A new approach to multivariate statistical process control and its application to wastewater treatment process monitoring

Osamu Yamanaka[1], Ryo Namba[2], Takumi Obara[3], and Yukio Hiraoka[4]

[1,2,3] *Toshiba Infrastructure Systems and Solutions Corporation, Fuchu-shi, Tokyo, 183-8511, Japan*
osamu2.yamanaka@toshiba.co.jp
ryo.namba@toshiba.co.jp
takumi.obara@toshiba.co.jp

[4] *Toshiba Infrastructure Systems and Solutions Corporation, Kawasaki-shi, Kanagawa, 212-8585, Japan*
yukio.hiraoka@toshiba.co.jp

Abstract

This paper presents a new process monitoring and fault diagnosis approach based on a modified Multivariate Statistical Process Control (MSPC) and evaluates its applicability to municipal wastewater treatment process monitoring. A conventional MSPC based on Principal Component Analysis (PCA) is firstly adjusted to have an easy-to-understand user interface and then a new yet simplified decomposable diagnostic model is introduced. The developed user interface is designed to seamlessly connect MSPC to existing process monitoring system adopting the so-called trend graphs. The proposed diagnostic model is derived in a constructive way by aggregating small size models with one input or two inputs to improve tractability of the diagnostic model. The effectiveness of the modified MSPC is illustrated through some off-line and on-line experiments by using a set of real multivariate process data at a municipal wastewater treatment plant.

1. Introduction

Process monitoring and fault diagnosis plays an important role for operation of social infrastructure systems such as power generation plants, water and wastewater treatment plants, railway and transportation systems, just to name a few. Supervisory Control And Data Acquisition (SCADA) system is often installed in such infrastructure systems where operators monitor time-series process data by the so-called trend graphs to keep process in control. The simplest yet often adopted fault diagnosis technique for stable plant operation is abnormality (out-of-control states) detection in process data by a pre-specified control limit for each single variable, which is similar to Statistical Process Control (SPC). To improve process performance and operational stability, however, earlier fault detection and cause localization will be important. It allows us to recognize how to improve process performance and how to avoid performance degradation.

Multivariate Statistical Process Control (MSPC) (Jackson et al. (1991), Wise et al. (1996)) is attractive data-driven approach for such purpose, which is suitable to monitor complex processes with high-dimensional data structure. The key idea of MSPC is subspace orthogonalization where Principal Component Analysis (PCA) is often utilized. High-dimensional process data is projected onto a subset of the subspaces and a few statistical indices for fault detection are constructed, which is effective in improving detection accuracy and cause localization ability. In addition, advanced MSPC methods have also been proposed (e.g., Choi et al. (2008), Uchida et al. (2022)) for further improvement. Despite such advances, the SPC-like monitoring is still popular and widely used as real-time process monitoring in many real plants, while various applications of MSPC have been reported (e.g., Camacho et al. (2016), Zhao et al. (2022)). Among various possible reasons, the difficulty of intuitive understanding of diagnostic results and the difficulty of PCA model handling in MSPC will be main reasons to hinder wide applications of real-time continuous monitoring by MSPC.

Based on this motivation, this paper tries to improve the applicability of MSPC to real-time process monitoring. To this end, this paper firstly introduces an improved user interface (UI) where the conventional trend-graph-based monitoring and MSPC-based diagnosis are combined, which will improve intuitive understanding by plant operators. Then, without changing the UI, this paper proposes a new yet simple MSPC algorithm to improve tractability and maintainability of PCA model used in MSPC. The proposed MSPC has some advantages over conventional PCA-based MSPC (PCA-

MSPC hereinafter) for real-time continuous applications of MSPC to real processes. The advantages include simple and clear dependence on a specific training data in modeling and easy replacement of input variables of a PCA model, which improves model tractability and possibly enables robust-and-interpretable modeling. While the proposed MSPC is simpler than the PCA-MSPC, it is confirmed that the MSPC effectively works for fault diagnosis by applying it to real process data in a municipal WasteWater Treatment Plant (WWTP).

