describe the basic theory of the survival function and grey model. The proposed system is analyzed in Section 3. In Section 4, Real trending data of low methane compressor acquired from condition monitoring routine are employed for evaluating the proposed method. Finally, the conclusions are presented in Section 5.

2. Theoretical Background

2.1 Survival Function

Survival function, denoted by $S(t)$, is defined as the probability that an individual or component survives up to time t. Let us first consider the simple case where all the components are observed to fail so that the survival times are exact and known. Let t_1, t_2, …, t_n be the exact survival times of the n components under study. So, the survival function can be estimated as [16]

$$S(t) = \frac{N - D_t}{N} \tag{1}$$

where N represents the total number of components, and D_t is the number of components failing until time t. On the other hand, the component or system does not undergo failure yet, the survival time is not also known yet. In this case, the survival function can be estimated by using the following equation.

$$S(t) = \exp[-(c\, e(t))] \tag{2}$$

Where c is a constant value, and $e(t)$ is called as the error indicator that can be obtained as

$$e(t) = \sqrt{\frac{\sum_{i=1}^{n}(d_i)^2}{n}} \tag{3}$$

where n represents the total number of data, and d is the deviation value between the limits of normal condition model and the data acquired from CM.

2.2 Grey model

2.2.1 GM(1,1)

The grey forecasting model uses the operations of accumulated generation to construct differential equations. Intrinsically speaking, it has the characteristics of requiring less data. The grey model GM(1,1), i.e., a single variable first-order grey model, is summarized as follows [11,14]:

Step 1: For an initial time sequence,

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2),\ldots, x^{(0)}(i),\ldots, x^{(0)}(n)\} \tag{4}$$

where $x^{(0)}(i)$ denotes the time series data at time ith.

Step 2: On the basis of the initial sequence $x^{(0)}$, a new sequence $x^{(1)}$ is set up through the accumulated generating operation in order to provide the middle message of building a model and to weaken the variation tendency, i.e.

$$x^{(1)} = \{x^{(1)}(1), x^{(1)}(2),\ldots, x^{(1)}(i),\ldots, x^{(1)}(n)\} \tag{5}$$

where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i) \qquad k = 1, 2, \ldots, n \tag{6}$$

Step 3: The first-order differential equation of grey model GM(1,1) is then the following

$$\frac{\mathrm{d}x^{(1)}}{dt} + ax^{(1)} = b \tag{7}$$

and its difference equation is

$$x^{(0)}(k) + aZ^{(1)}(k) = b \qquad k = 2, 3, \ldots, n \tag{8}$$

also take

$$Z^{(1)}(k+1) = ½\,(x^{(1)}(k) + x^{(1)}(k+1)) \qquad k = 1, 2,\ldots, (n–1) \tag{9}$$

where $Z^{(1)}(k+1)$ is the $(k+1)$th background value.

and from Eq. (7), it is easy to get

$$\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(1)}(n) & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \tag{10}$$

where a and b are the coefficients to be identified, then let

$$Y_n = [x^{(0)}(2), x^{(0)}(3), \ldots , x^{(0)}(n)]^{\mathrm{T}} \tag{11}$$

$$B = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(1)}(n) & 1 \end{bmatrix} \tag{12}$$

where Y_n and B are the constant vector and the accumulated matrix respectively, and

$$A = [a, b]^{\mathrm{T}} \tag{13}$$

Applying ordinary least-square method to Eq. (10) on the basis of Eqs. (11) – (13), coefficient A becomes

$$A = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}Y_n \tag{14}$$

Step 4: Substituting A in Eq. (7) with Eq. (14), the approximate equation becomes the following

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) – b/a)\, \mathrm{e}^{–ak} + b/a \tag{15}$$

where $\hat{x}^{(1)}(k+1)$ is the predicted value of $x^{(1)}(k+1)$ at time $(k+1)$. After the completion of an inverse accumulated generating operation on Eq. (15), $\hat{x}^{(0)}(k+1)$, the predicted value of $x^{(0)}(k+1)$ at time $(k+1)$ becomes available and therefore,

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \tag{16}$$

2.2.2 Modification of GM(1,1)

The above method can be modified in attempt to improve the accuracy of prediction. It can be done by the following procedure:

