

Combining Canonical Variate Analysis, Probability Approach and Support Vector Regression for Failure Time Prediction

Xiaochuan Li, Fang Duan
School of Engineering
London South Bank University
London, UK

Ian Bennett
Rotating Equipment Department
Royal Dutch Shell
Hague, Netherlands

David Mba
Faculty of Technology
De Montfort University
Leicester, UK

Abstract—Reciprocating compressors are widely used in oil and gas industry for gas transport, lift and injection. Critical reciprocating compressors that operate under high-speed conditions and compress hazardous gases are target equipment on maintenance improvement lists due to downtime risks and safety hazards. Estimating performance deterioration and failure time for reciprocating compressors could potentially reduce downtime and maintenance costs, and improve safety and availability. This study presents an application of Canonical Variate Analysis (CVA), Cox Proportional Hazard (CPHM) and Support Vector Regression (SVR) models to estimate failure degradation and remaining useful life based on sensory data acquired from an operational industrial reciprocating compressor. CVA was used to extract a one-dimensional health indicator from the multivariate data sets, thereby reducing the dimensionality of the original data matrix. The failure rate was obtained by using the CPHM based on historical failure times. Furthermore, a SVR model was used as a prognostic tool following training with failure rate vectors obtained from the CPHM and the one-dimensional performance measures obtained from the CVA model. The trained SVR model was then utilized to estimate the failure degradation rate and remaining useful life. The results indicate that the proposed method can be effectively used in real industrial processes to predict performance degradation and failure time.

Keywords—Condition Monitoring; Remaining Useful Life; Canonical Variate Analysis; Cox Proportional Hazard Model; Support Vector Regression

I. INTRODUCTION

Modern industrial facilities such as natural-gas processing plants are becoming increasingly complex and large-scale due to the use of machines of different nature. The complexity of large-scale industrial facilities makes it difficult to build first-principle dynamic models for health monitoring and prognostics [1]. Thus, existing condition monitoring approaches for industrial processes are typically derived from routinely collected system operating data. Due to the rapid growth and advancement in data acquisition technology, long-term continuous measurements can be taken with the different

sensors mounted on the machinery systems. The monitored data are easily stored and explored to extract important process condition information.

A number of methods have been developed to combine the multi-dimensional process variables for health monitoring and prognostics, such as state-space and filtering based models [2], time series analysis methods [3] and neural networks [4]. Some of the major challenges associated with the implementation of the above mentioned multivariate techniques are high-dimensional data, non-Gaussian distributions and dynamically varying operational conditions [5]. Recent developments of dimensionality reduction techniques have shown improvement for high-dimensional data analysis. Two examples of such techniques are principal component analysis (PCA) [6] and partial least-squares analysis (PLSA) [7]. Both PCA and PLSA assume that the monitored variables are time-independent. This assumption might not hold true for real industrial processes (especially chemical and petrochemical processes) because sensory signals affected by noises and disturbances often show strong correlation between the past and future sampling points [8]. Therefore, dynamic extensions of PCA and PLSA, so called dynamic PCA and dynamic PLSA, were developed later to solve the time-independency problem, making them more suitable for dynamic processes monitoring [9]. Aside from approaches derived from PCA and PLSA, the canonical variable analysis (CVA) is also a multivariate analysis tool. CVA is a state-space-based method which takes both time correlations and relationship between different variables into account, hence is particularly suitable for dynamic process modelling [8]. The performance of CVA has been tested by several researchers using simulated data [10] [11] and data obtained from small-scale experimental test rigs [12]. However, the effectiveness of CVA in real complex industrial processes has not been fully studied. In this investigation, CVA is adopted to transform the high-dimensional data from the sensors distributed over the machine to a one-dimensional matrix called the health indicator that can be used to indicate the health condition of a system.

In medical research field, the Cox Proportional Hazard Model (CPHM) has been widely used for analysing death rate or the probability of recurrence of a disease with censored survival data [13]. But its effectiveness in mechanical prognostic area has not been fully studied and only a limited number of publications have addressed its applicability for failure prediction of rotating machines [14] [15]. In this study, the CPHM model is utilized to address the failure probability of reciprocating compressor samples. To be specific, the failure rate as a degradation measure is calculated using the CPHM. Censored run-to-failure data is used to estimate the baseline survival function and the temperature measurements obtained at the fault location is assumed as a time-dependent covariate to investigate the failure time distribution.

Both the failure rate vectors calculated by the CPHM and the one-dimensional health indicators obtained from the CVA model are regarded as target vectors indicating the health condition of the compressor under study. Furthermore, a SVR model is employed as a degradation assessment method and then trained by the health indicators and failure rate vectors. After training, the SVR model is utilized to predict the failure degradation and failure time of individual failure sample given unseen values of input.

