A data-based technique for monitoring of wound rotor induction machines: A simulation study

Fouzi Harrou a,*, Jacques F. Ramahaleomiarantsoa b, Mohamed N. Nounou c, Hazem N. Nounou d

a King Abdullah University of Science and Technology (KAUST), Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, Thuwal 23955-6900, Saudi Arabia
b Université de Corse, U.M.R. CNRS 6134 SPE, BP 52, Corte, France
c Chemical Engineering Program, Texas A&M University at Qatar, Doha, Qatar
d Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar

ARTICLE INFO

Article history:
Received 24 October 2015
Revised 20 April 2016
Accepted 27 April 2016
Available online xxxx

Keywords:
Fault detection
Wound rotor induction machines
Principal components analysis
EWMA control scheme
Hotelling T² statistic
Q statistic

ABSTRACT

Detecting faults in induction machines is crucial for a safe operation of these machines. The aim of this paper is to present a statistical fault detection methodology for the detection of faults in three-phase wound rotor induction machines (WRIM). The proposed fault detection approach is based on the use of principal components analysis (PCA). However, conventional PCA-based detection indices, such as the $T^2$ and the $Q$ statistics, are not well suited to detect small faults because these indices only use information from the most recent available samples. Detection of small faults is one of the most crucial and challenging tasks in the area of fault detection and diagnosis. In this paper, a new statistical system monitoring strategy is proposed for detecting changes resulting from small shifts in several variables associated with WRIM. The proposed approach combines modeling using PCA modeling with the exponentially weighted moving average (EWMA) control scheme. In the proposed approach, EWMA control scheme is applied on the ignored principal components to detect the presence of faults. The performance of the proposed method is compared with those of the traditional PCA-based fault detection indices. The simulation results clearly show the effectiveness of the proposed method over the conventional ones, especially in the presence of faults with small magnitudes.

© 2016 The Authors. Publishing services by Elsevier on behalf of Karabuk University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The history of fault detection and diagnosis of electrical motors goes back to almost the date of their invention. Manufacturers of motors have been the first ones who have investigated simple protection techniques such as the over current, over voltage, and ground fault protection schemes [1]. The increase in the complexity and the importance of motors has generated a corresponding significant progress in the field of fault detection and diagnosis [2,3]. The problem of fault detection in electrical machines has been the subject of research and investigation in various applications, such as electric vehicles [4], wind turbines [5], and many others.

Previously, DC and synchronous machines were commonly used in industrial applications, and thus they were the focus of the reliability-related research. However, with technological and economic developments and the advancements in power electronics, the squirrel cage and the wound rotor induction machines have taken their place in several applications [6], such as transportation, energy production and electrical drives due to their robustness, reliability and lower costs. Although improvements have been made, these machines still remain subject to potential stator and rotor failures [7,8]. Thus, monitoring of these machines is essential for their proper and safe operation.

Proper system monitoring can help minimize their downtime, improve their safety of operation, and reduce their manufacturing costs. Monitoring can be defined as the set of actions carried out to detect and isolate faulty measurement sources and then remove these faults before they affect the process performance [9]. The role of detection is to identify any fault quantified by a change from the nominal behavior of the system. Fault isolation, on the other hand, determines the location of the detected fault. In this paper, the focus will be on fault detection and its application to wound rotor induction machines (WRIM). If faults in WRIM are not detected in time or if they are allowed to propagate further, they may lead to serious failures.

Over the past few decades, various monitoring techniques for induction machines were reported in the literature, and they can

http://dx.doi.org/10.1016/j.jestch.2016.04.008
2215-0986/© 2016 The Authors. Publishing services by Elsevier on behalf of Karabuk University.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

be broadly classified into three main categories: data-based or model-free techniques [7,10,11], model-based techniques [12–14], and knowledge-based techniques [15–17]. Knowledge-based fault detection is usually a heuristic process [18]. Model-based fault detection methods, on the other hand, rely on comparing the system measured variables with information obtained from a mathematical model, which is usually developed based on some fundamental understanding of the system under fault-free conditions. In contrast to the model-based approaches, where a priori knowledge about the inspected system is needed, in data-based methods, only the availability of historical process data is required [19]. Since they do not explicitly require process models, data-based methods are usually more attractive to practical applications with complex systems. However, the performance of data-based methods mainly depends on the availability and quality of the required data.

