**Development of clustering-based sensor fault detection and diagnosis strategy for chilled water system**

X.J. Luo1, K.F. Fong1,\*, Y.J. Sun1, M.K.H. Leung2

1Division of Building Science and Technology, College of Science and Engineering, City University of Hong Kong

2School of Energy and Environment, City University of Hong Kong

\*Corresponding author

Email address: bssquare@cityu.edu.hk

Postal address: Tat Chee Avenue, Kowloon, Hong Kong, China

**Abstract**

This paper presents a new clustering-based sensor fault detection and diagnosis (SFDD) strategy for chilled water system. For data clustering, *k*-means algorithm was used and the optimal quantity of clusters was determined by Davis-Bouldin value. With the cluster centroid dataset, the featuring centroid score (*CS*) was determined for the fault-free sensor reading dataset thus the threshold for fault detection could be set. The database for sensor fault detection was then formed. By characterizing the *CS* patterns of different types of sensor fault, the database for sensor fault diagnosis was generated accordingly. Various sensor fault types could be handled, including bias, drift, precision degradation and complete failure. In this study, the developed SFDD strategy was applied to the sensor of primary chilled water return temperature in a water-cooled chilled water system. With the databases of sensor fault detection and diagnosis, the real-time measured sensor readings can be examined. Once sensor fault is detected, the fault type can be confirmed within a day at soonest or 2 days at most. The smallest absolute bias value, absolute drifting rate and precision degradation error identified could be down to 0.25 °C, 0.025 °C/h and 0.1 °C respectively, demonstrating robustness of the proposed SFDD strategy.

Keywords: Fault detection and diagnosis; Sensor; Chilled water system; *k*-means clustering; Centroid score; Data mining.

**1. Introduction**

Heating, ventilating and air-conditioning system is commonly designed in modern high-rise commercial buildings. A centralized chiller plant hence a chilled water system is installed to provide central air-conditioning. The operating performance of chilled water system is affected by the changing ambient air conditions (primarily dry bulb air temperature and relative humidity), the indoor heat gains (occupancy, lighting, electrical equipment and any miscellaneous sources), the various equipment members, the control and operation strategy. In the chilled water system, reliable and accurate measurements are essential for monitoring chiller performance and deciding its operating condition and sequence. On the contrary, sensor faults would result in improper control on the chilled water system, and eventually result in higher energy consumption or thermal comfort problem. Generally, fault detection and diagnosis (FDD) strategies can be classified into three categories: quantitative model-based methods, qualitative model-based methods and process history-based methods [1-5].

In quantitative model-based methods [6-10], physical models were usually developed to simulate the steady and transient behavior of chilled water systems. The quantitative models were based on fundamental physical principles, engineering relationships and physical processes. The quantitative results from applying the mathematical models to different operating conditions set the baseline fault-free performance. The residuals, defined as the difference between measurements from the actual system and the baseline fault-free performance, would be used to detect system and sensor faults. The qualitative model-based methods mainly referred to rule-based models [11-12]. The rules were generally derived from knowledge of the fundamental physical processes occurring in the chilled water system. Fault detection was based on the condition that the measurements from actual operating system exceeded those prescribed upper or lower bounds of certain assessment criterial. Fault diagnosis was conducted by analyzing the changes in each measured variable. However, the quantitative models needed to be validated with experimental data for both fault-free and faulty operation. Thus, large amounts of experiments were generally needed. The total quantity of sensors was also increased due to additional measurements required. Meanwhile, the effectiveness of the qualitative model-based FDD method depended on the expertise of the model developer. More importantly, both quantitative and qualitative models were generally specific to certain equipment schedule of chilled water systems, which made it difficult to extend their application on other type of systems.

To overcome the deficits in quantitative and qualitative model-based FDD methods, process history-based FDD strategies have been investigated. In process history-based FDD method, data-driven black box models were generally formulated to find out the relationship between different operating variables of the chilled water system. Polynomial regression [13], association rule mining [14], principal component analysis (PCA) [15-17] and support vector data description (SVDD) [18-23] were commonly used to develop such data-driven models. For example, in [13], polynomial models were built using regression analysis to observe the operating condition of a centrifugal chiller. The inputs of the model were operating temperatures and cooling load, while the outputs were six different performance indices. The calculated residual values of performance indices were used for fault detection and diagnosis purposes. In PCA model-based sensor FDD strategy [17], Q-statistic, Q-contribution plot and squared prediction error minimization were used for fault detection, diagnosis and estimation, respectively. In SVDD model-based sensor FDD strategy [19], the distance-based D-statistic plot and the distance variation based DV-contribution plot were employed for sensor detection and diagnosis purposes, respectively. Overall speaking, FDD strategies of the aforesaid techniques were based on the premise that faulty sensor readings or system operating performance indices were quite different from the fault-free ones. For example, in [17], only 2 °C and 3°C sensor bias were demonstrated. In [19], the lowest absolute sensor bias and drifting rate was 1 °C and 0.09 °C/h, respectively. The precision degradation was found difficult to be detected in [19]. Moreover, for PCA models, the examined dataset of operating variables was assumed to follow Gaussian distribution, which was not actually applicable in the practical chilled water system.

To tackle the problems mentioned above, clustering-based FDD strategies were recently developed using the data-mining approach. Clustering analysis is able to directly extract the useful information from a given data set, thus it can categorize the data into clusters which share common characteristics [24]. The use of clustering analysis approaches for fault detection has been studied and reported in several research areas [25-33], but only four of them were applied on chilled water systems. To be more specific, in [25], through clustering, fault-free and faulty training datasets of the chilled water system were partitioned into different clusters. One cluster represented the fault-free datasets while others stood for datasets under different faults. The fault type and corresponding severity levels were detected and diagnosed when the actual measurements were the closest to one of the predetermined cluster centroids. In [31], the combined neural networks were developed to detect the abnormities in sensors of the chilled water system, while clustering analysis was used to classify the various faulty conditions. Through subtractive clustering analysis, the different faults could be separated into different clusters. In [23], based on its chilled water flow rate, the operating conditions of the chilled water system was classified into two clusters through clustering. After that, two sub-PCA models were built to conduct sensor FDD analyses. In [29], a clustering-based sensor fault detection strategy was proposed for the air handling units. The data points in different clusters were checked for temporal separation to detect sensor faults. From these research works, it is seen that the clustering-based methods were effective in sensor and system FDD purposes. However, in [25], all the fault-free datasets were partitioned into one cluster. In [26], clustering analysis was only adopted as an auxiliary for PCA model. The ability of clustering analysis was not fully utilized. In addition, only bias sensor fault was demonstrated in [26,31], and the lowest absolute bias value was 1 °C. The existing fault detection approach proposed in [29] was only useful for offline bias sensor fault.

Based on the above literature review, it is found that the research regarding FDD of chilled water system using clustering analysis was still not significantly progressed yet. Therefore, a new sensor fault detection and diagnosis (SFDD) strategy was proposed in this study accordingly, in order to detect and diagnose various sensor faults for proper operation and preventive maintenance of the chilled water system. The fault types would include bias, drift, precision degradation and complete failure. It is expected that the proposed SFDD strategy would handle a wide range of faults.

**2. Chilled water system under study**

*2.1 System description*

In this study, a typical water-cooled chilled water system installed in a high-rise commercial building in Hong Kong was used to evaluate the performance of the proposed SFDD strategy. The chilled water system consisted of six identical centrifugal chillers, and its total cooling capacity was 43,380 kW. Each chiller had an associated chilled water pump, cooling water pump and cooling tower. Differential pressure bypass valve was used to allow surplus chilled water passing through when the air handling units were under part-load operation. The schematic diagram of the chilled water system is shown in Fig. 1, and its key design parameters are summarized in Table 1.

Table 1. Key design parameters of the chilled water system.

|  |  |
| --- | --- |
| Rated cooling capacity of each chiller (kW) | 7,230 |
| Rated chilled water flow rate of each chiller (L/s) | 345 |
| Set-point of chilled water supply temperature (°C) | 6 |
| Design chilled water return temperature (°C) | 11 |

Water-cooled chiller 1

*Tchws,s*

*mbp*

*Tchwr,s*

*mchw1 Tchw1*

*mchw2 Tchw2*

*mchw3 Tchw3*

Water-cooled chiller 2

Water-cooled chiller 3

*mchw4 Tchw4*

*mchw5 Tchw5*

*mchw6 Tchw6*

Water-cooled chiller 4

Water-cooled chiller 5

Water-cooled chiller 6

Air handling units

(installed around the building)

*Tchws,p mchws,p*

*Tchwr,p*

Cooling tower

*mcw1 Tcw1*

*mcw2 Tcw2*

*mcw3 Tcw3*

*mcw4 Tcw4*

*mcw5 Tcw5*

*mcw6 Tcw6*

Chilled water pump

Cooling water pump

Differential pressure bypass valve

Fig. 1. Schematic diagram of chilled water system in a high-rise commercial building.

*2.2 Sequencing control and cooling load profiles*

The sequencing control of the chilled water system followed the common practice, which was based on the calculated cooling load *Qc* determined from the primary chilled water return temperature *Tchwr,p*, the primary chilled water supply temperature *Tchws,p* and the primary chilled water supply flow rate *mchws,p* according to Eq (1).

*Qc* = *mchws,p* *Cp* (*Tchwr,p*− *Tchws,p*) (1)

where,

*Cp*: specific heat capacity of water (kJ/kg/°C).

Then the chiller switch-on or switch-off scheme for the sequencing control would be implemented by Eqs (2)-(4):

*Qon* = *nch,0* *Qch γon* (2)

*Qoff* = (*nch,0* −1) *Qch* *γoff* (3)

 (4)

where,

*nch*: quantity of chiller in operation in next time step.

*nch,*0:quantity of chiller in operation in current time step.

*Qch*: rated cooling capacity of each chiller (kJ/h).

*Qoff*: cooling threshold for chiller switch-off (kJ/h).

*Qon*: cooling threshold for chiller switch-on (kJ/h).

*γoff*: switch-off coefficient, *γoff* = 0.90.

*γon*: switch-on coefficient, *γon* = 0.95.

The minimum up- and down-time constraint was set at 30 minutes since a chiller should not be switched on immediately after it was off, or vice versa. For this chiller plant, it was designated that the corresponding chilled water pump of an idle chiller would be operated according to the design mass flow rate. Then the chilled water supply of this idle chiller would have the temperature same as that of the chilled water return, and would be further mixed with the chilled water supply from the operating chillers before serving the building. As such, the chilled water supply temperature would be higher than the design set-point in some occasions.

7 representative days in 4 seasons (total 28 days) of cooling load profiles were used as the required cooling demand in the following test cases. Data information used in this study is summarized in Table 2. Based on the measurements of *Tchwr,p* *Tchws,p* and *mchws,p*, *Qc* was obtained in the 28 days as shown in Fig. 2.

