

# Study on modifications of PLS approach for process monitoring

Shen Yin\* Steven X. Ding\* Ping Zhang\*  
Adel Hagahni\* Amol Naik\*

*\* Institute of Automation and Complex Systems, University of  
Duisburg-Essen, Bismarckstrasse 81 BB, 47057 Duisburg, Germany  
(e-mail: shen.yin@uni-due.de)*

---

**Abstract:** Partial least squares (PLS) is an efficient approach for multivariate statistical process monitoring. Although it works in many industrial applications, Zhou et al. [2010] revealed that some properties of PLS algorithm may hamper overall efficiency of process monitoring scheme. To solve these problems, a modified approach is proposed in this paper. Compared with the existing PLS approaches, the new approach performs an orthogonal decomposition on regression variable space to eliminate the variations useless for output prediction. Based on the new approach, a complete process monitoring scheme is also developed. Finally, the effectiveness of the proposed approach is verified on an industrial benchmark of Tennessee Eastman process.

Keywords: Multivariate statistical process monitoring; Fault diagnosis; Partial least squares; Data-Driven methods; Tennessee Eastman process.

---

## 1. INTRODUCTION

Partial least squares (PLS) is a popular method for multivariate statistical process monitoring applications. Thanks to its efficiency in processing huge amount of highly correlated plant data, PLS is recognized as a powerful tool for data-driven model building, fault detection and diagnosis Khan et al. [2008], Kresta et al. [1991], Wise and Gallagher [1996]. The basic PLS algorithm can be found in Helland [1998], Hoskuldsson [1998]. Several modification of it, such as recursive PLS Qin [1998], multiblock PLS Lee et al. [2004], MacGregor et al. [1994], local PLS Kruger and Dimitriadis [2008] and dynamic PLS Choi and Lee [2005] are also proposed and successfully implemented in variety of industrial applications. Recently, a geometric explanation of PLS approach is proposed by Li et al. [2010].

The classic PLS algorithm divides subspace of regression variables,  $X$ , into two subspaces, i.e.  $\hat{X}$  and  $\tilde{X}$ , depending on their correlation with output variables,  $Y$ . The Hotelling's  $T^2$  and squared prediction error ( $SPE$ ) statistics are typically used for monitoring both the subspaces. Although classic PLS algorithm works in many cases, the complexity of algorithm, especially the iterative computation procedures caused by many PLS components, leads to the PLS model difficult to interpret. Recently, Zhou et al. [2010] revealed that the classic PLS approach may result in variations in  $\hat{X}$  orthogonal to  $Y$  thus, are useless to predict  $Y$ . In addition,  $\tilde{X}$  may contain large variability of  $X$  in it and is therefore, not suitable for monitoring as residual subspace by means of  $SPE$  statistic. To solve these problems, the authors proposed the so-called total projection technique, which performs classic PLS algorithm and further decomposes certain subspaces.

The drawback due to orthogonal variations amongst interacting subspaces comes from the nature of the PLS algorithm. To improve the performance in monitoring large-scale plants, we would like to review the standard PLS algorithm in depth, which is the primary objective of this paper. Based on our study, we shall propose a modified approach, which is computationally less expensive than the classic PLS and more importantly, avoids the drawback mentioned earlier.

According to the new approach, the complete process monitoring scheme is designed. To this end, only two  $T^2$  based statistical indices are dedicated to monitor subspace of  $X$ . It is more efficient compared to all other existing approaches. In addition, the residual subspace of  $Y$ , which is uncorrelated to  $X$ , can also be monitored by  $SPE$  index. Generally, this statistic serves as an off-line indicator to reflect the nature of the fault occurred in  $Y$ , since in practice product quality variables,  $Y$ , is always sampled slower than process variables,  $X$ .

The rest of this paper is organized as follows. In section 2, the classic PLS algorithm is introduced and the problem is formulated. Section 3 addresses the theoretical core of the modified approach. The complete process monitoring scheme is presented in section 4. The results obtained from industrial benchmark of Tennessee Eastman process are shown in section 5. Section 6 provides the concluding remarks of the work.