1. Adding the values of sequence in Eq. (15) with m value. m is the number of modified prediction value, and it is an arbitrary value. In this work, m of 5 to m of 11 was employed as presented in Table 2. So, Eq. (15) for $k = n$ becomes

$$\hat{x}^{(1)}(n+i) = (x^{(0)}(1) – b/a)\, \mathrm{e}^{–a(n+i\,-1)} + b/a \qquad i = 1,2,\ldots,m \tag{17}$$

2. Consequently, Eq. (17) can increase the predicted value to the number of m in Eq. (18.a). This equation was derived from Eq. (16). Eventually, the final result of the predicted value can be determined by taking the mean value of m predicted values as depicted in Eq. (18.b).

$$\hat{x}^{(0)}(n+i) = \hat{x}^{(1)}(n+i) - \hat{x}^{(1)}(n+i\,-1) \qquad i = 1,2,\ldots,m \tag{18.a}$$

$$\text{The predicted value of } x^{(0)}(n+1) = (1/m) \times [\hat{x}^{(0)}(n+1) + \hat{x}^{(0)}(n+2) + \ldots + \hat{x}^{(0)}(n+\mathrm{m})] \tag{18.b}$$

3. Methodology

The prognostics system must be able to predict the future performance condition of machine or equipment based on past and current condition. CM data can provide information on current working age and state of the system that may affect its future life. This information then can be used for prediction of remaining useful life of the system or components. This research introduces a method for predicting the future condition and assessing the RUL based on survival probability and grey model.

Fig. 1 depicts the proposed method in this work. It is summarized as follows:

Step 1: Modeling and Calculating. Build a model based on normal condition data of CM then calculate the error indicator by using Eq.(3). It is important to identify whether the system or machine is in good condition or not.

Step 2: Estimating. In order to express current performance condition of system or machine, survival function is employed. By using error indicator and Eq.(2), the survival probability can be estimated.

Step 3:Modeling, Predicting and Assessing. In this step, first build the grey model based on past and current state of survival probability then using grey model coupled with one-step-ahead technique to predict the future value of survival probability. Finally, assessing the RUL based on incipient and final failure threshold determination in survival probability prediction.

In order to evaluate the predicting performance, the root-mean square error (*RMSE*) and linear correlation (*R*) are utilized as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\tilde{y}_i)^2}{n}} \tag{19}$$

where *n* represents the total number of data points in the test set, y_i is the actual value, and $\tilde{y}_i$ represents the predicted value of the model.

$$R = \frac{Cov(y,\tilde{y})}{\sigma_y \sigma_{\tilde{y}}}, \quad Cov(y,\tilde{y}) = \frac{1}{N}\sum_{i=1}^{N}(y_i-\bar{y})(\tilde{y}_i-\bar{\tilde{y}}) \tag{20}$$

where $Cov(y,\tilde{y})$ is covariance between actual and predicted value. $\bar{y}$ is the mean of actual value and $\bar{\tilde{y}}$ is the mean of predicted value. The standard deviation of the actual and predicted values, σ_y and $\sigma_{\tilde{y}}$, respectively, can be calculated as

$$\sigma_y = \left[\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{y})^2\right]^{1/2}, \sigma_{\tilde{y}} = \left[\frac{1}{N-1}\sum_{i=1}^{N}(\tilde{y}_i-\bar{\tilde{y}})^2\right]^{1/2} \tag{21}$$

