Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Energy 82 (2005) 197–213
APPLIED ENERGY
www.elsevier.com/locate/apenergy

Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method

Shengwei Wang *, Jingtan Cui

Department of Building Services Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Received 27 August 2004; revised 26 October 2004; accepted 6 November 2004
Available online 4 February 2005

Abstract

An online strategy is developed to detect, diagnose and validate sensor faults in centrifugal chillers. Considering thermophysical characteristics of the water-cooled centrifugal chillers, a dozen sensors of great concern in the chiller-system monitoring and controls were assigned into two models based on principal-component analysis. Each of the two models can group a set of correlated variables and capture the systematic trends of the chillers. The *Q*-statistic and *Q*-contribution plot were used to detect and diagnose the sensor faults, respectively. In addition, an approach based on the minimization of squared prediction error of reconstructed vector of variables was used to reconstruct the identified faulty-sensors, i.e., estimate their bias magnitudes. The sensor-fault detection, diagnosis and estimation strategy was validated using an existing building chiller plant while various sensor faults were introduced.

Keywords: Centrifugal chiller; Sensor fault; Sensor bias; Fault detection; Fault diagnosis; Sensor estimation; Principal-component analysis

* Corresponding author. Tel.: +852 27665858; fax: +852 27746146.
E-mail address: beswwang@polyu.edu.hk (S. Wang).

0306-2619/$ - see front matter
doi:10.1016/j.apenergy.2004.11.002

Nomenclature

a	number of loading vectors retained in PCA model
c_{pw}	specific-heat of water (kJ/kg K)
c_α	confidence limit in normal distribution
Eff_{comp}	efficiency of compression process
$\mathbf{e}$	residual vector
f_i	estimate of fault magnitude in the direction $\boldsymbol{\beta}_i$
h	specific enthalpy (kJ/kg)
$\mathbf{I}$	unit matrix
LMTD	logarithmic mean temperature-difference (K)
M	water flow-rate (kg/s)
m	number of variables
n	number of observations
$\mathbf{P}$	loading matrix consists of a retained loading-vectors
P	pressure (Pa)
Q	load (kW)
$\mathbf{S}$	covariance matrix of $\mathbf{X}$
T	temperature (°C)
$\mathbf{U}$	matrix of eigenvectors specific volume (kg/m^3)
v	specific volume (kg/m^3)
W	rate of input work (kW)
$\mathbf{X}$	scaled sample matrix with m variables and n observations
$\mathbf{x}$	an observation (row) vector
$\hat{\mathbf{x}}$	estimate of $\mathbf{x}$
$\mathbf{x}^*$	true value of $\mathbf{x}$
$\bar{\mathbf{x}}$	best estimate of $\mathbf{x}^*$
$\mathbf{y}$	projection of $\mathbf{x}$ into the lower-dimension (a-dimension) space
$\boldsymbol{\Lambda}$	diagonal matrix of non-negative real eigenvalues with decreasing magnitude
α	confidence percentage
$\boldsymbol{\beta}$	detected fault direction
γ	mean isentropic coefficient of refrigerant
ΔT	temperature difference between chilled-water supply and return
λ	eigenvalue of $\mathbf{S}$

Subscripts

1,2,3,4	state points in pressure–enthalpy graph
cd	condenser
chw	chilled water
chwr	chilled-water return
chws	chilled-water supply
comp	compressor
cw	condenser water

ecw	entering condenser-water
elec	electrical
ev	evaporator
hg	hot-gas discharge
i	ith sensor in an observation vector
lcw	leaving condenser-water

1. Introduction

Chiller systems, especially centrifugal ones, accounts for a large portion of the energy consumption of HVAC systems. In large office buildings, it is estimated that the electricity consumption of chillers is typically 35–40% of the total building energy-consumption for commercial buildings in Hong Kong. During operation, chiller performance is affected by ever-changing environmental conditions, e.g., outdoor air temperature and humidity. In addition, chiller performance degrades naturally and different kinds of faults may occur, resulting in a great waste of energy, complaints from occupants and shortened equipment life. Therefore, there is a great need to develop optimal and robust controls and automatic fault-detection and diagnosis (FDD) of chiller systems to deal with these problems. There have been many publications pertinent to the FDD of chiller systems as presented in the review papers of Comstock et al. [1], Reddy et al. [2] and Dexter et al. [3]. Typical examples of recent investigations into the efficiency-centered controls of chiller systems include the works of Wang [4], Chan [5], and Chang [6].

It is worth pointing out that the performance of both the FDD methods and the optimal control strategies depend strongly on the quality (i.e., accuracy and reliability) of the measurements from the sensors in chiller systems as concluded in the IEA project – Annex 34 [7]. With the advance in networking technologies and standard communication protocols, chiller-control panels are now integrated into building management systems (BMSs) to provide measurements for chiller-system monitoring and controls. Moreover, accurate online measurements of temperature, flow rate, electrical power, refrigerant pressure, etc., are also essential to safety interlocks of the chiller system, quantification of effectiveness of energy-efficiency improvements [8], monitoring of chiller efficiency [9]. Therefore, it is essential to ensure the accuracy and robustness of sensors used by chiller systems and it is of great benefit for BMSs to perform online sensor monitoring and FDD of chillers.

Generally speaking, the sensor errors can be divided into two main categories including hard failure (complete failure) and bias error, which is the fixed or systematic component of the total error [10] if a sensor suffers from a soft fault. It is easy to handle hard failures although they are harmful to chiller monitoring and control, while not much research has addressed the sensor-bias error in chiller systems. The existence of sensor biases might result in great deficiency or even malfunctioning of chiller monitoring, control, optimization and FDD. For example, it is straightforward and convenient to measure chiller-cooling load by the product of the differential chilled-water temperature and flow rate ($\Delta T \times M_{\mathrm{chw}}$), which is widely adopted for chiller sequencing control. However, the differential chilled-water temperature, ΔT, is usually small with a design value of around 5 K only. A bias error of 1 °C in both supply and return water temperature sensors alone might result in up to a 40% error in the total cooling-load measurement, making no sense for incorporating an energy-efficient chiller sequencing control.