In this paper, a PCA-based exponentially weighted moving average (EWMA) fault detection scheme is proposed for monitoring wound rotor induction machines (WRIM). PCA is a well known data-based multivariate statistical technique and has received important attention in last few years [20–22]. PCA is a linear dimensionality reduction modeling technique, which is very helpful when dealing data sets having a high degree of cross correlation among the variables. The central idea of PCA is to reduce the dimensionality of highly correlated data, while retaining the maximum possible amount of variability present in the original data set [23]. This reduction is achieved by transforming correlated variables into the set of new uncorrelated variables which are called principal components (PCs), each of which is a linear combination of the original variables. The PCA by reducing the dimension of the process variables is able to eliminate noise and retain only important process information, and can be employed to compress noisy and correlated measurements into a smaller informative subspace for measurement data sets. PCA-based anomaly detection have been widely used in practice because they need no prior knowledge about the process model, the only information needed is a good historical database describing the normal process operation [24]. Unfortunately, the conventional PCA-based monitoring indices, such as $T^2$ and $Q$ statistics, often fail to detect small or moderate changes [25,26]. A key shortcoming of these conventional detection indices ($T^2$ and $Q$) is that they only use the information in the last observation thus they have a short memory. Consequently, these detection indices are relatively insensitive to small changes in the process variables, and thus may result in missed detections [26]. These shortcomings of the $T^2$ and $Q$ statistics motivate the use of other alternatives in order to mitigate these disadvantages.

This paper is aimed at presenting new indice to improve the detectability of conventional PCA-based methods such $T^2$ and $Q$ statistics. Indeed, the history data obtained before actual point contain useful information for process monitoring; but, conventional PCA-based monitoring indices ignore such information at all. The ability to detect smaller parameter shifts can be improved by using a chart based on a statistic that incorporates information from past samples in addition to current samples. Alternatively, the exponentially weighted moving average (EWMA) monitoring chart consider not only the last data point, but the entire past data [27]. It makes them more sensitive than the $T^2$ and $Q$ charts to small anomalies. The main contribution of this work is to exploit the advantages of the exponentially weighted moving average (EWMA) chart and those of PCA modeling for enhancing detection performances of conventional PCA, especially for detecting small faults in highly correlated multivariate data. Such a choice is mainly motivated by the greater ability of the EWMA metric to detect small fault in process mean, which makes it very attractive as anomaly detection. In fact, the objective is to extend the abilities of the univariate EWMA monitoring chart to deal with multivariate processes.

- In this approach, PCA is used to express a process data matrix as the sum of two matrices: approximate and residual. After a model is obtained using PCA, the EWMA control scheme is applied using the ignored principal components (which have smallest variances) to improve fault detection. The smallest ignored PCs are used as an indicator about the existence or absence of anomalies.

The remainder of this paper is organized as follows. The analytical modeling of WRIM and a description of the various possible faults in these machines are presented in Section 2. Then, a brief introduction to PCA and how it can be used in fault detection are presented in Section 3. Then, the EWMA control scheme is described in Section 4, followed by a description of the proposed PCA-based EWMA fault detection approach (which integrates PCA modeling and the EWMA control scheme) in Section 5. Then, in Section 6, the performance of the proposed PCA-based EWMA control scheme is illustrated through a simulated example using WRIM data. Finally, some concluding remarks are presented in Section 7.

### 2. Analytical modeling of wound rotor induction machines (WRIM)

Effective monitoring of wound rotor induction machines requires developing a model that can accurately describes the behavior of these machines. In this work, a three-phase model that is based on magnetically coupled electrical circuits is used. To develop such a model, some modeling assumptions need to be made, which are described next.

#### 2.1. Modeling assumptions

In the proposed approach, it's assumed that the:

- magnetic circuit is linear, and the relative permeability of iron is very large compared to the vacuum,
- skin effect is negligible,
- hysteresis and eddy currents are negligible,
- air gap thickness is uniform,
- magnetomotive force created by the stator and the rotor windings follows a sinusoidal distribution along the air gap,
- stator and rotor have the same number of turns in series per phase,
- coils have the same properties,
- WRIM stator and rotor coils are coupled in star configuration and connected to the considered balanced state grid.

#### 2.2. Dynamic modeling of the WRIM

Defining the voltage vectors ($[V_A, V_B, V_C]$), the current vectors ($[I_A, I_B, I_C]$) and the flux vectors ($[\phi_A, \phi_B, \phi_C]$) for the stator and the rotor as:

$$
[\text{[Vs]}] = \begin{bmatrix} V_A \\
V_B \\
V_C \end{bmatrix}, \quad [\text{[Is]}] = \begin{bmatrix} I_A \\
I_B \\
I_C \end{bmatrix}, \quad [\phi] = \begin{bmatrix} \phi_A \\
\phi_B \\
\phi_C \end{bmatrix}
$$
where \( V_j, I_j \) and \( \phi_j \) (\( j : A, B, C \) for the stator phases and \( a, b, c \) for the rotor phases) are the voltages, currents, and magnetic flux of the stator and the rotor phases, respectively, and \( \theta \) is the angular position of the rotor relative to the stator. The fluxes can be related to the voltages and currents as follows:

\[
[V_k] = [R_k][I_k] + \frac{d\phi_k}{dt},
\]

where \( R_k \) and \( \phi_k \) are the resistance matrices, \( I_k \) and \( \phi_k \) are the own inductance matrices, and \( M_{S_k} \) and \( M_{R_k} \) are the matrices of the mutual inductances between the stator and the rotor coils. From Eqs. (1)–(4), it follows that:

\[
[V_k] = [R_k][I_k] + \frac{d([L_k][I_k])}{dt} + \frac{d(M_{S_k}[I_k])}{dt},
\]

\[
[V_k] = [R_k][I_k] + \frac{d([L_k][I_k])}{dt} + \frac{d(M_{R_k}[I_k])}{dt}.
\]

A dynamic analysis of the rotor results in the following mechanical motion equation [28,29]:

\[
J_p \frac{d\Omega}{dt} + f_p \Omega = C_{em} - C_r,
\]

where,

\[
\Omega = \frac{d\theta}{dt}
\]

and

\[
c_{em} = \frac{1}{2}[|I|^2 \frac{d(|I|)}{d\theta} - |I|]\]

where \( J_p \) is the total inertia brought to the rotor shaft, \( \Omega \) is the shaft rotational speed, \( [I] = [I_A I_B I_C I_a I_b I_c]^T \) is the current vector, \( f_p \) is the viscous friction torque, \( C_{em} \) is the electromagnetic torque, \( C_r \) is the load torque, \( \theta \) is the angular position of the rotor relative to the stator, and \( [L] \) is the inductance matrix of the machine. Defining the cyclic inductances of the stator and the rotor as \( l_{s} = \frac{1}{2}l_{s} \) and \( l_{c} = \frac{1}{2}l_{c} \) (where, \( l_{s} \) is the inductance of each stator phase and \( l_{c} \) is the inductance of each rotor phase) and denoting the pole pair number as \( p \), the inductance matrix of the WRIM can be written as follow:

\[
[L] = \begin{bmatrix}
L_{s} & 0 & 0 & M_{s1} & M_{s2} & M_{s3} \\
0 & L_{s} & 0 & M_{s1} & M_{s2} & M_{s3} \\
0 & 0 & L_{s} & M_{s1} & M_{s2} & M_{s3} \\
M_{s1} & M_{s2} & M_{s3} & L_{c} & 0 & 0 \\
M_{s2} & M_{s3} & M_{s1} & 0 & L_{c} & 0 \\
M_{s3} & M_{s1} & M_{s2} & M_{s1} & 0 & L_{c}
\end{bmatrix},
\]

where,

\[
f_1 = \cos(p\theta).
\]

By choosing the state variables to be the stator and rotor currents, the shaft rotational speed, and the angular position of the rotor relative to the stator, the dynamic WRIM model becomes:

\[
[X] = [X]^{-1}([U] - [b][X]),
\]

where,

\[
[X] = [I_A \ I_B \ I_C \ I_a \ I_b \ I_c \ \Omega \ \theta]^T; \quad [A] = \begin{bmatrix}
[I] & 0 & 0 \\
0 & J_\theta & 0 \\
0 & 0 & 1
\end{bmatrix};
\]

\[
[U] = \begin{bmatrix}
-V_A & V_B & V_C \\
-V_a & V_b & V_c \\
0 & -C_r & 0 \\
0 & 0 & -1
\end{bmatrix};
\]

\[
[B] = \begin{bmatrix}
R + \Omega \frac{d\theta}{dt} & 0 & 0 \\
-\frac{1}{2}|I|^2 \frac{d(I)}{d\theta} & f_p & 0 \\
0 & 0 & -1
\end{bmatrix}.
\]

This dynamic WRIM model will be used to simulate the normal and faulty behavior of the winding rotor induction machines.

### 2.3. Types of WRIM faults

Designing lighter machines with a long lifetime is now possible due to recent advances in engineering and material science. Despite these advancements in the reliability of machines, various types of faults may still exist. Faults can be the result of a normal wear in certain parts of the machine, a poor design, a poor assembly (misalignment), an improper use, or a combination of these different causes. Generally, faults in induction machines can be classified into four main categories [30]:

- **Stator faults** can be found on the coils or breech. In most cases, the winding failure is caused by inter-turn faults, which can grow and cause different faults between coils, between phases, or between phase and earth points [8]. The breech of electrical machines is built with insulated thin steel sheets in order to minimize the eddy currents for a greater operational efficiency. In the case of high power machines, the core is compressed to minimize the vibrations of the rolling sheets and to maximize the thermal conduction. Core problems in WRIM do not usually occur very often, around 1% compared to winding problems [31].