Table 2. Sample periods of four seasons.

|  |  |  |
| --- | --- | --- |
| Day | Season | Sample week |
| 1-7 | Spring | 17th April 2012 to 23rd April 2012 |
| 8-14 | Summer | 16th June 2012 to 22nd June 2012 |
| 15-21 | Autumn | 2nd October 2013 to 8th October 2013 |
| 22-28 | Winter | 28th November 2013 to 4th December 2013 |

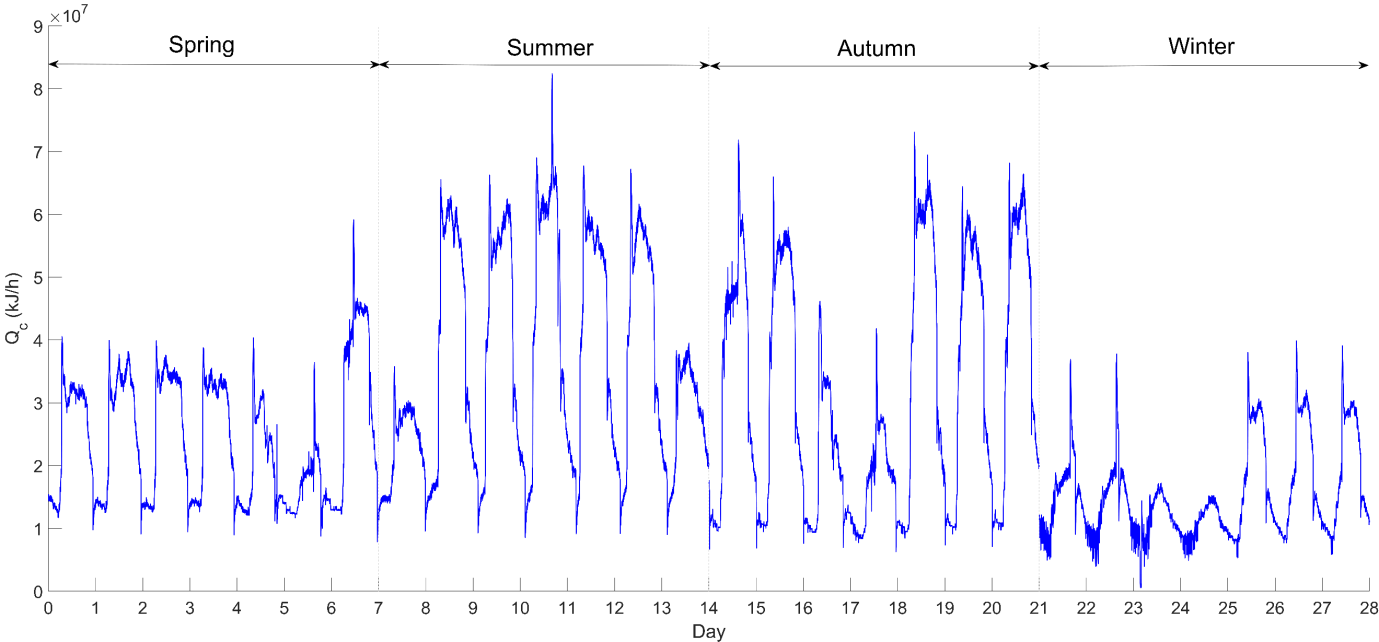


Fig. 2. Cooling load profiles of the commercial building on 7 representative days in 4 seasons.

*2.3 Validation of system model*

2.3.1 Initial validation

The entire chilled water system was simulated using TRNSYS 17 [34]. Types 666, 51 and 3 were used to simulate chiller, cooling tower and water pump, respectively. Type 9 was used to supply the outdoor air information data to the cooling tower. The equations of sequencing control as described in Eqs. (2-4) were programmed in MATLAB and embedded in TRNSYS 17 using component Type 155. In order to validate the system model, the simulation results of *Tchws,p*, *Tchwr,p* and *mchws,p* were compared with the measurement data as shown in Fig. 3.

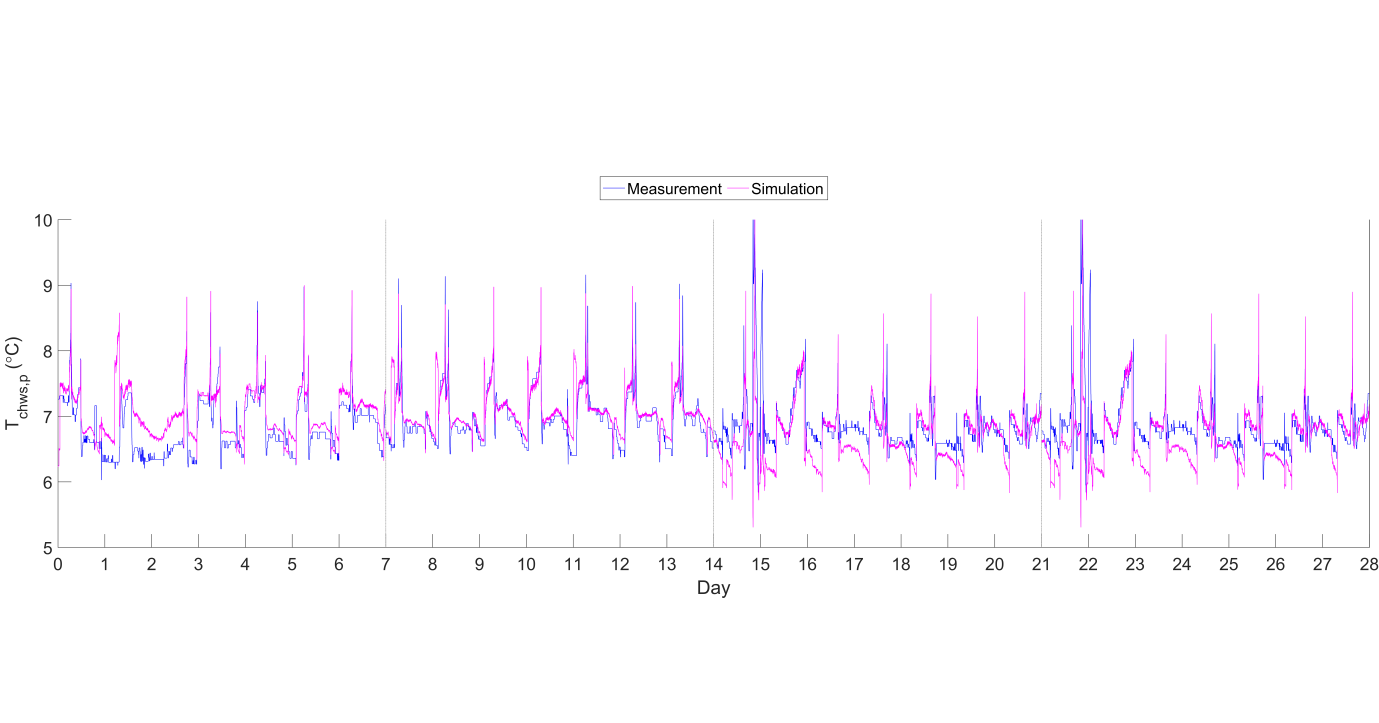
It is found that the root-mean-square errors of *Tchws,p* , *Tchwr,p* and *mchws,p* were 4.50%, 5.78% and 7.33% respectively, while their mean absolute errors were 0.36 °C, 0.14 °C and 6638.8 kg/h. Thus, the simulation results were acceptable, and the chilled water system model was deemed to be validated.

|  |
| --- |
| (a) *Tchws,p* |
| (b) *mchws,p* |
| (c) *Tchwr,p* |

Fig. 3. Measurement and simulation data on the representative 28 days.

2.3.2 Further validation

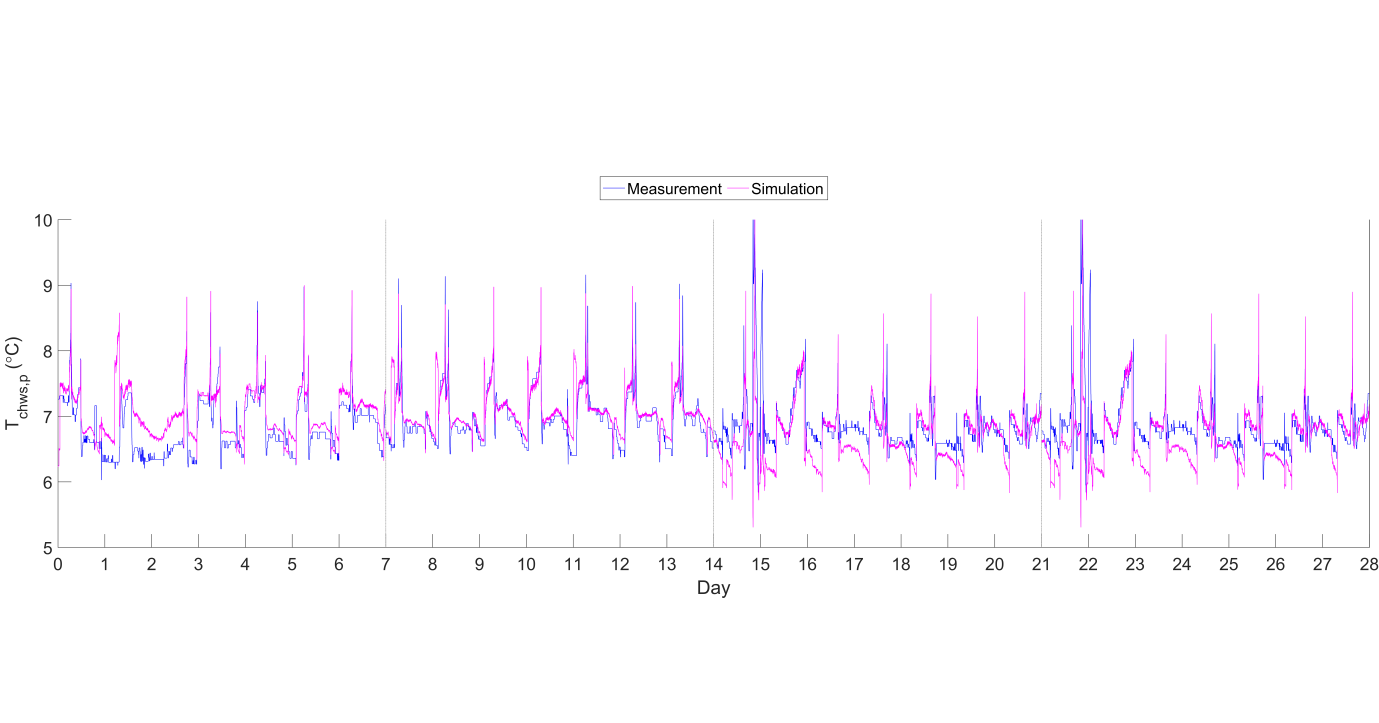
To further validate the simulation model, another 7 successive days in each season was randomly chosen. The simulation results of *Tchws,p*, *Tchwr,p* and *mchws,p* were compared with the measurement data as shown in Fig. 4. It is found that the root-mean-square errors of *Tchws,p* , *Tchwr,p* and *mchws,p* were 5.48%, 6.51% and 8.85% respectively, while their mean absolute errors were 0.01 °C, 0.37 °C and 94709 kg/h. As a result, the chilled water system model was reliable and it could be used to explore the system response due to various sensor faults. The 28 days formed by the representative week of each season would be used as the basis for ongoing development of the SFDD strategy.



(a) *Tchws,p*



(b) *mchws,p*



(c) *Tchwr,p*

Fig. 4. Measurement and simulation data on another 28 days.

*2.4 Fault test cases of critical sensor*

Since the operating sequence of chillers depended on *Tchws,p* (*T*1), *Tchwr,p* (*T*2) and *mchws,p* according to the chiller sequencing control algorithm, these three sensors were regarded as the essential sensors. Readings of these three sensors were collected every minute. Thus, the total time step is 28 × 24 × 60 = 40,320. To formulate the SFDD strategy, a total quantity of 40,320 × 3 fault-free sensor readings, which were generated from the chilled water system simulation model developed on TRNSYS, would be used to establish the baseline database for sensor fault detection. The fault-free sensor reading base case was named Case 0. Since the sensor fault of *T*1 and *mchws,p* can be detected according to the set-point of chilled water supply temperature and the operating quantity of chillers [10], the SFDD algorithm focused on *T*2, which became the critical sensor in this study. Through the chilled water system model again, different types of sensor fault were therefore added on *T*2 in order to evaluate the corresponding fault patterns. The mathematical description of sensor fault reading is shown in Eq. (5).

*T*2*,i,e* = *T*2*,i,o* + *ei* (5)

where,

*T*2*,i,e*: faulty sensor reading of *T*2 at time step *i*.

*T*2*,i,o*: fault-free sensor reading of *T*2 at time step *i*.

*ei*: the measurement difference between faulty and fault-free sensor at time step *i*.

The detailed information of the test cases of *T*2 sensor fault is summarized in Table 3.

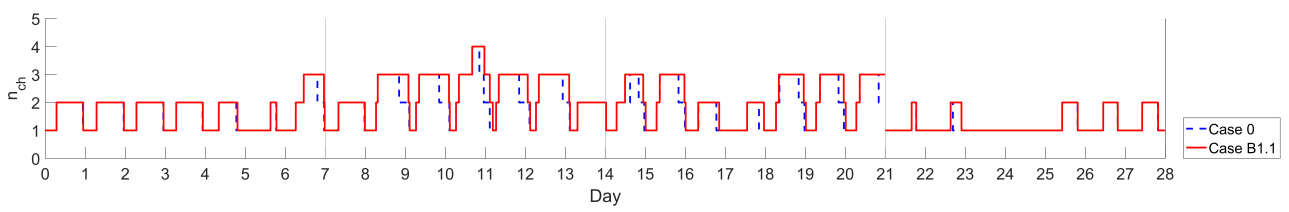
Table 3. Test cases of *T*2 sensor fault.