Sensor failures have long been a major concern in engine-performance monitoring [11]. Research on sensor FDD in engineering systems has been very active in recent years. The physical-redundancy method was adopted for detecting and diagnosing sensor faults in nuclear-power plants [12], where several sensors were used to measure the same physical quantity. Any serious discrepancy among the measurements indicates a sensor fault. Even though the physical-redundancy approach can often be effective, the cost and complexity of incorporating redundant sensors might make this approach unattractive to some extent in other less critical applications, e.g., HVAC and chemical fields. In the HVAC field, Stylianou et al. [13] studied a model-based method to detect soft sensor-faults, aiming at making measurements reliable when monitoring the performance of a laboratory chiller. Lee et al. [14] used general regression neural-network modes to investigate the detection and automatic recovery of a faulty supply air-temperature sensor in an AHU (air-handling unit). These methods aimed at detecting/diagnosing part of the sensors in the systems. Validating all or most of the sensors in a field is a much heavier and more complicated task. As for sensor faults in a building's chilling systems, e.g., a building's supply-water temperature, a building's return-water temperature and water flow rate in primary/secondary systems, a rule-based diagnosis and validation strategy was studied by Wang et al. [15,16]. Although the principal-component analysis (PCA) method has been used in many engineering practices [17,18], only recently the PCA method has entered the sensor FDD in HVAC systems [19] and proved to be an efficient means to generate useful residuals for the sensor FDD there.

This paper presents a PCA-based strategy, which uses the Q-statistic as the indexes of fault detection, and use the Q-contribution plot to isolate faults in chillers. Chiller data from a real-world chiller system are used to train the PCA model and evaluate the capability of the strategy to detect, diagnose and estimate the sensor faults. The sensor-fault detection, diagnosis and estimation (FDD&E) strategy and its validation as well as examples of field application are presented in this paper.

2. Outline of the PCA method in FDD

PCA is a multi-variate statistical-analysis technique, in which a group of correlated variables are transformed into a new group of variables which are uncorrelated or orthogonal to each other [20,21]. PCA separates the observation space into a subspace capturing the systematic variations of the process and a subspace containing random noise.

The principle of using the PCA method for FDD applications is illustrated on the basis of a training data set of n observations (samples) and m process variables. Since variables in engineering systems usually have different units, these data are transformed into standard units by subtracting from each observation its mean and dividing by its standard deviation. This makes the covariance and the correlation coefficient matrices the same. The transformed data set will be denoted by a sample matrix $\mathbf{X}(\mathbf{X} \in \Re^{n\times m})$. The eigenvalue decomposition of the sample covariance matrix $\mathbf{S}$ is obtained as shown in Eq. (1) where the $\mathbf{\Lambda}$ is a diagonal matrix of non-negative real eigenvalues with decreasing magnitude ($\lambda_1 > \lambda_2 > \cdots > \lambda_m$). $\mathbf{U}(\mathbf{U} \in \Re^{m\times m}$ and $\mathbf{U}\mathbf{U}^T = \mathbf{I})$ is a matrix whose columns are the corresponding eigenvectors.

$$\mathbf{S} = \frac{\mathbf{X}^T\mathbf{X}}{n-1} = \mathbf{U}\mathbf{\Lambda}U^T. \tag{1}$$

In order to optimally capture the variations of the data while minimizing the effect of random noise, only the eigenvectors in $\mathbf{U}$, which are associated with the first a largest eigenvalues, are retained in the PCA models. A loading matrix $\mathbf{P}(\mathbf{P} \in \Re^{m\times a})$, whose columns correspond to the retained eigenvectors, is constructed. A new observation vector, $\mathbf{x}$, can be projected into the lower-dimensional space whose direction is defined by the known loading matrix $\mathbf{P}(\mathbf{P} \in \Re^{m\times a})$, as shown in Eq. (2).

$$\mathbf{y} = \mathbf{x}\mathbf{P}. \tag{2}$$

The estimate of $\mathbf{x}$, namely $\hat{\mathbf{x}}$, can be obtained by projecting $\mathbf{y}$ back into the m-dimensional vector, as shown by Eq. (3). The difference between $\mathbf{x}$ and its estimate $\hat{\mathbf{x}}(\hat{\mathbf{x}} = \mathbf{x}\mathbf{P}\mathbf{P}^T)$ is the residual vector $\mathbf{e}$ as shown in Eq. (4), which describes the abnormal variation and noise and can be used as the basis for the FDD&E.

$$\hat{\mathbf{x}} = \mathbf{y}\mathbf{P}^T = \mathbf{x}\mathbf{P}\mathbf{P}^T, \tag{3}$$

$$\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}} = \mathbf{x}(\mathbf{I} - \mathbf{P}\mathbf{P}^T). \tag{4}$$

With respect to FDD applications of PCA, the Q-statistic [22], also known as the squared prediction error (SPE), can be used as an index of faulty conditions. The Q-statistic measures the total sum of variations in the residual vector. Therefore, faults can be detected by using the Q-statistic, as shown by Eq. (5).

$$Q\text{-statistic} = \text{SPE} = \mathbf{e}^T\mathbf{e} = \|\mathbf{x} - \hat{\mathbf{x}}\|^2 = \|\mathbf{x}(\mathbf{I} - \mathbf{P}\mathbf{P}^T)\|^2 \leqslant Q_\alpha, \tag{5}$$

where, Q_α is the threshold for the Q-statistic and can be statistically determined [20,21] according to measurements of the training matrix. When there is no fault, the Q-statistic will be less than Q_α, indicating that the correlations among the measurements of the variables remain unchanged. On the contrary, when faults exist, the correlations among the measurements of the variables will be destroyed and a higher value of the Q-statistic is obtained.

