MACHINE PERFORMANCE DEGRADATION ASSESSMENT AND REMAINING USEFUL LIFE PREDICTION USING PROPORTIONAL HAZARD MODEL AND SVM

Van Tung Tran[a,b], Hong Thom Pham[a], Bo Suk Yang[a], Tan Tien Nguyen[b]

[a]Department of Mechanical & Auotmotive Engineering, Pukyong National University, San 100, Yongdang-dong, Nam-gu, Busan 608-739, South Korea.

[b]Faculty of Mechanical Engineering, Hochiminh City University of Technology, 268 Ly Thuong Kiet St., Dist. 10, Ho Chi Minh City, Vietnam.

This paper proposes a three-stage method involved system identification techniques, proportional hazard model, and support vector machine for assessing the machine health degradation and forecasting the machine remaining useful life (RUL). In the first stage, only the normal operating condition of machine is used to create identification model to mimic the dynamic system behaviour. The machine degradation is indicated by degradation index which is the root mean square of residual errors. These errors are the difference between identification model and behaviour of system. In the second stage, the Cox's proportional hazard model is generated to estimate the survival function of the system. Finally, support vector machine, one of the remarkable machine learning techniques, in association with direct prediction method of time-series techniques is utilized to forecast the RUL. The data of low methane compressor acquired from condition monitoring routine are used for appraising the proposed method. The results indicate that the proposed method could be used as a potential tool to machine prognostics.

Keywords: **Prognostics, Performance degradation, Remaining useful life, Proportional hazard model, Support vector machine**.

1. INTRODUCTION

Machinery operation generally consists of a series of states which are transient, normal operating, and degrading before fault occurrences. Amongst these, even degrading state plays an important role in maintenance activity, it is not a root cause of the machine breakdown. However, it does decline the performance reliability and increase the potential for faults and failures. Once the failure degradation has been detected, the remaining lifetime of machinery needs to be predicted in order that system operators implement the timely maintenance actions to avoid the catastrophic failures.

In order to estimate the remaining lifetime or the remaining useful life (RUL), condition based maintenance (CBM), which is one of the efficient maintenance strategies, has been recently received considerable attention in modern industries. In CBM, prognostics is of necessity and gradually becomes a key component. It provides capability to foretell the remaining operation life, future machine state, or probability of reliable machine operation based on the data acquired from condition monitoring process [1]. Additionally, more advanced prognostics are being focused on performance degradation monitoring and assessment so that failure can be predicted and prevented. According to [2], three crucial steps are necessary to fulfill the goal of prognostics. Firstly, the defect or abnormality should be able to be detected at an early stage. Secondly, the machine performance degradation should be assessed robustly and tracked continuously. Finally, the RUL and possible failure mode of the machine should be effectively predicted. Therefore, the key challenges of implementing machine prognostics are machine performance degradation assessment and RUL prediction in which the former is a critical procedure to the latter.

Recently, several considerable efforts have been implemented to develop methods and tools for these challenges. Qiu et al. [2] developed a robust performance degradation method for rolling element bearing. This method based on a combination of wavelet filter and self-organizing map (SOM) for enhancing weak signal to fault identification and assessing performance degradation, respectively. Lee et al. [3, 4] applied a pattern discrimination model based on a cerebellar model articulation controller (CMAC) neural network to monitor and assess performance degradation in a robot. Lin et al. [5] combined traditional reliability modeling methods with vibration-based monitoring techniques and CMAC neural network in an integrated system to perform the machine reliability and determine the health status of machine. The results of CMAC were then verified by Weibull proportional hazard model. Xu et al. [6] proposed a fuzzy based extension of CMAC neural network to analyze two types of machine degradation severities: the network was trained by signals from different levels of machine degradation severity, and the network was only trained by signals of machine normal state. Other neural network based approaches to assess bearing performance degradation and forecast the RUL were presented by Gebraeel et al. [7] and Huang et al. [8], respectively. Liao and Lee [9] proposed a method which combined wavelet packet analysis, principal component analysis, and Gaussian mixture model to define performance index. This index was determined based on the overlap between the distribution of the baseline feature space and that of the testing feature space. An extension of support vector machine (SVM), namely least squares SVM, was used for assessment of machine performance degradation was also proposed in reference [10]. Using fuzzy c-means (FCM) clustering method is a current trend approach to assess the machine performance degradation. Huang et al. [11] proposed a combined model in which wavelet packet transform was used to process raw signals, the Fisher criterion was applied for future selection, and FCM was utilized to assess and classify the performance of machine. Pan et al. [12] proposed an FCM based method for bearing performance assessment. This method comprised lifting wavelet packet decomposition for feature extraction and FCM for assessing bearing performance degradation. In case of using model-based techniques, proportional hazard model (PHM) and logistic regression model (LRM) were applied for performance machine reliability indices and machine RUL prediction [13]. Other applications of LRM could be found in references [14-16].