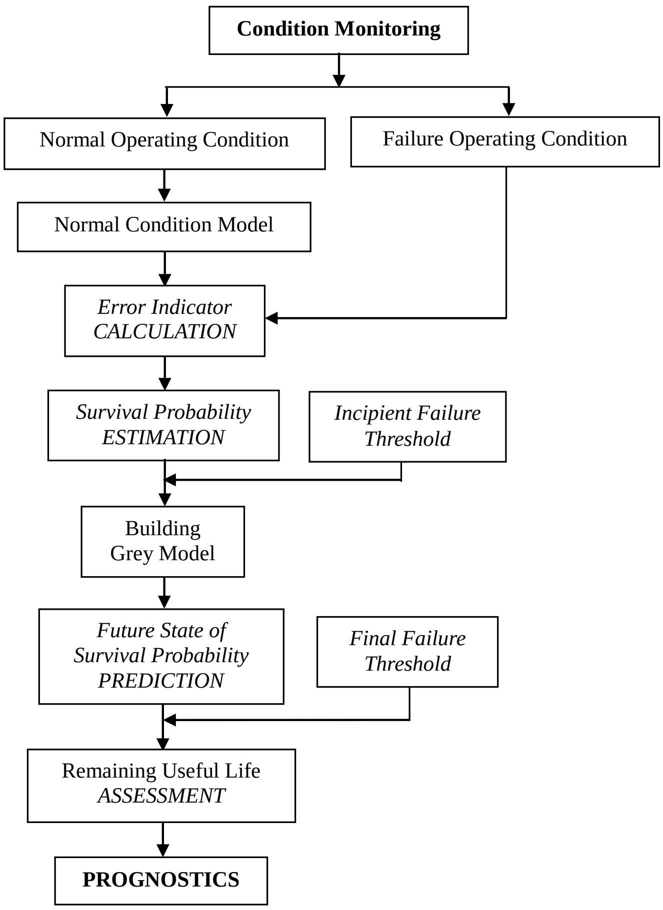


Fig. 1 The proposed method

4. Application and Result

4.1 Condition monitoring data

The data applied in this study contains information of machine history with respect to time sequence (vibration amplitude) as shown in Fig. 3 [17]. It is peak acceleration trending data of low methane compressor shown in Fig. 2. Other information of the system is summarized in Table 1.

The average sampling rate was 6 hr during the data acquisition process. The machine is in normal condition during the time correlated with the first 291 points. After that time, degradation condition has occurred. These faults were the damages of the main bearings of the compressor. The final failure occurred at time of 308, after that some maintenance actions were applied.

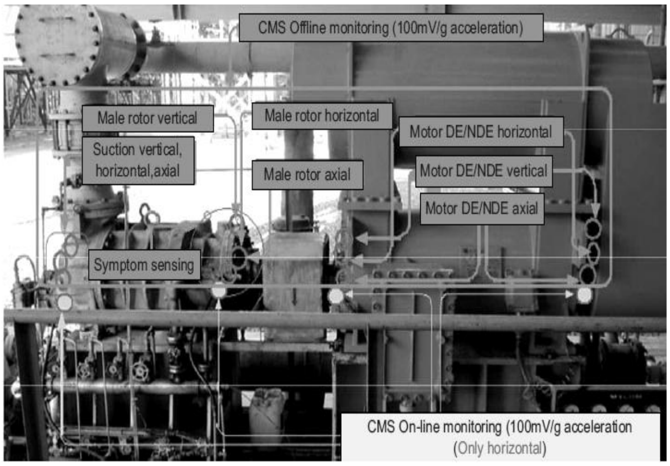


Fig. 2 Low methane compressor

Table 1 Description of system

Electric motor		Compressor	
Voltage	6600 V	Type	Wet screw
Power	440 kW	Lobe	Male rotor(4 lobes)
Pole	2 Pole		Female rotor(6 lobes)
Bearing	NDE: #6216, DE: #6216	Bearing	Thrust: 7321 BDB
RPM	3565 rpm		Radial : Sleeve type

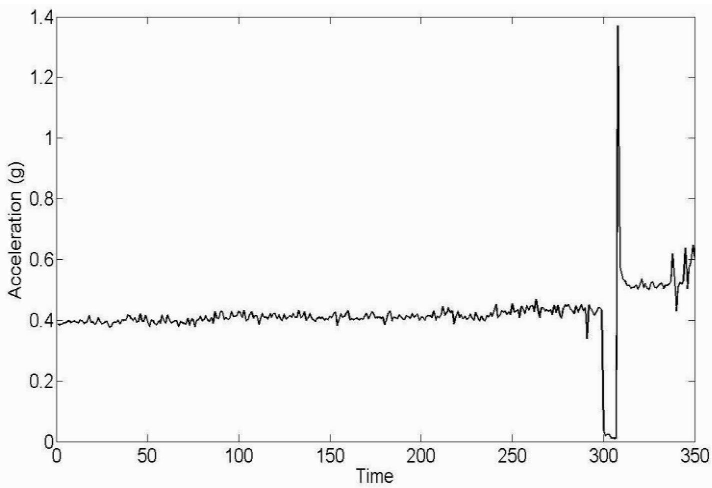
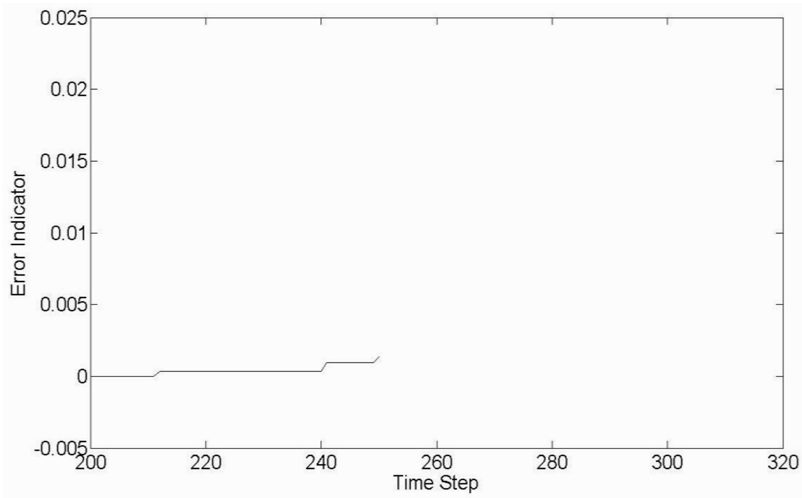