3. PCA-based chiller sensor FDD and bias-estimation strategy

3.1. Typical centrifugal-chiller systems and their sensors

Fig. 1 shows the schematic (a) of a typical water-cooled centrifugal chiller system and the corresponding pressure–enthalpy diagram (b). The condenser water from the

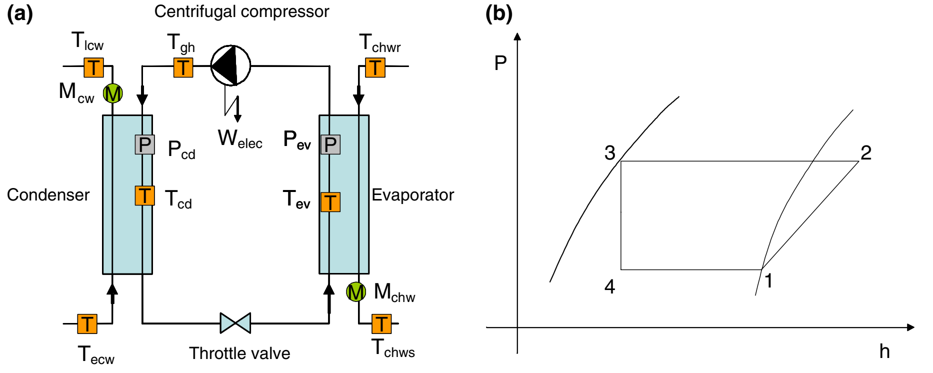

Fig. 1. Schematic of a typical centrifugal-chiller refrigeration cycle.

cooling tower enters into the condenser and absorbs the heat discharged by the refrigerant. The return chilled-water from the building enters the evaporator, where it is cooled by the evaporating refrigerant. For minimizing chiller-energy use, the entering condenser water set-point should be as low as possible. However, the control set-point of entering condenser water should be at or above the lowest temperature attainable by cooling tower at certain air (wet-bulb) temperatures to avoid the waste of fan energy trying to reach an unobtainable value. The chilled-water supply temperature is maintained at its set-point by regulating the inlet guide vane position. The vane-control motor opens or closes the inlet guide vane by a feedback controller (e.g., a proportional-integral controller) to maintain the preset chilled-water supply temperature using the difference between the set point and measured value as reference [23]. There are a great number of sensors implemented and the sensor instrumentation is different for different systems. In this study, the most essential sensors for chiller monitoring and control, which exist in most chiller systems, are included as listed in Table 1.

Table 1
Sensors in typical centrifugal-chiller systems

Measurement parameter	Description	Unit
M_{chw}	Chilled-water flow rate	L/s
M_{cw}	Condenser water-flow rate	L/s
P_{cd}	Condensing pressure	Pa
P_{ev}	Evaporating pressure	Pa
T_{cd}	Condensing temperature	°C
T_{chwr}	Chilled-water return temperature	°C
T_{chws}	Chilled-water supply temperature	°C
T_{ecw}	Entering condenser water temperature	°C
T_{ev}	Evaporating temperature	°C
T_{hg}	Hot-gas discharge temperature	°C
T_{lcw}	Leaving condenser water temperature	°C
W_{elec}	Chiller electrical-power input	kW

3.2. PCA models of chiller process

Two PCA models are developed in the strategy to capture the correlations among the sensors, as listed in Table 1, by dividing the sensors into two groups after testing various combinations.

The first PCA model, model **A** concerning energy performance, includes nine sensors (as shown in Eq. (6)) closely related to the performance of chillers.

$$\mathbf{A} = [T_{\mathrm{chws}}, T_{\mathrm{chwr}}, T_{\mathrm{ev}}, P_{\mathrm{ev}}, T_{\mathrm{lcw}}, T_{\mathrm{ecw}}, T_{\mathrm{cd}}, P_{\mathrm{cd}}, T_{\mathrm{hg}}], \tag{6}$$

where T_{chws}, T_{chwr}, T_{ev}, P_{ev}, T_{lcw}, T_{ecw}, T_{cd}, P_{cd} and T_{hg} are, respectively, chilled-water supply temperature, chilled-water return temperature, evaporating temperature, evaporating pressure, leaving condenser-water temperature, entering condenser-water temperature, condensing temperature, condensing pressure, and hot-gas discharge temperature.

For a specific chiller system with constant chilled water or cooling water flow rates, the temperatures of chilled water (T_{chws}, T_{chwr}) as well as those of cooling water (T_{ecw}, T_{lcw}) are deterministic variables for the operating condition and performance of the chiller system [23–25]. Among these nine variables, four are the deterministic variables of the operating condition. The rest are also variables in quantifying the performance of the chillers. These variables are used to calculate the performance indexes, which are usually employed in applications, as shown in Eqs. (7)–(9). It can be observed from the equations that the main variables needed to quantify the performance of chillers of constant water-flows are included in model **A**.

$$\mathrm{Eff}_{\mathrm{comp}} = \frac{\gamma P_1 v_1 [(P_2/P_1)^{(\gamma-1)/\gamma} - 1]}{(\gamma - 1)(h_2 - h_1)}, \tag{7}$$

$$\mathrm{LMTD}_{\mathrm{ev}} = \frac{T_{\mathrm{chwr}} - T_{\mathrm{chws}}}{\ln\left(\frac{T_{\mathrm{chwr}} - T_{\mathrm{ev}}}{T_{\mathrm{chws}} - T_{\mathrm{ev}}}\right)}, \tag{8}$$

$$\mathrm{LMTD}_{\mathrm{cd}} = \frac{T_{\mathrm{lcw}} - T_{\mathrm{ecw}}}{\ln\left(\frac{T_{\mathrm{lcw}} - T_{\mathrm{cd}}}{T_{\mathrm{ecw}} - T_{\mathrm{cd}}}\right)}, \tag{9}$$