II. METHODOLOGY

The flowchart of the combined prognostic method is shown in Fig. 1. The following sub-sections give detailed theoretical explanations of CVA, CPHM and SVR model.

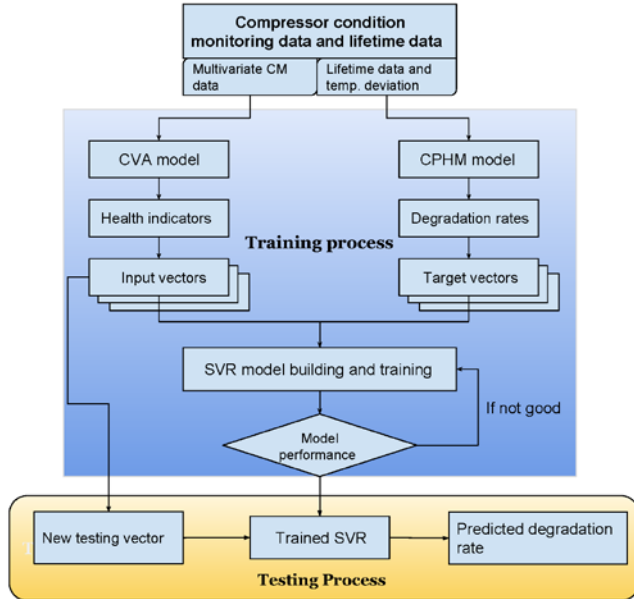


Figure 1. Schematic diagram of the proposed prognostic method

A. Canonical Variate Analysis

CVA is a dimension reduction technique to monitor the process by converting the multidimensional observed data into a one-dimensional health indicator. The objective of CVA is to maximize the correlation between two sets of variables [1]. For this purpose, the measurement vector $y_k \in \mathfrak{R}^m$ (m

indicates that measurement at each time instance containing m variables) is expanded at each time instance k by considering p number of past measurements and f number of future measurements to give the past and future sample vectors $y_{p,k} \in \mathfrak{R}^{mp}$ and $y_{f,k} \in \mathfrak{R}^{mf}$.

$$y_{p,k} = [y_{k-1}^T \ y_{k-2}^T \ \dots \ y_{k-p}^T]^T \in \mathfrak{R}^{mp} \quad (1)$$

$$y_{f,k} = [y_k^T \ y_{k+1}^T \ \dots \ y_{k+f-1}^T]^T \in \mathfrak{R}^{mf} \quad (2)$$

To avoid the domination of variables with larger absolute values, it is important to normalize the past vector $y_{p,k} \in \mathfrak{R}^{mp}$ future vector $y_{f,k} \in \mathfrak{R}^{mf}$ to zero mean vector of $\widehat{y}_{p,k}$ and $\widehat{y}_{f,k}$, respectively. Then the vectors calculated at different time points are arranged in columns to produce past and future matrices Y_p and Y_f :

$$Y_p = [\widehat{y}_{p,p+1}, \widehat{y}_{p,p+2}, \dots, \widehat{y}_{p,p+M}] \in \mathfrak{R}^{mp \times M} \quad (3)$$

$$Y_f = [\widehat{y}_{f,p+1}, \widehat{y}_{f,p+2}, \dots, \widehat{y}_{f,p+M}] \in \mathfrak{R}^{mf \times M} \quad (4)$$

where $M = n - f - p + 1$, and n represents the total number of samples for y_k . Y_p and Y_f are then processed by using the Cholesky decomposition to form a Hankel matrix [16]. The purpose of using Cholesky is to form a new correlation matrix with reduced dimensionality, thereby facilitating subsequent calculations. To find the linear combination that maximizes the correlation between the two sets of variables, the truncated Hankel matrix H is then decomposed by using Singular Value Decomposition (SVD):

$$H = \Sigma_{ff}^{-1/2} \Sigma_{pf} \Sigma_{pp}^{-1/2} = U \Sigma V^T \quad (5)$$

where Σ_{pp} , Σ_{ff} and Σ_{pf} represent the sample-based covariance and cross-covariance matrix of matrices Y_p and Y_f , respectively. The mp -dimensional past vector Y_p can be converted into the r -dimensional canonical variates z by:

$$z = J \cdot Y_p \quad (6)$$

where J represents the transformation matrix, and $J = V_r^T \Sigma_{pp}^{-1/2}$. The truncated matrix $V_r \in \mathfrak{R}^{r \times M}$ can be obtained by selecting the first r columns of V having the highest pairwise correlation with those of U [17]. The canonical variates matrix z consists of valuable information that is needed to construct health indicators. The health indicator adopted in this study is the Hotelling statistics T^2 (introduced by Hotelling in 1936 [18]), which is the locus on the ellipse-like confidence region in the canonical variate space [19]. The Hotelling health indicator can be calculated as:

$$T_k^2 = \sum_{i=1}^r z_{k,i}^2 \quad (7)$$

Process data acquired during normal operating conditions are used to identify optimal threshold values of the health indicator for fault detection. Since the Gaussian distribution doesn't hold true for non-linear processes, the actual probability density function of the health indicator is calculated by using a method named Kernel Density Estimation (KDE) [8]. Then machine faults will be considered every time when the health indicator exceeds the threshold. In this study, the initial and final failure point of individual data set are determined by the experience of site engineers.

B. Cox Proportional Hazard Model

Machinery fault degradation can be predicted by analyzing either condition monitoring measurements or historical lifetime data [20]. The CPHM, proposed by Cox [21],

attempts to use both types of information for prognostic analysis of machinery fault degradation and failure times. A lifetime data set consists of failure times T of the machine under study, recorded either at failure time or before the final failure. In some cases, maintenance actions may be taken prior to failure to prevent a device or component from failing. Then these cases are considered as censored since the actual failure time is unknown. In these cases, the recorded lifetime data is called censored data. The condition monitoring measurements used in CPHM can be any sensory signal that has a significant effect on the machine health condition.

CPHM assumes that the hazard rate or failure rate of a machine depends on two factors: the baseline hazard rate and the effects of covariates (condition measurements). Hence, the hazard rate of a machine at service time t can be written as:

$$h(t) = h_0(t) \exp\left(\sum_{k=1}^p \beta_k Z_k\right) \quad (8)$$

where $h_0(t)$ is called the baseline hazard function (It reflects the failure rate due to age meaning that it is determined by the lifetime data); $\exp\left(\sum_{k=1}^p \beta_k Z_k\right)$ is the covariate function that describes how the covariates Z_k influence health degradation. The covariates are weighted through the regression parameters β_k . The estimation of the regression parameters is achieved by using a method called partial likelihood approach, which was proposed by Cox in 1972 [21]. According to Cox's theory, the partial likelihood of β_k can be written as:

$$L(\beta) = \prod_{i=1}^n \frac{\exp\left(\sum_{k=1}^p \beta_k Z_{ik}(t_i)\right)}{\sum_{j \in R(t_i)} \exp\left(\sum_{k=1}^p \beta_k Z_{jk}(t_j)\right)} \quad (9)$$

Then the optimal regression parameters can be estimated by maximising the log likelihood of β_k :

$$\begin{aligned} \text{LL}(\beta) &= \sum_{i=1}^n \sum_{k=1}^p \beta_k Z_{ik}(t_i) \\ &- \sum_{i=1}^n \ln \left[\sum_{j \in R(t_i)} \exp \left(\sum_{k=1}^p \beta_k Z_{jk}(t_j) \right) \right] \end{aligned} \quad (10)$$

After model parameters are estimated, the hazard function can be calculated as:

$$\hat{h}_0(t; \hat{\beta}) = \frac{1}{\sum_{j \in R(t)} \exp\left(\sum_{k=1}^p \hat{\beta}_k Z_{jh}(t_j)\right)} \quad (11)$$

Then the cumulative hazard function and machine degradation rate can be approximated by formula (12) and (13), respectively:

$$\hat{H}(t) = \sum_{t_i \leq t} \hat{h}(t_i; \hat{\beta}) \quad (12)$$

$$\hat{S}(t) = \exp[-\hat{H}(t)] \quad (13)$$

C. Support Vector Regression

SVR is a supervised nonlinear regression approach. Application of the SVR model in the field of rotating machinery health monitoring and prognostics has been reported in [22] [23]. The target of SVR is to learn the dependency of an input vector $\{x_i\}_{i=1}^N$ on a target vector $\{y_i\}_{i=1}^N$ to make accurate forecast of y based on unseen values of x . When performing nonlinear regression, a kernel function is often chosen to map nonlinear inputs into a higher dimensional feature space, after which a minimum linear

margin fit can be found in that space to perform linear regression. The form of the model is given as:

$$y = f(x, w) = \sum_{i=1}^N w_i K(x, x_i) + b \quad (14)$$

where $w = (w_1, w_2, \dots, w_N)^T$ is a weight vector, which elucidates the links between the high dimensional space and the target output; and $K(x, x_i)$ denotes the kernel function, and b denotes the bias.

