Application of Cointegration Testing Method to Condition Monitoring and Fault Diagnosis of Non-Stationary Systems

Qian Chen†, Uwe Kruger‡, Yuyu Pan†

†College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, 29 Yudaoo Street, Nanjing 210016, P.R. China
‡Intelligent Systems and Control Research Group, Queen’s University Belfast, Belfast, BT9 5AH, U.K.

Abstract: This paper introduces cointegration testing method for monitoring multivariate non-stationary processes. Compared with Box-Jenkins models, cointegration testing method is able to model long-run equilibrium relationships and circumvent the effect of forecast recovery. Cointegration has been successfully utilised to describe such relationships in economics systems, which are non-stationary in nature. Cointegration method can identify a stationary process using a linear combination of the non-stationary processes. This stationary process describes a long-run equilibrium relation between the non-stationary process variables, which provides a possibility to apply the cointegration testing method to the modelling and monitoring of engineering systems. A simulation example from the literature shows the capabilities and benefits of cointegration testing method for monitoring non-stationary processes. Copyright © 2006 USTARTH

Keywords: Cointegration testing, non-stationary systems, statistical process control, fault detection.

1. INTRODUCTION

Non-stationary dynamic processes are a frequent occurrence in engineering systems. As highlighted in the research literature, non-stationary systems are difficult to monitor, where existing work predominantly relates to the application of Auto-Regression Integrated Moving Average (ARIMA) model or the Wavelet or Fourier analysis.

Box, et al, (1994) showed that (i) non-stationary sequence is characterised by a varying mean and (ii) differencing the original non-stationary sequences converts a non-stationary to a stationary sequence, i.e. a sequence that has a constant mean. Hence, non-stationary sequence can be modelled by an ARIMA(p,d,q) model, where a total of d differentiation have been performed and the stationary sequence is described by p lagged auto-regressive and q lagged moving average terms.

Berthouex and Box (1996) and Wikstrom et al, (1998) employed ARIMA models to generate a set of stationary signals for modelling a multivariate system for process monitoring. However, as pointed out by Superville and Adams (1994) and Apley and Jianjun (1999), ARIMA models affect the signature and magnitude of fault conditions, an effect known as “forecast recovery”.

Backshi and Stephanopoulos (1996) proposed an signal-based approach using the wavelet transformation and monitored parameter changes of Haar wavelets. Signal-based techniques, however, may be dif-
ficult to implement in a multivariate context if the process variables are interrelated, since such relations cannot be modelled using a single-variable analysis.

As summarised above, although non-stationary sequences can be described by multivariate ARIMA models, the inherent differentiation can have a profound impact upon the fault magnitude and signature. This paper utilises cointegration testing method for modelling the relationships between the process variables by identifying a cointegrating coefficients. These coefficients are utilised to form a linear combination of the non-stationary process variables. This linear combination generates a stationary sequence or equilibrium residual sequence that describes the long-run equilibrium relationship between the non-stationary sequences. The residual sequences can then be utilised in a statistical process control framework for process monitoring purpose.

Since cointegration testing method does not rely on the statistical analysis of differentiated sequences, the deficiency of multivariate ARIMA models, i.e. the isolation of the correct fault magnitude and signature, are circumvented. The concept of cointegration testing was developed by Engle and Granger (1987) and has been successfully applied in econometrics and, in particular, to model the relationships between power consumption and economic development (Chen and Hurvich, 2003; Hansen, 2003; Kaufman, 2004; Galindo, 2005; Hassler, et al., 2006).

To verify the applicability of the cointegration testing method to engineering systems, [Pan and Chen, 2006] employed it for a car engine condition monitoring and obtained some encouraging results. In this paper an example of a simulated nonstationary Fluid Catalytic Cracking Unit (FCCU) system was discussed. The results show that the cointegration testing method might have a big potential in nonstationary process monitoring and fault diagnosis.