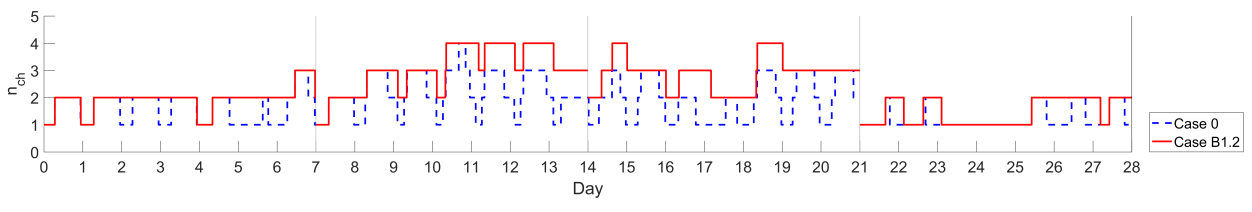
|  |  |  |  |
| --- | --- | --- | --- |
| Fault type | Fault description | Case | Fault value |
| Fault-free | *ei* = 0 | 0 (base case) | --- |
| Positive bias (+δ) | *ei* = + *δ* | B1.1 | *δ* = 0.25 °C |
| B1.2 | *δ* = 0.50 °C |
| B1.3 | *δ* = 0.75 °C |
| Negative bias (-δ) | *ei* = − *δ* | B2.1 | *δ* = 0.25 °C |
| B2.2 | *δ* = 0.50 °C |
| B2.3 | *δ* = 0.75 °C |
| Positive drift (+δ) | *ei+1* = *ei* + *δ* | D1.1 | *δ* = 0.025 °C/h drift |
| D1.2 | *δ* = 0.05 °C/h drift |
| D1.3 | *δ* = 0.1 °C/h drift |
| Negative drift (-δ) | *ei+1* = *ei* − *δ* | D2.1 | *δ* = 0.025 °C/h drift |
| D2.2 | *δ* = 0.05 °C/h drift |
| D2.3 | *δ* = 0.1 °C/h drift |
| Precision degradation (*μ*) | *ei* = G(0, *μ*) | P1 | *μ* = 0.1 °C |
| P2 | *μ* = 0.5 °C |
| P3 | *μ* = 1 °C |
| Complete failure (*δ*) | *T*2*,i,e* = *δ* | C1 | *δ* = 9 °C |
| C2 | *δ* = 10 °C |
| C3 | *δ* = 11 °C |

For bias, precision degradation and complete failure faults, the fault value remained during the whole testing period. For drift fault, the drifting rate increased from the beginning of each 7-day period (i.e. 1st, 8th, 15th and 22nd day). Based on the faulty sensor reading of *T2*, sensor readings of *T1*, *mchws,p* and the quantity of operating chillers *nch* were obtained from the system simulation model under each test case.

*2.5 Effect of sensor faults on chiller operation*

To study the effects of different sensor faults on chiller operation, the chiller operating sequence in each test case compared to the base case is shown in Fig. 5. It is seen that when sensor fault of *T2* occurred, the operating sequences of the chillers would deviate from the fault-free base case. For positive bias or positive drift sensor faults, more frequent chiller operation happened. For sensor fault of negative bias or negative drift, there was less frequent chiller operation. For complete failure, the quantity of operating chillers was flat in each 7-day period. When sensor reading of *T2* was fixed at 9 °C or 10 °C, the calculated cooling load *Qc* did not exceed the cooling thresholds (i.e. *Qon* and *Qoff*), thus the operating sequence under Cases C1 and C2 were the same. The over-provision of operating chiller in positive bias and drift would cause extra electricity consumption. On the other hand, the under-provision of chiller in negative bias and drift sensor faults resulted in insufficient supply of chilled water to the air side equipment, hence causing thermal comfort problem.





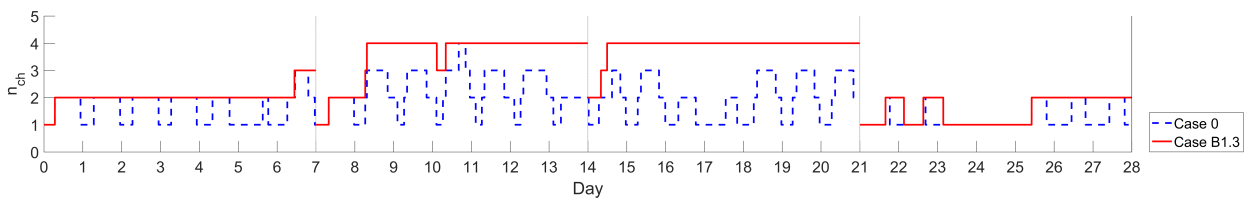
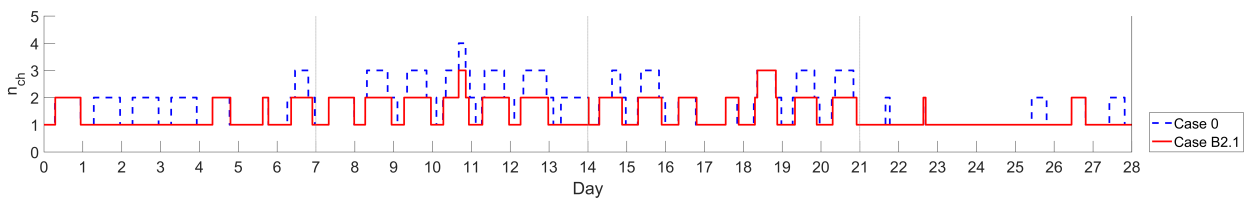
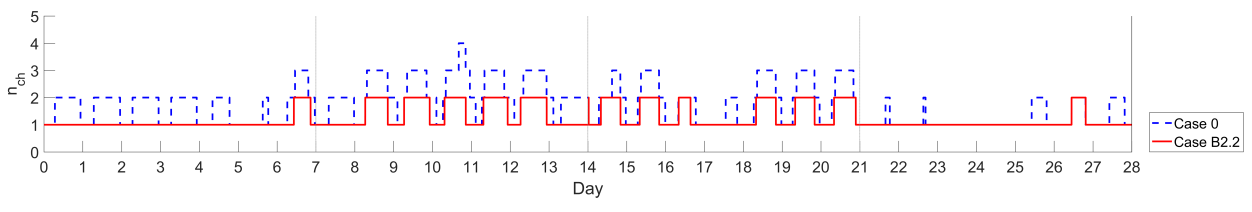


Fig. 5(a). Sequence of chiller operation in positve bias test cases.





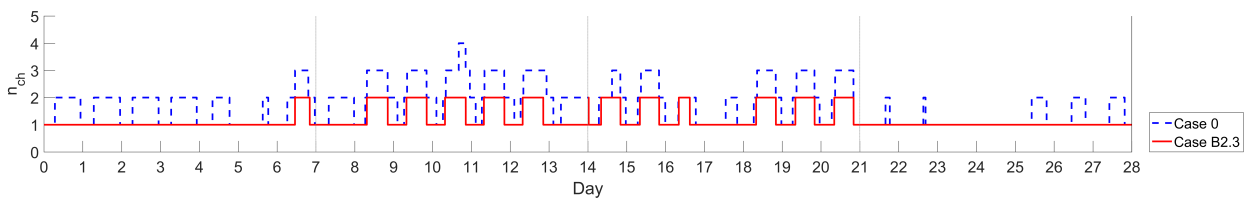
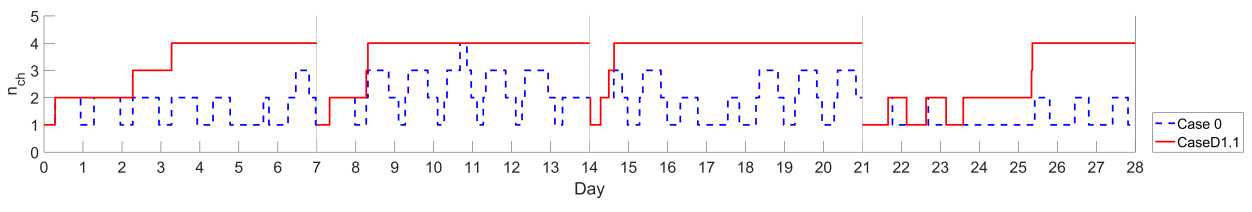
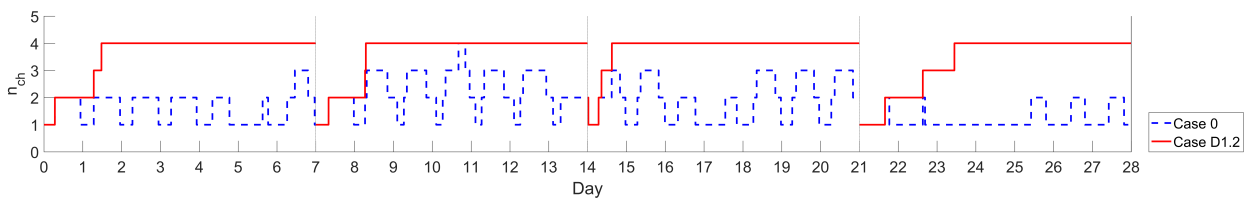


Fig. 5(b). Sequence of chiller operation in negative bias test cases.





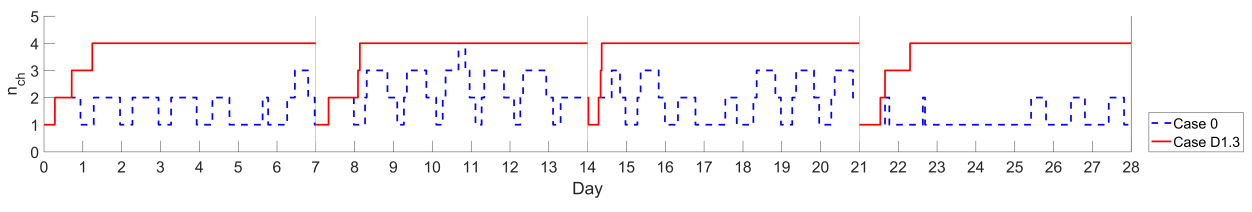
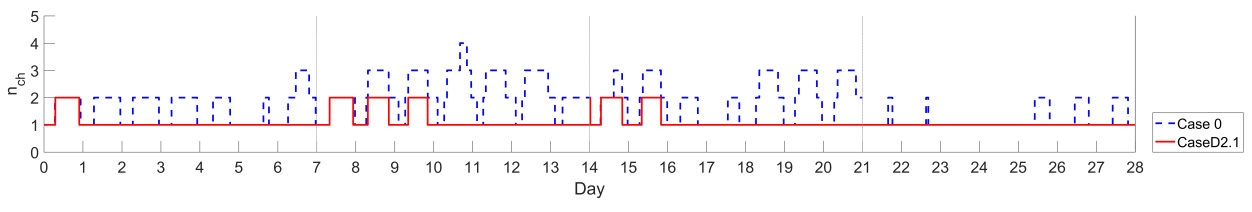
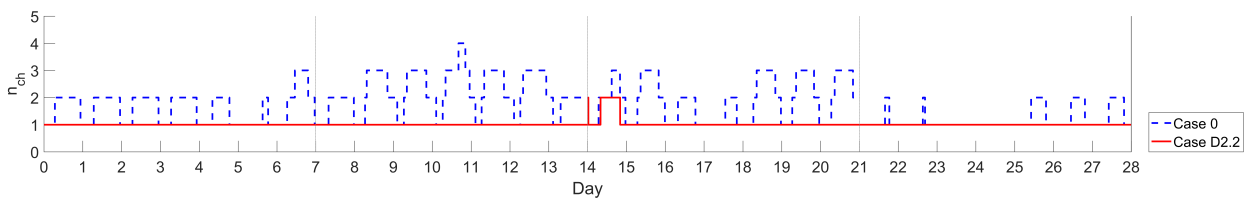


Fig. 5(c). Sequence of chiller operation in positive drift test cases.