2. CONVENTIONAL PCA-MSPC

This section overviews PCA-MSPC (Jackson et al. (1991); Wise et al. (1996); Camacho et al. (2016)) prior to presenting a modified MSPC. PCA-MSPC utilizes Hotelling's T^2 statistic (D statistic) and Q statistic (Squared Prediction Error) as fault indices together with their control limits and the variable contributions to them, which can be defined by using PCA.

PCA transforms an $n \times m$ data matrix $\boldsymbol{X}$ by combining the variables as a linear weighted sum, represented as

$$\begin{aligned}\boldsymbol{X} &= \boldsymbol{T}_{all}\boldsymbol{P}_{all}^T = \boldsymbol{T}\boldsymbol{P}^T + \boldsymbol{E} = \hat{\boldsymbol{X}} + \boldsymbol{E} \\ &= [\boldsymbol{t}_1\boldsymbol{p}_1^T + \cdots + \boldsymbol{t}_k\boldsymbol{p}_k^T] + [\boldsymbol{t}_{k+1}\boldsymbol{p}_{k+1}^T + \cdots + \boldsymbol{t}_m\boldsymbol{p}_m^T], \quad (1)\end{aligned}$$

where $\boldsymbol{p}_i$, $i = 1, \cdots, m$, are the principal component loadings, $\boldsymbol{P}_{all} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_m]$ and $\boldsymbol{P} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_k]$ are the loading matrices, $\boldsymbol{t}_i$, $i = 1, \cdots, m$, are the principal component scores, $\boldsymbol{T}_{all} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_m]$ and $\boldsymbol{T} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_k]$ are the score matrices, and $\boldsymbol{E} = [\boldsymbol{t}_{k+1}\boldsymbol{p}_{k+1}^T + \cdots + \boldsymbol{t}_m\boldsymbol{p}_m^T]$ is the residual matrix. n ,m, and k are the number of samples, that of variables, and the retained number of principal components by truncation, respectively. Superscript T denotes the transpose of a vector/matrix. It is usually assumed that the columns of $\boldsymbol{X}$ have been standardized to zero mean and unit variance by normalizing each column by its mean μ_i and standard deviation $\sigma_i, i = 1, 2, \cdots, m$. The principal component loadings are the direction vectors creating a hyperplane that is embedded inside the m-dimensional spaces and captures the maximum possible residual variance in the measured variables, while maintaining orthonormality with the other loading vectors. The loadings correspond to the eigenvectors of the covariance matrix of $\boldsymbol{X}$ and its eigenvalues indicate the variance captured by the corresponding eigenvector. The (sample) covariance matrix $\boldsymbol{\Sigma} = \frac{1}{n-1}\boldsymbol{X}^T\boldsymbol{X}$ can be described by

$$\boldsymbol{\Sigma} = \boldsymbol{P}_{all}\boldsymbol{\Lambda}_{all}\boldsymbol{P}_{all}^T = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T + \boldsymbol{F}, \quad (2)$$

where $\boldsymbol{\Lambda}_{all} = diag(\lambda_1, \lambda_2, \cdots, \lambda_m)$ is a diagonal matrix of the eigenvalues $\lambda_i, i = 1, 2, \cdots, m$ of $\boldsymbol{\Sigma}$, $\boldsymbol{\Lambda}$ is a partial matrix of $\boldsymbol{\Lambda}_{all}$ of which elements consist of k largest eigenvalues, and $\boldsymbol{F}$ is the residual of the covariance matrix $\boldsymbol{\Sigma}$. Note that the covariance matrix $\boldsymbol{\Sigma}$ is identical to the correlation matrix (denoted by $\boldsymbol{R}$ hereinafter) if each column of $\boldsymbol{X}$ is standardized. Using these matrices, the T^2 and Q statistics (of measurements at time t) are defined by

$$\begin{aligned}T^2(t) &:= \boldsymbol{t}^T(t)\boldsymbol{\Lambda}^{-1}\boldsymbol{t}(t) = \boldsymbol{x}^T(t)\boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T\boldsymbol{x}(t), \quad (3)\\ Q(t) &:= \boldsymbol{e}^T(t)\boldsymbol{e}(t) = \boldsymbol{x}^T(t)(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^T)\boldsymbol{x}(t), \quad (4)\end{aligned}$$

