Adaptive Canonical Variate Analysis for Performance Estimation with Application to Industrial Rotating Machines

Xiaochuan Li$^{(\boxtimes)}$ and David Mba

Faculty of Technology, De Montfort University, Leicester, UK
{xiaochuan.li,david.mba}@dmu.ac.uk

Abstract. An adaptive multivariate process modelling approach is developed to improve the accuracy of traditional canonical variate analysis (CVA) in predicting the performance of industrial rotating machines under faulty operating conditions. An adaptive forgetting factor is adopted to update the covariance and cross-covariance matrices of past and future measurements. The forgetting factor is adjusted according to the Euclidean norm of the residual between the predicted model outputs and the actual measurements. The approach was evaluated using condition monitoring data obtained from an operational industrial gas compressor. The results show that the proposed method can be effectively used to predict the performance of industrial rotating machines under faulty operating conditions.

1 Introduction

Multivariate statistical techniques such as principal component analysis (PCA) (Harrou et al. 2013), partial least squares (PLS) (Yacoub and Macgregor 2004) and canonical variate analysis (CVA) (Ruiz-Cárcel et al. 2016) have been widely applied for the detection of abnormalities in large industrial systems. Multivariate subspace identification models based on PCA, PLS or CVA have attracted attention over the past decades because they can be utilized for process monitoring, modelling and system identification. The authors of (Juricek et al. 2005) demonstrated that system-identification models based on CVA outperform modelling models based on regression methods such as PLS. The authors of (Ruiz-Cárcel et al. 2015) demonstrated that monitoring methods based on CVA are more suitable for systems working under changing operating conditions compared to models based on PCA and PLS. The literature provides examples of extensive application of CVA for industrial process modelling and health monitoring. Li et al. (2018) developed a prediction method based on CVA and support vector machine for modelling of industrial reciprocating compressors. The authors of (Larimore et al. 1993) proposed a state-space method using canonical variable states for modelling linear and nonlinear time series. Negiz and Cinar (1997) used a CVA-based subspace-identification approach to describe a high-temperature short-time milk-pasteurization process. CVA was utilized in (Li et al. 2018) to predict performance deterioration and estimate the behaviour of a system under faulty operating conditions. The authors illustrated the performance of the

proposed method in a large-scale 3-phase flow facility. Conventional multivariate subspace-identification approaches based on PCA or PLS are based on the assumption that the process variables are linearly correlated and are independent and identically distributed (IID) (Choi et al. 2006). The requirement that process variables be IID and linear tends to limit the scope of many subspace-identification methods to linear processes operating under steady state conditions. Occasionally, problems associated with the effectiveness of the modelling tools can arise when the underlying assumptions are violated; for instance, the presence of nonlinear distortions, time-dependency, system dynamics and varying operating conditions. Therefore, it is necessary to develop adaptive subspace-identification approaches for systems in which variations in the mode of operation and changes in the system dynamics are common. A number of recursive monitoring methods have been proposed to address these limitations. An extension to the modelling approaches based on the conventional PCA method was proposed by Lane et al. in (2003). The authors illustrated the performance of the proposed recursive PCA model in a polymer film-manufacturing process. Choi et al. (Choi et al. 2006) developed an adaptive multivariate statistical process monitoring (MSPC) for the monitoring of dynamic processes where variations in operating conditions are incurred. The authors of (Lee and Lee 2008) proposed a recursive state-space model based on CVA. In that study, the norm of the difference between consecutive measurements was used to adjust forgetting factors, and the calculation of the optimal values of the minimum and maximum forgetting factors was not detailed.

In this paper, we develop an adaptive monitoring tool based on CVA for the modelling of time-varying processes. We explore the ability of adaptive CVA to predict the behaviour of industrial rotating machines under slowly evolving faulty conditions. To obtain an accurate estimate of system outputs, forgetting factors calculated based on the residual between the model outputs and actual measurements are adopted to update the covariance and cross-covariance matrices of the system. The proposed method is validated on industrial data captured from an operational gas compressor.