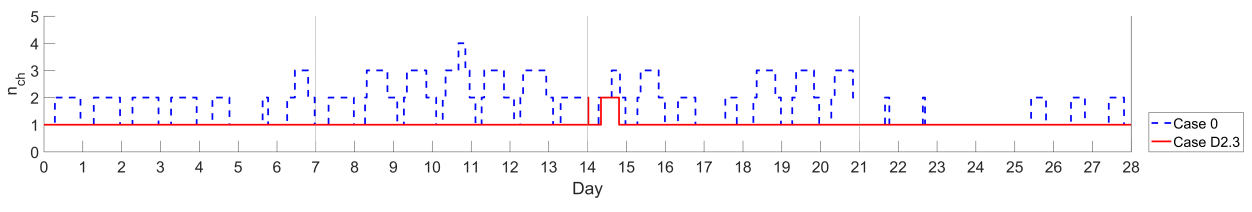
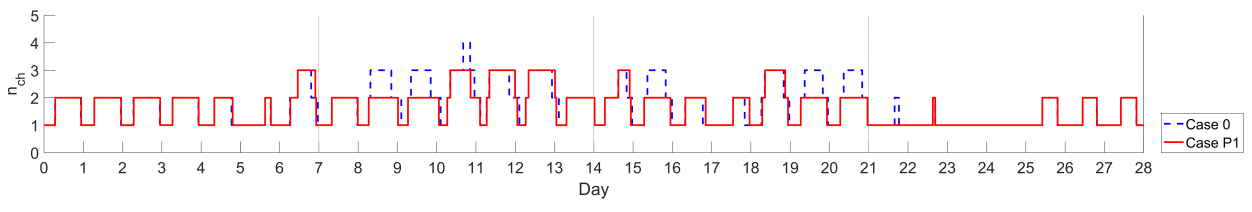


Fig. 5(d). Sequence of chiller operation in negative drift test cases.



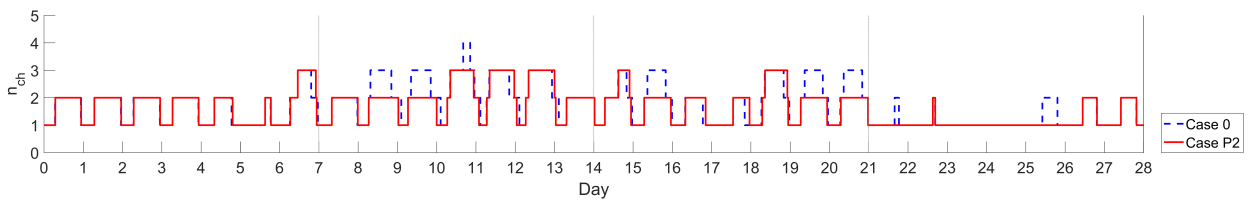
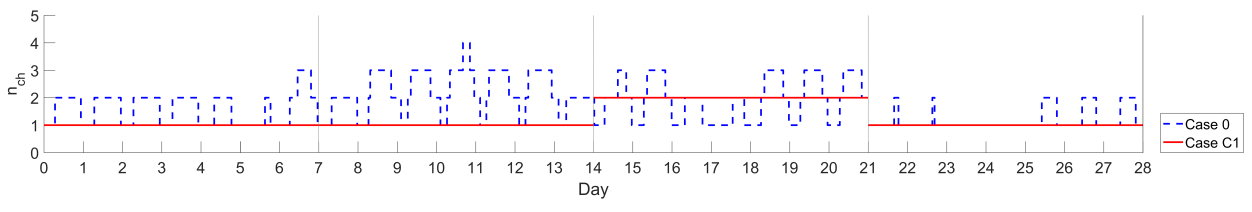
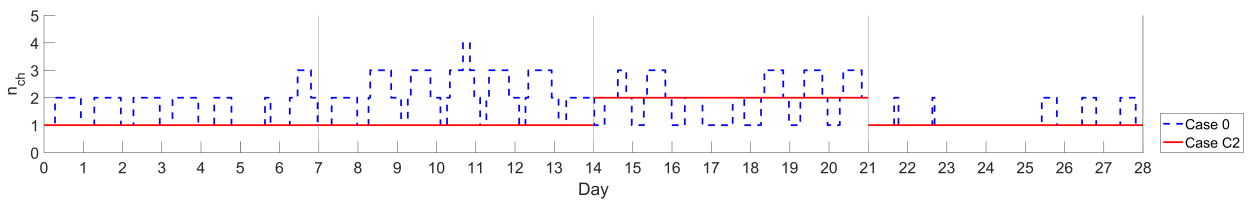




Fig. 5(e). Sequence of chiller operation in precision degradation test cases.





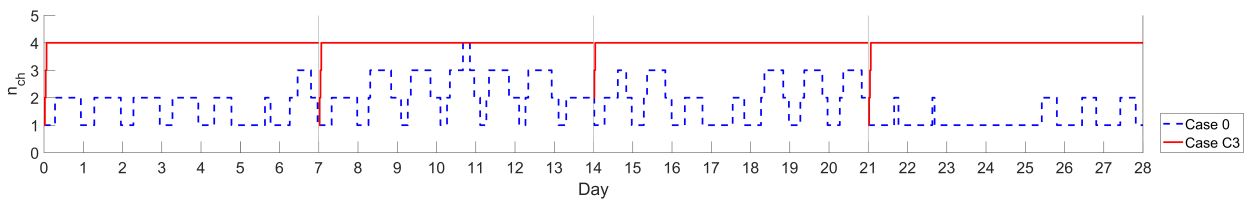


Fig. 5(f). Sequence of chiller operation in complete failure test cases.

**3. Formulation of the clustering-based SFDD strategy**

The objective of data clustering algorithm is to group dataset into different clusters so that the dataset in the same cluster have high similarity. In this study, *k*-means clustering algorithm was employed. And the cluster centroid was defined as the mean value of the dataset within the cluster. The *k*-means algorithm is an iterative process that assigns each dataset the cluster with the closest centroid by calculating the Euclidean distance between each dataset and its corresponding cluster centroid. Through this clustering-based method, the fault-free sensor readings were close to the centroids of cluster while those faulty ones would be far away. The proposed SFDD strategy consisted of database building for sensor fault detection; database building for sensor fault diagnosis; and online SFDD for measurement data.

*3.1 Database building for sensor fault detection*

To prepare the database for fault detection, *tmax× M* original fault-free sensor reading dataset **Yo** = {*yi,j,o* | *i* = 1, 2, …, *tmax*; *j* = 1, 2, …, *M*} was normalized to become the dataset **Xo** = {*xi,j,o* | *i* = 1, 2, …, *tmax*; *j* = 1, 2, …, *M*} using Eq. (6) as follows:

|  |  |
| --- | --- |
|  | (6) |

where,

*i*: time step, *i* = 1, 2, …, *tmax*.

*j*: sensor symbol or number, *j* = 1, 2, …, *M*.

*M*: total quantity of sensor (*M* = 3 for this study, as mentioned in Section 2.4).

*tmax*: maximum time step (*tmax* = 40,320 for this study, as mentioned in Section 2.4).

*xi,j,o*: normalized fault-free sensor reading.

*yi,j,o*: original fault-free sensor reading.

Next, by using the *k*-means clustering algorithm, the normalized dataset **Xo** was clustered with different quantity of clusters *Gk,o* (*k* = 1, 2, …, *N*) at each time step. Meanwhile, the cluster centroid subset of cluster *k*, *Ck* = {c*k,j* | *j* = 1, 2, …, *M*} was found. In order to evaluate the quality of clustering, Davis-Bouldin (*DB*) value [35] was employed to determine the optimal quantity *Nopt* of clusters as follows:

|  |  |
| --- | --- |
|  | (7) |

where,

*DBN*: DB value for *N* clusters.

*k*1, *k*2: cluster number, *k*1 = 1, 2, …, *N*; *k*2 = 1, 2, …, *N*; *k*1 ≠ *k*2.

: normalized sensor reading subset of cluster *k*1 at each time step, and = {*xi,j,o* | *j* = 1, 2, …, *M*}, where = **Xo**.

: quantity of data object included in cluster *k*1.

*Xi,o* : normalized sensor reading subset at each time step, and *Xi,o* = {*xi,j,o* | *j* = 1, 2, …, *M*}.

: Euclidean distance.

Lower value of *DBN* indicated better clustering quality, so the optimal quantity of clusters *Nopt*, was determined according to the lowest value. As such, the cluster centroid dataset **C** = = {c*k,j* | *k* = 1, 2, …, *Nopt*; *j* = 1, 2, …, *M*} could then be established. Meanwhile, the fault-free centroid score *CSi,j,o* of each sensor *j* at each time step *i* was calculated below:

 (8)

where,

*k*: cluster number, *k* = 1, 2, …, *Nopt*.

*c*: cluster centroid.

: normalized sensor reading subset of cluster *k* at each time step, where = **Xo**.

And the fault-free total centroid score *CSi,t,o* at time step *i* was calculated as:

 (9)

Let the centroid score sets ***CSj,o*** = {*CSi,j,o* | *i =* 1, 2, …, *tmax*} and ***CSt,o*** = {*CSi,t,o* | *i =* 1, 2, …, *tmax*}. The confidence interval *CI* of the 99th, 97th and 95th percentiles of ***CSj,o*** were determined as thresholds *CI99,j*, *CI97,j* and *CI95,j*, respectively. Meanwhile, the *CI* of the 99th, 97th and 95th percentiles of ***CSt,o*** were found as thresholds *CI99,t*, *CI97,t* and *CI95,t*, respectively.

As a result, the cluster centroid dataset **C** and the thresholds *CIj* and *CIt* would constitute the database for sensor fault detection.

*3.2 Database building for sensor fault diagnosis*

The different faulty sensor readings described in Table 3 were used as the training dataset to build the database for sensor fault diagnosis. To begin with, the faulty sensor reading dataset **Ye** = {*yi,j,e* | *i* = 1, 2, …, *tmax*; *j* = 1, 2, …, *M*} was normalized using the corresponding maximum and minimum values found with the fault-free sensor readings, leading to **Xe** = {*xi,j,e* | *i* = 1, 2, …, *tmax*; *j* = 1, 2, …, *M*} as follows:

|  |  |
| --- | --- |
|  | (10) |

After that, the normalized faulty sensor reading subset *Xi,e* = {*xi,j,e* | *j* = 1, 2, …, *M*} at each time step *i* was assigned into corresponding clusters based on the cluster centroid dataset **C** generated from the previous stage for sensor fault detection. And the centroid score of the faulty sensor readings *CSi,j,e* and the total centroid score *CSi,t,e* were calculated as follows:

 (11)

 (12)

where,

: normalized faulty sensor reading subset of cluster *k* at each time step, where = **Xe**.

By obtaining the *CSi,j,e* and *CSi,t,e* patterns of each type of sensor fault and comparing against those of the base case, the characteristics of different fault types could be identified. As a result, the database for sensor fault diagnosis could be generated.

*3.3 Online SFDD for measurement data*

After generating both databases for sensor fault detection and diagnosis, it is ready for practical application in the chilled water system. At each time step, the online real-time measurements *Yi,m* = {*yi,j,m* | *j* = 1, 2, …, *M*} are normalized to *Xi,m* = {*xi,j,m* | *j* = 1, 2, …, *M*} as follows:

 (13)

Then the centroid score of the measured sensor readings *CSi,j,m* and the total centroid score *CSi,t,m* can be determined below:

 (14)

 (15)

where,

: normalized real-time measured sensor reading subset of cluster *k* at each time step, where = *Xi,m*.

In online SFDD, the primary criterion is whether *CSi,j,m*, which is directly associated to the measured sensor reading *T*2, is higher than the threshold *CIj*. Another criterion of sensor fault detection is established from *CSi,t,m*, which is related to all the operating variables concerned (i.e. *T*1, *T*2 and *mchws,p*). If the measured sensor reading of *T*2 were faulty, it would affect the operating sequence of the chilled water system, causing *T*1 and *mchws,p* abnormal as well. Then, *CSi,t,m* would be higher than the threshold *CIt*. As a result, sensor fault is detected if “*CSi,j,m > CIj* or *CSi,t,m > CIt*” as shown in Fig. 6. Then, the patterns developed by *CSi,j,m* and *CSi,t,m* can be used to compare with those in the database of fault patterns for sensor fault diagnosis as obtained from the previous stage.

The entire SFDD strategy, comprising database building for sensor fault detection; database building for sensor fault diagnosis; and online SFDD for measurement data, is consolidated in Fig. 6.

**SFDD strategy**

**Database building for sensor fault detection**

***CSj,o***, ***CSt,o***

**Online SFDD for measurement data**

Data normalization: Eq. (13)

**Sensor**

**fault detection**

*CSi,j,m* > *CIj*

or *CSi,t,m* > *CIt*

?