where $\boldsymbol{e}(t)$, $\boldsymbol{t}(t)$, and $\boldsymbol{x}(t)$ are the residual vector at time t, the score vector at time t, and the measurement data vector (sample) at time t, respectively. $\boldsymbol{I}$ is the identity matrix of size m. The T^2 statistic defined by the sum of normalized squared scores is a measure of the variation within the PCA model, while the Q statistic indicates how well each sample conforms to the PCA model and is a measure of the amount of variation not captured by the k principal components retained in the model. Note that "PCA model" means the created hyperplane by PCA here, but it also denotes the set of loadings (eigenvectors), eigenvalues, and means and standard deviations of all input variables for PCA in the following.

The T^2 statistic and the Q statistic are complementary used for fault detection with their (upper) control limits that are often approximately expressed by

$$\begin{aligned}T_\alpha^2 &= \frac{k(n+1)(n-1)}{n(n-k)}F_\alpha(k, n-k), \quad (5)\\ Q_\alpha &= \theta_1\left(\frac{c_\alpha\sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2}\right)^{\frac{1}{h_0}}, \quad (6)\end{aligned}$$

where $F_\alpha(k, n-k)$ is the value (upper limit) at $100(1-\alpha)\%$ confidence level of the F distribution with $(k, n-k)$ degrees of freedom, $\theta_i = \sum_{j=k+1}^m \lambda_j^i, i = 1, 2, 3, h_0 = 1 - \frac{2\theta_1\theta_2}{3\theta_3^2}$, c_α is the $100(1-\alpha)\%$ standardized normal percentile. Note that T_α^2 can be approximated by $\chi_\alpha^2(k)$, the value at $100(1-\alpha)\%$ confidence level of the χ^2 distribution with k degrees of freedom, if n is sufficiently large since $\frac{(n+1)(n-1)}{n(n-k)} \to 1$ and $kF_\alpha(k, n-k) \to \chi_\alpha^2(k)$ as $n \to \infty$. The contributions of i_{th} component $x_i(t)$ of the data vector $\boldsymbol{x}(t)$ for the Q statistic and for the T^2 statistic could be defined as

$$\begin{aligned}Q_i(t) &:= (\boldsymbol{x}^T(t)(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^T)\boldsymbol{e}_i)^2, \quad (7)\\ T_i^2(t) &:= (\boldsymbol{x}^T(t)(\boldsymbol{P}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{P}^T)\boldsymbol{e}_i)^2, \quad (8)\end{aligned}$$

where $\boldsymbol{e}_i$ is the i_{th} column of the identity matrix $\boldsymbol{I}$ of size m. It should be noted that the variable contributions are defined so that $Q(t) = \sum_{i=1}^m Q_i(t)$ and $T^2(t) = \sum_{i=1}^m T_i^2(t)$ hold.

The Q and T^2 statistics with the contributions make it possible to detect faults earlier and to localize cause variables.

3. MODIFIED MSPC FOR REAL-TIME MONITORING

Intuitive design and tractability of monitoring systems will play an important role to enhance the applicability of MSPC to real-time monitoring in existing plants where SCADA is installed. To this end, this section firstly introduces an easy-to-understand UI which seamlessly connects MSPC and existing trend-graph-based monitoring. Then, we propose a novel yet simplified MSPC algorithm without changing the UI, which improves tractability of MSPC model in real time and also may improve intuitive understanding by operators.

3.1. Adjusted-MSPC with Improved User Interface

Figure 1 shows an example of the developed UI consisting of three parts, which will be easy-to-understand.