Generally, these above approaches can indicate the machine degradation as well as assist in forecasting the machine RUL. However, some existing disadvantages lead to difficulty of applying these approaches to industrial equipment. FCM, SVM, and back propagation neural network need to be trained by historical data including both normal and failure operation to generate the assessment models. This is expensive or inapplicable in industrial area. SOM and CMAC neural network assessment models could be created by using only normal operation data which is more beneficial. Nevertheless, they are either not intuitive enough to reflect degradation degree or seriously influenced by some parameters defined by user, which makes it unpractical [12]. Moreover, these approaches only focused on the machine performance degradation assessment without RUL prediction which is the ultimate goal of prognostics. Performance degradation and RUL prediction methods based on PHM and LRM also need data of whole machine life. This is impossible in case of newly installed machine in normal operating condition. Yan et al [15] proposed a solution for this problem by using technician's experience for acceptable level or unacceptable level to predefine the probability thresholds. However, these levels are non-scientific and human intuitiveness.

In this study, a three-stage method for both targets involving machine performance degradation assessment and RUL prediction is proposed. In the first stage, autoregressive moving average (ARMA) model [17], which is one of the system identification techniques, is generated by using only normal operating data to identify the behavior of the complex system. Degradation index defined as the root mean square of residual errors is then used to indicate the machine degradation. The residual errors are the different outputs between identification model and behavior of system. Based on this degradation index, operators or maintainers could define the failure threshold for the system. In the second stage, the Cox's PHM [18] is established to estimate the system survival function. In the last stage, support vector machine for regression [19, 20], which is one of the remarkable machine learning techniques, in association with multi-step ahead prediction method of time-series forecasting techniques is utilized to forecast the RUL.

2. PROPOSED METHOD

The proposed method employed to assess the machine performance degradation and forecast the RUL is depicted in Figure 1. In order to apply this method, machine condition monitoring process is firstly carried out to acquire the machine condition. These data are then used as the inputs for the proposed method. The role of each step is summarized as follows:

Step 1 (System behavior identification): in this step, the data obtained from the normal operating condition of system is used to generate an identification model. This model will mimic the behavior of the system in the future states. Based on the residual errors which are the difference between the real system behavior and identification model, the degradation index is defined. The failure threshold is also determined by operators or maintainers. This step will be continued until occurrence of that the degradation index is higher than the failure threshold.

Step 2 (Survival function estimation): The Cox's PHM is created to estimate the survival function of system once the degrading mode occurred.

Step 3 (RUL prediction): By using the survival time data, SVM in association with prediction method of time-series techniques are trained and forecast the system RUL.