- **Rotor faults** can be bar breaks, coils faults, or rotor eccentricities.

- **Bearings faults** can be caused by a poor choice of materials during the manufacturing process, rotation problems within the breech, or chipped or cracked bearing.

- **Other faults** which are usually due to flange or shaft faults. These faults are generally due to manufacturing defects.

Just to give an idea about the relative possibility of these faults, Fig. 1(left) and (right) present the fault distribution in WRIM machines produced by a German company. Fig. 1 shows the fault distribution in low and medium power machines (50-200 kW), and Fig. 1 (right) shows a similar distribution in high power machines (above 200 kW) [32,33]. Fig. 1 (left) and (right) show that the most commonly encountered fault are the stator faults...
in low/medium power machines and the stator and mechanical faults in high power machines.

Another investigation related to the generators and converters reliabilities in wind turbines has been done in [33]. The results carried out by this study, which illustrate the fault distribution of WRIMs used in wind turbines, are shown in Fig. 2. This study showed that the most occurred failures are those related to the bearings and the second most common failure in wind turbines generators are the rotor failures.

2.4. Considered faults

In this paper, faults that are related to resistances in stator and rotor will be considered. Insulation degradation of coils can cause short circuits. Also, overloading the machine increases the temperature, which increases the resistances of the windings. In normal operation, a change in the resistance from its nominal value (at the ambient temperature, 25 °C) is considered a machine fault, which can be due to machine overload or coils degradation [29]. The relationship between the resistance and temperature can be expressed as:

\[ R = R_0 (1 + \alpha \delta T), \]

where, \( R_0 \) is the resistance value at \( T_0 = 25 \) °C, \( \alpha \) is the temperature coefficient of the resistance, and \( \delta T \) is the temperature variation.

The aim of this paper is to develop a PCA-based EWMA fault detection algorithm and then apply it to detect electrical faults in three-phase wound rotor induction machines. Since the PCA-based EWMA fault detection algorithm relies on PCA as a modeling framework, a brief introduction to PCA, and how it can be used in fault detection, is presented next.

3. Principal component analysis (PCA)

PCA is a linear dimensionality reduction modeling method, which can be helpful when handling data with a high degree of cross correlation among the variables. The main idea behind PCA is briefly introduced in this section, and more details can be found in [35,23].

3.1. PCA modeling

Let us consider the following raw data matrix

\[ X = [x_1, \ldots, x_n]^T \in \mathbb{R}^{n \times m} \]

consisting of \( n \) observations and \( m \) correlated variables. The data are collected when the monitored process is under normal operating condition so that the PCA’s model that will be built represents a reference of the normal process behavior. Before computing the PCA model, the raw data matrix \( X \) is usually pre-processed by scaling every variable to have zero mean and unit variance.
variance. Let $\tilde{X}$ denote the autoscaled matrix of $X$. By using singular value decomposition (SVD), PCA transforms the data matrix $X$ into a new matrix $T = [t_1, t_2, \ldots, t_m] \in \mathbb{R}^{n \times m}$ of uncorrelated variable called score or principal components. Each new variable is a linear combination of the original variables, so that $T$ is obtained from $X$ by an orthogonal transformations (rotations) designed by $P = [p_1, p_2, \ldots, p_m] \in \mathbb{R}^{m \times m}$ which is given as following:

$$X = TP^T,$$

(16)

where the column vectors $p_i \in \mathbb{R}^m$ of the matrix $P \in \mathbb{R}^{n \times m}$ (also known as the loading vectors) are formed by the eigenvectors associated with the covariance matrix of $X$, i.e., $\Sigma$. The covariance matrix, $\Sigma$, is defined as follows:

$$\Sigma = \frac{1}{n-1} X^T X = P\Lambda P^T \quad \text{with} \quad \Lambda = P^T P = I_n,$$

(17)

where, $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$ is a diagonal matrix containing the eigenvalues of covariance matrix in a decreasing order, and $I_n$ is the identity matrix $[36]$. In the case of collinear process, the dimensionality reduction of the $m$-dimensional space is obtained by retaining only the first $(l)$ largest principal components which are corresponding to the largest eigenvalues of the covariance matrix. The first $(l)$ largest principal components normally describe the most of the variance of the data. On the other hand, the smallest principal components are considered as a noise contributor. An important step in the building of PCA model is to determine the number of PCs, $l$, that are required to adequately capture the major variability in the data sets. Several techniques have been proposed to select the number of PCs including Scree plot $[37]$, cumulative percent variance (CPV), parallel analysis, sequential tests, resampling, profile likelihood $[37]$, and cross validation $[38]$. In this study, the CPV technique will be used to determine the number of PCs for PCA model. The CPV is defined as $\text{CPV}(l) = \sum_{i=l+1}^{m} \lambda_i \times 100$. Once the number of principal components $l$ is determined, the data matrix $X$ can be represented using PCA as the sum of two orthogonal parts: an approximated data matrix $\tilde{X}$ and a residual data matrix $E$ (see Fig. 3), i.e.,