2   Methodology

Given system input time-series u_t and output time-series y_t, a linear state-space model can be built as follows (Qin 2006):

$$x_{t+1} = Bx_t + Cu_t + Ke_t \tag{1}$$

$$y_t = Dx_t + Eu_t + e_t \tag{2}$$

where u_t, y_t and x_t are system inputs, system outputs and state vectors; B, C, D, E and K are model coefficient matrices; and e_t is zero-mean and normally distributed independent white noise.

The objective of CVA is to maximize the correlation of two sets of variables (Russell et al. 2000). To generate two data matrices from the measurements, the measurement vector is expanded at each sampling time by including a, the number of

previous samples, and b, the number of future samples, to construct the past and future sample vectors $z_{a,t} \in \mathcal{R}^{(n_y+n_u)\cdot a}$ and $u_{b,t} \in \mathcal{R}^{n_u \cdot b}$ (n_y and n_u are the number of output variables and input variables).

$$z_{a,t} = \begin{bmatrix} y_{t-1} \\ u_{t-1} \\ y_{t-2} \\ u_{t-2} \\ \vdots \\ y_{t-a} \\ u_{t-a} \end{bmatrix} \in \mathcal{R}^{(n_y+n_u)\cdot a} \tag{3}$$

$$u_{b,t} = \begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+b-1} \end{bmatrix} \in \mathcal{R}^{n_u \cdot b} \tag{4}$$

The observations can be expanded at each sampling time t by including $a+1$ observations to form the extended past vectors $z_{a+1,t} \in \mathcal{R}^{(n_y+n_u)\cdot(a+1)}$:

$$z_{a+1,t} = \begin{bmatrix} y_{t-1} \\ u_{t-1} \\ y_{t-2} \\ u_{t-2} \\ \vdots \\ y_{t-a-1} \\ u_{t-a-1} \end{bmatrix} \in \mathcal{R}^{(n_y+n_u)\cdot(a+1)} \tag{5}$$

To avoid the domination of variables with large absolute values, the past and future sample vectors are normalized to the zero-mean vectors $\hat{z}_{a,t}$ and $\hat{u}_{b,t}$. Then, the vectors $\hat{z}_{a,t}$ and $\hat{u}_{b,t}$ at different sampling times are rearranged to produce the reshaped matrices $\hat{Z}_a$ and $\hat{U}_b$:

$$\hat{Z}_a = \left[\hat{z}_{a,t+1}, \hat{z}_{a,t+2}, \ldots, \hat{z}_{a,t+N} \right] \in \mathcal{R}^{(n_y+n_u)a\times N} \tag{6}$$

$$\hat{U}_b = \left[\hat{z}_{b,t+1}, \hat{z}_{b,t+2}, \ldots, \hat{z}_{b,t+N} \right] \in \mathcal{R}^{n_u a\times N} \tag{7}$$

where $N = l - a - b + 1$, and l represents the total number of samples for measurements y_t. Cholesky decomposition is then applied to the past and future matrices $\hat{Z}_a$ and $\hat{U}_b$ to configure a Hankel matrix $\mathcal{H}$. To find the linear combination that maximizes the correlation of the two sets of variables, the truncated Hankel matrix $\mathcal{H}$ is decomposed using singular value decomposition (SVD):

$$\mathcal{H} = \Sigma_{b,b}^{-1/2}\Sigma_{b,a}\Sigma_{a,a}^{-1/2} = U\Sigma V^T \tag{8}$$

where $\Sigma_{a,a}$ and $\Sigma_{b,b}$ are the sample covariance matrices and $\Sigma_{b,a}$ denotes the cross-covariance matrix of $\hat{Z}_a$ and $\widehat{U}_b$. $\Sigma_{a,a}$, $\Sigma_{b,b}$ and $\Sigma_{b,a}$ are calculated as follows (Odiowei and Yi 2010):