where $\mathrm{Eff}_{\mathrm{comp}}$ is the efficiency of the compression process, $\mathrm{LMTD}_{\mathrm{ev}}$ and $\mathrm{LMTD}_{\mathrm{cd}}$ are, respectively, the logarithmic mean-temperature differences of the evaporator and condenser, v_1 is the specific volume of the refrigerant at the inlet of the compressor and determined by the evaporating pressure and temperature, γ is the mean isentropic coefficient of the refrigerant, P_2 is the refrigerant pressure at the compressor outlet (approximating the condensing pressure P_{cd}) and P_1 is the refrigerant pressure at the compressor inlet (approximating the evaporating pressure P_{ev}), and h_2 is the specific enthalpy of the refrigerant at the compressor outlet, which is primarily determined by the hot-gas discharge temperature T_{hg}. T_{ev} and T_{cd} are the evaporating and condensing temperatures which correspond to the P_{ev} and P_{cd}, respectively.

The second PCA model, i.e., model **B** concerning the energy balance of the chillers, involves seven variables: W_{elec}, M_{chw}, T_{chws}, M_{cw}, T_{chwr}, T_{ecw} and T_{lcw}, as shown in Eq. (10). This model is based on the fact there exist strong correlations among these seven variables due to the energy balance as represented in Eqs. (11) and (12). The model allows that the faults of these seven sensors can be detected and estimated.

$$\mathbf{B} = [T_{\mathrm{chws}}, T_{\mathrm{chwr}}, M_{\mathrm{chw}}, T_{\mathrm{lcw}}, T_{\mathrm{ecw}}, M_{\mathrm{cw}}, W_{\mathrm{elec}}], \tag{10}$$

$$W_{\mathrm{comp}} + Q_{\mathrm{ev}} - Q_{\mathrm{cd}} = 0, \tag{11}$$

$$W_{\mathrm{elec}} + M_{\mathrm{chw}} c_{\mathrm{pw}} (T_{\mathrm{chwr}} - T_{\mathrm{chws}}) - M_{\mathrm{cw}} c_{\mathrm{pw}} (T_{\mathrm{lcw}} - T_{\mathrm{ecw}}) \approx 0, \tag{12}$$

where W_{elec}, M_{chw}, M_{cw}, Q_{ev}, Q_{cd} and c_{pw}, are, respectively, the electrical power input to the compressor, the chilled-water flow rate, the condenser-water flow rate, the evaporator load, the condenser load and the water specific-heat.

3.3. Structure of sensor FDD strategy

The structure of the PCA-based sensor FDD strategy is shown in Fig. 2. It basically consists of two major groups of tasks, i.e., the development and training of PCA models, the online sensor FDD and bias estimation. The PCA models are developed and trained prior to implementation for the online application.

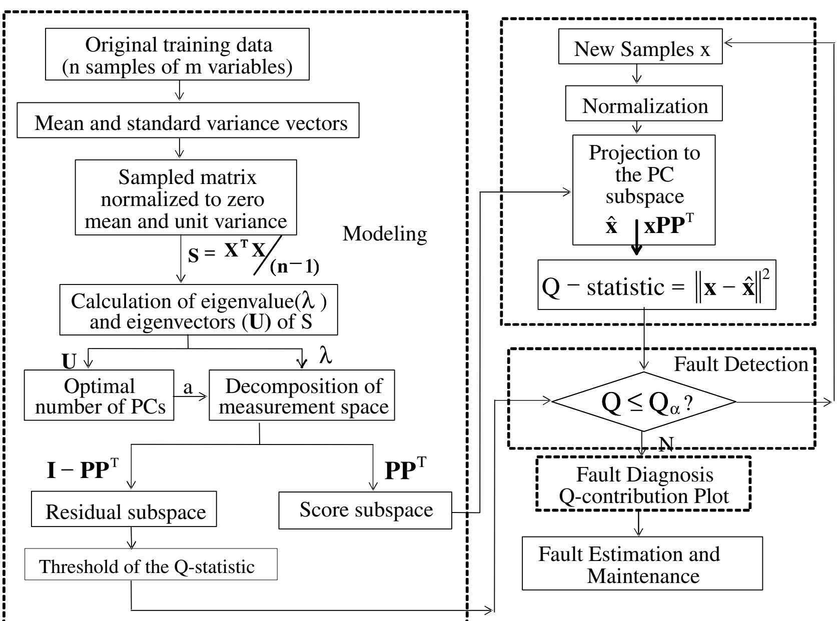

Fig. 2. Flow chart of PCA method in sensor FDD.

3.3.1. Development and training of PCA models

The development and training of PCA models have three main steps: decomposing of the covariance matrix of training matrixes, retaining loading vectors, and determining Q-statistics. In order to obtain the PCA models, including the retained loading vectors P and threshold of Q-statistic, two training matrixes, **A** and **B** corresponding to two models described above, are constructed from measurements when the sensors are in healthy conditions. Matrix **A** is used to train the PCA model **A**. Matrix **B** is used to train PCA model **B**. Q-contribution plot can be used to diagnose the fault. The eigenvalues and eigenvectors of the covariance matrixes of the training matrix **A** and **B** can be obtained through their eigenvalue decompositions, respectively. The numbers of the loading vectors retained in the two PCA models are determined by analyzing the proportions of the traces of the relevant covariance matrixes, respectively. The thresholds of the Q-statistic are determined based on the analysis of the residual subspace.

3.3.2. Online sensor FDD and bias estimation

The sensor fault FDD and bias estimation, at each sampling step, involve four main steps: calculation of Q-statistics of the two models, sensor fault detection by comparing Q-statistic against its threshold, sensor fault diagnosis based on Q-contribution plot, and estimation of sensor bias by minimizing the SPE of the reconstructed sensor measurements.

When the Q-statistic of a new sample goes beyond the threshold, the sensor fault is detected and a fault report/warning is generated after a certain number of consecutive fault detections.