$$\tilde{X} = TP = [\tilde{T} P]^T \in \mathbb{R}^{n \times l}$$

(18)

where $\tilde{T} \in \mathbb{R}^{l \times l}$ and $P \in \mathbb{R}^{l \times (m-l)}$ are matrices containing the $l$ retained principal components and the $(m-l)$ ignored principal components, respectively, and the matrices $\tilde{T} \in \mathbb{R}^{n \times l}$ and $\tilde{P} \in \mathbb{R}^{l \times (m-l)}$ are matrices containing the $l$ retained eigenvectors and the $(m-l)$ ignored eigenvectors, respectively.

### 3.2. Fault detection with PCA

In conventional PCA-based fault-detection, a PCA model is first constructed using faultless data representing the normal operation of the process, and then the PCA model is used to detect faults using one of snapshot Shewhart type monitoring charts, such as the $T^2$ or $Q$ statistics. The $T^2$ statistic describes the variability captured by the PCA model while the $Q$ statistic computes the residuals that are missed by the model. More specifically, the $T^2$ statistic is used to detect anomalies associated with abnormal variations within a PCs subspace. The $T^2$-based PCA chart based on the first $l$ PCs is defined as $[39]$

$$T^2 = \sum_{i=1}^{l} \frac{t_i^2}{\lambda_i},$$

(19)

where, $\lambda_i$ is the $i$th eigenvalue of the covariance matrix $\Lambda$. For new testing data, when the value of $T^2$ exceeds the value of the threshold, $T_{lim}$ given in $[39]$, a fault is declared. In a previous study $[40]$, the authors have shown that the $T^2$ statistic can result in false negatives (missed detection) due to the latent space sometimes being insensitive to small process upsets, which is because each latent variable is a combination of all process variables. The main disadvantage of using PCs in process monitoring is the lack of physical interpretation $[41]$. Additionally, the disadvantage of $T^2$ statistic is that a faults in the process mean that are orthogonal to the first PCs cannot be detected by using the $T^2$ $[42]$.

The $Q$ statistic, on the other hand, which is defined as $[24]$

$$Q = \|P\tilde{T}\|^2$$

(20)

measures the projection of a data sample on the residual subspace, which provides an overall measure of how the data sample fits the PCA model. When a vector of new data is available, the $Q$ statistic is calculated and compared with the threshold value $Q_{\text{lim}}$ given in $[36]$. If the confidence limit is violated, then a fault is declared. Fig. 4 provides an example of a simple dataset in which one observation has a large $Q$ value while the other has a large $T^2$. The $Q$ statistic is usually more preferred than $T^2$ in fault detection because it is more sensitive to fault with smaller magnitudes.

Unfortunately, the $T^2$ and $Q$ statistics use only the observed data at the current time point alone for making decision about the process performance at the current time point. They take into account only the present information of the process thus they have a short memory. Consequently, these detection indices are relatively insensitive to small changes in the process variables, and thus may result in missed detections $[26]$. To overcome these limitations of fault detection using PCA, in this paper, an alternative fault detection approach is developed, in which PCA is used as a modeling framework for fault detection using a EWMA control scheme.
More details about EWMA, and how it can be used in fault detection are presented next.

4. Quality control using EWMA

The aim of statistical process control is to monitor a process to detect abnormal behavior. SPC has been widely used in various production systems for monitoring processes and improving product quality [26]. Statistical controls charts (also referred to as monitoring charts) are one of the most commonly used tools in SPC and have been extensively used in quality engineering as a monitoring tool to detect the presence of possible anomalies in the mean or variance of process measurements. Control charts play a crucial role in detecting whether a process is still working under normal operating conditions (usually termed in-control) or not (out-of-control) [26]. Numerous control charts have been developed to monitor a mean of process variable over time, and include the Shewhart chart, the cumulative summation (CUSUM) chart [43], and the EWMA [44–46]. The Shewhart control charts was the first proposed method of testing the hypothesis, which are very popular in statistical process control, can be effectively used to detect large shifts in the process mean [47]. A key disadvantage of Shewhart charts, however, is that they only use the last data sample about inspected process and does not carry a memory of the previous data (i.e., ignore any potential information contained in past samples) [26]. These shortcomings motivate the use of other alternatives, such as EWMA and CUSUM charts, which are better suited to detecting smaller shifts in process mean [26]. This is because the CUSUM and EWMA charts take into account the information in the entire process history. The CUSUM chart gives equal weights to the entire process history observations when it accumulates all useful information in the historical data [26]. However, since EWMA uses a weighted average of all past and current observations, it is a lot less sensitive to violating the normality assumption than CUSUM charts [48]. Also, CUSUM is relatively slow to respond to large shifts. Therefore, EWMA-based charts are an appropriate monitoring scheme to be adopted when dealing with individual observations [26]. According to the literature, EWMA is one of the most frequently used control charts for process monitoring because of its flexibility and sensitivity to small shifts [26].