$$\Sigma_{a,a} = \hat{Z}_a\hat{Z}_a^T/(N-1) \tag{9}$$

$$\Sigma_{b,b} = \widehat{U}_b\widehat{U}_b^T/(N-1) \tag{10}$$

$$\Sigma_{b,a} = \widehat{U}_b\hat{Z}_a^T/(N-1) \tag{11}$$

U, V and $\sum$ have the following form:

$$U = [u_1, u_2, \ldots, u_r] \in \mathcal{R}^{n_u\mathrm{a}\times n_u\mathrm{a}}$$

$$V = [v_1, v_2, \ldots, v_r] \in \mathcal{R}^{(n_y+n_u)\mathrm{a}\times(n_y+n_u)\mathrm{a}}$$

$$\Sigma = \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_r \end{bmatrix} \in \mathcal{R}^{n_ua\times(n_y+n_u)\mathrm{a}}$$

The columns of $U = [u_1, u_2, \ldots, u_r]$ and the columns of $V = [v_1, v_2, \ldots, v_r]$ are called the left-singular and right-singular vectors of $\mathcal{H}$. Σ is a diagonal matrix, and its diagonal elements are called singular values and depict the degree of correlation between the corresponding left-singular and right-singular vectors. The right-singular vectors in V corresponding to the largest q singular values are retained in the truncated matrix $V_q = [v_1, v_2, \ldots, v_q] \in \mathcal{R}^{(n_y+n_u)\mathrm{a}\times q}$. This matrix is used later to perform dimension reduction on the measured data.

With the truncated matrix V_q, the (n_y+n_u) dimensional past vector $\hat{Z}_a \in \mathcal{R}^{(n_y+n_u)\mathrm{a}\times N}$ is further converted into a reduced q-dimensional matrix $\Phi \in \mathcal{R}^{q\times N}$ (the columns of Φ are z_t, which are called canonical state variates) by the following:

$$\Phi = [z_{t=1}, z_{t=2}, \ldots, z_{t=N}] = K \cdot \hat{Z}_a = V_q^T\Sigma_{a,a}^{-1/2} \cdot \hat{Z}_a \tag{12}$$

where $K = V_q^T\Sigma_{a,a}^{-1/2} \in \mathcal{R}^{q\times(n_y+n_u)\mathrm{a}}$ is the projection matrix that maps the past observations into the canonical variate space. In this investigation, the number of q is determined in the same way as the traditional CVA model. According to the literature (Odiowei and Yi 2010), if the number of retained states q is no less than the actual order of the system, we can substitute the state variates x_t with the canonical state variates z_t. Therefore, the state variables are defined as a linear combination of the past measurement vector $\hat{Z}_a$ (Lee and Lee 2008):

$$\hat{x}_{t+1} = [K \quad 0]z_{a+1,t} \tag{13}$$

$$\hat{x}_t = [0 \quad K]z_{a+1,t} \tag{14}$$

where $0 \in \mathcal{R}^{q \times (n_y + n_u)}$ with all zero entries. According to the literature (Shang et al. 2015), after the estimates of the state variates are calculated, the matrices B, C, D, E and K are calculated from the measurements through linear least-squares regression as follows:

$$[D \quad E] = Y_{(:,1:N-a-1)} \begin{bmatrix} \hat{X}_{(:,1:N-a-1)} \\ U_{(:,1:N-a-1)} \end{bmatrix}^{+} \tag{15}$$

$$[B \quad C \quad K] = \hat{X}_{(:,2:N-a)} \begin{bmatrix} \hat{X}_{(:,1:N-a-1)} \\ U_{(:,1:N-a-1)} \\ \hat{E}_{(:,1:N-a-1)} \end{bmatrix}^{+} \tag{16}$$

where $\hat{E} = Y - D\hat{X} - EU$.