2. COINTEGRATION PRELIMINARIES

According to Box et al. (1994), if a non-stationary process variable \( \eta_t \) becomes stationary after being differentiated \( d \) times it is said to be integrated of order \( d \), denoted by \( \eta_t \sim I(d) \). Engle and Granger, (1987) considered a set of such \( d \)-integrated non-stationary variables to possess in long-run equilibrium if and only if:

\[
\beta_1 y_{1t} + \beta_2 y_{2t} + \cdots + \beta_d y_{dt} = \xi_t
\]

(1)

Here, the equilibrium residual is defined as

\[
\xi_t = \beta'y_t
\]

(2)

where \( \beta = (\beta_1, \beta_2, \cdots, \beta_d)' \in \mathbb{R}^d \) and \( y_t = (y_{1t}, y_{2t}, \cdots, y_{dt})' \in \mathbb{R}^d \). In general, \( y_t \) are an integrated non-stationary stochastic variables and \( \xi_t \) is a stationary sequence that is assumed to describe the departure of \( y_t \) from its equilibrium. This implies that there exists a vector \( \beta \) that, according to Eqn. (1), produces a \( \xi_t \) that is of order 0, i.e. \( \xi_t \sim I(0) \). More precisely, the Engle-Granger method identifies a specific coefficient vector \( \beta \) such that the residuals of the equilibrium relationship are stationary and stochastic.

2.1 Definition of cointegration

The definition of cointegration is, according to Engle and Granger (1987), that the non-stationary sequences, stored in \( y_t \), are cointegrated of order \( (d, b) \), denoted by \( y_t \sim I(d-b) \) if and only if:

1. All variables in \( y_t \) are integrated of order \( d \).
2. There exists a vector \( \beta \) such that the linear combination \( \beta'y_t \) is integrated of order \( I(d-b) \), where \( d \geq b > 0 \) and \( b \) being the number of reduced order integration.

Here, \( \beta \) is known as cointegrating vector and if \( b = d \), then \( \beta'y_t \) is, according to Eqns. (1) and (2) stationary.

If the elements in \( y_t \) are integrated of order \( d \) and cointegrated of order \( (d,b) \), there may be as many as \( n-1 \) linearly independent cointegrating vectors. If multiple cointegrating relationships are found, it may be difficult to identify a behavioural relationship. For statistical process monitoring, a cointegrating vector corresponding to a most stationary residual series should be selected. This can be accomplished using the Dickey-Fuller (DF) statistic (Dickey and Fuller, 1979). This hypothesis test relies on values of the DF statistic, which are compared with a given criteria, usually with a significance of 5%.

2.2 Cointegration and common trends

Cointegrated variables share common trends. For simplicity, consider only two variables

\[
y_{1t} = w_{1t} + \epsilon_{1t}
\]

(3a)

\[
y_{2t} = w_{2t} + \epsilon_{2t}
\]

(3b)

Each of the \( w_t \)'s is a random walk and each of the \( \epsilon_t \)'s is a stochastic and stationary variable. If variables \( y_{1t} \) and \( y_{2t} \) are cointegrated of order \( (1,1) \) there must be nonzero values of \( \beta_1 \) and \( \beta_2 \) for which the linear combination \( \beta_1 y_{1t} + \beta_2 y_{2t} \) is stationary. Thus

\[
\beta_1 y_{1t} + \beta_2 y_{2t} = \beta_1 (w_{1t} + \epsilon_{1t}) + \beta_2 (w_{2t} + \epsilon_{2t})
\]

(4)

\[
= (\beta_1 w_{1t} + \beta_2 w_{2t}) + (\beta_1 \epsilon_{1t} + \beta_2 \epsilon_{2t})
\]

is stationary if and only if \( \beta_1 w_{1t} + \beta_2 w_{2t} \) vanishes:

\[
(\beta_1 w_{1t} + \beta_2 w_{2t}) = 0, \quad \forall t
\]

(5)

Hence, if \( \beta_1 w_{1t} + \beta_2 w_{2t} \neq 0 \) then \( \beta_1 w_{1t} = -\beta_2 w_{2t} \). Since \( w_{1t} \) and \( w_{2t} \) being random walks, i.e. non-stationary stochastic trends, they (i) are equal up to a scaling factor \(-\beta_2/\beta_1 \), (ii) stochastic non-stationary \( I(1) \) sequences that have the same stochastic trend and (iii) cointegrated of order \( (1,1) \). The parameters of \( \beta \) must consequently suppress the non-stationary trend encapsulated in the linear combination of \( y_{1t} \) and \( y_{2t} \).