The EWMA control scheme was first introduced by Roberts [49], and has been extensively used in time series analysis. In the EWMA control scheme, the moving average is calculated by multiplying the historical observations by a weight that decays exponentially with time [26]. The EWMA decision statistic is described by the following recursive formula:

$$z_t = \lambda x_t + (1 - \lambda) z_{t-1},$$

where $\lambda$ is a weighted parameter, with $0 < \lambda \leq 1$, and $x_t$ is the value of the supervised variable at time $t$. The starting value $x_0$ is set equal to the process in-control mean, $\mu_0$. Generally, smaller values of $\lambda$ increase the chart’s sensitivity to smaller shifts in the process mean, while larger values of $\lambda$ increase its sensitivity to larger shifts [48]. The standard deviation of $z_t$ is defined as

$$\sigma_z = \sigma_0 \sqrt{\frac{1}{2\lambda} (1 - (1 - \lambda)^2)}.$$  

where $\sigma_0$ is the standard deviation of the fault-free or preliminary data set. The EWMA control scheme declares an anomaly when the value of $z_t$ falls outside of the interval between the control limits. The upper and lower control limits, UCL and LCL, are set as [26]:

$$UCL = LCL = \mu_0 + \sigma_0 \sqrt{\frac{2}{\lambda}},$$

where $L$ is a multiplier of EWMA standard deviation $\sigma_z$. $L$ and $\lambda$ are two parameters need to be set carefully [26,46].

In the next section, the EWMA control scheme is integrated with PCA to help enhance the ability of PCA in detecting small faults occurring in the mean of system measurements.

5. Fault detection using a PCA-based EWMA control scheme

In this section, PCA is integrated with EWMA to develop a new fault detection scheme with a higher sensitivity to small faults in the data. Towards this end, PCA is used to represent a matrix of the system measurements as the sum of two orthogonal parts (an approximated data matrix and a residual data matrix) as shown in Eq. (16). In PCA model, the principal components associated with large eigenvalues capture most of the variations in the data, while, ones associated with small eigenvalues mostly represent noise and are sensitive to the observations that are inconsistent with the correlation among the variables [50,51]. Therefore, the smallest principal components (i.e., associated with small eigenvalues) should be useful in fault detection. The smallest PCs can be used as an indicator about the existence or absence of faults. When the monitored system is under normal operating conditions (no faults), the least important principal components are close to zero. However, when a fault occurs, the they tend to largely deviate from zero indicating the presence of a new condition that is significantly distinguishable from the normal faultless working mode. In this paper, EWMA is used to enhance process monitoring through its integration with PCA. Because of the ability of the EWMA control scheme to detect small changes in the data, this technique is appropriate to improve the detection of small faults. Thus, this work exploits the advantages of the EWMA control scheme to improve fault detection over the conventional PCA-based methods. Towards this end, the EWMA control scheme is used to monitor the ignored principal components, which correspond to the small eigenvalues of the PCA model. Of course, this approach can only provide detection of faults (i.e., no isolation).

5.1. PCA-based EWMA process monitoring algorithm:

In the PCA-based EWMA fault detection algorithm, the EWMA monitoring scheme is applied using the principal components ignored from the PCA model. As shown in Eq. 16, the matrix $X$ can be written using PCA as follows:

$$X = \begin{bmatrix} T \circ P \circ P \circ P \circ \ldots \circ P \circ P \circ P \circ T \circ T \circ \ldots \circ T \circ T \circ T \end{bmatrix} = TP^T + TP^T.$$

Defining the matrix of ignored principal components as $\tilde{T} = [t'_{1}, \ldots, t'_{1}, \ldots, t'_{n}]$, where $t'_{1} \in \mathbb{R}^{m}$, i.e., $t' = [t'_{1}, \ldots, t'_{1}, \ldots, t'_{n}]$, then the EWMA decision function can be computed using the residuals of the jth principal component as follows:

$$z'_j = \lambda t'_{j} + (1 - \lambda) z'_{j-1}, \quad j \in [1, m].$$

In this case, since the EWMA control scheme is applied on the ignored $m - I$ principal components, $m - I$ EWMA decision functions will be computed to monitor system. However, this approach can only detect the presence of faults, i.e., it can not determine their locations. This approach is summarized in Table 1 and is schematically illustrated in Fig. 5.