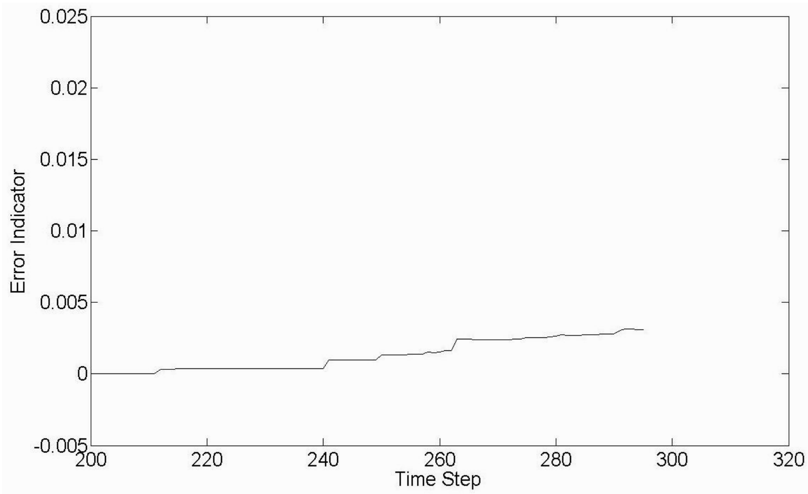
Fig. 3 The peak acceleration data of low methane compressor

4.2 Result

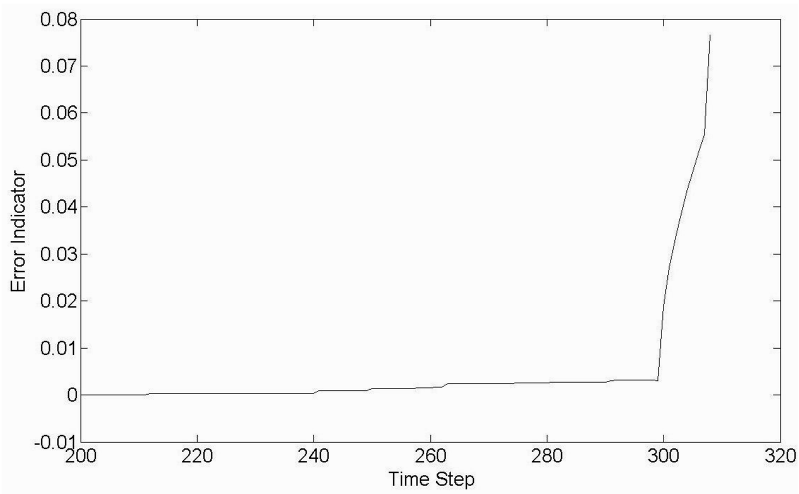
The first step of the proposed method is to build a normal condition model and calculate the error indicator. In this step, the model was built based on the first 200 data, from time of 1 to 200, and the limits of normal condition model are taken at acceleration values of 0.4 ± 0.039. The acceleration value related to each time is analyzed whether inside of the limits or not. If it is not inside of the limits, its deviation value must be obtained. Subsequently, the error indicator can be obtained by using Eq. (3) as shown in Fig. 4.