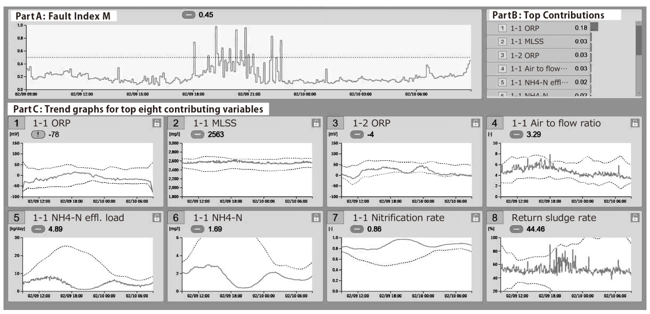

Figure 1. Developed user interface combining MSPC-based monitoring and trend-graph-based monitoring

The part A shows the time-series of a fault index M derived from the T^2 and Q statistics. The index M is defined by

$$M(t) \;:=\; 1 - \exp(-\ln(2)C(t)) \tag{9}$$

$$C(t) \;:=\; \frac{1}{2}\left(\frac{Q(t)}{Q_\alpha} + \frac{T^2(t)}{T^2_\alpha}\right), \tag{10}$$

where C is a scaled combined statistic of the T^2 and Q statistics with its control limit 1. The M is defined so that its range is from 0 to 1 and the control limit becomes 0.5. The reason for introducing M is as follows. Firstly, distinguishing the T^2 and the Q hampers intuitive understanding by plant operators since these two are just a fault index for operators. While introducing the combined index C is sufficient for this purpose, frequently observed outliers of process data generate unnecessary extremely large values of the C and hence the index M is introduced to bounding the range.

The part B is a contribution plot of the M of which contribution M_i is defined by

$$M_i(t) \;:=\; M(t)\frac{C_i(t)}{C(t)}, \tag{11}$$

$$C_i(t) \;:=\; \frac{1}{2}\left(\frac{Q_i(t)}{Q_\alpha} + \frac{T_i^2(t)}{T^2_\alpha}\right), \tag{12}$$

Note that the contribution M_i is defined so that $M(t) = \sum_{i=1}^{m} M_i(t)$ holds. The variables whose contribution is in the top eight at a present time are listed from the top to the eighth so that operators can easily notice abnormal events.

The part C is the time-series (trend graphs) of the top eight variables with their normal operating range calculated by other univariate statistics. Each trend graph is linked to another detail one (not shown) by clicking it. As process monitoring based on trend graphs is common in reality, MSPC can be connected to conventional process monitoring in this way. The PCA-MSPC with this UI is referred to as "Adjusted-MSPC" in the following.

3.2. Modular-MSPC for Improvement of Tractability

While defining Normal Operating Condition (NOC) is crucial for applications of MSPC to real plants, it is sometimes difficult, particularly for non-stational and/or disturbance driven processes such as WWTPs whose performance heavily depends on the ambient temperature and uncontrolled sewer and storm water. To adapt such varying conditions and disturbances, operational conditions should also be adjusted. Thus, no unique NOC can be defined in reality. In addition, sensor failures and/or replacements are the rule rather than the exception in real plants like WWTPs. Thus, PCA model should be updated appropriately in real-time monitoring but when and how to update is difficult in general. The difficulty will be partly caused by complex and strong dependence of PCA models on training data, which in turn makes diagnostic results difficult to interpret uniformly and intuitively. In addition, adding and deleting of input variables for MSPC is not easily handled since PCA treats multivariate data collectively, which also makes model update process more tedious. To cope with it, we propose a simplified decomposable MSPC that is referred to as "Modular-MSPC" , which is more simply and clearly dependent on training data. The main idea of the Modular-MSPC is to define a new combined statistic named S, instead of the C statistic, by aggregating T^2 statistics of all m single variables and pairwise Q statistics for all possible $\binom{m}{2} = \frac{m(m-1)}{2}$ combinations of variables. In the Modular-MSPC, each T^2 or pairwise Q statistic is considered as the basic building block and thus variable contributions are also defined by partly aggregating these basic building blocks, which allows us to add and delete input variables easier. In addition, dependence on training data of the Modular-MSPC is clear and simple, which is illustrated below.

Firstly, the T^2 statistic for single i_{th} variable is exactly same as the square of the univariate t statistic, which is just the square of (a sample) of i_{th} variable if the data is already standardized, which is represented as

$$T_i^2(t) = (t_i(t))^2 = (x_i(t))^2, \tag{13}$$

where subscript i stands for i_{th} variable.