Practically, more detailed time series structures, including a constant shift and/or deterministic trends, can be employed. For example, Watson (1994) utilised AR(2) models of the form:

\[
y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \epsilon_t
\]

(6a)

\[
y_t = \alpha + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \epsilon_t
\]

(6b)

\[
y_t = \alpha + \mu t + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \epsilon_t
\]

(6c)

For industrial processes, the measured variables often show common stochastic non-stationary and/or
For simplicity consider a non-stationary time series defined by Eqn (6a) which has no deterministic trend. The AR(2) model of the time series is

\[ y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t \] (7a) 

or

\[ y_t = (\rho_1 + \rho_2) y_{t-1} - \rho_2 \Delta y_{t-1} + \varepsilon_t \] (7b)

where, \( \varepsilon_t \sim N(0, \sigma^2) \) is assumed to be a white noise sequence and the parameters \( \rho_1 \) and \( \rho_2 \) are estimated using Ordinary Least Square (OLS). Assuming that \( |\rho_2| < 1 \), the hypothesis \( H_0: \rho_1 + \rho_2 = 0 \) can be tested to determine whether \( y_t \) has a unit root. If so, Eqn (7b) becomes

\[ \Delta y_t = -\rho_2 \Delta y_{t-1} + \varepsilon_t \] (8)

since \( |\rho_2| < 1 \), \( \Delta y_t \) is a stationary, zero mean sequence. More precisely, \( y_t \) is an I(1) non-stationary sequence. If \( H_0 \) is rejected, carry on the test on the AR(2) model using \( z_t = \Delta y_t \) and construct Eqsns (7a) and (7b) using \( z_t \) instead of \( y_t \). If \( H_0 \) is now accepted, then \( y_t \) is an I(2) nonstationary sequence. Conversely, if \( H_0 \) is rejected again repeat this test using higher order differentials of \( y_t \) until \( H_0 \) is accepted. The hypothesis whether \( \rho_1 + \rho_2 = 1 \) can be tested using the \( t \)-statistic:

\[ t_0 = \frac{\hat{\rho}_1 + \hat{\rho}_2}{s.e(\hat{\rho}_1 + \hat{\rho}_2)} \] (9)

where \( s.e(\cdot) \) refers to the standard deviation for the estimated parameters. As discussed in Dickey and Fuller (1979), this statistic follows a DF distribution. Based on Eqn. (9), the hypothesis for the \( t \)-statistic is tested with a confidence of 1%, 5% or 10% and associated confidence limits can be obtained from a DF table.

3.2 Augmented Dickey-Fuller test

Since it can not be guaranteed that \( \varepsilon_t \) is a white noise sequence, the parameter estimation is biased. This was addressed by proposing an augmented DF (ADF) test. Consider the deterministic trend described by the AR(2) model in Eqn. (6c)

\[ \Delta y_t = \alpha + \mu - \rho_2 \Delta y_{t-1} + (\rho_1 + \rho_2) y_{t-1} + \varepsilon_t \] (10)

After estimating the parameters of the model and obtaining the \( t \)-statistic, the ADF test can be carried out to test the hypotheses \( H_0: \rho_1 + \rho_2 = 1, H_0: \mu = 0 \) and \( H_0: \alpha = 0 \). Associated confidence limits to evaluate the deterministic trends for \( \rho_1 + \rho_2, \mu \) and \( \alpha \) can be obtained from an ADF criterion table.

4. COINTEGRATION TESTS

4.1 Engle-Granger test

For a bivariate case, \( y_{1t}, \) and \( y_{2t}, \) Engle and Granger (1987) develop a two-step testing method to examine if they are cointegrated.

Step1: Estimate the parameter of the regression of \( y_{1t} \) on \( y_{2t} \) using OLS

\[ y_{1t} = \hat{\beta} y_{2t} + \varepsilon_t \] (11)

The prediction of \( y_{1t} \) and the estimation of \( \beta \) are obtained as

\[ \hat{y}_{1t} = \hat{\beta} y_{2t} \] (12)

where the prediction error is given by

\[ \hat{\varepsilon}_t = y_{1t} - \hat{y}_{1t} = y_{1t} - \hat{\beta} y_{2t} \] (13)

Step2: Test \( \hat{\varepsilon}_t \sim i(0). \) If \( H_0 \) is accepted, \( y_{1t} \) and \( y_{2t} \) are cointegrated.

4.2 Johansen test

If the number of variable exceeds two, the above Engle-Granger test cannot be applied. Johansen, (1988) and Johansen and Juselius (1990) introduced a cointegration test method based on Vector Auto-Regression (VAR) model applicable in a general multivariate context. This test should be performed after confirming that the variables are non-stationary and \( I(1) \). Given a the data vector \( y_t = (y_{1t}, y_{2t}, \ldots, y_{mt})^T \in \mathbb{R}^m \sim i(1) \), in which the elements \( y_{it} \) are represented by non-stationary stochastic variables, the Johansen test is based on an unrestricted VAR model. This implies that an unrestricted VAR model must be constructed prior to multivariate cointegration tests and the identification of the corresponding cointegration coefficients.