(a)



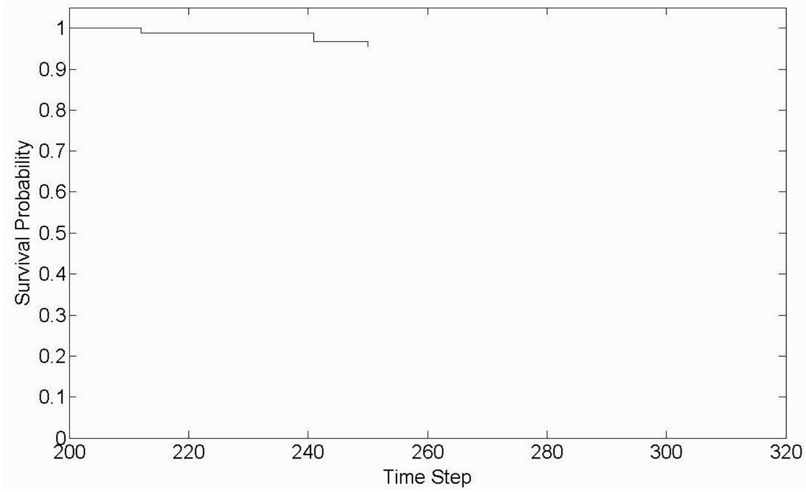
(b)



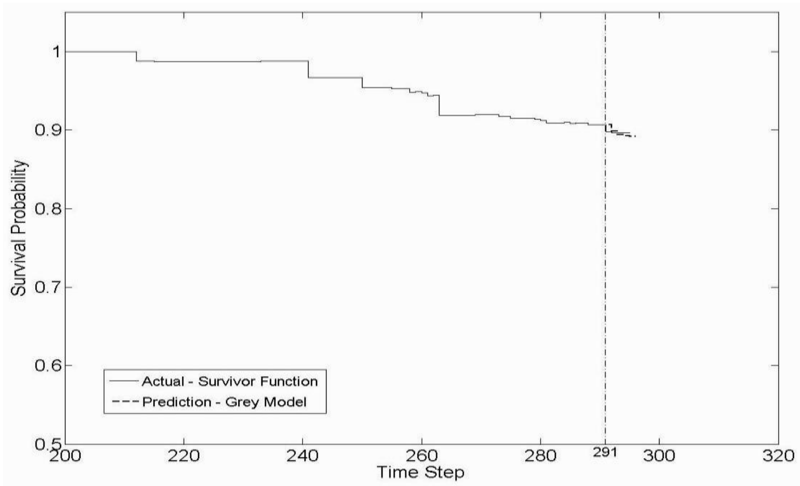
(c)

Fig. 4 Error indicator; (a) from time of 200 to 250 (b) from time of 200 to 295, and (c) from time of 200 to 308

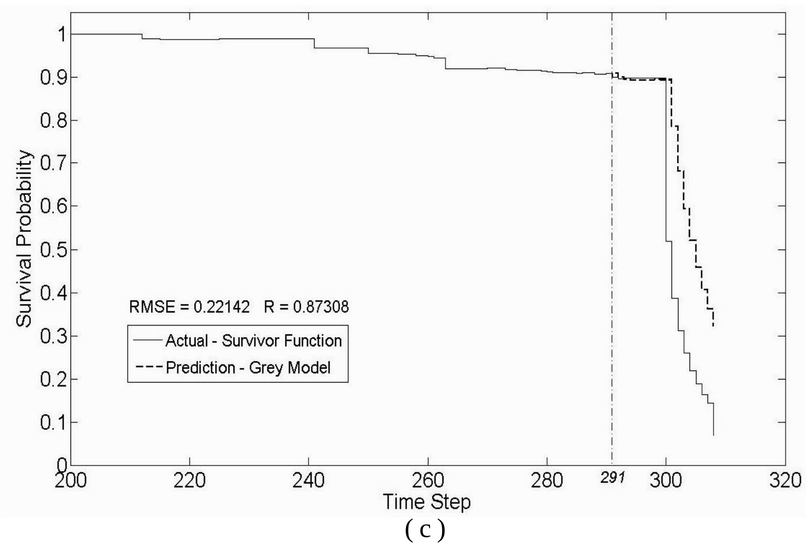
The second step of this prognostics method is to estimate the survival probability of the machine. In the beginning of the working condition of a machine the survival probability value is equal to one, which means the machine is in good condition or has no faults. After that, according to time the survival probability value will be less than one and toward to zero. It means machine is in a degradation state or there are some faults in the machine. Based on the error indicator from previous step, the survival probability can be estimated by using Eq. (2) which c of 35 is taken. Eventually, Fig. 5 depicts the survival probability of the machine.