Due to non-stationary process behaviour, many industrial processes have time-varying characteristics that may cause rapid changes in state variates over time. The sample covariance matrices $\Sigma_{a,a}$ and $\Sigma_{b,b}$ and the cross-covariance matrix $\Sigma_{a,b}$ change according to the change in operating conditions. Constant covariance and cross-covariance matrices may not be able be fully capture the system dynamics. Therefore, the exponential weighted moving-average method is employed in this investigation to update the matrices $\Sigma_{a,a}$, $\Sigma_{b,b}$ and $\Sigma_{b,a}$:

$$\Sigma_{a,a(t)} = (1 - \beta)z_{a,t}z_{a,t}^T + \beta\Sigma_{a,a(t-1)} \tag{17}$$

$$\Sigma_{b,b(t)} = (1 - \beta)u_{b,t}u_{b,t}^T + \beta\Sigma_{b,b(t-1)} \tag{18}$$

$$\Sigma_{b,a(t)} = (1 - \beta)u_{b,t}z_{a,t}^T + \beta\Sigma_{b,a(t-1)} \tag{19}$$

where β is the forgetting factor, which is calculated according to the Euclidean norm of the residual between the predicted model outputs and actual measurements. The initial values of $\Sigma_{a,a(t)}$, $\Sigma_{b,b(t)}$ and $\Sigma_{b,a(t)}$ are determined by the traditional CVA model. Tracking time-varying parameters is an important problem in subspace modelling. A constant forgetting factor is not suitable for tracking time-varying parameters and therefore cannot fully reflect the dynamics of a process under nonstationary conditions (Leung and So 2005). Therefore, the forgetting factor must be changed according to the rate of process change to yield satisfactory predictive results in time-varying environments. In this investigation, the forgetting factor is adjusted based on the Euclidean norm of the residual between the predicted model outputs and the actual measurements as adopted in (Shang et al. 2015). When the forgetting factor is small, it gives more weight to present observations to reduce the impact of past observations on the current model. As the value of the forgetting factor approaches unity, it gives more weight to

past measurements, thereby permitting long-term memory of the model. The forgetting factor used in this study is calculated as follows:

$$\beta_t = A e^{-(\|e_{t-1}\|)/\sigma_1} \tag{20}$$

where A is a constant. The empirical parameter selection procedure proposed by (Choi et al. 2006; Shang et al. 2015) is adopted in this study to determine the value of A. Typically, a value between 0.999 and 0.9 is selected. $\| e_{t-1} \|$ denotes the Euclidean norm of the residual of the actual measurements and model outputs. The tuning parameter σ_1 controls the sensitivity of the model to prediction errors. The larger σ_1 is, the less sensitive the model is to prediction error. The value of the forgetting factor β_t is adjusted at every time instance when new measurements are available.

After the forgetting factor is determined, weighted recursive least squares (WRLS) with adaptive forgetting factor (Leung and So 2005; Turksoy et al. 2014) can be used to update model coefficient matrices B, C, D, E and K. The system described by Eqs. 1 and 2 is rewritten as follows:

$$y_t = \Theta_t^y \emptyset_t + e_t \tag{21}$$

$$\hat{x}_{t+1} = \Theta_t^x \Psi_t \tag{22}$$

where $\Theta_t^y = [\, D_t \quad E_t \,], \emptyset_t = [\, \hat{x}_t^T \quad u_t^T \,]^T, \Theta_t^x = [\, B_t \quad C_t \quad K_t \,], \Psi_t = [\, \hat{x}_t^T \quad u_t^T \quad e_t^T \,]^T$. Θ_t^y can be calculated by using the recursive least squares (RLS) (Turksoy et al. 2013):

$$\Theta_t^y = \Theta_{t-1}^y + \left(y_t - \Theta_{t-1}^y \emptyset_t\right) \emptyset_t^T P_t \tag{23}$$

$$P_t = 1/\beta \left(P_{t-1} - \frac{P_{t-1} \emptyset_t \emptyset_t^T P_{t-1}}{\beta + \emptyset_t^T P_{t-1} \emptyset_t} \right) \tag{24}$$

The innovation noise sequence is defined as:

$$e_t = y_t - \Theta_t^y \emptyset_t \tag{25}$$

Similarly, Θ_t^x is calculated as follows:

$$\Theta_t^x = \Theta_{t-1}^x + \left(\hat{x}_{t+1} - \Theta_{t-1}^x \Psi_t\right) \Psi_t^T Q_t \tag{26}$$

$$Q_t = 1/\beta \left(Q_{t-1} - \frac{Q_{t-1} \Psi_t \Psi_t^T Q_{t-1}}{\beta + \Psi_t^T Q_{t-1} \Psi_t} \right) \tag{27}$$

The procedures for subspace identification and performance estimation using the model described above are summarized as follows:

Step 1: Calculate model coefficient matrices using the traditional CVA model.