The working of this test is best motivated using the following example. When dealing with a trivarial economic system based on the logarithm of nominal wages, prices and labour productivity, two cointegrating relationships could intuitively be isolated: determining (i) the employment equation and (ii) the wage equation.

In general, a cointegration model of \( n \) nonstationary variables is given by
\[ \xi_t = \beta^T y_t, \]  
(14)

where \( \xi_t \) is a stationary residual. A Vector Error Correction Model (VECM) forms the basis for the Johansen testing. Consider the VAR(\( p \)) model of a vector of \( n \) non-stationary variables (Engle and Granger, 1987):

\[ y_t = \sum_{i=1}^{n} \Pi_i y_{t-i} + \epsilon_t, \]  
(15)

To derive the VECM, subtract \( y_{t-1} \) from both sides of Eqn. (15) to produce:

\[ \Delta y_t = \Pi y_{t-1} + \sum_{k=1}^{p-1} \Phi_k \Delta y_{t-k} + \epsilon_t, \]  
(16)

where

\[ \Pi = -I_r + \sum_{i=1}^{p} \Pi_i = -\Pi(1) \]  
(17a)

\[ \Phi_i = -\sum_{j=1}^{n} \Pi_j, \quad k = 1, 2, \ldots, p - 1 \]  
(17b)

Now, \( \Pi(1) \) is decomposed into a product of three matrices, i.e.

\[ \Pi(1) = U(1) M(1) V(1) \]  
(18)

Here, \( M(1) \) is a diagonal matrix of rank \( r \). Furthermore, \( \Pi = -\Pi(1) \) which is also of rank \( r \). Let \( \beta \) denote an \( n \) by \( r \) matrix whose columns construct a basis of the row space for \( \Pi \), that is \( \Pi = \delta \beta^T \), where \( \delta \) is an \( n \) by \( r \) matrix with full column rank. Eqn. (16) becomes

\[ \Delta y_t = \delta \beta^T y_{t-1} + \sum_{k=1}^{p-1} \Phi_k \Delta y_{t-k} + \epsilon_t, \]  
(19a)

or

\[ \Delta y_t = \delta \xi_{t-1} + \sum_{k=1}^{p-1} \Phi_k \Delta y_{t-k} + \epsilon_t, \]  
(19b)

Here, \( \xi_t = \beta^T y_t \) and \( \xi_{t-k} \) is given by

\[ \xi_{t-k} = [\delta \delta^T]^{-1} \delta^T \left[ \Delta y_t - \sum_{i=1}^{k} \Phi_i \Delta y_{t-i} - \epsilon_t \right]. \]  
(20)

From Eqn. (19b), it can be concluded that the elements of \( \xi_{t-k} \) are \( k(0) \) and consequently represent stationary sequences. Thus, the scaled sum of the \( k(1) \) elements in \( y_t \) produce the stationary \( k(0) \) sequence \( \xi_t \). Furthermore, the column vectors of the matrix \( \beta \) are defined as the cointegrating vectors.

The VECM imposes \( k < n \) unit roots in the VAR model by including first differentials of \( y_{t-1} \) and \( r = n - k \) linear combinations of the non-stationary variables. The zero hypotheses should be

\[ H_0: \text{Rank}(\Pi) \leq r \quad \text{or} \quad \Pi = \alpha \beta^T \]  
(21)

The case \( \beta^T y_t = 0 \) could be interpreted as the “equilibrium” of a dynamic system and \( \xi_t \) as the “equilibrium errors”. In fact, Eqn. (20) describes the self correcting mechanism of the system from its “non-equilibrium” state.

The identification of \( \delta, \beta \) and \( \Pi \) is constraint such that \( \delta \) and \( \beta \) can be individually identified. The matrix \( \beta \) can be identified on the basis of the maximum likelihood (ML) method using the probability density function of \( \Delta y_t \) at instance \( t \)

\[ f(\Delta y_t) = \prod_{i=1}^{n} f(\Delta y_t) = \]  
(22)

\[ (2\pi)^{-n/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2} \epsilon_t^T \Omega^{-1} \epsilon_t \right\} \]

where \( \epsilon_t \sim \mathcal{N}(0, \Omega) \). \( \Omega \) is also a parameter matrix to be identified for the VAR model. Including a total of \( T \) samples Eqn. (22) becomes

\[ f(\Delta y_T) = \prod_{i=1}^{T} f(\Delta y_t) = \]  
(23)

\[ f(\Delta y_T) = (2\pi)^{-Tn/2} |\Omega|^{-T/2} \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T} \epsilon_t^T \Omega^{-1} \epsilon_t \right\} \]