(a)



(b)



(c)

Fig. 5 Survival probability; (a) from time of 200 to 250, (b) from time of 200 to 295, and (c) from 200 to 308

The third step of this new method is to build a grey model; predict the future value of survival probability; and assess the RUL of the system. When the survival probability decreases and then reaches the incipient failure threshold, the future values of survival probability must predict immediately. Thus, the RUL can be known before the final failure occurs.

Figure 5c shows the survival probability from time of 200 to 308. Especially, from time of 292 to 308, it shows both actual and prediction values. This prediction result indicates that the model which *RMSE* of 0.22142 and *R* of 0.87308 cannot reach well the actual values. The final failure, based on the real condition, occurs at time of 308 for survival probability of 0.1. Whereas the final failure based on the prediction, does not occur at time of 308 because its value of survival probability is more than 0.1.

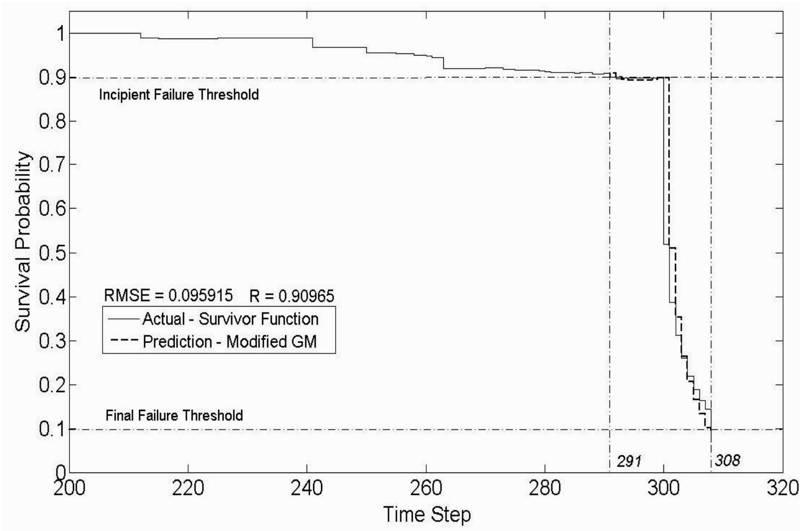


Fig. 6 Survival probability after modification

Table 2 Result of *RUL* prediction by using modified grey model for *m* of 5 to 11

m	Time of Final Failure prediction	*RUL* prediction (hour)	Accuracy of *RUL* prediction (%)
5	309	108	94.1
6	309	108	94.1
7	308	102	100
8	308	102	100
9	308	102	100
10	307	96	94.1
11	307	96	94.1

The final failure problem is very important for RUL assessment. Therefore, modifying the grey model is needed to increase the accuracy of prediction. Modification procedures as explained in section 2.2.2 are applied to the previous model. Finally, the accuracy of the modified grey model is very satisfactory, because the modified model can anticipate the sudden decrease of survival probability. Figure 6 depicts the prediction which using modified grey model with *m* of 9. Moreover, the modified grey model with a small error of 0.095915 and close correlation of 0.90965 is able to predict the final failure occurring at time of 308. It is in accordance with the degradation state based on real condition.

Based on the survival probability predictions, RUL of the machine can be assessed. Specifically, the RUL is the time interval between incipient failure threshold and final failure threshold. This study applies the incipient and final failure threshold are at survival probability of 0.9 and 0.1 respectively, and the incipient failure occurs at time of 291. The RUL and accuracy can be assessed using the followings:

$$\text{RUL} = [(\text{final failure time}) - (\text{incipient failure time})]\text{x 6 hrs} \tag{22}$$

$$\text{Accuracy} = \left(1 - \frac{|\text{actual time} - \text{predicted time}|}{\text{predicted time}}\right)\text{x }100\% \tag{23}$$

Based on the above formula, the actual RUL is 102 hours ((308-291)x 6 = 102). Table 2 describes the RUL prediction and the accuracy of prediction which using *m* of 5 to 11.

5. Conclusion

This paper presents a new method for machine condition prognostics based on grey model and survivor probability. Performance conditions of a machine were expressed by survival probability values acquired from the error indicator. The RUL can be clearly observed by obtaining the time deviation between incipient failure time and final failure time. This proposed method also presents that a modified grey model has been made, in order to improve the accuracy of prediction. The proposed method has been validated by predicting the RUL of a low methane compressor. However, this work has only been able to predict the RUL which the incipient and final failure threshold are acquired. Hence, a more in-depth investigation into the incipient and final failure threshold is needed in future studies.