## 2. PRELIMINARIES AND PROBLEM FORMULATION

### 2.1 PLS based process monitoring scheme

Given a measurement data matrix (regression variables)  $X$ , which records  $N$  samples of  $n$  process variables and

$Y$  consisting of  $N$  samples of  $m$  product quality variables (outputs), i.e.

$$X = \begin{bmatrix} x_1^T \\ \vdots \\ x_N^T \end{bmatrix} \in \mathcal{R}^{N \times n}, Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix} \in \mathcal{R}^{N \times m}$$

$$x_i \in \mathcal{R}^n, y_i \in \mathcal{R}^m, i = 1, \dots, N$$

then PLS involves projecting  $X$  and  $Y$  onto the so-called latent variables,

$$T = [t_1 \dots t_\gamma] \in \mathcal{R}^{N \times \gamma}$$

such that the correlation model of  $X$  and  $Y$  becomes

$$X = TP^T + \tilde{X} = \hat{X} + \tilde{X} \quad (1)$$

$$Y = TQ^T + E_y = XM + E_y \quad (2)$$

where  $\gamma$  is the number of latent variables and  $P \in \mathcal{R}^{n \times \gamma}, Q \in \mathcal{R}^{m \times \gamma}$  are loading matrices of  $X$  and  $Y$ , respectively.  $M \in \mathcal{R}^{n \times m}$  is the matrix of regression coefficients.  $\hat{X} = TP^T$  is highly correlated with  $Y$ .  $\tilde{X}$  and  $E_y$  are residual subspaces and assumed to be uncorrelated with  $Y$  and  $X$ , respectively. From the PLS model (1)-(2), the score matrix  $T$  and coefficient matrix  $M$  can be calculated as

$$T = XR, M = RQ^T$$

with

$$P^T R = R^T P = I_{\gamma \times \gamma}, R \in \mathcal{R}^{n \times \gamma}.$$

The predicted output  $\hat{Y}$  can be calculated by

$$\hat{Y} = XM.$$

Algebraically, PLS is implemented with nonlinear iterative partial least squares algorithm (NIPALS) which can be summarized as follows Dayal and MacGregor [1997], Hoskuldsson [1998]:

- (1) Center columns of  $X$  and  $Y$  to zero mean and scale them to unit variance
- (2) Perform  $\gamma$  times following iterative computations:  
for  $i = 1, \dots, \gamma$

$$(w_i^*, q_i^*) = \arg \max_{\|w_i\|=1, \|q_i\|=1} w_i^T X_i^T Y q_i, X_1 = X$$

$$t_i = X_i w_i^*, p_i = \frac{X_i^T t_i}{\|t_i\|^2}, X_{i+1} = X_i - t_i p_i^T$$

$$r_1 = w_1^*, r_i = \prod_{j=1}^{i-1} (I_{n \times n} - w_j^* p_j^T) w_i^*, i > 1$$

where  $\gamma$  is determined by applying a known criteria, e.g. cross validation Wold et al. [2001].

- (3) Compute  $P, Q, R, T$  and  $M$  matrices as

$$P = [p_1 \dots p_\gamma], T = [t_1 \dots t_\gamma]$$

$$Q = [q_1 \dots q_\gamma], R = [r_1 \dots r_\gamma]$$

$$M = RQ^T.$$

The original idea behind the PLS is to identify the correlation model (1)-(2) based on the covariance matrix  $cov(x, y)$  and to predict  $y$  using the (online) observation  $x$ . On the assumption that  $x \sim \mathcal{N}(0, \Sigma_x)$  and  $y \sim \mathcal{N}(0, \Sigma_y)$ , fault

detection can be achieved using suitable test statistics based on  $\hat{x} = PR^T x$  and its residual  $\tilde{x} = (I_{n \times n} - PR^T) x$ . To this aim, Hotelling's  $T^2$  statistic and  $SPE$  statistic are popularly used to detect changes in  $\hat{x}$  and  $\tilde{x}$ , i.e.

$$T^2 = x^T R \left( \frac{T^T T}{N-1} \right)^{-1} R^T x \quad (3)$$

$$SPE = \|\tilde{x}\|^2 = \|(I_{n \times n} - PR^T) x\|^2 \quad (4)$$

The threshold for fault detection can be calculated as:

$$J_{th, T^2} = \frac{\gamma(N^2 - 1)}{N(N - \gamma)} F_\alpha(\gamma, N - \gamma) \quad (5)$$

$$J_{th, SPE} = g \chi_{h, \alpha}^2 \quad (6)$$

where  $F_\alpha(\gamma, N - \gamma)$  is  $F$  distribution with  $\gamma$  and  $N - \gamma$  degrees of freedom with significance level  $\alpha$ .  $g \chi_{h, \alpha}^2$  is the  $\chi^2$  distribution under significance level  $\alpha$  with scaling factors  $g = S/2\mu$  and  $h = 2\mu^2/S$ , where  $\mu$  and  $S$  are sample mean and variance of  $SPE$  statistic Nomikos and MacGregor [1995], Tracy et al. [1992].