Applying the natural logarithm function of Eqn. (23) produces

\[ L(\Phi_1, \ldots, \Phi_p; \Pi, \Omega) = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln|\Omega| - \frac{1}{2} \sum_{t=1}^{T} \epsilon_t^T \Omega^{-1} \epsilon_t \]  
(24)

It should be noted that the ML estimation of \( \Pi \) is cumbersome. Computationally beneficial is an estimation of \( \Pi \) by a nonlinear least squares regression of \( \Delta y_t \) using lagged values. In contrast, the maximisation of Eqn. (24) to identify \( \Phi_1, \Phi_2, \ldots, \Phi_p; \Pi \) is a computationally inexpensive OLS regression of \( \Delta y_t = \alpha \beta^T y_{t-1} \). This produces the residuals \( \hat{\epsilon}_t \), followed by a regression of \( y_{t-1} \) on the lagged differentials of \( \Delta y_t \) generating the residuals \( \hat{\epsilon}_0 \). It can be shown that another regression model

\[ \hat{\epsilon}_0 = \Pi \hat{\epsilon}_1 + \epsilon_t \]  
(25)

can be obtained based on the theorem of uniqueness of vector orthogonal projection. Substituting Eqn. (25) into Eqn. (24) leads to

\[ L(a, \beta, \Omega) = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln|\Omega| - \frac{1}{2} \sum_{t=1}^{T} (e_t - \alpha \beta^T \epsilon_t)^T \Omega^{-1} (e_t - \alpha \beta^T \epsilon_t) \]  
(26)

Following the preceding discussion, Johansen (1995) showed that determining the rank of \( \hat{\Pi} \) is equivalent to the test for the number of canonical correlations between \( \hat{\epsilon}_0 \) and \( \hat{\epsilon}_1 \) that are different from zero. This can be conducted using one of the following two test statistics

\[ \hat{\lambda}_p (r) = -T \sum_{j=r+1}^{T} \ln(1 - \hat{\lambda}_j) \quad (27a) \]

\[ \hat{\lambda}_{\max} (r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1}) \]  
(27b)

where \( \hat{\lambda}_j \) are the eigenvalues of the matrix \( S_{10} \) with respect to the matrix \( S_{11} \), arranged in descending order \( 1 > \hat{\lambda}_1 > \cdots > \hat{\lambda}_r > 0 \), where

\[ S_{ij} = \frac{1}{T} \sum_{t=1}^{T} \hat{\epsilon}_{ij} \hat{\epsilon}_{ij}, \quad i, j = 0, 1. \]  
(28)

The eigenvalues can be obtained based on the determinant of the following equation

\[ \det[S_{11} - S_{10} S_{00}^{-1} S_{01}] = 0 \]  
(29)

The statistic in Eqns. (27), known as the trace statistic, tests the null hypothesis that the number of
cointegrating vectors is less than or equal to $r$ against a general alternative. Knowing that $\ln(1) = 0$ and $\ln(0) \to \infty$, both statistics are equal to zero if $\hat{\lambda}_i$ is zero, and the larger the eigenvalues the smaller $\ln(1 - \hat{\lambda}_i)$ becomes. Likewise, the statistic in Eqn. (26), known as the maximum eigenvalue statistic, can be tested to obtain $r$, $H_0$, against the alternative, $H_1$, of $r + 1$ cointegrating vectors. As above, if $\hat{\lambda}_{i,1}$ is close to zero, the statistic produces a small value.

Furthermore, if the null hypothesis is accepted, the $r$ cointegrating vectors contained in matrix $\beta$ can be estimated as the first $r$ columns of matrix $\hat{V} = (\hat{v}_1, \cdots, \hat{v}_r)$ which contains the eigenvectors associated to the eigenvalues in Eqn. (29):

$$\left( \hat{\lambda}_i S_{11} - S_{10} S_{00}^{-1} S_{01} \right) \hat{v}_i = 0, \quad i = 1, 2, \cdots, n \quad (30)$$

Once an estimate of $\beta$ has been obtained, the estimates of $\delta$, $\Pi_i$, and $\xi_t$ in Eqn. (19a) can be obtained by inserting this estimate in the corresponding OLS formulation. However, if only the cointegration vectors are of interest rather than the complete VECM model, it is sufficient to compute the eigenvectors of Eqn. (30).