Step 2: Calculate the forgetting factor β_t as per Eq. 20.

Step 3: Compute the updated covariance and cross-covariance matrices $\Sigma_{a,a(t)}$, $\Sigma_{b,b(t)}$ and $\Sigma_{b,a(t)}$ according to Eqs. 17–19.

Step 4: Update the Hankel matrix $\mathcal{H}$ as per Eq. 8.

Step 5: Estimate the state vectors $\hat{x}_{t+1}$ and $\hat{x}_t$ as per Eqs. 13–14.

Step 6: Update the model coefficient matrices via Eqs. 23–27.

Step 7: Estimate the model outputs y_t according to Eqs. 1–2.

Step 8: Update the forgetting factor β_t based on the residual between the estimated outputs and actual measurements.

Step 9: Repeat step 1 – step 8 iteratively.

3   Case Study

Rotating machines that operate at high speed and under high pressure are subject to performance degradation and failures. If a fault occurs and the fault evolution is slow, the machine operator may choose to keep the machine running until repair facilities and spare parts are available at the plant. In such a case, the proposed adaptive ACVA model can be used to estimate how the system will behave under faulty operating conditions given future system inputs. In this subsection, the proposed method is applied to an operational industrial centrifugal compressor to predict the performance of the machine during bearing degradation.

The measured time series from compressor A consisted of 368 observations and 13 variables. For this study, all data were captured at a sampling rate of one sample per hour. Table 1 summarizes all measured variables for this compressor. As shown in Fig. 1, the compressor is operated under healthy conditions during the first 320 samples. The readings of the four different bearing-temperature sensors start to rise at around the 321th sampling point; the machine continued to run until the 368th sampling point. At that time, site engineers shut down the compressor for inspection and maintenance. To compare the performance of the developed ACVA approach with that of the traditional CVA model, the first 240 sampling points of the monitored time series were utilized to build an offline CVA model. Then the constructed CVA model was fed with the speed set points used throughout the degradation process to estimate how the system was affected by the fault. On the other hand, the developed adaptive CVA approach was employed to update the constructed model iteratively according to steps 2–8 described in Sect. 2. The predicted outputs obtained from the adaptive CVA model were compared with those obtained from the traditional CVA model to evaluate the performance of the proposed adaptive monitoring method.

**Table 1.** Measured variables of compressor A

ID	Variable name	ID	Variable name	ID	Variable name
1	Speed	6	Radial vibration overall X 1	11	Radial bearing temperature 2
2	Suction pressure	7	Radial vibration overall Y 1	12	Active thrust bearing temperature
3	Discharge pressure	8	Radial bearing temperature 1	13	Inactive thrust bearing temperature
4	Discharge temperature	9	Radial vibration overall X 2		
5	Actual flow	10	Radial vibration overall Y 2		
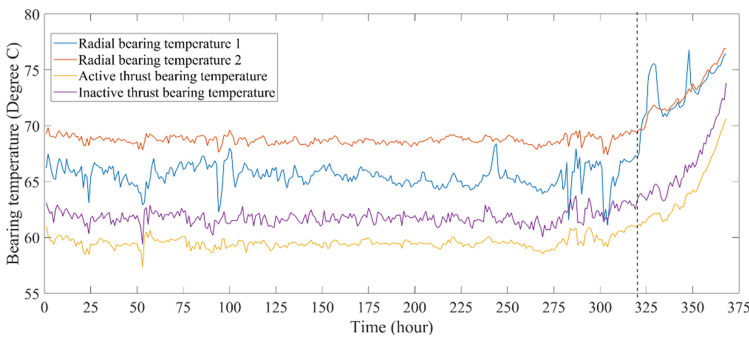
**Fig. 1.** Trend of four different bearing temperature sensor measurements of compressor A.