III. APPLICATION ON RECIPROCATING COMPRESSOR CONDITION MONITORING DATA

A. Data Acquisition

Reciprocating compressors are widely used in oil and gas industry for gas transport, lift and injection. They are typically operating under high rotating speed, high pressure and high load conditions, and are therefore subject to performance degradations. These machines are highly automated with various sensors being mounted all over the system, and signals from different sensors can be stored and accessed through an e-maintenance system. The data used in this study were gathered from a two-stage, four-cylinder, double-acting reciprocating compressor used in a refinery in Europe. The compressor experienced twelve valve failures from July 2013 to December 2014, with all failures took place at cylinder 4. Machine inspections revealed that the failure mode under study was valve leakage caused by broken valve plate, as shown in Fig. 2. The failed valves were either the head end or the crank end discharge valve. A total of 12 failure samples were obtained from the site engineer and each sample was a multivariate time series consisting of 40 variables. The sampling rate was 1Hz and the failure degradation duration for each sample was different.



Figure 2. Compressor valve plate damage caused by corrosion (left) and fatigue (right) [24]

B. Determination of Incipient and Final Failure Time

Since the failure mode under study is head end/crank end valve damage took place in cylinder 4, the method employed to determine the incipient and final failure time, as suggested by the site engineers, is to look at the difference between crank end (CE) discharge temperature and head end (HE) discharge temperature in cylinder 4. To be specific, during healthy operating conditions and after failure propagation, as shown in Fig. 3, the temperature difference between CE and HE is relatively constant. However, the temperature difference grows continuously when the valve is failing.

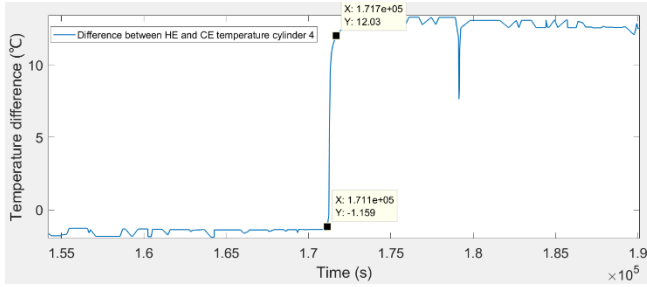


Figure 3. Difference between CE and HE discharge temperature in cylinder 4 – failure sample No. 2

TABLE I. DEGRADATION DURATION FOR ALL FAILURE SAMPLES

Sample No.	Degradation Length (s)
6	171
11	191
3	231
1	371
13	381
10	391
5	401
8	441
2	451
4	501
12	601
9	641

As such, the incipient failure is identified when the value of temperature difference starts to increase, as shown in Fig.3, the incipient failure happens at 171190s. Meanwhile, the final failure time is identified when the temperature difference goes back to stable. In Fig.3 for example, the final failure happens at 171640s. Therefore, the degradation duration for failure case 2 is 451s. The degradation duration for all failure samples can be found in Table 1.

C. CVA Model Building

The CVA model was firstly built and trained in order to transform the multivariate condition monitoring data into a one-dimensional health indicator. This process can be considered as a data fusion and dimensionality reduction procedure as it incorporates the information from all the measured 40 variables to generate a health indicator that reflects the health condition of the system. In each sample data set, data from health status were used to train the CVA algorithm to obtain the normal operating limits of T_k^2 , and data from the incipient failure and onwards were used to construct a health indicator. Besides, to build a CVA model as described in (1) to (6), three tuning parameters need to be determined, namely, the number of time lags p and f , and the number of dimensions retained r . According to [8], p and f are determined by computing the autocorrelation function of the summed squares of all measurements. The autocorrelation function indicates how long the signal is correlated with itself, and thus can be used to determine the maximum number of significant lags. By looking at the autocorrelation function of the training data set, the number of p and f was set to 25. For example, Fig. 4 shows the autocorrelation function of sample No. 1, and it is obvious that after 25 lags, the self-correlation of training data become negligible. The optimal number of dimensions

retained r is determined by considering the dominant singular values in the matrix D [25]. After several tests, $r = 3$ was finally adopted to represent the order of the system, leading to the minimum false alarm rate. Then the values of the health indicators were calculated according to (7).