In the next section, the performance of the proposed PCA-based EWMA fault detection method will be evaluated and compared to that of the conventional PCA fault detection scheme through their application to monitor would rotor induction machines.

6. Illustrative example

In this section, the developed PCA-based EWMA fault detection algorithm is utilized to improve the detection of stator and rotor winding faults in three-phase induction machines (WRIM). The performance of the developed method is compared to that of the conventional PCA.
To assess the performance of the proposed PCA-based EWMA fault detection scheme, the WRIM model presented in Section 2.2 will be used to simulate both normal and abnormal operations of the WRIM. The nominal values of the WRIM model parameters are presented in the Table 2. Also, a mechanical load torque of 10 Nm is applied to the machine at t = 2 s. The WRIM model is used to generate 10,000 measurements over a 4 s period. The data is obtained. Practically, even small faults in the WRIM may result in efficiency reduction, increases in temperature (which can reduce insulation lifetime), and increases in vibrations (which can reduce the bearing lifetime).

### 6.1. Data generation

To assess the performance of the proposed PCA-based EWMA fault detection scheme, the WRIM model presented in Section 2.2 will be used to simulate both normal and abnormal operations of the WRIM. The nominal values of the WRIM model parameters are presented in the Table 2. Also, a mechanical load torque of 10 Nm is applied to the machine at t = 2 s. The WRIM model is used to generate 10,000 measurements over a 4 s period. The data include nine WRIM state variables (m = 9), three stator phase currents \(I_s, I_r, I_c\), three rotor phase currents \(I_s, I_r, I_c\), the shaft rotational speed \(n\), the angular position \(\theta\), and the electromagnetic torque \(T_e\). The faults, which are changes in the values of the stator and rotor resistances, are introduced starting at time equal to 2 s. To assesses the abilities of the various fault detections methods, three different levels of increases (3%, 10% and 30%) in the coils of both resistances (stator and rotor) will be considered in this simulated example.

The time evolution in the stator’s current, rotor's current, shaft rotational speed, angular position, and electromagnetic torque of the WRIM in the healthy and faulty cases are shown in Fig. 6. Also, Fig. 7 shows the variation in electromagnetic torque versus the shaft rotational speed of the WRIM again in both the healthy and faulty cases. From the Fig. 6, it can be observed that the stator current remains insensitive to the considered faults. The Figure clearly shows that it is difficult to visualize small changes in the stator current of the monitored WRIM. Fig. 6 also shows that the rotor current, the shaft rotational speed, and the electromagnetic torque are the variables which provide more information in the presence of faults. Fig. 7, on the other hand, shows that as the faults increase from 3%, 10% to 30%, more deviation in the electromagnetic torque is obtained. Practically, even small faults in the WRIM may result in efficiency reduction, increases in temperature (which can reduce insulation lifetime), and increases in vibrations (which can reduce the bearing lifetime).

### 6.2. Training of WRIM model

The fault-free data set (which includes 9 variables and 10,000 samples) described in Section 6.1 is used as a training data set to construct the PCA model. The data matrix, which has 10,000 rows and 9 columns, is scaled (to be zero mean with a unit variance) before constructing the model.

In PCA models, most variations in the data are captured by the few principal components (which are associated with large eigenvalues), while the remaining principal components represent mainly noise. An important issue in PCA model building is the selection of the number of retained principal components. The cumulative percent variance (CPV) method, which has been usually used in the literature for determining the number of PCs, is used in this work. In this study, the threshold of cumulative variance value is chosen to be 90%. In our PCA model, this results in 14.79%, 14.45%, and 12.78% of the total variations) as shown in retaining the first six PCs (which capture 20.24%, 18.87%, 16.52%, 14.79%, 14.45%, and 12.78% of the total variations) as shown in Fig. 8.

Indeed, the EWMA chart is based on the assumptions that the measurements are normally and independently distributed. Therefore, it is necessary to check whether the ignored principal components distribution follows a Gaussian distribution. The residual normality hypothesis was verified in this study by examining the histogram. Checking the normality of the residuals can be done by visually checking the histograms of these three PCs vectors, which are shown in Fig. 9. These histograms indicate that the
normality assumption appears to be a reasonable one. Next, we check the absence of autocorrelation of ignored PC (specifically, the absence of autocorrelation), which is assumed to be uncorrelated. If the assumption is satisfied, the autocorrelation function (ACF) of the ignored PCs will have no significant spikes at any non-zero lags. Fig. 10 indicates that the ignored PC is not significantly correlated.