N

No fault

Y

**Sensor fault diagnosis** Fig. 15 (Section 4.3.5)

*Yi,m*

*Xi,m*

*CSi,j,m, CSi,t,m*

**Database building for sensor fault diagnosis**

Data normalization: Eq. (6)

Determination of

*CIj* and *CIt*

Data clustering

**Yo**

*Nopt*

**C**

*CIj* and *CIt*

Determination of *CS*:

Eqs. (8) & (9)

*Gk,o*, **C**

***Database for***

***sensor fault detection***

Finding *Nopt*: Eq. (7)

**Xo**

***CSj,e,*** ***CSt,e***

Data normalization: Eq. (10)

Data clustering

**Ye**

Determination of *CS*:

Eqs. (11) & (12)

**Xe**

Characterization of

*CSi,j,e* and *CSi,t,e* patterns

for various fault types

Data clustering

Determination of *CS*:

Eqs. (14) & (15)

***Database of fault patterns for sensor fault diagnosis***

*Gk,e*

*Gk,m*

Fig. 6. Flow chart of SFDD strategy.

**4. Implementation of SFDD strategy on the chilled water system**

The SFDD strategy was applied in order to generate the database of sensor fault detection and diagnosis. Based on the practical chilled water system described in Section 2, the optimal quantity of clusters *Nopt* was firstly determined according to the *DB* value. Meanwhile, the cluster centroid dataset **C** was obtained and the various thresholds were determined for sensor fault detection purpose. After that, the characteristics of *CSi,T2,e* and *CSi,t,e* patterns in each fault type would be identified and served for sensor fault diagnosis purpose.

*4.1 Determination of optimal quantity of cluster and cluster centroid dataset*

The normalized dataset **Xo** of Case 0 was clustered by *k*-means clustering algorithm using 2-50 clusters, hence the values from *DB*2 to *DB*50 were obtained to find out the optimal quantity of clusters, as shown in Fig. 7. In order to balance problem complexity and computational time, it is reasonable to consider *N* up to 50. It is seen that *DB* had the lowest value when *N* was 42, hence *Nopt* = 42.

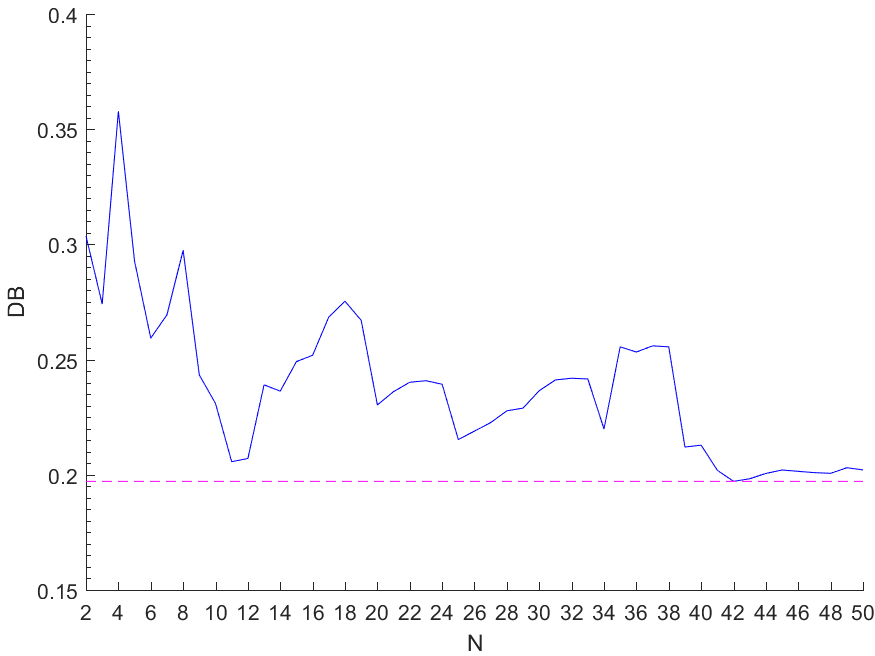


Fig. 7. *DB* of different quantity of cluster

As a result, *Gk,o* and *Ck* could be determined based on *Nopt*. Meanwhile, **C** was established accordingly. To further understand the performance of the clustering algorithm, the clustering result when *Nopt* = 42 is shown in Fig. 8. It is seen that the normalized subset *Xi,o* from the fault-free sensor readings at different time steps along 28 days were assigned into the 42 clusters *Gk,o*, as represented by the corresponding color types in Fig. 8.

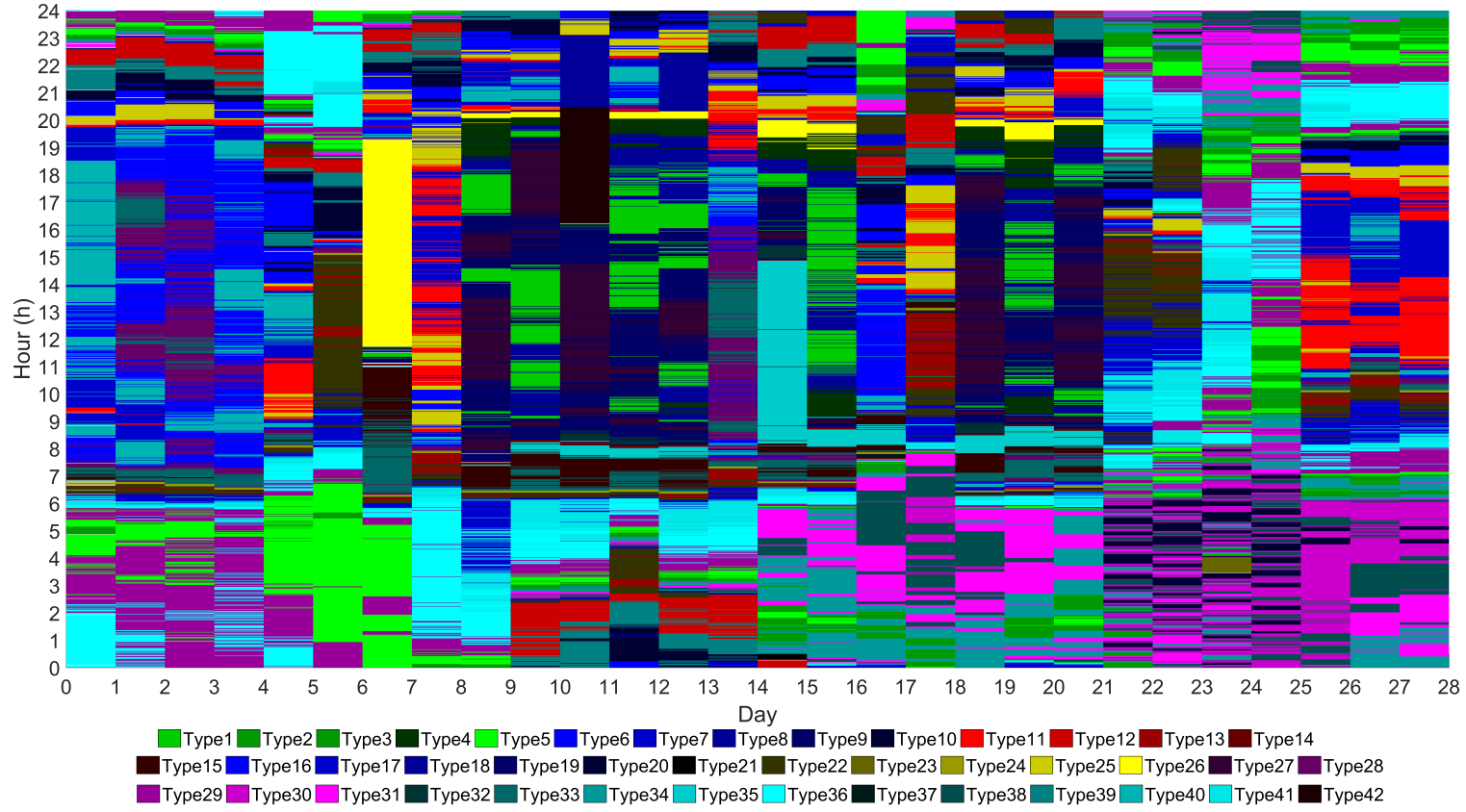


Fig. 8. Temporal contours of k-means clustering for *Nopt* = 42.

*4.2 Determination of thresholds for* ***CST2,o*** and***CSt,o*** *for sensor fault detection*

Based on the clustering result of the fault-free sensor reading, the probability distributions of***CST2,o*** and***CSt,o*** are shown in Fig. 9.

|  |  |
| --- | --- |
| (a) ***CST2,o*** | (b) ***CSt,o*** |

Fig. 9. Histograms of *CS*and threshold values determined by 3 confidence intervals.

According to the CI of the 99th, 97th and 95th percentiles of***CST2,o*** and***CSt,o***, the thresholds *CI99,T2*, *CI97,T2*, *CI95,T2* and *CI99,t*, *CI97,t*, *CI95,t* were determined for fault detection purpose, as summarised in Table 4.

Table 4. Thresholds for***CST2,o*** and***CSt,o***.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Threshold name | *CI99,T2* | *CI97,T2* | *CI95,T2* | *CI99,t* | *CI97,t* | *CI95,t* |
| Threshold value | 2.76 | 2.22 | 2.08 | 4.82 | 3.54 | 3.18 |

*4.3 Characterization of CSi,T2 and CSi,t patterns*

When different types of sensor faults were added to*T2*, the normalised subset *Xi,e* of each test case at each time step was partitioned into *Gk,o* (*k* =1, 2, …, 42) according to the cluster centroid subset *Ck*. Thus, the *CSi,T2,e* and *CSi,t,e* patterns under different sensor faults (i.e. positive bias, negative bias, positive drift, negative drift, precision degradation and complete failure) were compared with those of *CSi,T2,o* and *CSi,t,o* in Case 0, so as to identify the corresponding distinguishable features for sensor fault diagnosis purpose. Characterization of *CSi,T2,e* and *CSi,t,e* patterns for the various fault types is conducted in the following subsections.

4.3.1 Bias sensor reading fault

As shown in Figs. 10 and 11, *CSi,T2,e* and *CSi,t,e* patterns at different positive bias test cases (i.e. Cases B1.1, B1.2 and B1.3) and negative bias test cases (i.e. Cases B2.1, B2.2 and B2.3) were plotted against fault free test case (i.e. Case 0), In each figure, the red dots represent *CSi,T2,e* and *CSi,t,e* at faulty test cases, while the blue dots stand for *CSi,T2,e* and *CSi,t,e* at the base case. Most of the *CSi,T2,o* and *CSi,t,o* values were under thresholds *CI99,T2* and *CI99,t*. On the contrary, more than half of the *CSi,T2,e* and *CSi,t,e* values would exceed the thresholds *CI99,T2* and *CI99,t*, respectively, which clearly showed the existence of sensor fault. As such, *CI*99 were chosen among the three percentiles of *CI* for this fault detection purpose.

Owing to the cooling load profile shown in Fig. 2, there was generally a wavy pattern on each day when bias fault of *T*2 occurred. And the negative bias fault would be more regular in different seasons. At the first and last few hours on each day, although the quantity of operating chillers was the same as that in Case 0, *Xi,e* was partitioned into different clusters against *Xi,o* due to the faulty sensor reading of *T*2. During the middle hours on each day, different quantities of operating chillers, together with different *T*2 and *T*1, caused the sensor reading subset *Xi,e* grouped into different clusters.

From Figs. 10 and 11, *CSi,T2,e* and *CSi,t,e* would increase with the rise of absolute bias values, since larger absolute bias values resulted in larger Euclidean distance from *Xi,e* to the corresponding cluster centroid *Ck*. Although the fault patterns of positive and negative bias might be similar, in particular Cases B1.1 and B2.1 with the smallest absolute bias value, they were distinguishable. Due to the different quantities of operating chillers at certain short period as shown in Fig. 5(a), there was “vertical-dot” feature in the positive bias. With such characteristic, these two types of bias could be differentiated.

|  |
| --- |
|  |
|  |
| (a) Cases 0 and B1.1. |
|  |
| (b) Cases 0 and B1.2. |
|  |
| (c) Cases 0 and B1.3.  Fig. 10. *CSi,T2,e* and *CSi,t,e* patterns in positive bias test cases. |
|  | |
|  | |
| 1. Cases 0 and B2.1. | |
|  | |
| 1. Cases 0 and B2.2. | |
|  | |
| 1. Cases 0 and B2.3.   Fig. 11. *CSi,T2,e* and *CSi,t,e* patterns in negative bias test cases. | |

4.3.2 Drift sensor reading fault

*CSi,T2,e* and *CSi,t,e* at different drifting rates of both positive drift sensor faults (i.e. Cases D1.1, D1.2 and D1.3) and negative drift sensor faults (i.e. Cases D2.1, D2.2 and D2.3) were plotted against base case (Case 0) in Figs. 12 and 13 respectively. For *T2* drift fault, the measurement difference *ei* between faulty and fault-free sensors gradually increased along the 7-day period, thus the increasing wavy patterns of both *CSi,T2,e* and *CSi,t,e* could be observed. In the same fault period, *ei* was larger at higher absolute drifting rate. The larger Euclidean distance from *Xi,e* to the corresponding cluster centroid *Ck* thus resulted in larger increasing rate of *CSi,T2,e* and *CSi,t,e*.