If a sensor fault is detected, the contribution of the individual variables to the Q-statistic is estimated. In principle, the larger the contribution of a variable makes to the Q-statistic, the higher the possibility of a fault in that variable. The Q-contribution plot is therefore generated and used to reduce the possible fault sources and thus focus on a smaller range of measurements among all original variables. The sensor is diagnosed to be faulty if its Q-contribution value is the largest in the Q-contribution plot.

Sensor-bias estimation is achieved on the basis of the reconstruction of faulty sensors. The method can estimate the true value, $\mathbf{x}^*$, of an observation vector using the observation vector $\mathbf{x}$, the constructed PCA model and the detected-fault direction as described below.

As presented by Dunia et al. [26], the reconstructed vector, $\bar{\mathbf{x}}$, shown in Eq. (13), can be obtained by correcting x in the direction where a fault is detected, where, $\boldsymbol{\beta}_i$ indicates the detected fault direction. For example, $\boldsymbol{\beta}_1 = [1,0,0,0,\cdots 0]$ means there is a failure associated with the first variable in an observation: f_i is the estimate of the bias magnitude in the direction $\boldsymbol{\beta}_i$, i.e., the ith sensor in the observation vector. The difference between reconstructed vector, $\bar{\mathbf{x}}$, and its projection on the score subspace is characterized similarly by the SPE of $\bar{\mathbf{x}}$, as shown in Eq. (14).

$$\bar{\mathbf{x}} = \mathbf{x} - f_i\boldsymbol{\beta}_i, \tag{13}$$

$$\mathrm{SPE}_{\bar{x}} = \|\bar{\mathbf{x}}(\mathbf{I} - \mathbf{P}\mathbf{P}^T)\|^2. \tag{14}$$

If $\bar{\mathbf{x}}$ is the best estimate of true value of the measured variable, it should lead to the minimization of SPE of $\bar{\mathbf{x}}$. According to the reconstruction method developed by Dunia et al. [26], the reconstructed observation, $\bar{\mathbf{x}}$, can be estimated as shown in Eq. (15).

$$\bar{\mathbf{x}} = \left(\mathbf{I} - \frac{\boldsymbol{\beta}_i^T \tilde{\boldsymbol{\beta}}_i}{\|\tilde{\boldsymbol{\beta}}_i\|^2}\right)\mathbf{x}, \tag{15}$$

where, $\tilde{\boldsymbol{\beta}}_i = \boldsymbol{\beta}_i(\mathbf{I} - \mathbf{P}\mathbf{P}^T)$.

4. Validation of sensor FDD&E strategy for a real chiller-system

4.1. Chiller system involved

Chiller data were collected, during a period of a month, from a 5400kW York seawater cooled centrifugal chillers in a large commercial building in Hong Kong. The condenser of the chiller is directly cooled by seawater and the water-flow rates through the condenser and evaporator are all constant. The chilled water and cooling-water flow rates, temperatures, pressures and compressor power were collected every minute from the built-in control panels of the chillers, which are integrated into the BMS via a gateway. Water temperatures were measured by thermistors, water-flow rates by electromagnetic flow meters and electric powers by three-phase power transducers. Pressures are measured by pressure transmitters.

4.2. Training of PCA models

The HVAC system of the building operates all day. The data of two days (i.e., 11th and 12th July) were used in training the PCA models. Data from the last days were added with biases to test the fault detectability and isolability of the PCA models.

During a one-day operating period, the deterministic variables vary in a small range, when the working conditions do not change significantly as a result of the chiller-system control. Therefore, it might be concluded that PCA models are capable of capturing the trend of the chiller system as well as the correlations among the measurements when these determinant variables do not vary significantly.

It is worth pointing out that data used for the model training need to go through a steady-state filter to filter out the data of significant dynamics similarly as the data used by the FDD schemes. The number of original samples under normal sensor-conditions was 1362. After passing through the steady-state filter and outlier detector, 494 samples remained. These samples were used to construct two training matrixes: training matrix **A** (494×7) corresponding to PCA model **A** and training matrix **B** (494×9) corresponding to PCA model **B**. The loading vectors and their corresponding eigenvalues can be obtained by solving an eigenvalue decomposition of the covariance matrix **S** of the training data set.

Table 2
Proportion of trace explained by PCs (PCA model **A**)

The *i*th principal component	Eigen value	Variance explained (%)	Cumulative variance explained (%)
1	6.0633	67.3695	67.3695
2	1.8610	20.6773	88.0468
3	0.6817	7.5744	95.6212
4	0.3153	3.5032	99.1244
5	0.0356	0.3961	99.5205
6	0.0171	0.1899	99.7104
7	0.0153	0.1697	99.8801
8	0.0097	0.1083	99.9884
9	0.0010	0.0116	100

It is commonly understood that the loading vectors associated with the larger eigenvalues describe most of the systematic or state variations occurring in the process, and the loading vectors associated with the smaller eigenvalues describe the random noise [27]. In order to decouple the systematic variations from the random variations, the appropriate number of loading vectors should be maintained in the PCA model. The larger the maintained number is, the better the fitness of the PCA model. The smaller the number is, the simpler the model. To determine the optimal number of maintained loading vectors (i.e., the number of the principal components), the method related to the proportion of the trace of the covariance matrix, which is explained by the principal components, is adopted. This method is easy to understand and manipulate.

With respect to training matrix **A**, three principal components corresponding to the three largest eigenvalues could explain 95.62% of the total variance of the system, as shown in Table 2. Therefore the three PCs were retained in the PCA model. The detection threshold of Q-statistic of 99.7% confidence level was calculated to be 3.9485. Fig. 3 shows the Q-statistic plot of the model in normal sensor-conditions.