## 2.2 Problem formulation

Although PLS based process monitoring performs excellently in many applications, it has been proven in Li et al. [2010], Zhou et al. [2010] that PLS perform an oblique decomposition on input space thus  $\hat{X}$  may contain variations orthogonal to  $Y$  that is not useful for prediction, while  $\tilde{X}$  may have large variations that hamper overall efficacy of the process monitoring scheme presented by (3)-(6). As a result, it is possible that (i) monitoring subspace  $\hat{X}$  may provide fault information unrelated to  $Y$  and, (ii)  $\tilde{X}$  is unsuitable for process monitoring as a residual subspace. To counter this problem, the following two issues must be resolved:

- a complete decomposition of  $X$  and  $Y$  spaces, such that

$$X = \hat{X} + \tilde{X} \quad (7)$$

$$Y = \hat{Y} + E_y \quad (8)$$

$$\hat{Y} = XM = \hat{X}M \quad (9)$$

where the subspaces  $\hat{Y}$  and  $E_y$  are correlated and uncorrelated with  $X$ , respectively.  $\hat{X}$  and  $\tilde{X}$  denote an orthogonal decomposition on input space such that  $\tilde{X}$  has no contribution for outputs prediction, while  $\hat{X}$  is fully responsible for predicting  $Y$  thus does not contain variations orthogonal to  $Y$ , and

- development of statistical indices for process monitoring based on the above decomposition.

To deal with the first issue, a modified approach is derived in the next section.

## 3. A MODIFIED APPROACH

### 3.1 A complete decomposition of $Y$ space

Based on the discussion so far, it is possible to write following desired relation as:

$$Y = XM + E_y = \hat{Y} + E_y \quad (10)$$

where  $M$  is a coefficient matrix and contains correlation information between  $X$  and  $Y$ .  $E_y$  is the residual part of  $Y$  which is uncorrelated with the input signal, i.e.

$$\text{cov}(e_y, x) = \mathcal{E}\{e_y x^T\} = 0$$

where  $x^T$  and  $e_y^T$  are row vector in  $X$  and  $E_y$ , respectively. Without the loss of generality,  $M$  is assumed a full column-rank matrix and  $1 \leq m < n$ .

According to (10), in case of  $N \gg \max\{n, m\}$ , it follows that

$$\frac{1}{N} Y^T X = \frac{1}{N} M^T X^T X + \frac{1}{N} E_y^T X \approx M^T \frac{X^T X}{N}$$

Thus, in case of inversable  $X^T X$ ,  $M$  can be easily calculated as:

$$M = (X^T X)^{-1} X^T Y. \quad (11)$$

Note that, in case the product,  $X^T X$ , is not full rank matrix, it follows that

$$M = (X^T X)^\dagger X^T Y. \quad (12)$$

The pseudo-inverse is calculated as:

$$(X^T X)^\dagger = P_{x,pc} \Lambda_{x,pc}^{-1} P_{x,pc}^T \quad (13)$$

where

$$X^T X = P_x \Lambda P_x^T = [P_{x,pc} \ P_{x,res}] \begin{bmatrix} \Lambda_{x,pc} & 0 \\ 0 & \Lambda_{x,res} \end{bmatrix} \begin{bmatrix} P_{x,pc}^T \\ P_{x,res}^T \end{bmatrix}$$

$$\Lambda_{x,res} = 0$$

It is now easier to show that  $Y$  is completely decomposed into two parts as desired in (10).

### 3.2 Orthogonal decomposition of $X$ space

So far, we have achieved complete decomposition of  $Y$ . Based on this result, we may proceed to decompose  $X$  into two parts, i.e.  $\hat{X}$ ,  $\tilde{X}$ , such that  $\hat{X}$  does not contains variations orthogonal to  $Y$  thus has the fully contribution in predicting  $Y$ , while  $\tilde{X}$  is orthogonal to  $\hat{X}$  and gives no contribution.