In order to determine the optimal number of retained states q, the trained offline CVA model was first utilized to predict system outputs for the data captured during the early stages of degradation. The predicted outputs are compared with the actual measurements and the mean absolute error (MAE) of all output variables are plotted against different values of retained state in Fig. 2. It can be observed from the figure that $q = 1$ gives the lowest prediction error; therefore, q was finally set to 1 to obtain the optimal model that gives the highest predictive accuracy. The value of A was set to 0.97 according to the empirical parameter-selection procedure proposed by (Choi et al., 2006; Shang et al. 2015). The value of the forgetting factor β_t can be updated at each time instance based on the difference between the predicted system outputs and the actual measurements. The value of β_t decreases to achieve faster identification with short memory when the residual of the system outputs is larger. When the residual is small, using more information about the past improves the prediction accuracy of the model. The tuning parameter σ_1 was set to 3.5 in this study, which is the minimum value that can ensure the convergence of the model while maximizing the sensitivity of the model to prediction errors.
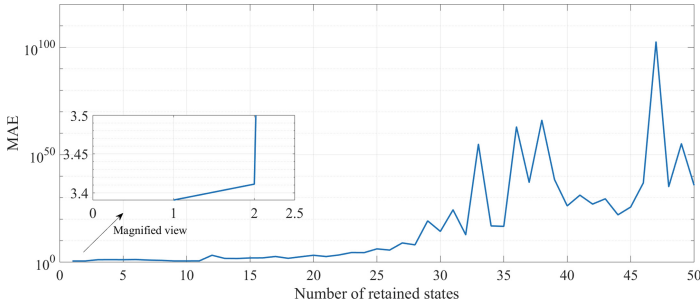
**Fig. 2.** MAE for all output variables for different values of retained state q

Figures 3, 4, 5 and 6 show the forecasted outputs of adaptive CVA and CVA model. The adaptive CVA model can track changes in bearing temperature measurements more accurately than the traditional CVA method. Table 2 summarises the mean absolute percentage error (MAPE) of the developed ACVA model and conventional CVA model. These results imply that the proposed method takes advantage of recursive state-space modelling to reveal the correlation between system input and output signals, thereby increasing the sensitivity of the adaptive CVA to bearing degradation compared to traditional CVA models.
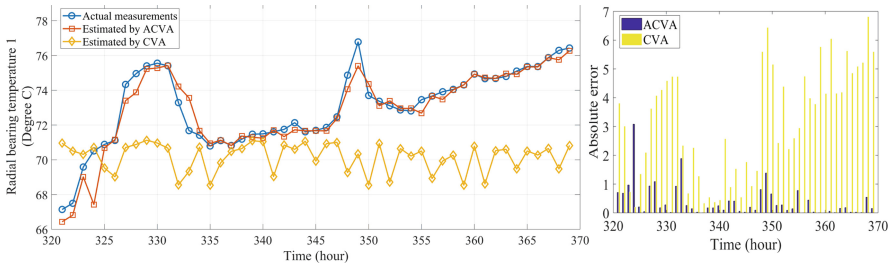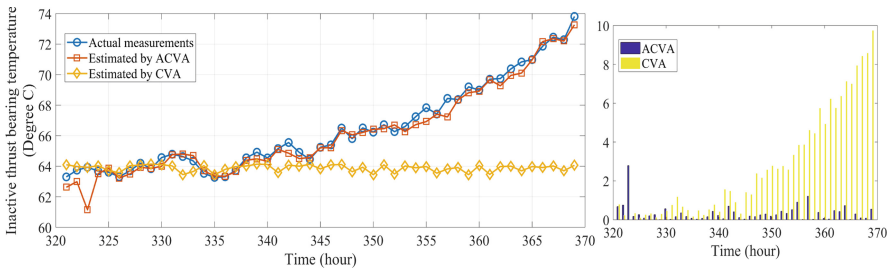