In the same fault period and at the same absolute drifting rate, the increasing rates of *CSi,T2,e* and *CSi,t,e* of negative drifts were larger than those of the positive ones. The *CSi,T2,e* and *CSi,t,e* at the negative drift of -0.1 °C/h (Case D2.3) could reach a “flat peak” from the end of the 4th day in each sample week. This was because a lower bound of *T2* was set in the system simulation model to prevent chilled water from freezing. During such period, the low *T2* had a dominant effect in determing the Euclidean distance, thus most of *Xi,e* were partitioned in the same cluster and resulted in relatively constant *CSi,T2,e*.

Although both the patterns of positive and negative drifts were generally increasing, they could be differentiated in due course. It is seen that the trendline of *CSi,T2,e* was continuously increasing in the cases of negative drift, while there was a drop in those of positive drift by the end of the second day. Therefore, the sensor fault of positive or negative drift can be diagnosed in 2-day time.

|  |
| --- |
|  |
|  |
| (a) Cases 0 and D1.1. |
|  |
| (b) Cases 0 and D1.2. |
|  |
| (c) Cases 0 and D1.3.  Fig. 12. *CSi,T2,e* and *CSi,t,e* patterns in positive drift test cases. |
|  | |
|  | |
| (a) Cases 0 and D2.1. | |
|  | |
| (b) Cases 0 and D2.2. | |
|  | |
| (c) Cases 0 and D2.3.  Fig. 13. *CSi,T2,e* and *CSi,t,e* patterns in negative drift test cases. | |

4.3.3 Precision degradation sensor reading fault

*CSi,T2,e* and *CSi,t,e* at sensor fault of different precision degradation errors (Cases P1, P2 and P3) were plotted against base case Case 0, as shown in Fig. 14. It is seen that the patterns of precision degradation were not like those of the bias or drift faults, since *CI99,T2* and *CI99,t* would be largely or even fully filled up by *CSi,T2,e* and *CSi,t,e* respectively. Since the measurement difference *ei* between faulty and fault-free *T2* followed Gaussian distribution, the *CSi,T2,e* and *CSi,t,e* patterns were scattered.

It was found that the pattern feature of precision degradation was “vertical shoot”: a shoot within very short period of time in Cases P1 and P2, or a shoot spreading a longer period in Case P3. Both *CSi,T2,e* and *CSi,t,e* had similar patterns, and they would increase with the rise of precision distribution error for *T2*. In Cases P1 and P2, there was not much *CSi,T2,e* and *CSi,t,e* exceeding the corresponding thresholds because such faults only occasionally affected the operating sequences of chillers, as found in Fig. 5(e) before. The occurrence of “vertical shoot” was due to the chiller operating sequence deviated from the fault-free one during certain short periods.

|  |
| --- |
|  |
|  |
| (a) Cases 0 and P1. |
|  |
| (b) Cases 0 and P2. |
|  |
| (c) Cases 0 and P3.  Fig. 14. *CSi,T2,e* and *CSi,t,e* patterns in precision degradation test cases. |

4.3.4 Complete failure sensor reading fault

*CSi,T2,e* and *CSi,t,e* at different complete failure test cases (i.e. Cases C1, C2 and C3) of sensor faults were plotted against Case 0 in Fig. 15. From Fig. 4(f), the complete failure of sensor reading of *T*2 would result in distinct different sequence of chiller operating. Therefore, most of the *CSi,T2,e* and *CSi,t,e* values exceeded the *CI99,T2* and *CI99,t* thresholds. Since the sensor reading of *T*2 was fixed at certain value, it has a dominant effect in determining which cluster *Xi,e* should be partitioned into. Most of the *Xi,e* were partitioned into the same cluster, thus there were horizontal line segments of *CSi,T2,e* during those time periods in Cases C1 and C2, and even almost a continuous horizontal line found in Case C3. Such unreasonably constant behaviour of complete failure is distinctive from the other fault types.

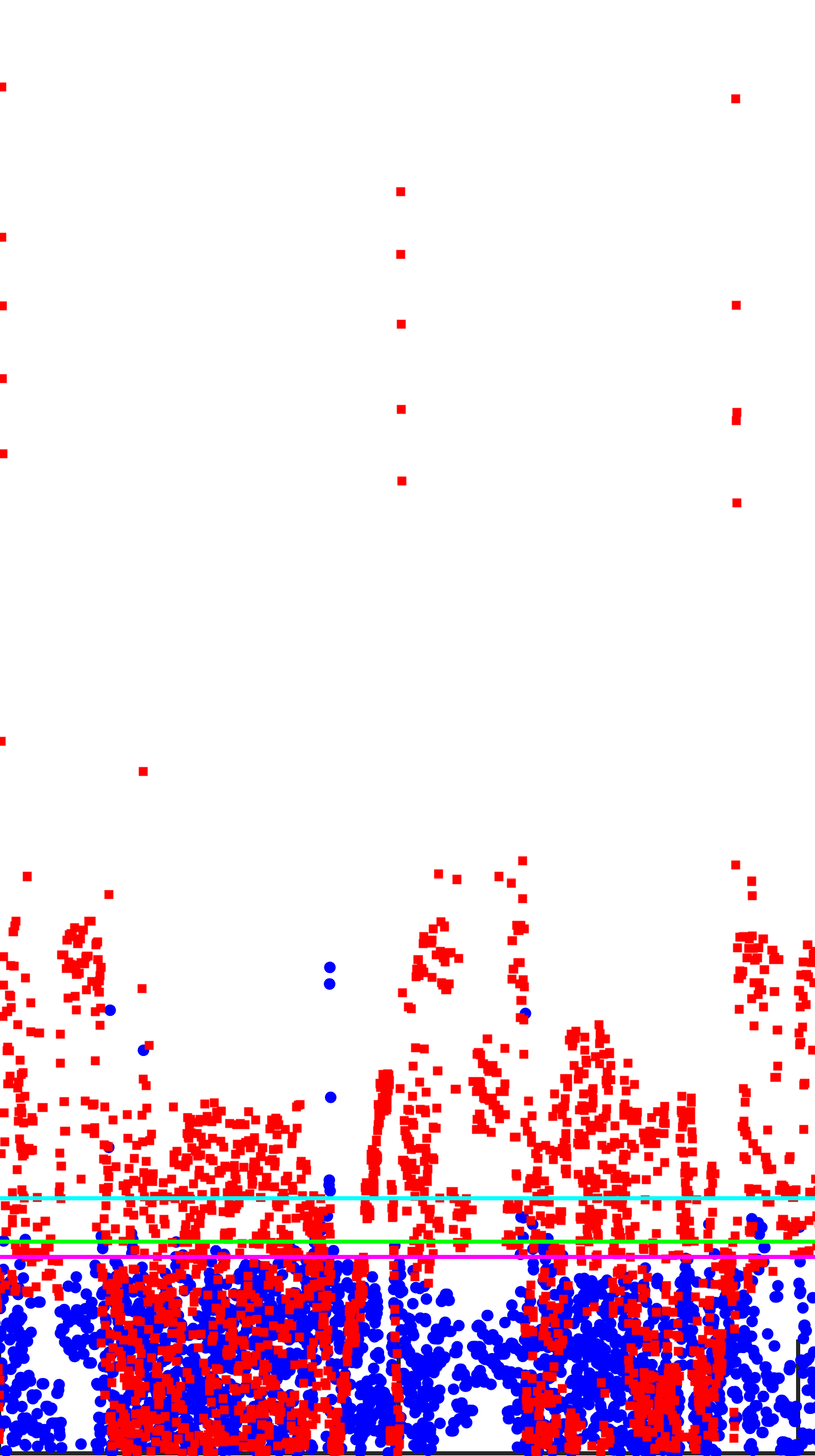
|  |
| --- |
|  |
|  |
| (a) Cases 0 and C1. |
|  |
| (b) Cases 0 and C2. |
|  |
| (c) Cases 0 and C3.  Fig. 15. *CSi,T2,e* and *CSi,t,e* patterns in complete failure test cases. |

4.3.5 Database establishment and algorithm development for sensor fault diagnosis

After thorough investigation of the *CSi,T2,e* and *CSi,t,e* patterns of various fault types for *T2*, they had highly distinguishable features from each other. In the chilled water system under study,

* the bias fault of *T2* sensor generally had a wavy pattern, while the negative bias would have a daily regular pattern and the positive one would have a vertical-dot one;
* the drift fault had initial increasing wavy pattern, while the negative drift would have trendline still increasing in 2 days but the positive drift would have it drop;
* the precision degradation had “vertical shoot” with solid fill within *CI*; and
* the complete failure had the easily noticed pattern of horizontal line segments.

As a result, all the six fault types could be identified by appropriate pattern recognition and the database for sensor fault diagnosis was established. In fact, the sensor fault diagnosis algorithm can be conducted from the ease of recognition, starting from complete failure, then precision degradation, followed by drift, finally bias. Consequently, the flow of such diagnosis in online SFDD application can be proceeded as illustrated in Fig. 16.



with vertical-dot pattern

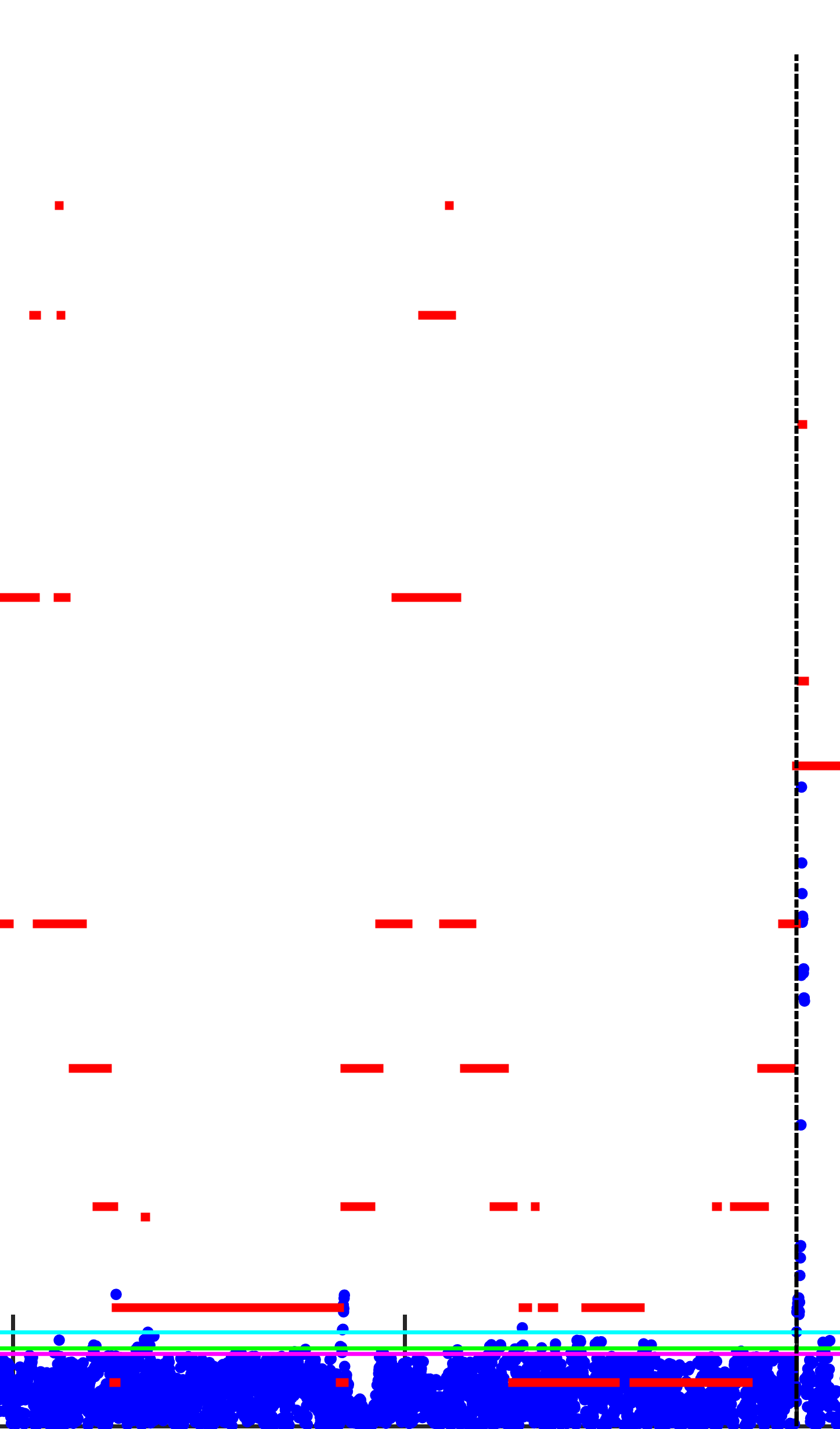
Start of sensor fault diagnosis

*T*2unreasonably constant?