A simple way to perform aforementioned decomposition is achieved by projecting  $X$  orthogonally onto  $\text{span}\{M\}$  and  $\text{span}\{M\}^\perp$ . Consequently,

$$\tilde{X} M = 0 \quad (14)$$

$$\hat{Y} = X M = \hat{X} M. \quad (15)$$

$$\hat{X} \equiv \text{span}\{M\}, \tilde{X} \equiv \text{span}\{M\}^\perp \quad (16)$$

For orthogonal projection, following steps are necessary:

- Perform SVD on matrix  $M M^T$

$$M M^T = [P_M \ \tilde{P}_M] \begin{bmatrix} \Lambda_M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_M^T \\ \tilde{P}_M^T \end{bmatrix}. \quad (17)$$

where  $P_M \in \mathcal{R}^{n \times m}$ ,  $\tilde{P}_M \in \mathcal{R}^{n \times (n-m)}$ ,  $\Lambda_M \in \mathcal{R}^{m \times m}$ .

- Construct  $\Pi_M$ ,  $\Pi_M^\perp$ , which are the orthogonal projectors on  $\text{span}\{M\}$  and  $\text{span}\{M\}^\perp$  respectively, i.e.

$$\Pi_M = P_M P_M^T, \Pi_M^\perp = \tilde{P}_M \tilde{P}_M^T.$$

Table 1. Description of subspaces

Subspace	Description
$\hat{X}$	subspace of $X$ that is fully responsible for predicting $Y$
$\tilde{X}$	subspace of $X$ that is orthogonal to $\hat{X}$ and has no contribution for prediction of $Y$
$\hat{Y}$	subspace of $Y$ that is correlated to $X$ and represents the prediction of outputs
$E_y$	subspace of $Y$ that is uncorrelated to $X$

- Decompose  $X$  into two orthogonal subspaces

$$X = \hat{X} + \tilde{X}$$

$$\hat{X} = X \Pi_M = X P_M P_M^T \in S_{\hat{x}} \equiv \text{span}\{M\} \quad (18)$$

$$\tilde{X} = X \Pi_M^\perp = X \tilde{P}_M \tilde{P}_M^T \in S_{\tilde{x}} \equiv \text{span}\{M\}^\perp. \quad (19)$$

Table 1 provides the descriptions of different subspaces and the complete algorithm is summarized in Table 2.

### Remarks

We would like to point out that

- compared to classic PLS, the proposed approach provides a desired decomposition of  $X$  and  $Y$  as shown in (7)-(9), thus avoids the drawbacks of classic PLS algorithm and,
- the new approach provides a clearer interpretation of the correlation model and has less computational overload than classic PLS algorithm.

## 4. PROCESS MONITORING SCHEME

### 4.1 Monitoring subspaces $\hat{X}$

Since  $\hat{X}$  and  $\hat{Y}$  are mutually correlated, it is easy to prove that both the subspaces are equivalent in the sense of change detection. Thus, in this subsection, we only focus on the monitoring based on  $\hat{X}$ , which enables detecting faults in  $X$  space that are related to  $Y$ . To this aim, consider the quadratic form of vector  $\hat{x}$ , which is recommended to design the test statistic:

$$\hat{x}^T \hat{x} = x^T P_M P_M^T P_M P_M^T x = x^T P_M P_M^T x. \quad (20)$$

Since

$$\text{rank}\{P_M^T X^T\} = m$$

$P_M^T x \in \mathcal{R}^m$  is a suitable candidate for  $T^2$  statistic for monitoring  $\hat{X}$ . Thus, the  $T^2$  statistic follows

$$T_{\hat{x}}^2 = x^T P_M \left( \frac{P_M^T X^T X P_M}{N-1} \right)^{-1} P_M^T x. \quad (21)$$

The threshold for  $T_{\hat{x}}^2$  is given by

$$J_{th, T_{\hat{x}}^2} = \frac{m(N^2-1)}{N(N-m)} F_{\alpha}(m, N-m). \quad (22)$$

where  $F_{\alpha}(m, N-m)$  is  $F$  distribution with  $m$  and  $N-m$  degrees of freedom and  $\alpha$  is user-specified significance level.