Nomenclature

e	Error indicator	S	Survival probability
m	The number of modified prediction value	t	Time
R	Linear correlation	$x^{(0)}$	Initial time sequence
$RMSE$	Root-mean square error	$\hat{x}^{(0)}(k+1)$	Predicted value of $x^{(0)}(k+1)$ at time $(k+1)$
σ_y	Standard deviation		

References

[1] Yang, B. S., and Widodo, A., 2010, "Introduction of Intelligent Machine Fault Diagnosis and Prognosis," Nova Science Publishers, New York, USA.

[2] Jardine, A. K. S., Lin, D., and Banjevic, D., 2006, "A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance," Mechanical Systems and Signal Processing, Vol. 20, pp. 1483-1510.

[3] Widodo, A., and Yang, B. S., 2011, "Machine Health Prognostics using Survival Probability and Support Vector Machine," Expert System with Application, Vol. 38, pp. 8430-8437.

[4] Peng, Y., Dong, M., and Zuo, M. J., 2010, "Current Status of Machine Prognostics in Condition-based Maintenance: a Review," International Journal of Advanced Manufacturing Technology, Vol. 50, pp. 297-313.

[5] Byington, C. S., Watson, M., and Edwards, D., 2004, "Data-Driven Neural Network Methodology to Remaining Life Predictions for Aircraft Actuator Components," Proceedings of the IEEE Aerospace Conference, Vol. 6, pp. 3581-3589.

[6] Gebraeel, N. Z., Lawley, M. A., Li, R., and Ryan, J. K., 2005, "Residual-Life Distributions from Component Degradation Signals: a Bayesian approach," IIE Transactions, Vol. 37, pp. 543-557.

[7] Huang, R. Q., and Xi, L. F., 2007, "Residual Life Predictions for Ball Bearing Based On Neural Networks," Chinese Journal of Mechanical Engineering, Vol. 43, No. 10, pp. 137-143.

[8] Gebraeel, N. Z., and Lawley, M. A., 2008, "A Neural Network Degradation Model for Computing and Updating Residual Life Distributions," IEEE Transactions on Automation Science and Engineering, Vol. 5, No. 1, pp. 387-401.

[9] Baruah, P., and Chinnam, R. B., 2005, "HMMs for Diagnostics and Prognostics in Machining Processes," International Journal of Production Research, Vol. 43, No. 6, pp. 1275- 1293.

[10] Liao, H., Zhao, W., and Guo, H., 2006, "Predicting Remaining Useful Life of an Individual Unit Using Proportional Hazards Model and Logistic Regression Model," IEEE 1-4244-0008-2/06, pp. 127-132.

[11] Ku, L. L., and Huang, T. C., 2006, "Sequential Monitoring of Manufacturing Processes: an Application of Grey Forecasting Models," International Journal of Advanced Manufacturing Technology, Vol. 27, No. 5-6, pp. 543-546.

[12] Gu, J., Vichare, N., Ayyub, B., and Pecht, M., 2010, "Application of Grey Prediction Model for Failure Prognostics of Electronics," International Journal of Performability Engineering, Vol. 6, No. 5, pp. 435-442.

[13] Deng, J. L., 1989, "Introduction to Grey System Theory," Journal of Grey System, Vol. 1, No. 1, pp. 1-24.

[14] Mao M., and Chirwa, E. C., 2006, "Application of Grey Model GM(1,1) to Vehicle Fatality Risk Estimation," Technological Forecasting and Social Change, Vol. 73, pp. 588-605.

[15] Si, X. S., Wang, W., Hu, C. H., and Zhou, D. H., 2011, "Remaining Useful Life Estimation-A Review on the Statistical Data Driven Approaches," European Journal of Operational Research, Vol. 213, pp. 1-14.

[16] Lee, E. T., and Wang, J. W., 2003, "Statistical Methods for Survival Data Analysis, 3rd Ed.," John Wiley & Sons, Inc, New Jersey, USA.

[17] Tran, V. T., Yang, B. S., Oh, M. S., and Tan, A. C. C., 2008, "Machine Condition Prognosis Based on Regression Trees and One-step-ahead Prediction," Mechanical Systems and Signal Processing, Vol. 22, pp. 1179-1193.