Y

Horizontal line segments of *CSi,T2,m* during

some time periods.



**Complete failure**

Y

**Precision**

**degradation**

Increasing

wavy pattern of *CSi,T2,m* and *CSi,t,m*?

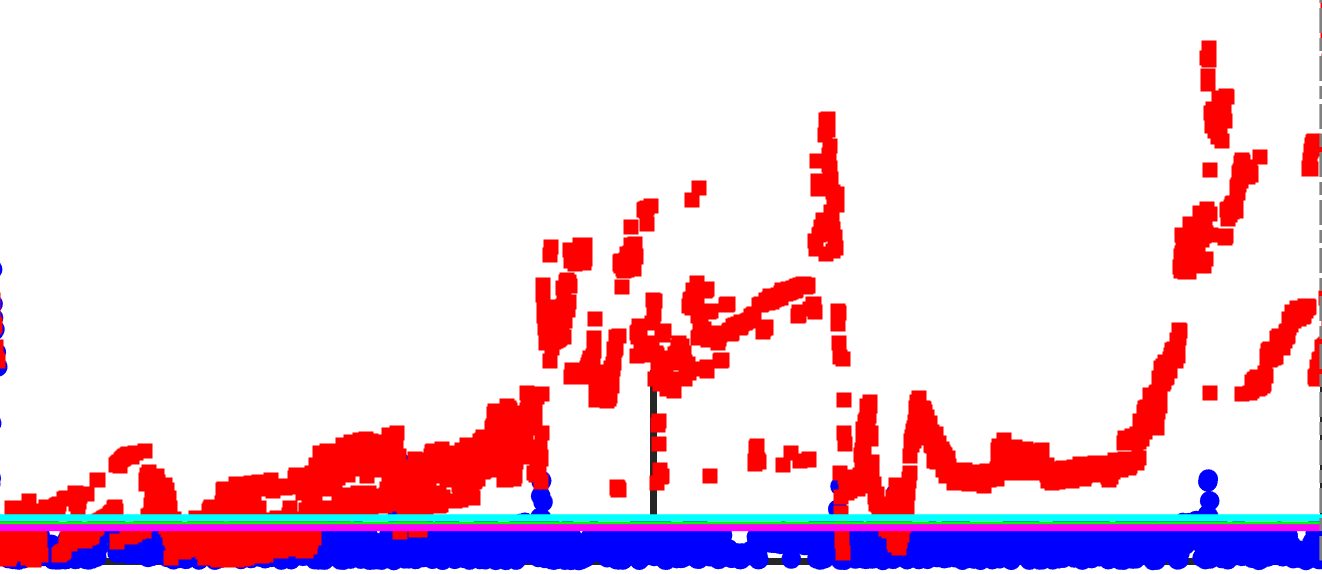
Y

In two

successive days,

trendline of *CSi,T2,m*

still increasing?



**Negative drift**

N

**Positive drift**

N: Bias

Y: Drift

Y

N

In two

successive days,

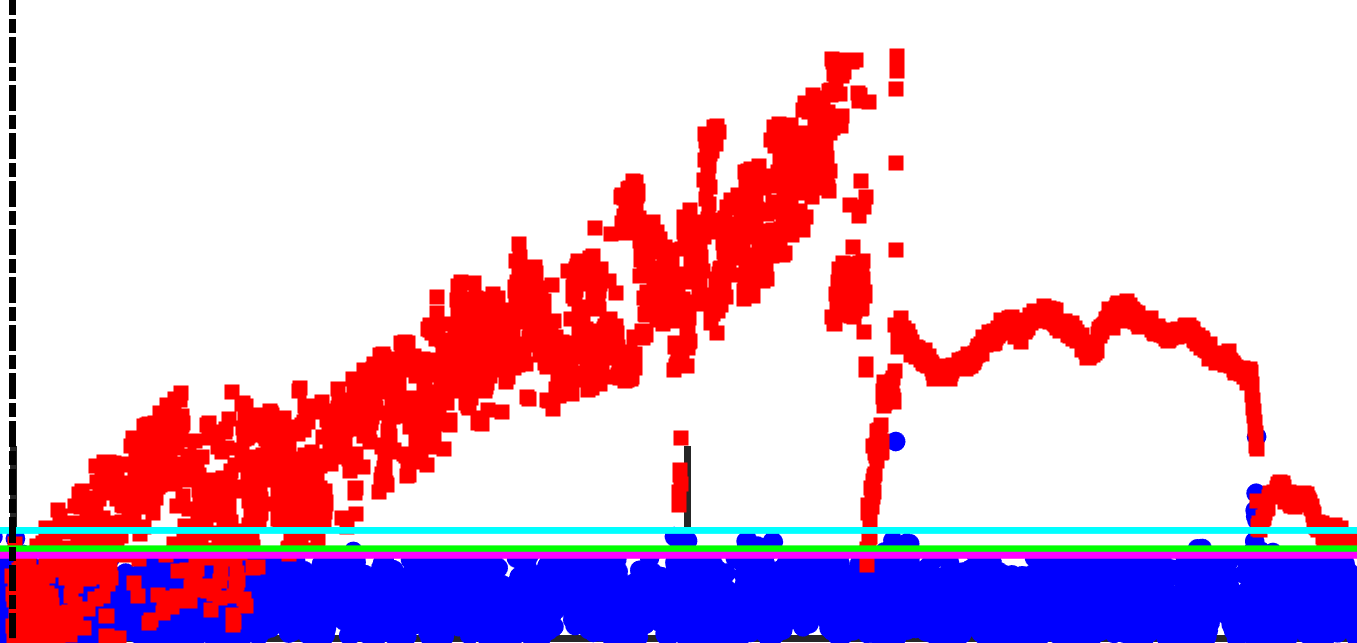
*CSi,T2,m* distribution like:

without vertical-dot pattern?

Drop of trendline

of *CSi,T2,m* by

the end of 2nd day



**Positive bias**

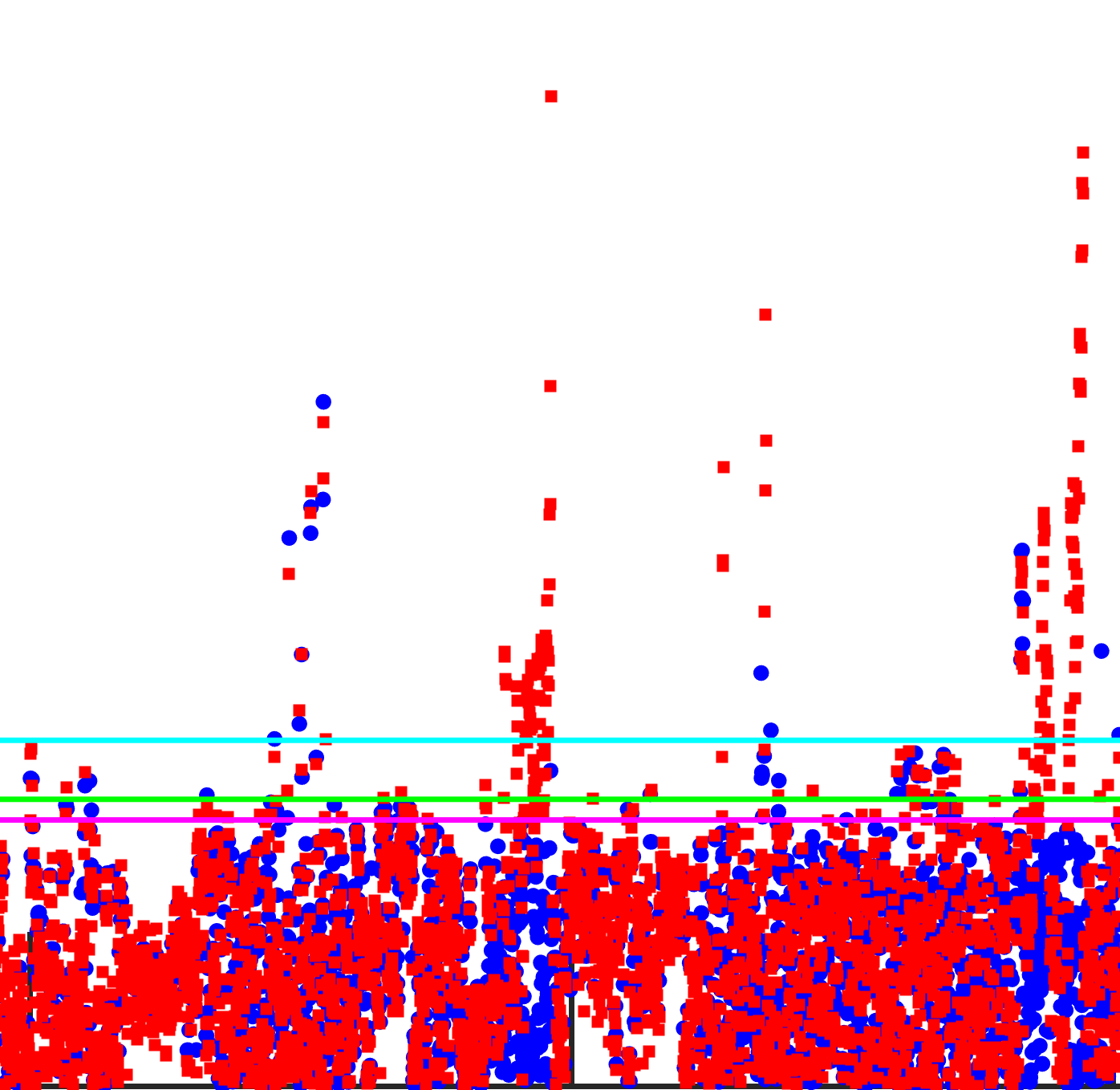
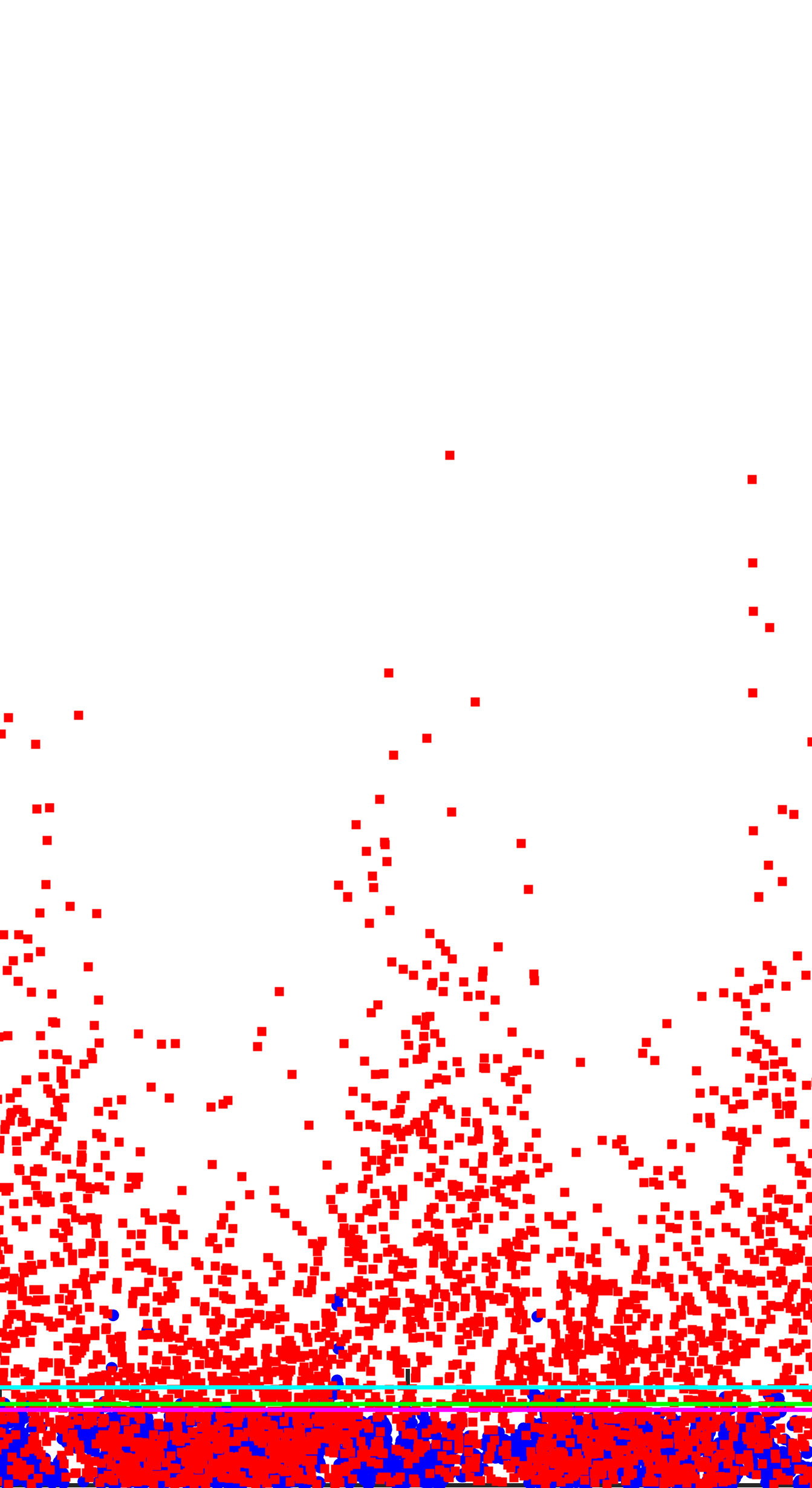
**Negative bias**