Table 2. Algorithm of modified approach

Based on the normalized data $X$ and $Y$
1. Calculated the coefficient matrix $M$ : $M = (X^T X)^{\dagger} X^T Y$ .
2. $\hat{Y} = XM$ , $E_y = Y - \hat{Y}$
3. Do SVD on $MM^T$ : $MM^T = [P_M \tilde{P}_M] \begin{bmatrix} \Lambda_M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_M^T \\ \tilde{P}_M^T \end{bmatrix}$ .
4. Calculate $\Pi_M = P_M P_M^T$ , $\Pi_M^{\perp} = \tilde{P}_M \tilde{P}_M^T$ .
5. $\hat{X} = X \Pi_M$ , $\tilde{X} = X \Pi_M^{\perp}$

The detection logic is:

$$T_{\tilde{x}}^2 > J_{th, T_{\tilde{x}}^2} \implies \text{faulty in } \hat{X}, \text{ otherwise fault-free.}$$

#### 4.2 Monitoring subspaces $\tilde{X}$

As mentioned in the previous section,  $\tilde{X}$  represents subspace in  $X$  that is not correlated to  $Y$  and has no contribution on predicting  $Y$ . Therefore, similar to (20), the quadratic form of  $\tilde{x}$  can be written as:

$$\tilde{x}^T \tilde{x} = x^T \tilde{P}_M \tilde{P}_M^T \tilde{P}_M \tilde{P}_M^T x = x^T \tilde{P}_M \tilde{P}_M^T x. \quad (23)$$

Note that

$$\text{rank} \left\{ \tilde{P}_M^T X^T \right\} = n - m$$

therefore  $\tilde{P}_M^T x \in \mathcal{R}^{n-m}$  can be used as  $T^2$  statistic for monitoring  $\tilde{X}$ . Thus, it follows

$$T_{\tilde{x}}^2 = x^T \tilde{P}_M \left( \frac{\tilde{P}_M^T X^T X \tilde{P}_M}{N-1} \right)^{-1} \tilde{P}_M^T x. \quad (24)$$

The threshold for  $T_{\tilde{x}}^2$  is

$$J_{th, T_{\tilde{x}}^2} = \frac{(n-m)(N^2-1)}{N(N-n+m)} F_{\alpha}(n-m, N-n+m). \quad (25)$$

where  $F_{\alpha}(n-m, N-n+m)$  is  $F$  distribution with  $n-m$  and  $N-n+m$  degrees of freedom with significance level  $\alpha$ . The fault detection logic is similar as shown in previous subsection.

#### 4.3 Monitoring subspace $E_y$

In this subsection, monitoring and change detection based on  $E_y$  is investigated. It is known that  $E_y$  is not correlated to  $X$  and can be represented by

$$E_y = Y - \hat{Y} = (\hat{X} + \tilde{X})M + E_y - \hat{X}M \quad (26)$$

which shows that  $E_y$  is not influenced by any change in  $X$ . In other words, only the faults in  $Y$ , e.g. sensor failures, can be detected by monitoring  $E_y$ . Generally,  $e_y$  is assumed small in magnitude and contains insignificant variance information. Based on the available  $y$ , a  $SPE$  statistic can be designed as

$$SPE_y = \|\hat{y}\|^2 = \|y - \hat{y}\|^2 = \|y - M^T x\|^2 \quad (27)$$

The related threshold is

$$J_{th, SPE_y} = g_y \chi_{h_y, \alpha}^2 \quad (28)$$

where  $g_y \chi_{h_y, \alpha}^2$  is  $\chi^2$  distribution with scaling factor  $g_y = S/2\mu$  and  $h_y = 2\mu^2/S$  under given significance level  $\alpha$ ,  $\mu$  and  $S$  are sample mean and variance of  $SPE_y$  statistic (27). The fault detection logic is given as:

$$SPE_y > J_{th, SPE_y} \implies \text{faulty in } Y, \text{ otherwise fault-free.}$$

Note that, Ding et al. [2010] have recently proposed a new  $SPE$  statistic which is similar to Hawkins  $T_H^2$  statistic, but without the numerical drawback. Compared to standard  $SPE$  statistic, the threshold selection for the new statistic is computationally simpler. Moreover, if  $e_y$  is large and retains significant variance, the test statistic can be simply extended to a combined index to ensure a high fault detectability. In this case, the combined index becomes

$$T_{comb}^2 = \tilde{y}^T P_{\tilde{y}} \Xi P_{\tilde{y}}^T \tilde{y} \quad (29)$$

where

$$\frac{1}{N-1} \tilde{Y}^T \tilde{Y} = P_{\tilde{y}} \Lambda_{\tilde{y}} P_{\tilde{y}}^T, \Lambda_{\tilde{y}} = \begin{bmatrix} \Lambda_{\tilde{y}, pc} & 0 \\ 0 & \Lambda_{\tilde{y}, res} \end{bmatrix}$$

$$\Xi = \text{diag} \left( \frac{\sigma_m^2}{\sigma_1^2}, \dots, \frac{\sigma_m^2}{\sigma_{m-1}^2}, 1 \right)$$

with

$$\sigma_1^2 \geq \dots \geq \sigma_{\beta}^2 \gg \sigma_{\beta+1}^2 \geq \dots \geq \sigma_m^2 > 0.$$