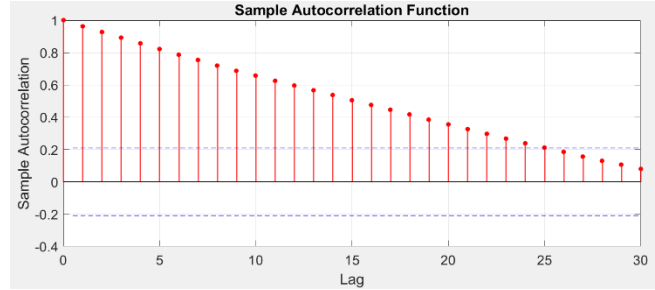


Figure 4. Autocorrelation of the summed squares of all variables in training data

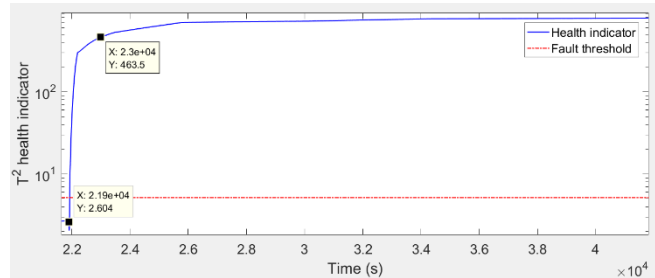


Figure 5. Untruncated health indicator of failure sample No. 5

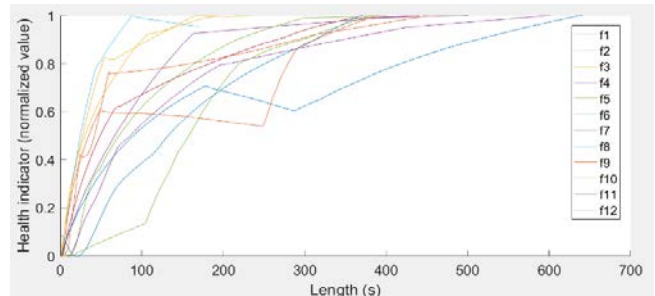


Figure 6. Truncated health indicators of all failure cases

Fig. 5 shows the constructed health indicator for failure sample no. 5. It can be observed that the value of T_k^2 exceeds the fault threshold at around 22000s, meaning that the failure was detected by the CVA model at 22000s. However, the incipient and final failure time in this study were determined by expert knowledge as explained in section B, and using CVA for failure detection is out of the scope of this study. Therefore, health indicator of each failure sample was then truncated to its corresponding degradation interval (i.e. truncated from the incipient failure time to final failure time). Only the truncated health indicators will be adopted for subsequent analysis. Fig.6 depicts the truncated health indicators for all 12 failure cases. They will be used hereafter as target vectors for SVR training.

D. CPHM Model Building

In order to build a CPHM model (as in (8)), lifetime data of 12 samples were used to estimate the baseline hazard function. In addition, the difference between CE discharge temperature and HE discharge temperature in cylinder 4 was assumed as a covariate and the regression parameter β_k was calculated as per (9) and (10) for each failure case. For example, Fig.7 shows the calculated degradation rate of failure case 9.

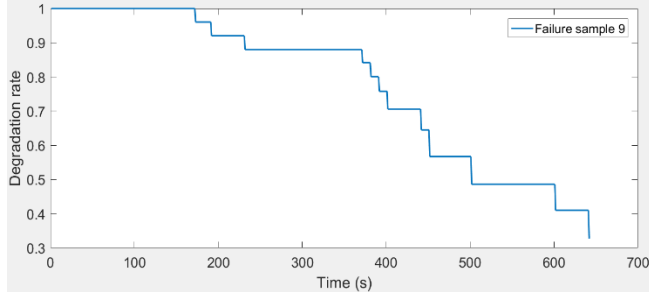


Figure 7. Hazard rate of failure sample No. 9

E. SVR Model Building and Testing

Furthermore, health indicators and failure rate vectors obtained previously will be used to train a SVR model. Then the trained SVR will be employed as a prognostic method to predict the failure degradation of individual failure case. To build a SVR model, we utilized a Radial Basis Function (RBF) kernel function to map input vectors into the high-dimensional feature space. The RBF kernel parameter γ and the soft margin parameter C were determined using grid search [26] together with 5-fold cross validation. For grid search, parameter γ and C take the following values:

$$\gamma = 2^{\{-10, -9, -8, \dots, 10\}}$$

$$C = 2^{\{-10, -9, -8, \dots, 10\}}$$

The health indicator and degradation curve of sample no. 10 was firstly utilized to train a SVR model and the optimal parameters resulting from grid search were 1024 and 64 for γ and C , respectively. The optimal parameters were determined by searching for the minimum Root-Mean-Squared Error (RMSE) between the actual degradation rate and the estimated degradation for each combination of γ and C candidates (as shown in Fig.8). Moreover, the health indicator of failure sample no. 13 was used as a testing vector to test the performance of the trained model. The predicted survival probability of sample no. 13 is depicted in Fig.9. It can be observed that the predicted failure time is 381s.