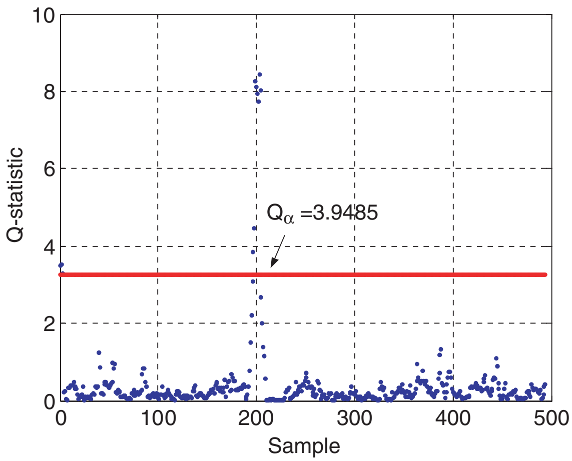

Fig. 3. Q-statistic plot of training data using PCA model **A**.

Table 3
Magnitudes/effective periods of biases introduced and estimated biases – Test 1

	T_{ev}	P_{ev}	T_{cd}	P_{ev}	T_{hg}
Effective period (samples)	21–40	41–60	61–80	81–100	101–121
Bias introduced	2 K	50 kPa	3 K	100 kPa	5 K
Bias estimated	2.11 K	52.42 kPa	3.27 K	105.7 kPa	5.43 K
Relative error	5.4%	4.8%	9.1%	5.7%	8.6%

Table 4
Magnitudes/effective periods of biases introduced and estimated biases – Test 2

	T_{chwr}	T_{lcw}	W_{elec}
Effective period (samples)	21–60	61–100	101–140
Bias introduced	2 K	3 K	10.0 kW
Bias estimated	2.17 K	2.92 K	7.51 kW
Relative error	8.3%	2.7%	24.9%

The Q-statistic of 98.6% samples was below the threshold value. The total number of samples detected to be abnormal was 7 (1.4%). This shows that the PCA model well captures the major correlation and variance among the system variables.

As for the training matrix **B**, the number of principal components ($a = 4$) and the detection threshold of the Q-statistic of 95% confidence level ($Q_\alpha = 1.0312$) were determined in the same way. The Q-statistic of 93.6% samples was below the threshold value. The total number of samples detected to be abnormal was 7 (6.4%). This means that the PCA model captured the major correlation and variance among system variables.

4.3. Introducing sensor-faults

Many tests were conducted by adding bias to the measurements of different sensors retrieved from the BMS in a few days after 12th July. In the FDD test presented in this paper, the 156 samples (after filtering) from 13th July were corrupted artificially with pre-determined biases in order to generate faulty sensors for testing the FDD strategy. Two validation tests are presented in this paper, where bias was added on the measurements of a few sensors in the each test for the tight presentation of the paper. In Test 1, biases were added to the measurements (measured on 13th July) of the five sensors closely related to the PCA model **A**. In Test 2, biases were added to the measurements (measured on the same day) of three sensors closely related to the PCA model **B**. Tables 3 and 4 show the magnitudes of the biases and corresponding sample range of the biases in the two tests, respectively. Note, only one sensor set was biased at a time, because the chance of simultaneous multiple sensor faults is low. In the tests, only the sensors not used for feedback control were added with biases as it is impossible to simulate the real operation of the chiller system by simply adding a bias used in the feedback control-loops without simulating its effects on the other variables. Faults on chilled and cooling-water flow rates are not included in the test presented in this paper as they are nearly constant in operation and their faults can be detected and diagnosed.

4.4. Results and analysis

Fig. 4 shows the Q-statistic plot of the PCA model **A** in Test 1. It is observed that nearly all the Q-statistics values were above the threshold, while one of the five predetermined biases existed in the corresponding sensor. The sensor faults were successfully detected. Fig. 5 shows the results of a sensor-fault diagnosis using the

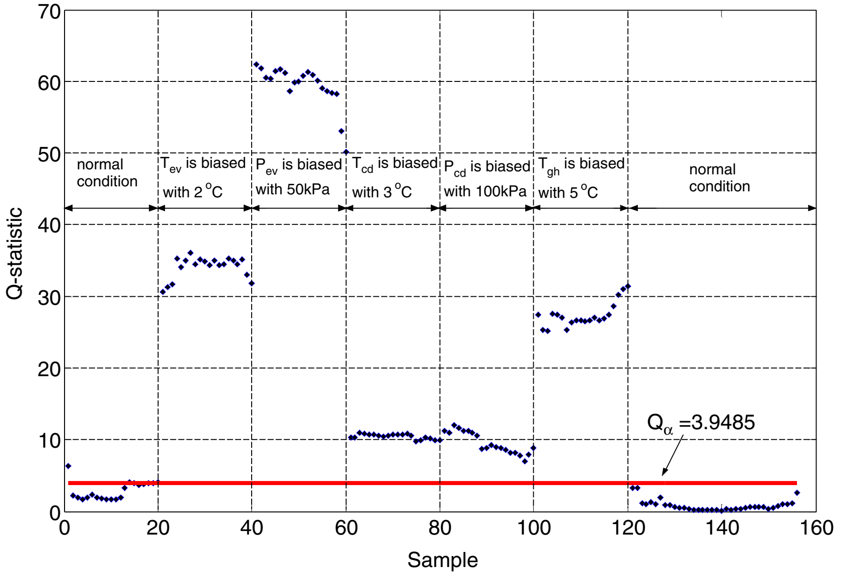

Fig. 4. Q-statistic plot of PCA model **A** – Test 1.