Then pairwise Q statistic can be derived by conducting PCA for two variables. Fortunately, $\boldsymbol{P}_{all}$ and $\boldsymbol{\Lambda}_{all}$ can be derived explicitly, which is expressed as

$$\begin{aligned}\boldsymbol{\Sigma} &= \boldsymbol{R} = \boldsymbol{P}_{all}\boldsymbol{\Lambda}_{all}\boldsymbol{P}_{all}^T \\ &= \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 1+r & 0 \\ 0 & 1-r \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} \end{bmatrix}, \end{aligned} \quad (14)$$

where r means the correlation coefficient r_{ij} of i_{th} and j_{th} variables under consideration. The equation (14) shows the following interesting properties as shown in Figure 2.

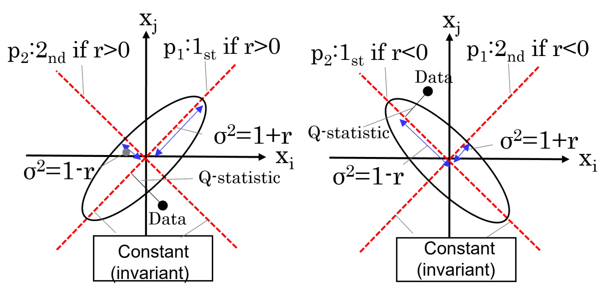

Figure 2. Two dimensional PCA to define pairwise Q

Firstly, $\boldsymbol{P}_{all}$ does not depend on training data, which implies that the direction vector never changed and can be fixed irrespective of training data. Next, diagonal elements of $\boldsymbol{\Lambda}_{all}$ can be characterized by the correlation coefficient r only, which means that the variances along the loadings are explicitly related to the correlation coefficient. Finally, the first principal component loading becomes $\boldsymbol{p}_1 = [\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]^T$ if $r > 0$, and $\boldsymbol{p}_2 = [\frac{\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}]^T$ if $r < 0$, thus one can distinguish the first and the second by just checking the sign of r. By noting the nice properties and assuming that the loading $\boldsymbol{P}$ consists of the first loading, the pairwise Q statistic can be derived as

$$Q_{ij}(t)(=Q_{ji}(t)) = \frac{(x_i(t) - sign(r_{ij})x_j(t))^2}{2}, \quad (15)$$

where Q_{ij}(=Q_{ji}) denotes the pairwise Q statistic between i_{th} variable and j_{th} variable ($i \neq j$). We prefer to rescale Q_{ij} by dividing by its variance $1 - |r_{ij}|$, which is redefined as

$$Q_{ij}(t)=(q_{ij}(t))^2 = \left(\frac{(x_i(t) - sign(r_{ij})x_j(t))}{\sqrt{2(1-|r_{ij}|)}} \right)^2, \quad (16)$$

where $q_{ij}(t)$ is defined as the square root of $Q_{ij}(t)$.This rescale allows us to consider that $Q_{ij}(t)$ and $T_i^2(t)$ (for all $i, j = 1, 2, \cdots, m, i \neq j$) follow $\chi^2(1)$ if n is sufficiently large under the standard assumption for deriving the control limits (5) and (6).

By using (13) and (16), non-scaled S statistic denoted S_0 and the i_{th} variable contribution S_{0i} are defined as

$$S_0(t) \;:=\; \sum_{i=1}^{m} \left(T_i^2(t) + \frac{1}{2} \sum_{j=1, j\neq i}^{m} Q_{ij}(t) \right), \quad (17)$$

$$S_{0i}(t) \;:=\; T_i^2(t) + \frac{1}{2} \sum_{j=1, j\neq i}^{m} Q_{ij}(t), \quad (18)$$

where $S_0(t) = \sum_{i=1}^m S_{0i}(t)$ holds. Note that $S_0(t)$ can also be simply expressed as $S_0(t) = \boldsymbol{z}^T(t)\boldsymbol{z}(t)$ by defining $\boldsymbol{z}(t) = [t_1(t), \cdots, t_m(t), q_{12}(t), \cdots, q_{(m-1)m}(t)]^T$ of size $\frac{m(m+1)}{2} \times 1$ vector. To obtain the scaled S statistic, the control limit of S_0 should be decided. A proper control limit can be derived as