Fig. 3. Radial bearing temperature 1 under faulty operating conditions predicted by adaptive CVA and the CVA model
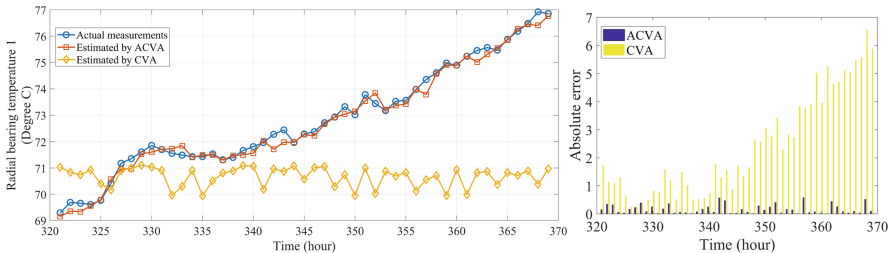
Fig. 4. Radial bearing temperature 2 under faulty operating conditions predicted by adaptive CVA and the CVA model
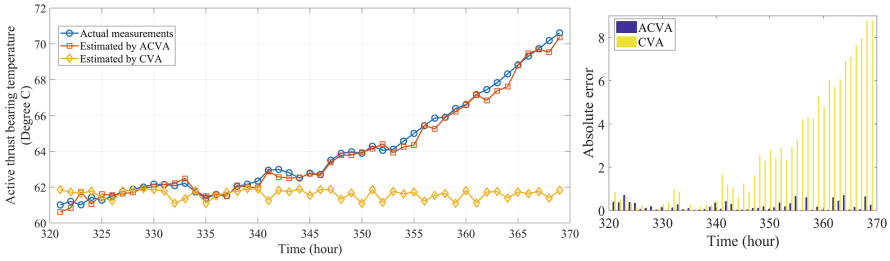
Fig. 5. Active thrust bearing temperature under faulty operating conditions predicted by adaptive CVA and the CVA model

Fig. 6. Inactive thrust bearing temperature under faulty operating conditions predicted by adaptive CVA and the CVA model

Table 2. Mean absolute percentage error (100%) for different output variables

MAPE	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}
CVA	1.5	0.2	4.8	3.7	11.7	18.9	1.2	7.3	3.5	0.2	1.7
ACVA	0.1	0	1	0.1	0.03	0.9	0.7	0.1	0.2	0.1	0.2

4   Conclusion

This paper proposes an adaptive CVA modelling tool to improve the predictive accuracy of traditional CVA methods. A variable forgetting factor was adopted to update the model coefficient matrices and covariance and cross-covariance matrices according to the residuals of the model outputs. The proposed model tracks rapid changes in system outputs due to the use of the adaptive forgetting factor. Condition-monitoring data captured from an operational industrial compressor were used to test the validity of the proposed method. The predicted outputs generated by adaptive CVA highly coincide with the actual measurements. The proposed method takes advantage of recursive state-space modelling to enhance the CVA prognostic performance and increase its sensitivity to bearing deterioration. This method can be used to provide site engineers with more reliable and robust performance estimates of systems operating

under varying and abnormal conditions. The information provided by the proposed method can be used to forecast the impact of a fault on the operational process and to develop appropriate production plans and optimal maintenance strategies, thereby making plant operations more safe, productive and profitable.

References

Choi, S.W., Martin, E.B., Morris, A.J., Lee, I.: Adaptive multivariate statistical process control for monitoring time-varying processes. Ind. Eng. Chem. Res. **45**(9), 3108–3118 (2006)

Harrou, F., Nounou, M.N., Nounou, H.N., Madakyaru, M.: Statistical fault detection using PCA-based GLR hypothesis testing. J. Loss Prev. Process Ind. **26** (2013). https://doi.org/10.1016/j.jlp.2012.10.003

Juricek, B., Larimore, W., Juricek, B.C., Seborg, D.E., Larimore, W.E.: Process control applications of subspace and regression-based identification and monitoring methods. In: American Control Conference, pp. 2341–2346 (2005). https://doi.org/10.1109/ACC.2005.1470316