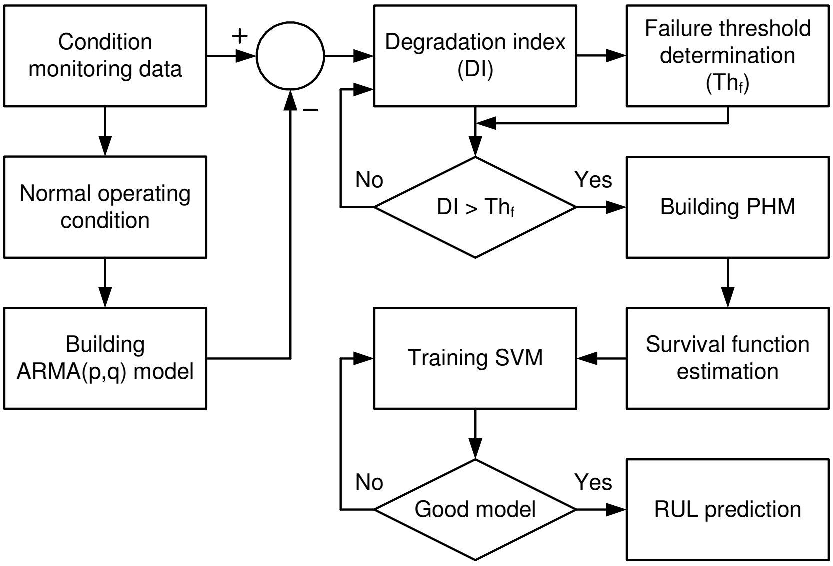

Figure 1. Schematic diagram of proposed method

3. APPLICATION AND RESULTS

4.1. Data acquisition

The proposed method is applied to forecast the RUL based on the data acquired from condition monitoring routine of a low methane compressor of a petrochemical plant. The compressor shown in Figure 2 is driven by a 440 kW motor, 6600 volt, 2 poles and operating at a speed of 3565 rpm. Other information of the system is summarized in Table 1

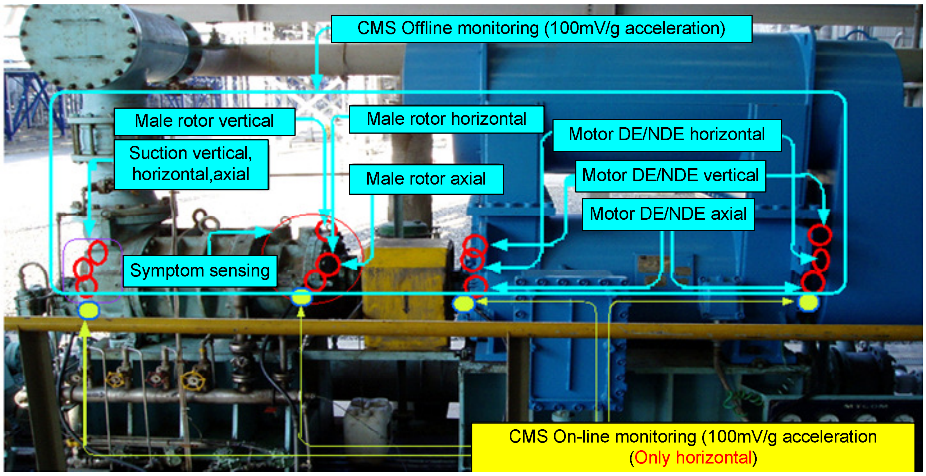

Figure 2. Low methane compressor

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

To monitor the machine condition, several kinds of signals could be used e.g. vibration, acoustic, oil analysis, temperature, pressure, moisture, etc. Amongst these, vibration signal is the most commonly acquired data due to the easy-to-measure signals and analysis. That is the reason why vibration signal was also applied for the condition monitoring system of this compressor. The vibration data include peak and envelope accelerations which were recorded from August 2005 to

November 2005. The average recording duration was approximately 6 hours during the data acquisition process. Each data record consisted of 1200 data points as shown in Figure 3 and 4. Consequently, these data contain information of machine history with respect to time sequence.