The corresponding threshold is

$$J_{th, T_{comb}^2} = \sigma_m^2 \chi_{\alpha}^2(m) \quad (30)$$

In the extreme case, the covariance matrix is singular, an SVD yields

$$\frac{1}{N-1} \tilde{Y}^T \tilde{Y} = [P_{\tilde{y}} \ P_{\tilde{y}}^{\perp}] \begin{bmatrix} \Lambda_{\tilde{y}} & 0 \\ 0 & 0 \end{bmatrix} [P_{\tilde{y}} \ P_{\tilde{y}}^{\perp}]^T$$

$$\Lambda_{\tilde{y}} = \text{diag}(\sigma_1^2, \dots, \sigma_m^2) \in R^{\bar{m} \times \bar{m}}, \bar{m} < m.$$

The fault detection can be achieved using (29)-(30) with suitable dimension  $\bar{m}$ . The corresponding detection logic is

$$T_{comb}^2 > J_{th, T_{comb}^2} \implies \text{faulty in } Y, \text{ otherwise fault-free.}$$

Remarks

To the end of this section, we would like to point out that

- the modified approach based process monitoring scheme employs two  $T^2$  indices to monitor entire

Table 3. Subspaces and related test statistics

Subspace	Test statistics	Threshold
$\hat{X}$	$T_{\hat{x}}^2$	$\frac{m(N^2-1)}{N(N-m)} F_{\alpha}(m, N-m)$
$\tilde{X}$	$T_{\tilde{x}}^2$	see (25)
$E_y$	$SPE_y$	$g_y \chi_{h_y, \alpha}^2$ or
	$T_{comb}^2$	$\tilde{y}^T P_{\tilde{y}} \Xi P_{\tilde{y}}^T \tilde{y}$

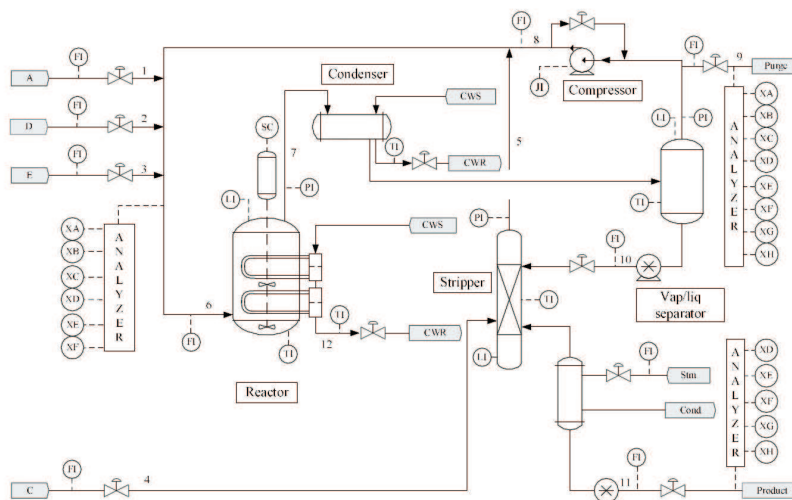


Fig. 1. The Tennessee Eastman process

subspace of  $X$ . Additionally, as against four test statistics to monitor  $X$  space as proposed in Zhou et al. [2010], the new approach based monitoring works more efficiently. The test statistics related to different subspaces are listed in Table 3 and,

- in practice the output vector  $y$ , e.g. which serves as product quality variable, is not always available on time. In this case, monitoring subspace  $E_y$  by means of  $SPE_y$  or  $T^2_{comb}$  will induce time delay, therefore it is best used in off-line analysis.