$$S_{0\alpha} = \sqrt{\frac{\kappa_2}{2k_0}} \left(\chi_\alpha^2(k_0) - k_0 \right) + \kappa_1, \quad (19)$$

where, $\kappa_i = 2^{i-1}(i-1)! \sum_{j=1}^{\frac{m(m+1)}{2}} \gamma_j^i, i = 1, 2, 3,\ k_0 = 8\kappa_2^3/\kappa_3^2$. $\gamma_j, j = 1, 2, \cdots, m(m+1)/2$ are the eigenvalues of $\frac{1}{n}\boldsymbol{Z}^T\boldsymbol{Z}$ where $\boldsymbol{Z}$ is the $n \times m(m+1)/2$ data matrix defined by $\boldsymbol{Z} := [\boldsymbol{z}(1), \boldsymbol{z}(2), \cdots \boldsymbol{z}(n)]^T$. The control limit can be derived under the assumption that $Q_{ij}(t)$ and $T_i^2(t)$ follow $\chi^2(1)$. The Hall-Buckley-Eagleson approximation for a weighted sum of $\chi^2(1)$ random variables (Bodenham et al. (2016)) is applied after transforming the vectors $\boldsymbol{z}(t), t = 1, 2, \cdots, n$ to their scores. The detail for deriving the control limit (19) including efficient algorithm for computing the eigenvalues γ_j is omitted due to space limitation.

By using the control limit (19), the S statistic and the variable contributions can be simply defined as $S(t) := S_0(t)/S_{0\alpha}$ and $S_i(t) := S_{0i}(t)/S_{0\alpha}$, respectively. One can adopts the S and S_i instead of using the C and C_i without changing the UI presented above. Moreover, interpretability and tractability of the Modular-MSPC will be improved since it was derived in a constructive way and simply depends on training data only through the correlation coefficients r_{ij} in the correlation matrix $\boldsymbol{R}$ of all input variables. A possible drawback is less detectability and diagnosability since the Modular-MSPC is simpler than the PCA-MSPC, but it actually works well for a real municipal WWTP, as presented below.

4. Application of modified MSPC to WWTP

4.1. Wastewater Treatment Plant

The modified MSPC was applied to a part (called the 1_{st} train) of a municipal WWTP in Japan, while the conventional PCA-MSPC had already been evaluated at the same WWTP as a part of a national project called B-DASH (B-DASH (2016)). Figure 3 shows the outline of the plant layout of the train. Influent wastewater is firstly stored in the primary settler where solid wastes are removed. Then liquid wastes such as organic matters (COD), nitrogen (NH_4-N,NO_3-N),

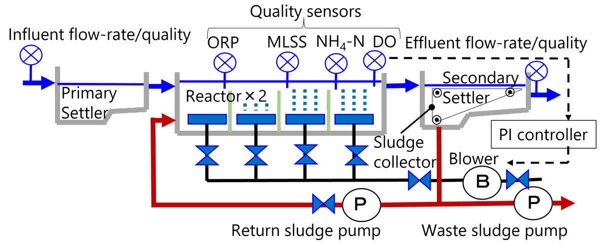

Figure 3. Plant layout for the first train of WWTP

and phosphorus (PO_4-P) are treated biologically in the reactors with (partial) aeration by the blowers. It is called activated sludge process since the sludge containing microorganisms are growing while treating organic matters. Finally, the activated sludge is separated into solids (sludge) and liquid in the secondary settler. A part of sludge is removed as waste sludge and the remainder is returned to the (pair of) bioreactors connected to common primary/secondary settlers. While a large amount of data for about a thousand variables are collected every 1 minute by a SCADA, we have chosen 82 variables that are relevant to process status of the 1_{st} train as the candidates of the input variables of MSPC. The variables are influent and effluent flow-rates, water levels, flow-rates of pumps and blowers, process control indices such as SRT (Sludge Retention Time), water quality indices such as MLSS (activated sludge concentration), DO (dissolved oxygen), NH_4-N,NO_3-N,PO_4-P and so on (B-DASH (2016)).