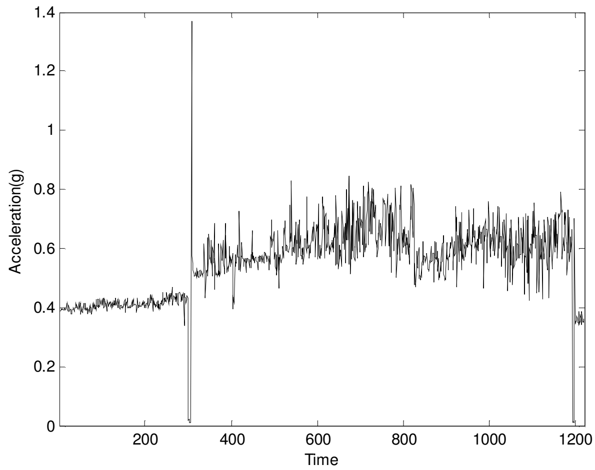

Figure 3. The entire peak acceleration data

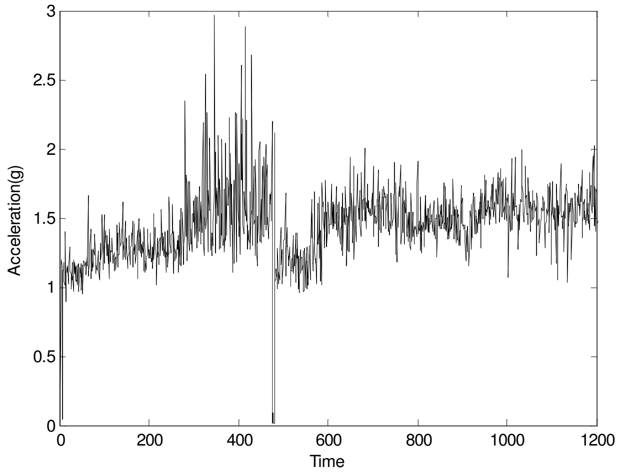

Figure 4. The entire envelope acceleration data

4.2. Results

As indicated in Figure 3, the machine was obviously in normal operating condition during the first 300 points of the time sequence. At the 291th point, the machine condition significantly reduced in comparison with other points. After the 300th point, the machine suddenly changed the condition. Due to the lack of the expert knowledge and/or historical fault data for this complex system to assist operators in determining the necessary activities, this compressor was wholly broken down at the 308th point. The occurred fault is the damage of main journal bearings because of insufficient lubrication which lead to the surfaces of these bearings to be overheated and delaminated [21]. However, this is difficult to determine the abnormal condition by using envelope acceleration data (Figure 4). Thus, peak acceleration data is employed for reveal the machine performance degradation in this study.

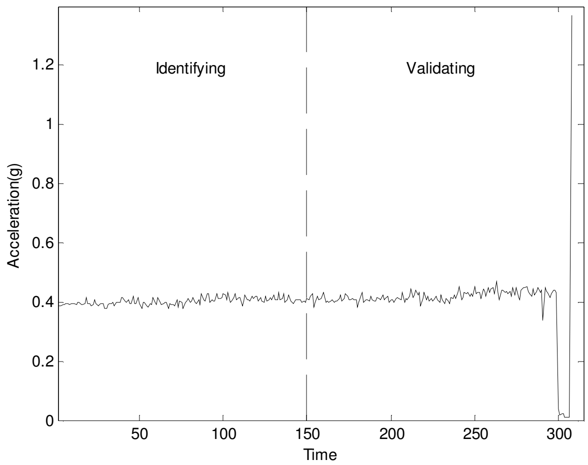

Figure 5. Data used for proposed method

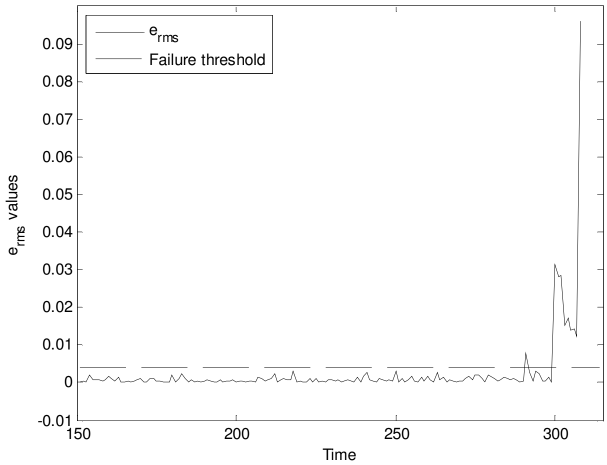

Figure 6. Degradation index

As mentioned in the previous section, the first step of proposed method is to create the identification model, which ARMA model is applied here, based on normal operating condition to mimic the behavior of system. For that reason, 308 points of machine condition included normal operating state and degrading state are split into two parts, namely identifying and validating, as shown in Figure 5. In indentifying process, 150 points of peak acceleration data which are taken from 300 points in normal operating state are used to generate ARMA model. A problem frequently encountered in this process is to determine the orders of ARMA model. Using the method proposed by Broersen [22-23], ARMA (4, 3) model is obtained. The AR and MA coefficients are $\varphi = \{1, -0.2408, -0.2667, -0.6316\}$ and $\phi = \{1, -0.058, -0.2726\}$, respectively. Afterward, this identification model is used to estimate the dynamic behavior of machine for the remaining points (158 points). The residual errors which are the difference between the machine behavior (peak acceleration values) and identification model are then gained. The degradation index depicted in Figure 6 is defined as the root mean square of residual errors, denoted by e_{rms}:

$$e_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_{ih})^2}\,,\ i = 1,...,N$$

where N is total number of observations, yi is the ith value of machine behavior, and yih is the ith value obtained from ARMA model.