## 5. BENCHMARK STUDY

In this section, the results obtained with an industrial benchmark of Tennessee Eastman process are presented to illustrate the performance of the modified approach based monitoring scheme. The model is a realistic simulation program of a chemical plant which is widely accepted as a benchmark for control and monitoring studies. The process is described in Downs and Fogel [1993] and the FORTRAN code of the process is available over internet. The code and simulation data used for our studies are downloaded from <http://brahms.scs.uiuc.edu>. Fig. 1 shows the flow diagram of the process with five major units namely, reactor, condenser, compressor, separator and stripper. The process has two products and four reactants. Additionally, an inert and a by-product are also present making a total of 8 components. The process allows total 53 measurements out of which 41 are of process variables and 12 are manipulated variables. More detailed discussion can be found in Downs and Fogel [1993].

The major objective here is to implement modified approach based process monitoring scheme to detect 21 predefined faults mentioned in Chiang et al. [2001] as shown by Table 4. 22 process measurements (XMEAS(1-22)) and 11 manipulated variables (XMV(1-11)) are taken into account as  $X$ . The analyzer for component G (XMEAS(35)) is treated as product quality variable,  $y$ . Both PLS and modified approach based process monitoring schemes are designed based on 480 samples obtained from the normal process operation. The number of latent variables is selected as 6 based on the cross validation tests. Since the faults in TE may occur in different subspaces, which

are generally unknown in practice, a reasonable process monitoring logic is based on joint use of the related test statistics, i.e. for modified approach by means of  $T^2_{\hat{x}}$ ,  $T^2_{\hat{x}}$  and PLS using  $T^2$ ,  $SPE$  indices.

Table 5 summarizes the results of the experiments for all the faults in TE process. It can be seen that only in the case of IDV(9), classic PLS shows slightly higher FDR over modified approach, but in all the other cases, the modified approach outperforms classic PLS technique in term of FDR. Especially, for IDV(5), IDV(10-11), IDV(16-17) and IDV(19-21), the modified approach proves an excellent choice.

If the nature of the fault is known in advance, the false alarm rate for the fault unrelated to  $y$  is another critical index for evaluating both the methods. In order to classify the faults, the criteria presented in Zhou et al. [2010] is used such that the faults can be divided into two categories, i.e. (a) faults affecting  $y$  and, (b) faults having no influence on  $y$ . Particularly, the fault is said to affect  $y$ ,

Table 4. Descriptions of process faults in TE

Fault number	Location
IDV(1)	A/C feed ratio, B composition constant
IDV(2)	B composition, A/C ration constant
IDV(3)	D feed temperature
IDV(4)	Reactor cooling water inlet temperature
IDV(5)	Condenser cooling water inlet temperature
IDV(6)	A feed loss
IDV(7)	C header pressure loss-reduced availability
IDV(8)	A,B,C feed composition
IDV(9)	D feed temperature
IDV(10)	C feed temperature
IDV(11)	Reactor cooling water inlet temperature
IDV(12)	Condenser cooling water inlet temperature
IDV(13)	Reaction kinetics
IDV(14)	Reactor cooling water valve
IDV(15)	Condenser cooling water valve
IDV(16)	Unknown
IDV(17)	Unknown
IDV(18)	Unknown
IDV(19)	Unknown
IDV(20)	Unknown
IDV(21)	The valve fixed at steady state position

Table 5. Fault detection rates (FDR) (%) given by modified approach and PLS

Fault number	modified app. ( $T_{\hat{x}}^2$ or $T_{\hat{y}}^2$ )	PLS ( $T^2$ or $SPE$ )
IDV(1)	99.88	99.75
IDV(2)	98.75	98.50
IDV(3)	18.73	14.23
IDV(4)	99.88	98.38
IDV(5)	99.88	33.58
IDV(6)	99.88	99.88
IDV(7)	99.88	99.88
IDV(8)	98.50	97.75
IDV(9)	12.11	14.48
IDV(10)	91.01	82.52
IDV(11)	83.15	78.53
IDV(12)	99.75	99.13
IDV(13)	95.38	95.13
IDV(14)	99.88	99.88
IDV(15)	23.22	22.97
IDV(16)	94.26	68.29
IDV(17)	97.00	94.13
IDV(18)	91.14	90.64
IDV(19)	94.13	25.97
IDV(20)	91.26	62.67
IDV(21)	72.66	59.80

if 10% or more samples of  $y$  exceed the threshold after the fault occurs. Otherwise, the faults are assumed to have no influence on  $y$ . The threshold for  $y$  is determined based on the normal operation data.