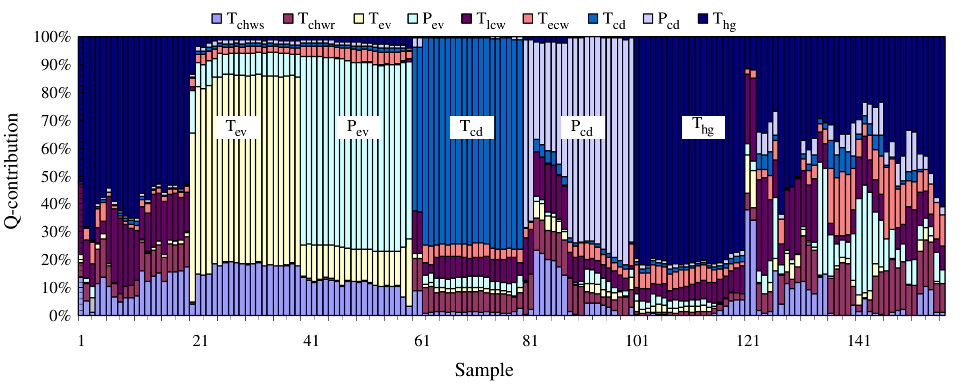

Fig. 5. Q-contribution plot of PCA model **A** – Test 1.

Q-contribution plot of the PCA model **A** in Test 1. The biased sensors were found using the *Q*-contribution plot. For example, when a bias of +2 °C existed in the evaporating temperature-sensor, the *Q*-contribution of the evaporating temperatures from samples 21 to 40 was more than 67%, which was much higher than those of other sensors in the same period. The biases of the faulty sensors were also estimated correctly by the strategy and the results are shown in Table 3.

The results of Test 2 using the PCA model **B** are shown in Figs. 6 and 7. As expected, nearly all the *Q*-statistics values in this test were above the threshold when a

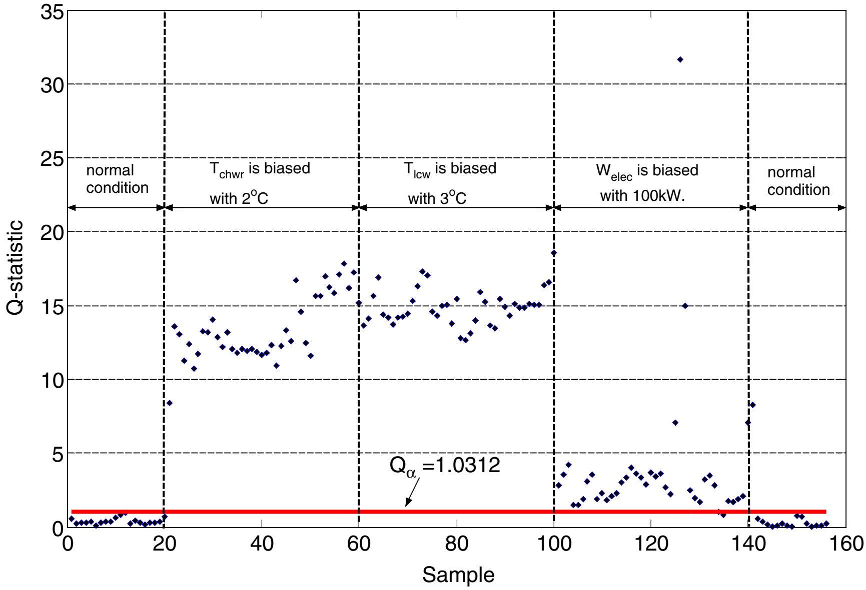

Fig. 6. *Q*-statistic plot of PCA model **B** – Test 2.

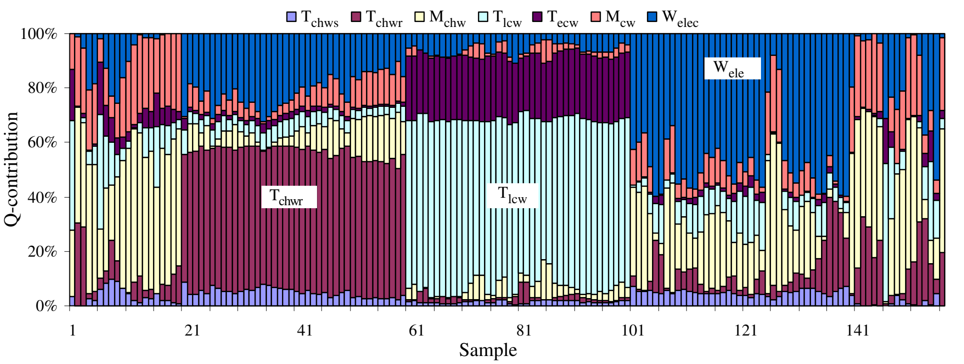

Fig. 7. *Q*-contribution plot of PCA model **B** – Test 2.

pre-determined bias existed in one of the three sensors. However, it is worth pointing out that the sensor faults of T_{chwr} and T_{lcw} could be detected by Q-statistic plots of both PCA model **A** and model **B** since these two variables are both included in the two models. Only the results in relation to PCA models are presented in this paper. Simultaneously, the biases of the three faulty sensors in this test were estimated correctly as shown in Table 4.

It can be found from the test results that most of the introduced sensor-biases can be estimated accurately with relative estimation errors lower than 10%, except for the electrical power input whose relative error was 24.9%. This might be explained as follows. The electrical-power input is, to a great extent, affected by variables, such as water temperatures and flow rates, and tends to experience a wider range of variation compared with these variables themselves. Therefore, if there is a bias error with the measurement of the electrical power input, it is more difficult for the FDD&E strategy to accurately estimate the magnitude of this bias. However, in practical applications, the measurement of electrical power is much more accurate and reliable compared with other thermal physical measurements. In this regard, the main concerns are on the monitoring and FDD of the thermal physical measurements. Up to now, it has been validated that the proposed sensor FDD strategy based on PCA for centrifugal chiller system is capable of detecting, diagnosing faulty-sensors as well as estimating the magnitudes of the bias errors.