4.2. Evaluation Method

Evaluation was carried out in the following two phases.
In the first phase, basic performance of the Modular-MSPC was evaluated by applying it to historical data, where four diagnosis models with different input variables were adopted. Detectability and diagnosability (cause localization ability) of the Modular-MSPC were compared with those of the Adjusted-MSPC for abnormal events that had been occurred at the WWTP in the past. An abnormal event of a sludge collector failure is focused here due to space limitation, while such comparisons were also conducted for the events reported in the B-DASH project (B-DASH (2016)). The failure occurred at a recorded time t_0 and was noticed by an operator about 24 hours later. During the period, the sludge concentration MLSS decreased and thereby water quality such as NH_4-N was deteriorated gradually. Detectability was assessed in terms of detection time and distinguishability of normal and abnormal states. The detection time was measured by the elapsed time from t_0 to the detected time when the value of the index M reached 0.5. The distinguishability was measured by the difference of the M-values before and after the event happening. The former is the value at the last minute to t_0 and the latter is the maximum value within 6 hours from t_0. Diagnosability was qualitatively evaluated in terms of the adequacy of process operation.

In the second phase, detectability and diagnosability of the Modular-MSPC was evaluated qualitatively by applying it in real-time. The plant operators were asked to monitor by a prototype system of the Modular-MSPC with the developed UI for about two months in real-time. In this real-time experiment, eleven diagnosis models with different combinations of input variables were applied simultaneously and the models were updated automatically every two weeks by using the process data of the latest two weeks. After finishing the experiment, we have investigated meaningful detected abnormal events, by interviewing the operators about real process status and operating conditions.

4.3. Evaluation Results

Regarding the detectability in the first phase, the detection time of the Modular-MSPC was slower than that of the Adjusted-MSPC, as was expected. The average detection time of the Modular-MSPC with four different combinations of variables was 152 minutes while that of the Adjusted-MSPC was 46 minutes. Meanwhile, distinguishability of the Modular-MSPC was better than that of the Adjusted-MSPC where the difference of the M-values of the former was 0.71 on average while that of the latter was 0.48. Figure 4 is an example of the result where the M-values and trend graphs for top three contributions are shown since the developed UI was only available on the on-line prototype. As can be seen,

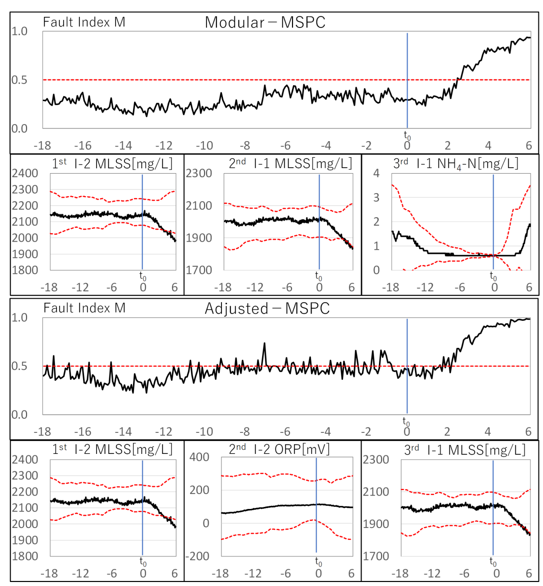

Figure 4. Example of fault diagnosis by Modular-MSPC and Adjusted-MSPC during sludge collector failure

earlier detection by the Adjusted-MSPC was achieved at the cost of decreasing the distinguishability as the M-value of the Adjusted-MSPC had been already just below the threshold 0.5. Note that the same value for the significance level $\alpha = 0.023$ was used for the control limits. As for the diagnosability, while similar contributions were obtained in both the MSPCs, the Modular -MSPC may be slightly better than the Adjusted-MSPC by the following reason. The MLSS of the two reactors – the most directly relevant variable to the event – were ranked as the top two during almost all the time by the Modular-MSPC, whereas incorrect variables probably not relevant to the event often accounted for a part of high ranks of the Adjusted-MSPC, as was pointed out in Camacho et al. (2016). Taking Figure 4 as an example, the top two of the Modular-MSPC are the MLSS and the third is also a relevant variable to the event, but the ORP not relevant to the event was ranked as the second by the Adjusted-MSPC despite it was not fluctuated.