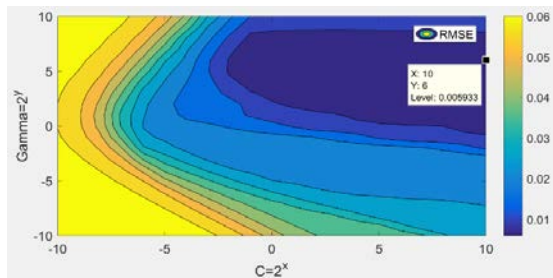


Figure 8. RMSE for various values of γ and C model parameters

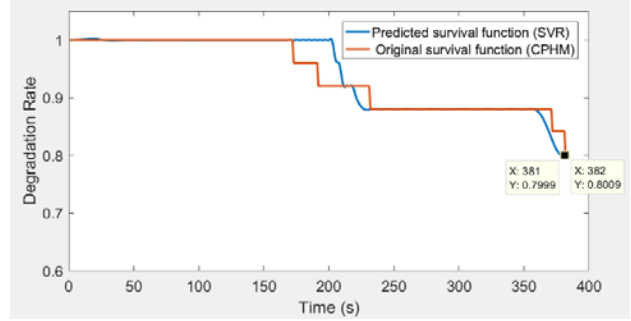


Figure 9. The result of SVR prediction for sample no. 13

Besides, in order to fully capture the dynamics of the compressor, a SVR model was then trained by 8 failure samples (f1, f13, f10, f5, f8, f4, and f12). The input vectors in this training process were obtained using the CVA method. In addition, the target vectors were acquired by an estimation of the degradation rate by means of CPHM. The optimal RBF parameters γ and C resulting from grid search were 128 and 256 respectively. Fig. 10 depicts the RMSE between the actual and the estimated target vectors for each combination of γ and C candidates. Moreover, the trained SVR model was utilized to predict the hazard rate of sample No. 2, and the predicted result is shown in Fig. 11. As shown in the picture, the predicted failure time is 449s while the actual failure happens at 452s.

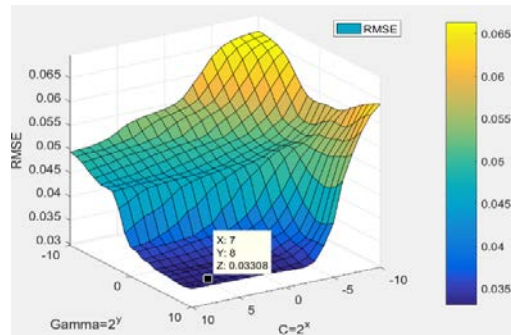


Figure 10. RMSE for various values of γ and C model parameters (using f1, f13, f10, f5, f8, f4, and f12 for training)

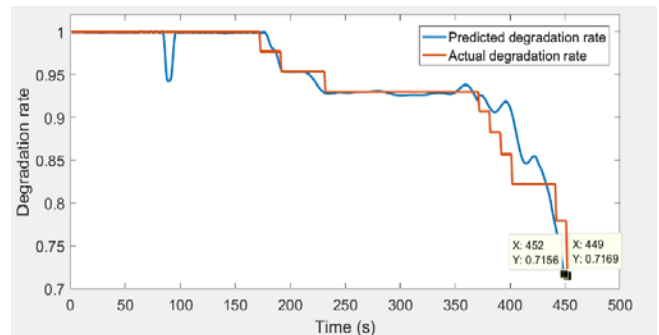


Figure 11. Predicted failure rate of sample No. 2

The performance of the prognostic model can be assessed using the following metrics, namely Accuracy introduced in

[19], root mean squared error (RMSE), mean absolute error (MAE) and Pearson's correlation coefficient (R). Formulae of the above metrics are listed as follows:

$$Accuracy = \left(1 - \frac{T_{actual} - T_{predicted}}{T_{actual}}\right) \times 100\% \quad (15)$$

$$RMSE = \left[\sum_{i=1}^N (S(t)_{actual,i} - S(t)_{predicted,i})^2 / N\right]^{1/2} \quad (16)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |S(t)_{actual,i} - S(t)_{predicted,i}| \quad (17)$$

$$R = \frac{\sum_{i=1}^N (S(t)_{act,i} - \bar{S}(t)_{act})(S(t)_{pre,i} - \bar{S}(t)_{pre})}{\sqrt{\sum_{i=1}^N (S(t)_{act,i} - \bar{S}(t)_{act})^2} \sqrt{\sum_{i=1}^N (S(t)_{pre,i} - \bar{S}(t)_{pre})^2}} \quad (18)$$

The higher the value of Accuracy, the better the prediction is. Meanwhile, the higher the value of RMSE/MAE, the lower the prediction accuracy is. A high Pearson's correlation coefficient means a high accordance between the actual and predicted degradation rate. The performance of the predictive model, based on four metrics, is summarized in table 2.