Lane, S., Martin, E.B., Morris, A.J., Gower, P.: Application of exponentially weighted principal component analysis for the monitoring of a polymer lm manufacturing process. Trans. Inst. Measur. Control **1**, 17–35 (2003)

Larimore, W.E., Drive, F., Baillieul, J.: Identification and filtering of nonlinear systems using canonical variate analysis. In: Aerospace Control Systems, vol. 1, pp. 837–841 (1993)

Lee, C., Lee, I.B.: Adaptive monitoring statistics with state space model updating based on canonical variate analysis. Korean J. Chem. Eng. **25**(2), 203–208 (2008). https://doi.org/10.1007/s11814-008-0037-y

Leung, S.H., So, C.F.: Gradient-based variable forgetting factor RLS algorithm in time-varying environments. IEEE Trans. Sig. Process. **53**(8), 3141–3150 (2005). https://doi.org/10.1109/TSP.2005.851110

Li, X., Duan, F., Loukopoulos, P., Bennett, I., Mba, D.: Canonical variable analysis and long short-term memory for fault diagnosis and performance estimation of a centrifugal compressor. Control Eng. Practice **72**(January), 177–191 (2018). https://doi.org/10.1016/j.conengprac.2017.12.006

Li, X., Duan, F., Mba, D., Bennett, I.: Combining canonical variate analysis, probability approach and support vector regression for failure time prediction. J. Intell. Fuzzy Syst. 746–752 (2018). https://doi.org/10.1109/sdpc.2017.146

Negiz, A., Cinar, A.: Statistical monitoring of multivariable dynamic processes with state-space models. AIChE J. **43**(8), 2002–2020 (1997). https://doi.org/10.1002/aic.690430810

Odiowei, P.E.P., Yi, C.: Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. IEEE Trans. Ind. Inform. **6**(1), 36–45 (2010). https://doi.org/10.1109/TII.2009.2032654

Qin, S.J.: An overview of subspace identification. Comput. Chem. Eng. **30**(May), 1502–1513 (2006). https://doi.org/10.1016/j.compchemeng.2006.05.045

Ruiz-Cárcel, C., Jaramillo, V.H., Mba, D., Ottewill, J.R., Cao, Y.: Combination of process and vibration data for improved condition monitoring of industrial systems working under variable operating conditions. Mechanical Systems and Signal Processing, pp. 1–17 (2015)

Ruiz-Cárcel, C., Lao, L., Cao, Y., Mba, D.: Canonical variate analysis for performance degradation under faulty conditions. Control Eng. Practice **54**, 70–80 (2016). https://doi.org/10.1016/j.conengprac.2016.05.018

Russell, E.L., Chiang, L.H., Braatz, R.D.: Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. Chem. Intell. Lab. Syst. **51**(1), 81–93 (2000). https://doi.org/10.1016/S0169-7439(00)00058-7

Shang, L., Liu, J., Turksoy, K., Min Shao, Q., Cinar, A.: Stable recursive canonical variate state space modeling for time-varying processes. Control Eng. Practice **36**, 113–119 (2015). https://doi.org/10.1016/j.conengprac.2014.12.006

Turksoy, K., Bayrak, E.S., Quinn, L., Littlejohn, E., Cinar, A.: Adaptive multivariable closed-loop control of blood glucose concentration in patients with type 1 diabetes. In: 2013 American Control Conference (ACC), pp. 2905–2910. Washington, D.C. (2013)

Turksoy, K., Quinn, L., Littlejohn, E., Cinar, A.: Multivariable adaptive identification and control for artificial pancreas systems. IEEE Trans. Biomed. Eng. **61**(3), 883–891 (2014)

Yacoub, F., Macgregor, J.F.: Product optimization and control in the latent variable space of nonlinear PLS models. Chem. Intell. Lab. Syst. **70**, 63–74 (2004). https://doi.org/10.1016/j.chemolab.2003.10.004