From this figure, the degradation index can be obviously recognized to increase suddenly from the 300th point. This appropriates to the sudden change of peak acceleration mentioned above as well as indicates the abnormal value at the 291th point. Therefore, degradation index is adequate to assess the machine degradation. Furthermore, based on the degradation index, operators or maintainers can easily determine the failure threshold for machine. This value is chosen after considering all values of degradation index during the normal operation state, for example 0.004 in this study. Furthermore, this threshold is also employed to attain the censored data which is used for generating the PHM. This data consists of a series of "0" and "1" values indicating the normal condition and failure condition, respectively.

Once the degradation index is higher than the failure threshold, Cox's PHM is established bases on the censored data obtained from previous step. In this study, two kinds of features $z_1(t)$ and $z_2(t)$ are used and denoted as peak and envelope features, respectively. The parameters of PHM are estimated as β_1 = -0.8042 and β_1 = 0.1062. Hazard rate and survival function estimation are depicted in Figure 7 and 8, respectively. In Figure 7, the hazard rate gradually increases with respect to time though the peak acceleration values still approximately remain in the same average value. This could be the increment of envelope that effects the hazard rate function. From the 300th point, the hazard rate significantly changes because of the rapid growth of peak acceleration values. Thus, the more the hazard rare increases, the less the reliability is.