TABLE II. MODEL PERFORMANCE BASED ON FOUR STATISTICAL INDEXES

Sample No.	Accuracy	RMSE	MAE	R
13	99.74%	0.02	0.0082	0.9485
2	99.33%	0.0076	0.0482	0.933

The predicted degradation rate of failure case No. 2 seems overestimated between 370s and 430s and underestimated between 431s to 449s, yielding a relatively high MAE value. But the accuracy is 99.33%, which is admissible for constructing the prognostic model.

F. Comparison with previous studies

The results presented in [27] indicated that CVA can be effectively used to identify machine abnormalities and predict future values of process measurements based on real-life data acquired from industrial compressors. The combined method proposed in this study further illustrates that CVA-based methodology can also be used for estimating performance deterioration and failure time of a real-world operating reciprocating compressor. However, efforts need to be made to allow plant operators to predict machine remaining useful life based on real-time online measurements. Moreover, the results of [22] have demonstrated that the combination of probability approach and regression method can successfully predict fault degradation of bearing based on univariate bearing vibration signals. The method presented in this investigation allows the prediction of the failure of complex real facilities by considering the possible synergy among different sensor signals. One drawback of the combined method is that the scarcity of lifetime data of real machines may limit its applications.

IV. CONCLUSION

In this study, CVA combined with CPHM and SVR were applied for the first time on data collected from an operational industrial reciprocating compressor to perform prognostics. Firstly, CVA was used to build a health indicator of each failure sample. Moreover, CPHM was utilized to calculate failure rate of individual failure data set, and then the calculated failure curves were regarded as the target vectors for training the SVR model. Grid search and 5-fold cross

validation were used to determine the optimal SVR model parameters during the training process. Furthermore, the trained SVR was employed to predict degradation rate and failure time of the compressor given unseen input health indicators. Four metrics were utilized to evaluate the accuracy of the proposed schemes. The results illustrate that the prognostic performances were satisfied. Compared with previous studies, the proposed method outperforms the traditional CVA-based fault detection approaches by providing estimations of machine degradation rate and failure time. Additionally, the combined method advances the traditional probability and regression approaches by including multivariate condition monitoring data into prognostic analysis.

Although, the results of this study clearly show the superior performance of the proposed method for failure prediction, some aspects require further investigation are listed as follows. Firstly, only the system dynamics captured in the retained space, namely the truncated r -dimensional canonical variates matrix in (7), were used to build a health indicator. As a result of doing this, the constructed health indicator might not fully represent the system variations as some useful information in the residual space is eliminated. Therefore, future work should be focused on incorporating the complete system dynamics in the health indicator. Secondly, due to the approximative nature of hazard function (shown in (11)), the degradation curves used in this investigation are stair functions with jumps at failure times. Thus, a degradation curve might not truly reflect the deterioration process when the number of historical failures is small, leading to inaccurate degradation prediction. Hence, techniques should be developed to calculate machine degradation rates accurately regardless of the scarcity of lifetime data.

REFERENCES

- [1] E. Russell, L. Chiang and R. Braatz, "Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis", *Chemometrics and Intelligent Laboratory Systems*, vol. 51, no. 1, pp. 81-93, 2000.
- [2] J. Sun, H. Zuo, W. Wang and M. Pecht, "Application of a state space modeling technique to system prognostics based on a health index for condition-based maintenance", *Mechanical Systems and Signal Processing*, vol. 28, pp. 585-596, 2012.
- [3] A. Negiz and A. Çinar, "PLSA, balanced, and canonical variate realization techniques for identifying VARMA models in state space", *Chemometrics and Intelligent Laboratory Systems*, vol. 38, no. 2, pp. 209-221, 1997.
- [4] G. Niu and B. Yang, "Intelligent condition monitoring and prognostics system based on data-fusion strategy", *Expert Systems with Applications*, vol. 37, no. 12, pp. 8831-8840, 2010.
- [5] C. Ruiz-Cárcel, L. Lao, Y. Cao and D. Mba, "Canonical variate analysis for performance degradation under faulty conditions", *Control Engineering Practice*, vol. 54, pp. 70-80, 2016.
- [6] W. Ku, R. Storer and C. Georgakis, "Disturbance detection and isolation by dynamic principal component analysis", *Chemometrics and Intelligent Laboratory Systems*, vol. 30, no. 1, pp. 179-196, 1995.
- [7] R. Muradore and P. Fiorini, "A PLSA-Based Statistical Approach for Fault Detection and Isolation of Robotic Manipulators", *IEEE Transactions on Industrial Electronics*, vol. 59, no. 8, pp. 3167-3175, 2012.