Based on this criteria, the process faults IDV(3-4), IDV(9,11), IDV(14-16) and IDV(19) have almost no influence on  $y$ . According to the earlier discussion,  $T_{\hat{x}}^2$  for the modified approach and  $T^2$  statistic for classic PLS determine false alarm rate for the faults unrelated to  $y$ . Table 6 gives the FAR of both methods and it can be seen that the modified approach gives significantly lower FAR than PLS in most of the cases, except comparable FAR in IDV(3) and IDV(19).

## 6. CONCLUSION

In this paper, we have studied on modifications of PLS approach for process monitoring. The associated computation cost for the modified approach is considerably simpler compared to the standard technique and avoids the drawbacks of classic PLS algorithm. Based on the proposed approach, a process monitoring scheme related to different subspaces is further developed. Note that only two  $T^2$  statistics are used to monitor the entire  $X$  space, which provides further efficiency compared to existing approaches. The proposed process monitoring scheme is tested on an industrial benchmark of Tennessee Eastman process which shows superior performance.

Table 6. False alarm rate (FAR) (%) given by modified approach and PLS

Fault number	modified approach ( $T_{\hat{x}}^2$ )	PLS ( $T^2$ )
IDV(3)	13.61	11.49
IDV(4)	10.99	63.05
IDV(9)	7.62	11.36
IDV(11)	10.24	64.27
IDV(14)	9.99	99.75
IDV(15)	10.49	20.58
IDV(16)	45.82	58.55
IDV(19)	6.99	6.62

## REFERENCES

- L.H. Chiang, E. Russell, and R.D. Braatz. *Fault Detection and Diagnosis in Industrial Systems*. Springer, London, 2001.
- S.W. Choi and I.B. Lee. Multiblock pls-based localized process diagnosis. *Journal of Process Control*, 15:295–306, 2005.
- B.S. Dayal and J.F. MacGregor. Improved pls algorithms. *Journal of Chemometrics*, 11:73–85, 1997.
- S. Ding, P. Zhang, E. Ding, S. Yin, A. Naik, P. Deng, and W. Gui. On the application of pca technique to fault diagnosis. *Tsinghua Science and Technology*, 15:138–144, 2010.
- J. Downs and E. Fogel. A plant-wide industrial process control problem. *Computers and Chemical Engineering*, 17:245–255, 1993.
- I.S. Helland. On the structure of partial least squares regression. *Communications in Statistics-Simulation and Computation*, 17:581–607, 1998.
- A. Hoskuldsson. Pls regression methods. *Journal of Chemometrics*, 2:211–228, 1998.
- A. Khan, J. Moyne, and D. Tilbury. Virtual metrology and feedback control for semiconductor manufacturing processes using recursive partial least squares. *Journal of Process Control*, 18:961–974, 2008.
- J.V. Kresta, J.F. MacGregor, and T.E. Marlin. Multivariate statistical monitoring of process operating performance. *Canadian Journal of Chemical Engineering*, 69: 35–47, 1991.
- U. Kruger and G. Dimitriadis. Diagnosis of process faults in chemical systems using a local partial least squares approach. *AIChE Journal*, 54:2581–2596, 2008.
- G. Lee, C.H. Han, and E.S. Yoon. Multiple-fault diagnosis of the tennessee eastman process based on system decomposition and dynamic pls. *Industrial and Engineering Chemistry Research*, 43:8037–8048, 2004.
- G. Li, S.J. Qin, and D.H. Zhou. Geometric properties of partial least squares for process monitoring. *Automatica*, 46:204–210, 2010.
- J.F. MacGregor, C. Jaeckle, C. Kiparissides, and M.Koutoudi. Process monitoring and diagnosis by multiblock pls methods. *AIChE Journal*, 40:826–838, 1994.
- P. Nomikos and J.F. MacGregor. Multivariate spc charts for monitoring batch processes. *Technometrics*, 37:41–59, 1995.
- S.J. Qin. Recursive pls algorithms for adaptive data modeling. *Computers and Chemical Engineering*, 22: 503–514, 1998.
- N.D. Tracy, J.C. Young, and R.L. Mason. Multivariate control charts for individual observations. *Journal of Quality Technology*, 24:88–95, 1992.
- B.M. Wise and N.B. Gallagher. The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, 6:329–348, 1996.
- S. Wold, M. Sjostroma, and L. Eriksson. Pls-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58:109–130, 2001.
- D.H. Zhou, G. Li, and S.J. Qin. Total projection to latent structures for process monitoring. *AIChE Journal*, 56: 168–178, 2010.