“Vertical shoot”

of *CSi,T2,m* and *CSi,t,m*,

with solid fill

within *CI99,t*?

N

N

Fig. 16. Flow chart of sensor fault diagnosis in online SFDD for measurement data.

**6. Conclusion**

This paper presents a new SFDD strategy for the chilled water system using *k*-means clustering algorithm. Based on the fault-free sensor readings from a practical chilled water system, the cluster centroid dataset and the thresholds for different *CS* values were used to generate the database for sensor fault detection. After that, the characterized *CS* patterns of various fault types were adopted to establish the database for sensor fault diagnosis. Both databases would be applied to online SFDD for actual measurement data during the operation of the chilled water system. The proposed SFDD strategy was found effective in detecting and diagnosing various types of sensor faults, including positive bias, negative bias, positive drift, negative drift, precision degradation and complete failure. The features of this SFDD strategy are highlighted as follows:

* When the fault occurred in a sensor and resulted in improper operation of the chilled water system, the *CS* values would exceed the *CI* thresholds, thus sensor faults could be detected. The smallest absolute bias value, drifting rate and precision degradation error able to be detected by the proposed SFDD strategy were 0.25 °C, 0.025 °C/h and 0.1 °C, respectively. the proposed SFDD strategy has significant improvement in FDD for sensor in chilled water system application.
* Once the fault was detected, the characteristics of the *CS* patterns were recognized from the sensor fault diagnosis algorithm. The sensor fault could be confirmed to be complete failure, precision degradation, drift or bias within one day. If it was a drift or bias fault, it could be diagnosed and confirmed to be positive or negative by two days. There was no previous works regarding how to distinguish different types of sensor faults, but the SFDD strategy can handle this effectively.
* The SFDD strategy is also a methodology to find the relationship in sensor reading dataset through data clustering. Besides the water-cooled chilled water system with differential pressure bypass under this study, there are potential future works for theSFDD strategy to be applied in more quantity of sensors, coincident multiple sensor faults, as well as other types of chilled water systems. Such approach can also be further developed for more complicated systems, like the air side system or the other building services systems.
* When the SFDD strategy is incorporated into the building management system for chiller plant, FDD for sensor can be adopted for online application, rather than offline use by the previous methods.

**Acknowledgement**

This work described in this paper is fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 11200315). Thanks also go to Mr. Yue MA for his help in plotting the special graphs in this manuscript.

**Nomenclature**

*c* Cluster centroid

**C** Cluster centroid dataset

*CI* Confidence interval

*Ck*Cluster centroid subset of cluster *k*

*Cp* Specific heat capacity of water (kJ/kg/°C)

***CS*** Centroid score set

*CS* Centroid score

*DB* Davis-Bouldin value

 Normalized sensor reading subset of cluster *k*

G Gaussian distribution

*i* Time step

*j* Sensor symbol or number

*k, k1, k2* Cluster number

*M* Total quantity of sensor

*N* Quantity of cluster

*nch*Quantity of chiller in operation

 Quantity of data object included in cluster *k*1

*Q* Cooling capacity (kJ/h)

*T* Temperature (°C)

*tmax* Maximum time step

*x* Normalized sensor reading at each time step

*X* Normalized sensor reading subset at each time step

**X** Normalized sensor reading dataset

*y* Original sensor reading before normalization at each time step

*Y*Original sensor reading subset before normalization at each time step

**Y** Original sensor reading dataset before normalization

Euclidean distance

Greek symbols

*γ* Switch-on/off coefficient

Subscripts

*bp* Bypass

*c* Calculated

*ch* Chiller

*chw*  Chilled water

*chwr* Chilled water return

*chws* Chilled water supply

*cw*  Cooling water

*e* Faulty

*i* Time step

*j* Sensor symbol or number

*m* Measured

*max* Maximum

*N* Quantity of cluster

*o* Fault-free

*off* Chiller switched-off

*on* Chiller switched-on

*opt* Optimal

*p* Primary

*s* Secondary

*t* Total

0 Current time step

1 Primary chilled water supply

2 Primary chilled water return

99 99th percentiles

97 97th percentiles

95 95th percentiles

**Abbreviations**

FDD Fault detection and diagnosis

PCA Principal component analysis

SFDD Sensor fault detection and diagnosis

SVDD Support vector data description

**References**

[1] ASHRAE Handbook: HVAC Systems and Equipment. 2016.

[2] Venkatasubramanian V, Rengaswamy R, Yin K and Kavuri SN. A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. Computers & Chemical Engineering. 27(2003)293-311.

[3] Venkatasubramanian V, Rengaswamy R and Kavuri SN. A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies. Computers & Chemical Engineering. 27(2003)313-326

[4] Venkatasubramanian V, Rengaswamy R and Kavuri SN. A review of process fault detection and diagnosis: Part III: Process history based methods. Computers & Chemical Engineering. 27(2003)327-346.

[5] Kim W and Katipamula S. A review of fault detection and diagnostics methods for building systems. Science and Technology for the Built Environment. 24(2018)3-21.

[6] Thumati BT, Feinstein MA, Fonda JW, Turnbull A, Weaver FJ, Calkins ME and Jagannathan S. An online model-based fault diagnosis scheme for HVAC systems. IEEE International Conference: In Control Applications. 2011.

[7] Gordon J M, Ng KC and Chua HT. Centrifugal chillers: thermodynamic modelling and a diagnostic case study. International Journal of Refrigeration. 18(1995)253-257.

[8] Liang J and Du R. Model-based fault detection and diagnosis of HVAC systems using support vector machine method. International Journal of Refrigeration. 30(2007)1104-1114.

[9] Bonvini M, Sohn MD, Granderson J, Wetter M and Piette MA. Robust on-line fault detection diagnosis for HVAC components based on nonlinear state estimation techniques. Applied Energy. 124(2014)156-166.

[10] Sun YJ, Wang SW and Huang GS. Online sensor fault diagnosis for robust chiller sequencing control. International Journal of Thermal Sciences. 49(2010)589-602.

[11] Kim M and Kim MS. Performance investigation of a variable speed vapor compression system for fault detection and diagnosis. International Journal of Refrigeration. 28(2005)481-488.

[12] Tassou SA and Grace IN. Fault diagnosis and refrigerant leak detection in vapour compression refrigeration systems. International Journal of Refrigeration. 28(2005)680-688.

[13] Cui J and Wang S. A model-based online fault detection and diagnosis strategy for centrifugal chiller systems. International Journal of Thermal Sciences. 44(2005)986-999.

[14] Hou ZJ, Lian ZW, Yao Y and Yuan XJ. Data mining based sensor fault diagnosis and validation for building air conditioning system. Energy Conversion and Management. 47(2006)2479-2490.

[15] Wang SW, Qiang Z and Fu X. A system-level fault detection and diagnosis strategy for HVAC systems involving sensor faults. Energy and Buildings. 42(2010)477-490.

[16] Beghi A, Brignoli R, Cecchinato L, Menegazzo G, Rampazzo M and Simmini F. Data-driven fault detection and diagnosis for HVAC water chillers. Control Engineering Practice. 53(2016)79-91.

[17] Wang SW and Cui JT. Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method. Applied Energy. 82(2005)197-213.

[18] Yang Z, Wang SW and Xiao F. Pattern recognition-based chillers fault detection method using support vector data description (SVDD). Applied Energy. 112(2013)1041-1048.

[19] Li GN, Li HR, Hu M, Guo YB, Shi SB and Hu WJ. A sensor fault detection and diagnosis strategy for screw chiller system using support vector data description-based D-statistic and DV-contribution plots. Energy and Buildings. 133(2016)230-245.

[20] Li D, Zhou YX, Hu GQ, and Spanos CJ. Fault detection and diagnosis for building cooling system with a tree-structured learning method. Energy and Buildings. 127(2016)540-551.

[21] Li GN, Hu YP, Chen HX, Shen LM, Li HR, Hu M, Liu JY and Sun KZ. An improved fault detection method for incipient centrifugal chiller faults using the PCA-R-SVDD algorithm. Energy and Buildings. 116 (2016)104-113.

[22] Zhao Y, Xiao F, Wen J, Lu Y and Wang S. A robust pattern recognition-based fault detection and diagnosis (FDD) method for chillers. HVAC&R Research. 20(2014)798-809.

[23] Turner WJN, Staino A and Basu B. Residential HVAC fault detection using a system identification approach. Energy and Buildings, 151(2017)1-17.

[24] Tan PN, Steinbach M, Kumar V. Introduction to Data Mining. Pearson Addison Wesley, USA 2006

[25] Li D, Hu GQ, and Spanos CJ. A data-driven strategy for detection and diagnosis of building chiller faults using linear discriminant analysis. Energy and Buildings. 128 (2016)519-529.

[26] Li GN and Hu YP. Improved sensor fault detection, diagnosis and estimation for screw chillers using density-based clustering and principal component analysis. Energy and Buildings. 173(2018)502-516.

[27] Yiakopoulos CT, Gryllias KC, Antoniadis IA. Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. Expert System. 38(2011)2888-2911.

[28] Baraldi P, Di Maio F, Rigamonti M, Zio E, Seraoui E. Clustering for unsupervised fault diagnosis in nuclear turbine shut-down transients. Mechanical Systems and Signal Processing. 58(2015)160-178

[29] Yan R, Ma Z, Kokogiannakis G and Zhao Y. A sensor fault detection strategy for air handling units using cluster analysis. Automation in Construction. 70(2016)77-88.

[30] Wei Z, Wang Y, He S and Bao J. A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection. Knowledge-Based Systems. 116(2017)1-12.

[31] Du Z, Fan B, Jin X and Chi J. Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis. Building and Environment 73(2014)1-11.

[32] Al-Dahidi S, Di Maio F, Baraldi P, Zio E and Seraoui R. A framework for reconciliating data clusters from a fleet of nuclear power plants turbines for fault diagnosis. Applied Soft Computing. 69(2018)213-231.

[33] Dey M, Rana SP and Dudley S. Smart building creation in large scale HVAC environments through automated fault detection and diagnosis. Future Generation Computer Systems. 2018.

[34] TRNSYS 17: A Transient System Simulation Program. Solar Energy Laboratory University of Wisconsin, Madison, USA.

[35] Davies DL and Bouldin DW. A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence. (1979)224-227.