- [8] P. Odiwei and Yi Cao, "Nonlinear Dynamic Process Monitoring Using Canonical Variate Analysis and Kernel Density Estimations", *IEEE Transactions on Industrial Informatics*, vol. 6, no. 1, pp. 36-45, 2010.
- [9] H. X. Pan, X. Y. Wei, and J. Y. Huang. "Fault feature extraction based on KPCA optimized by PSO algorithm," In *Proceedings of Industrial Informatics Conference on*, pp. 102-107, 2010.
- [10] B. Jiang, D. Huang, X. Zhu, F. Yang and R. Braatz, "Canonical variate analysis-based contributions for fault identification", *Journal of Process Control*, vol. 26, pp. 17-25, 2015.
- [11] H. Linzhe, C. Yuping, T. Xuemin and D. Xiaogang, "A Nonlinear Quality-relevant Process Monitoring Method with Kernel Input-output Canonical Variate Analysis", *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 611-616, 2015.
- [12] C. Ruiz Cárcel, Y. Cao and D. Mba, "A benchmark of Canonical Variate Analysis for fault detection and diagnosis", in *Control (CONTROL), 2014 UKACC International Conference on*, Loughborough, 2014.
- [13] B. Zupan, J. Demšar, M. Kattan, J. Beck and I. Bratko, "Machine learning for survival analysis: a case study on recurrence of prostate cancer", *Artificial Intelligence in Medicine*, vol. 20, no. 1, pp. 59-75, 2000.
- [14] A. Jardine, P. Anderson and D. Mann, "Application of the weibull proportional hazards model to aircraft and marine engine failure data", *Quality and Reliability Engineering International*, vol. 3, no. 2, pp. 77-82, 1987.
- [15] A. Jardine, D. Banjevic, M. Wiseman, S. Buck and T. Joseph, "Optimizing a mine haul truck wheel motors' condition monitoring program Use of proportional hazards modeling", *Journal of Quality in Maintenance Engineering*, vol. 7, no. 4, pp. 286-302, 2001.
- [16] D. Hardoon, S. Szedmak and J. Shawe-Taylor, "Canonical correlation analysis; An overview with application to learning methods", *Egham*, 2003.
- [17] R. T. Samuel and Y. Cao, "Kernel Canonical Variate Analysis for Nonlinear Dynamic Process Monitoring", in *9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2015*, Whistler, 2015.
- [18] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321-327, Dec. 1936.
- [19] J. Jackson, "Quality Control Methods for Several Related Variables", *Technometrics*, vol. 1, no. 4, p. 359, 1959.
- [20] X. Li, F. Duan, D. Mba and I. Bennett, "Multidimensional prognostics for rotating machinery: A review", *Advances in Mechanical Engineering*, vol. 9, no. 2, p. 168781401668500, 2017.
- [21] D. R. Cox, "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 34, No. 2, pp. 187-220, 1972.
- [22] W. Caesarendra, A. Widodo, P. Thom, B. Yang and J. Setiawan, "Combined Probability Approach and Indirect Data-Driven Method for Bearing Degradation Prognostics", *IEEE Transactions on Reliability*, vol. 60, no. 1, pp. 14-20, 2011.
- [23] Y. Qian and R. Yan, "Remaining Useful Life Prediction of Rolling Bearings Using an Enhanced Particle Filter", *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 10, pp. 2696-2707, 2015.
- [24] S. Senada, "Peek Material for Compressor Valve", *Slideshare*, 2017. [Online]. Available: <https://www.slideshare.net/mandalika3/peek-for-material-compressor-valve>.
- [25] A. Negiz and A. Çınar, "Monitoring of multivariable dynamic processes and sensor auditing", *Journal of Process Control*, vol. 8, no. 5-6, pp. 375-380, 1998.
- [26] Z. Liu, M. J. Zuo, and H. Xu, "Parameter selection for Gaussian radial basis function in support vector machine classification," In *Proceedings of 2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, Chengdu, pp. 576-581, 15-18 June, 2012.
- [27] X.C. Li, F. Duan, D. Mba and I. Bennett, "Canonical Variable Analysis for Fault Detection, System Identification and Performance Estimation," In *Proceedings of 7th International Congress Design and Modelling of Mechanical Systems*, Hammamet, pp. 1-11, 27-29 March, 2017.