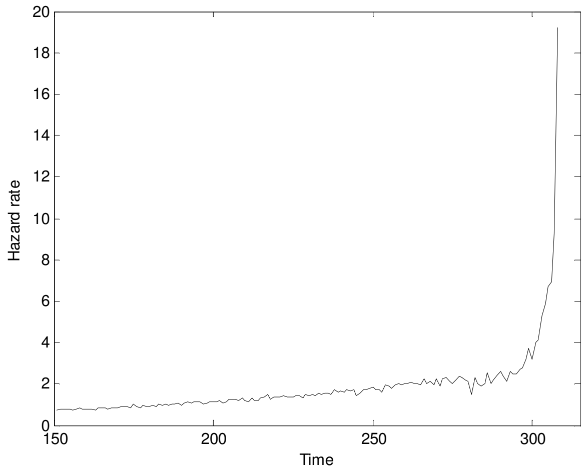

Figure 7. Hazard rate estimation

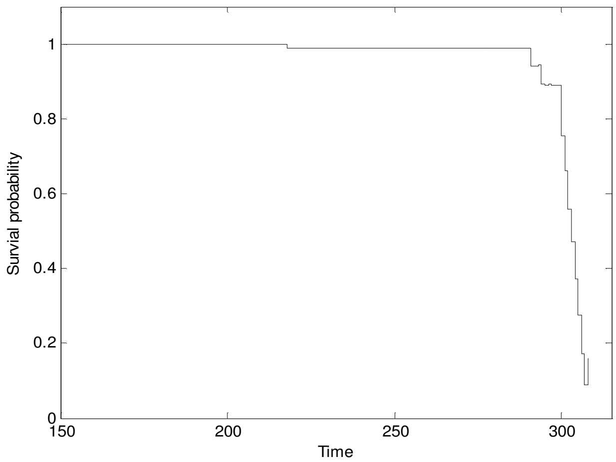

Figure 8. Survival function estimation

After attaining the survival function, the process of training and forecasting by using SVM in association with time-series forecasting techniques is carried out. The multi-step ahead direct prediction method [24] of time-series forecasting techniques is applied for this study. The values of survival function from the 151th to 291th point are used to train SVM model in which the Gaussian kernel $K(x, y) = \exp(-|x - y|^2 /(2\sigma)^2)$ is employed. The other predefined parameters ε and C are set to 0.001 and 500, respectively. Moreover, 5-fold cross validation is also applied to choose the best SVM model. The predicted results of 17 points from the 292th point where the machine commences degrading the state are depicted in Figure 9. Even though the multi-step ahead is employed, the predicted results can track the reduction of reliability with the root mean square error (RMSE) of 0.33674. From these results, the RUL of machine could be estimated by using the predefined survival probability for total fault. For example, if this value is chosen as 0.2, the point where predicted result reaches for 0.2 is 308th. This means the predicted RUL of machine after degrading state occurred is 96 hours ((308 - 292) × 6 = 96) while the actual RUL is 84 hours (the 306th point). The accuracy can be evaluated using the follow formula:

$$Accuracy = \left(1 - \frac{\left|t_{predicted} - t_{actual}\right|}{t_{predicted}}\right) \times 100\% = \left(1 - \frac{96-84}{96}\right) \times 100\% = 87.5\%$$

Even if the accuracy is optimistic and acceptable in industrial application, the predicted RUL is lag behind actual RUL due to the implementation of several-step prediction. It is not only a difficult task but also challenging task for time series prediction problem. However, several-step prediction and improvement in forecasting results is of necessity for real application of machine prognostics. For further improvement of predicting accuracy, the predicted value has been moved backward with an interval of the average lag as shown in Figure 10. As a result, the predicted results and actual values of survival function is closely resemble with the accuracy of 100%, which shows the feasibility of the proposed method.

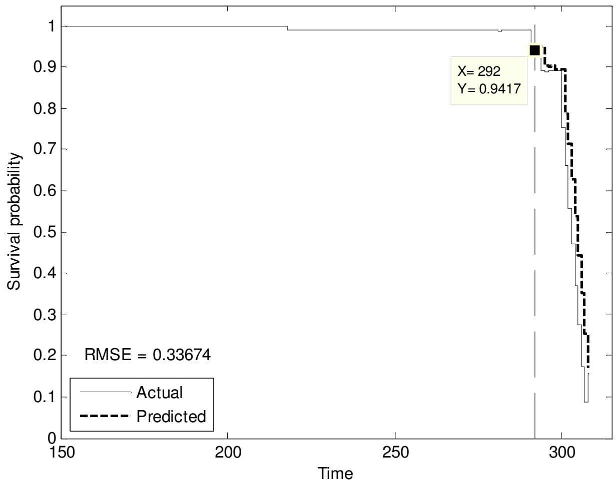

Figure 9. Predicted results

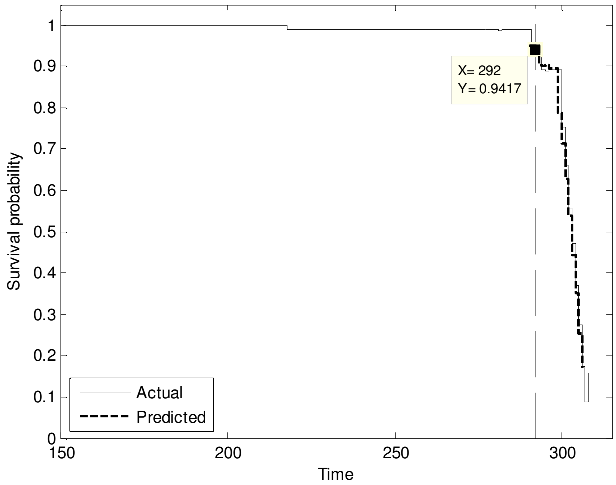

Fig. 10 Predicted results after improvement

4. CONCLUSION

In this study, the method for assessing the machine health degradation and forecasting the RUL is proposed, where ARMA (4, 3) identification model, Cox's PHM, and SVM are combined in an integrated tool. By using this method, the machine performance degradation assessment could be attained. Furthermore, the machine RUL is also obtained from the combination of the predicted results and predefined fault probability. The data of low methane compressor acquired from condition monitoring routine is used for appraising the proposed method. The high accuracy of predicted results indicates that the proposed could be used as a reliable tool to machine prognostics.

5. REFERENCES

1 Heng A., Zhang S., Tan A.C.C., Mathew J. (2009) Rotating machinery prognostics: state of the art, challenges and opportunities. *Mechanical Systems and Signal Processing*, 23, 724-739.