As outlined above, only one cointegrating vector, i.e. one of the eigenvectors of Eqn. (30), is required to construct a monitoring model. To identify the most suitable candidate, the ADF test can be performed on lagged values of $\hat{\varepsilon}_t$.

### 4.3 Identifying cointegration models

The identification of a cointegration model, that is the identification of $\beta$, involves the following steps:

1. Use ordinary least squares (OLS) to identify an AR model for each variable to be tested.
2. Carry out a unit root test for each model to examine whether these are integrated of order $I(1)$.
3. Use OLS to fit two VAR models for $\Delta y_1$ and $y_1$ on lagged values of $\Delta y_1$ to obtain two residual vectors to form the eigenvalue equation.
4. Solve the eigenvalue problem to obtain the cointegrating vectors $\hat{\beta}_i$ to generate the equilibrium residuals $\hat{\xi}_i^T$, and examine the residuals to determine the best cointegrating vector.

After establishing a cointegration model using the above steps, the application of this model for fault detection and diagnosis is as follows:

1. Apply the established model to new data and monitor the residuals to detect any abnormal variations in the equilibrium residual $\hat{\xi}_i$.
2. If anomalous behaviour has been detected, diagnose this event by invoking reduced cointegration models as discussed next.

### 5. APPLICATION STUDY

To demonstrate the working of cointegration testing method in a condition monitoring context, an application study to the simulation example in Ku et al. (1995) is summarised below. This example includes two input and two output variables, $u_1 = v_1 + e_1$ and $y_1 = x_4 + f_4$, respectively. Here, $u_1$ and $y_1$ are the measured input and output variables, $x_4$ is calculated by a $1^\text{st}$ order ARX model using $x_{1,1}$ and $v_{1,1}$ and $e_3$ and $f_3$ are independently distributed stationary white noise sequences. Furthermore, the sequences $v_1$ were non-stationary and produced from a multivariate ARIMA model.

From the above process, two data sets of 1000 samples each were generated, where the first set served as a reference set for determining the cointegration model and the second one included a step type fault of magnitude 1.8 superimposed to the $1^\text{st}$ output variable after 600 samples were simulated.

For fault detection, a cointegration model was established using the steps summarised in Section 4.3. The identified cointegration vector was

$$\hat{\beta} = (1 \quad -0.17 \quad 0.0148 \quad -1.1519)^T \quad (31)$$

Defining the vector $z = (y_1 \ y_2 \ u_1 \ u_2)^T$, the stationary residual variable, $\hat{\xi}_i$, was given by $\hat{\xi}_i = z_i - 0.17 z_2 + 0.0148 z_3 + 1.1519 z_4$. The residual vector was now determined for both data sets and Fig. 1 shows that the step-type fault could correctly be detected.

For fault diagnosis, the following 4 reduced models were obtained to isolate the impact of the fault to a single variable.
\[\begin{align*}
\xi_{t1} &= z_t - 0.4749z_{t-1} + 0.5247z_{t-2} \\
\xi_{t2} &= z_t + 0.0014z_{t-5} + 0.9959z_{t-4} \\
\xi_{t3} &= z_t - 0.3717z_{t-1} + 1.3649z_{t-4} \\
\xi_{t4} &= z_t + 2.1614z_{t-2} + 1.1655z_{t-3}
\end{align*}\] (32)

After the fault was detected, these reduced models were applied in turn. Fig. 2 shows the resultant residual sequences for each model and clearly illustrates that omitting \(y_1\) no fault information is manifested in the associated residual sequence. In contrast, each of the 3 remaining reduced models, as well as the original model, produced an offset of 1.8 after 600 samples.

6. CONCLUSIONS

This paper introduces cointegration testing method for condition monitoring of multivariate nonstationary processes. This technique establishes a linear combination of the nonstationary variables such that a stationary residual sequence is produced that can be integrated into the statistical process control framework for abnormality detection in process.

An application example to a simulation from the literature showed that cointegration testing method can detect a fault condition and correctly identify the variable affected by the fault. In contrast to existing work, forecast recovery and the limitations of signal-based techniques are therefore overcome.

It has been found that different cointegration models based on the \(r\) cointegrating vectors have different sensitivities to the faults. Therefore, it should be investigated how to utilise a total of \(r\) cointegration models for process monitoring and diagnose more complex fault scenarios that affect several variables.

REFERENCES