6.3. Simulation results

Now, the fault detection abilities of the conventional PCA and the proposed PCA-based EWMA fault detection algorithm will be assessed using the WRIM data. Three case studies representing different levels of faults are used to achieve this purpose. In the first case, an increase of 30% in both the rotor and stator resistances are
introduced, in the second case, a smaller increase of 10% in both resistances and in third case an increase of 3% is used to test the sensitivities of the various monitoring methods.

6.3.1. Case study 1: Detecting faults due to 30% increases in the rotor and stator resistances

In this case study, faults of 30% in the values of the rotor and stator resistances are introduced in the WRIM data. The $Q$ and $T^2$ statistics for this case are shown in Figs. 11 and 12, respectively. The dashed lines represent a 95% confidence interval used to identify possible faults. The results using the $Q$ statistic, which are plotted in Fig. 11, show that it could recognize this fault only for a short period of time after which it failed to recognize the fault. The Hotelling’s $T^2$ statistic, on the other hand, completely failed to detect the fault as shown in Fig. 12. This result can be explained by the fact that the $T^2$ statistic provides a measure of the variations in the PCs that are of greatest importance to the normal system operation. Thus, the normal operating region defined by the $T^2$ control limits is usually larger than the one defined by the $Q$ control limits. Therefore, faults with moderate magnitudes can easily exceed the $Q$ threshold, but not the $T^2$ threshold, which makes the $Q$ statistic usually more sensitive than $T^2$ for this type of faults. On the other hand, the application of the proposed PCA-based EWMA fault detection algorithm (using the three smallest principal components) resulted in the detection of the simulated fault as illustrated in Fig. 13. The smoothing parameter $\lambda$ and the control limit $L$ used are 0.3 and 3, respectively. The simulation results show the ability of the proposed method to detect the presence of the fault with the exception of small regions of missed detection.

6.3.2. Case study 2: Detecting faults due to 10% increases in the rotor and stator resistances

In this case study, a smaller change in the rotor and stator resistances is incorporated in the WRIM data to test their abilities to detect small faults. The performances of the $Q$ and $T^2$ statistics are demonstrated in Figs. 14 and 15, respectively. These results show that the conventional PCA based methods ($Q$ and $T^2$) are completely unable to detect this small simulated fault. This is because these conventional PCA based fault detection metrics only take into account the information provided by the present data samples in the decision making process, which makes these metrics not very powerful in detecting small changes. The results of PCA-based EWMA fault detection algorithm, however, which are shown in Fig. 16, clearly indicate the ability of this proposed method to detect this small fault without false alarms, but some small regions of missed detection. This case study clearly shows the advantage of the PCA-based EWMA method over the conventional PCA approach methods, especially in the case of small faults. Again, for EWMA control scheme, the values used for the parameters $\lambda$ and $L$ are the same as in the first case study.

6.3.3. Case study 3: Detecting faults due to 3% increases in the rotor and stator resistances

In the third case study, the testing data contain a smaller change in the rotor and stator resistances. This small fault is undetectable by the $Q$ and $T^2$ charts (see Figs. 17 and 18) The results
using the PCA-based EWMA chart (shown in Fig. 19) show that it could successfully detect this small fault. In summary, the proposed PCA-based EWMA fault detection method showed a satisfactory performance compared with the conventional PCA methods (e.g., $Q$ and $T^2$) through their application to monitor a three-phase wound rotor induction machine (WRIM). The results of two simulated case studies show an even clearer advantage for the PCA-EWMA method in the presence of smaller faults. These results are encouraging especially when it is of interest to detect faults with small magnitudes.

7. Conclusion

Induction machines are commonly used in the industry because of their ruggedness, simplicity of design, and low manufacturing cost. The necessity for having reliable electric machines is more important than ever and the trend continues to increase. However, there is still a great potential for failures in these machines. In this paper, a fault detection approach that is based on PCA is proposed to improve monitoring of WRIM. PCA has been used in this work as a modeling framework for fault detection using EWMA. The greater ability of the EWMA scheme to detect small faults makes it very attractive compared to the conventional PCA monitoring...
statistics. The main contribution of this work is to integrate PCA modeling with the EWMA control scheme to improve fault in WRIM especially in the presence of small faults. To achieve this objective, the EWMA control scheme is applied on the ignored principal components from the PCA model constructed using the WRIM data. The simulation results obtained using two case studies demonstrate the advantage of the proposed PCA-based EWMA fault detection method (over the conventional methods, such as Q and $T^2$) in detecting faults in WRIM data, especially faults with small magnitudes.

In future works, the performance of the proposed approach in detecting faults in wound rotor induction machines will be tested and validated on experimental data.

References


[38] B. Li, J. Morris, E. Martin, Model selection for partial least squares regression, Chemometrics Intelligent Lab. Syst. 64 (1) (2002) 79–89.