2 Qiu H., Lee J., Lin J., Yu G. (2003) Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Advanced Engineering Informatics*, 17, 127-140.

3 Lee J. (1996) Measurement of machine performance degradation using a neural network model. *Computer in Industry*, 30, 193-209

4 Lee J. & Kramer B. M. (1992) Monitor machine degradation using an enhanced CMAC neural network. *IEEE International Conference on Systems, Man and Cybernetics 2*, pp.1010-1015.

5 Lin C. C., Tseng H. Y. (2005) A neural network application for reliability modeling and condition-based predictive maintenance. *The International Journal of Advanced Manufacturing Technology*, 25, 174-179

6 Xu R. Z., Xie L., Zhang M. C. (2008) Machine degradation analysis using fuzzy CMAC neural network approach. *The International Journal of Advanced Manufacturing Technology*, 36, 765-772.

7 Gebraeel N., Lawley M., Liu R., Parmeshwaran V. (2004) Residual life prediction from vibration-based degradation signals: a neural network approach. *IEEE Transactions on Industrial Electronics*, 5, 1694-700.

8 Huang R., Xi L., Li X., Liu C. R., Qiu H., Lee J. (2007) Residual life predictions for ball bearing based on self-organizing map and back propagation neural networks methods. *Mechanical Systems and Signal Processing*, 21, 193-207.

9 Liao L., Lee J. (2009) A novel method for machine performance degradation assessment based on fixed cycle feature test. *Journal of Sound and Vibration*, 326, 894-908.

10 Xu Y., Deng C., Wu J. (2009) Least Squares Support Vector Machines for performance degradation modeling of CNC equipments. *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*. pp. 201-206

11 Huang Y., Liu C., Zha X. F., Li Y. (2010) A lean model for performance assessment of machinery using second generation wavelet packet transform and Fisher criterion. *Expert Systems with Applications*, 37, 3815-3822.

12 Pan Y., Chen J., Li X. (2010) Bearing performance degradation assessment based on lifting wavelet packet decomposition and fuzzy c-means. *Mechanical Systems and Signal Processing*, 24, 559-566.

13 Liao H., Zhao W., Guo H. (2006) Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model. *Reliability and Maintainability Symposium*, pp. 127-132.

14 Yan J., Lee J. (2004) Degradation assessment and fault modes classification using logistic regression. *Journal of Manufacturing Science and Engineering-Transactions of the ASME*, 127(4), 912-207.

15 Yan J., Koç M., Lee J. (2004) A prognostic algorithm for machine performance assessment and its application. *Production Planning and Control*, 15(8), 796-801.

16 Caesarendra W., Widodo A., Yang B. S. (2010) Application of relevance vector machine and logistic regression for machine degradation assessment. *Mechanical Systems and Signal Processing*, 24(4), 1161-1171.

17 Box G.E.P., Jenkins G. (1970) Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, CA.

18 Cox D. R. (1972) Regression models and life-table. *Journal of the Royal Statistical Society*, Series B 34(2), 187-220.

19 Vapnik V. N. (1995) The Nature of Statistical Learning Theory. Springer, New York.

20 Vapnik V. N., Golowich S., Smola A. (1996) Support vector machine for function approximation regression estimation and signal processing. *Advances in Neural Information Processing Systems*, 9, 281-287.

21 Tran V. T., Yang B. S., Oh M. S., Tan A. C. C. (2008) Machine condition prognosis based on regression trees and one-step-ahead prediction. *Mechanical Systems and Signal Processing*, 22, 1179-1193.

22 Broersen P. M. T. (2002) Automatic spectral analysis with time series models. *IEEE Transactions on Instrumentation and Measurement*, 51, 211-216

23 Broersen P. M. T., Matlab Toolbox ARMASA. Available: http://www.dcsc.tudelft.nl/Research/Software/

24 Tran V.T., Yang B.S., Tan A.C.C (2009) Multi-step ahead direct prediction for the machine condition prognosis using regression trees and neuro-fuzzy systems. *Expert Systems with Applications*, 36, 9378-9387

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support from Brain Korea (BK) 21 and The Vietnam National Foundation for Science and Technology Development (NAFOSTED) for this study.