5. Conclusions

Sensors are of great significance for chiller-performance monitoring, fault diagnosis and optimal control of chilling plants. A sensor FDD&E strategy based on the PCA method is developed and presented. The results of tests on a real centrifugal chiller plant confirmed that the sensor FDD&E strategy can effectively detect, diagnose and estimate sensor faults in centrifugal chillers successfully. Two PCA models, each of which contains a set of variables with strong correlations, are capable of effectively decoupling the systematic variations of centrifugal chillers from random noises. The uses of Q-statistic, Q-contribution plot and SPE minimization, respectively, as tools of fault detection, diagnosis and estimation were tested to be effective in the applications concerning sensors in chiller plants.

PCA models provide a good means of generating useful residuals for sensor-fault detection and diagnosis. Another advantage of PCA models in chiller-sensor FDD application is that the thresholds for fault detection can be easily calculated, with a specified confidence level, rather than empirically designated. With the help of a well-designed time-limiting program to eliminate false outputs, the fault diagnosis based on Q-contribution can be capable of accurately identifying faulty sensors in online applications. It can be expected that the combined use of the sensor FDD&E along with chiller-efficiency monitoring, component FDD and optimal control can significantly improve the reliability of the BMS as well as enhance the overall energy efficiency of chiller plants and buildings.

Acknowledgement

The research presented in this paper was financially supported by a grant from the Research Grants Council (RGC) of the Hong Kong SAR.

References

[1] Comstock MC, Chen B, Braun JE. Literature review for application of fault detection and diagnostic methods to vapor-compression cooling equipment. Ray W. Herrick Laboratories Report HL99-19, Report #4036-2. School of Mechanical Engineering, Purdue University, W. Lafayette, IN; 1999.

[2] Reddy TA, Niebur D, Gordon J, Seem J, Cabrera G, Jia Y, et al. Development and comparison of on-line model training techniques for model-based FDD methods Applied to vapor-compression chillers: evaluation of mathematical models and training techniques. Final report, ASHRAE Research Project 1139; 2001.

[3] Dexter Arthur, Jouko Pakanen. Fault detection and diagnosis methods in real buildings. International Energy Agency, Energy conservation in buildings and community systems, Annex 34: Computer-aided evaluation of HVAC system performance; 2001.

[4] Wang SW, Burnett J. Online adaptive control for optimizing variable-speed pumps of indirect water-cooled chilling systems. Appl Thermal Eng 2001;21(11):1083–103.

[5] Chan KT, Yu FW. Applying condensing-temperature control in air-cooled reciprocating water chillers for energy efficiency. Appl Energy 2002;72(3–4):565–81.

[6] Chang Yung-Chung. A novel energy conservation method – optimal chiller loading. Electric Power Syst Res 2004;69(2–3):221–6.

[7] Annex 34 – computer-aided evaluation of HVAC system performance: the practical application of fault detection and diagnosis techniques in real buildings. International Energy Agency; 1997.

[8] Phelan J, Brandemuehl M, Krarti M. In-situ performance testing of chillers for energy analysis. ASHRAE Trans 1997;103(1):290–302.

[9] Hartman TB. Instrumentation issues for monitoring chiller plant efficiency. ASHRAE Trans 2001;107(Part 1):407–14.

[10] Ian BD. McIntosh. A model-based fault detection and diagnosis methodology for HVAC subsystems. PhD thesis, University of Wisconsin, Madison; 1999.

[11] Ogaji SOT, Singh R, Probert SD. Multiple-sensor fault-diagnoses for a 2-shaft stationary gas-turbine. Appl Energy 2002;71(4):321–39.

[12] Dorr R, Kratz F, Ragot J, Loisy F, Germain JL. Detection, isolation, and identification of sensor faults in nuclear power plants. IEEE Trans Control Syst Technol 1997;5(1):42–60.

[13] Stylianou M, Nikanour D. Performance monitoring, fault detection, and diagnosis of reciprocating chillers. ASHRAE Trans 1996;102(1):615–27.

[14] Lee Won-Yong, House John M, Kyong Nam-Ho. Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks. Appl Energy 2004;77(2):153–70.

[15] Wang SW, Wang JB. Law-based sensor fault diagnosis and validation for building air-conditioning systems. Int J HVAC&R Res 1999;5(4):353–80.

[16] Wang SW, Wang JB. Automatic sensor evaluation in BMS commissioning of building refrigeration-systems. Automat Constr 2002;11(1):59–73.

[17] Lennox J, Rosen C. Adaptive multi-scale principal-components analysis for online monitoring of wastewater treatment. Water Sci Technol 2002;45(4–5):227–35.

[18] Pranatyasto TN, Qin SJ. Sensor validation and process fault diagnosis for FCC units under MPC feedback. Control Eng Pract 2001;9(8):877–88.

[19] Wang SW, Xiao F. AHU sensor fault diagnosis using principal-component analysis method. Energy Buildings 2004;36(2):147–60.

[20] Jackson J Edward. A user's guide to principal components. John Wiley & Sons Inc; 1991.

[21] Jolliffe IT. Principal component analysis. New York Inc: Springer-Verlag; 1986.

[22] Edward Jackson J, Govinod S Mudholkar. Control procedures for residual associated with principal component analysis. Technometrics 1979;21(3):341–9.
[23] Jia Yongzhong. Model-based generic approaches for automated fault detection, diagnosis, evaluation (FDDE) and for accurate control of field operated centrifugal chillers. PhD thesis, Drexel University; 2002.
[24] Braun JE. Methodologies for the design and control of central-cooling plant. PhD thesis, University of Wisconsin, Madison; 1988.
[25] Pacific Gas and Electric. CoolTools: a Toolkit to optimize chilled-water plants. San Francisco, CA; 2001.
[26] Dunia R, Qin SJ. Joint diagnosis of process and sensor faults using principal component analysis. Control Eng Pract 1998;6(4):457–69.
[27] Wise BM, Ricker NL, Velkamp DJ, Kowalski BR. A theoretical basis for the use of principal-component models for monitoring multivariate processes. Process Control Qual 1990;1:41–51.