In the second phase, the M-values sometimes exceeded the threshold 0.5. From such cases, we have identified the 26 important events including process faults and intentional changes of operating condition by interviewing the operators. The detected faults were sensor failures of wastewater quality and flow-rate, large and sudden influent disturbances, bad control performances such as the so-called hunting of PI control for DO concentration. The operating condition changes were the distribution rate change of inflow flow-rate to each reactor and the set-point changes of the waste and return sludge flow-rates to adjust the MLSS.

Figure 5 is an example of the screenshot of the developed UI when the influent flow-rate sensor failure occurred. The measured flow-rate fluctuated up and down while real flow-rate would never vary in such a way. This phenomenon was caused by the failure of the circuit board used to the flow-rate sensor, which was confirmed by the interview after the experiment. The details for other detected events are omitted due to space limitation. In addition, we have received positive opinions from the operators that the developed UI is easy-to-understand and helpful for process monitoring.

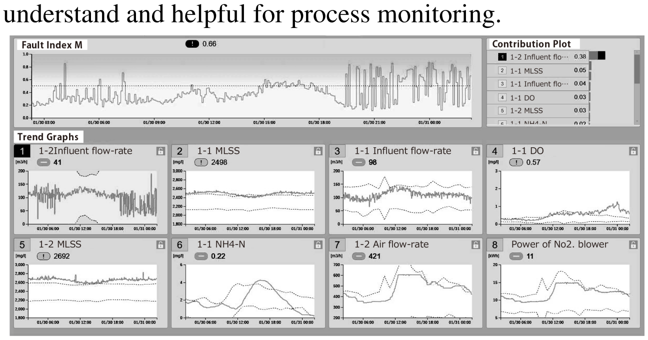

Figure 5. Screenshot of UI during flow-rate sensor failure

5. CONCLUSION

This paper has presented a practical monitoring method based on a modified MSPC and has illustrated its effectiveness by applying it to a real municipal WWTP in Japan in both an off-line manner and an on-line manner. Firstly, a novel user interface by combining the conventional PCA-MSPC with existing SPC-like process monitoring has been developed. Then, a simplified decomposable PCA-MSPC named "Modular-MSPC" has been proposed. In addition to improvement of model tractability by adopting the Modular-MSPC, it has been shown, through the WWTP application, that the Modular-MSPC works well in terms of detectability, diagnosability, and interpretability. Future works will be: long-term evaluation of the Modular-MSPC for various real plants, and development of effective model update algorithm and a useful automatic input variable selection method.

REFERENCES

B-DASH project (2016). Guideline for introducing ICT-based advanced process control and remote diagnosis technology for efficient wastewater treatment plant operation,Technocal Note of NILIM, No.939, http://www.nilim.go.jp/lab/bcg/siryou/tnn/tnn0939.htm

Bodenham, D. A., and Adams, N. M. (2016). A comparison of efficient approximations for a weighted sum of chi-squared random variables. Statistics and Computing, 26(4), 917-928.

Camacho, J., Pérez-Villegas, A., García-Teodoro, P., and Maciá-Fernández, G. (2016). PCA-based multivariate statistical network monitoring for anomaly detection. Computers & Security, 59, 118-137.

Choi, S. W., Morris, J., and Lee, I. B. (2008). Nonlinear multiscale modelling for fault detection and identification. Chemical engineering science, 63(8), 2252-2266.

Jackson, J.E. (1991). A user's guide to principal components, New York:Wiley.

Uchida, Y., Fujiwara, K., Saito, T., and Osaka, T. (2022). Process Fault Diagnosis Method Based on MSPC and LiNGAM and its Application to Tennessee Eastman Process. IFAC-PapersOnLine, 55(2), 384-389.

Wise, B.M., Gallagher, N.B. (1996). The process chemometrics. approach to process monitoring and fault detection, J. Proc. Cont, 6(6), 329–348.

Zhao, L. T., Yang, T., Yan, R., and Zhao, H. B. (2022). Anomaly detection of the blast furnace smelting process using an improved multivariate statistical process control model. Process Safety and Environmental Protection, 166, 617-62.