Vidal-Puig, S.; Vitale, R.; Ferrer, A. (15-0). Data-driven supervised fault diagnosis methods based on latent variable models: a comparative study. Chemometrics and Intelligent Laboratory Systems. 187:41-52. https://doi.org/10.1016/j.chemolab.2019.02.006

Data-driven supervised fault diagnosis methods based on latent variable models: a comparative study

Santiago Vidal-Puig[a,*], Raffaele Vitale[a,b,c,*], Alberto Ferrer[a]

[a]*Grupo de Ingeniería Estadística Multivariante, Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain*
[b]*Molecular Imaging and Photonics Unit, Department of Chemistry, Katholieke Universiteit Leuven, Celestijnenlaan 200F, B-3001, Leuven, Belgium*
[c]*Laboratoire de Spectrochimie Infrarouge et Raman - UMR 8516, Université de Lille - Sciences et Technologies, Bâtiment C5, 59655, Villeneuve d'Ascq, France*

Abstract

A comparison among widely used multivariate latent variable-based techniques for supervised process fault diagnosis was carried out.
In order to assess their overall performance several diagnosis criteria were proposed (C_1: most suspected fault assignment; C_2: threshold-based fault assignment). Additionally, it was evaluated i) how the size of the training set used to build the latent variable models affected the diagnosis ability of the methods under study, ii) how they behaved under new types of failures not included in the original list of fault candidates and iii) which of them were more suitable for either early or late diagnosis.
To accomplish all these objectives, the approaches were tested in different scenarios. Two datasets were analysed: the first was generated by a Simulink-based model of a binary distillation column, while the second relates to a pasteurisation process performed in a laboratory-scale plant.

Keywords: supervised process fault diagnosis, fault reconstruction, fault signature, Partial Least Squares Discriminant Analysis (PLS-DA), sensitivity, specificity

1. Introduction

In the last decades, the employment of techniques such as Principal Component Analysis (PCA) and Partial Least Squares regression (PLS) [1–3] for the implementation of Latent Variable-based Multivariate Statistical Process Control (LV-MSPC) schemes, which can quickly, easily and efficiently recognise possible failures occurring during the manufacturing campaign, has become rather popular [4]. A fundamental step in LV-MSPC is to diagnose ongoing faults and identify the

*Corresponding authors:
Telephone number: +33769476654
Email address: svidalp@eio.upv.es (Santiago Vidal-Puig); rvitale86@gmail.com (Raffaele Vitale)
These authors have equal contributions

phenomena generating the respective out-of-control signals in order to minimise process down-time, increase safety of plant operations and reduce costs [5–10]. Fault diagnosis can be carried out in two different ways:

- unsupervised approach: the variables characterised by an abnormal evolution with respect to an in-control situation are pointed out by tools like contribution plots [11, 12]. Notice that unsupervised methods do not highlight the root causes of a failure which have to be determined by the process engineers afterwards;
- supervised approach: the characterisation of ongoing failures is addressed by comparing them with a database of previously observed ones. This strategy permits to directly focus on the reasons behind the deviations from Normal Operating Conditions - NOC - (provided they have historically been experienced), conceivably reducing the intervention time needed to fix them [13].

Several methodologies resorting to the principles of fault reconstruction [13–15] and fault signature extraction [16] as well as classification techniques such as Partial Least Squares Discriminant Analysis (PLS-DA) [17, 18] or Soft Independent Modelling of Class Analogies (SIMCA) [19, 20] have been successfully applied for latent variable-based supervised fault diagnosis [21–23].
The main aim of this article is to assess the potential of the most representative of these strategies:

1. SPE-based fault reconstruction (SPE-FR) [14]
2. Combined index-based fault reconstruction (CI-FR) [15]
3. Fault signature (FS) [16]
4. Partial Least Squares Discriminant Analysis (PLS-DA) [17, 18]

Their overall performance will be tested and compared through the analysis of two datasets: the first was generated by a Simulink-based model of a distillation tower, while the second relates to a pasteurisation process performed in a laboratory-scale plant. Additionally, it will be evaluated i) how the size of the training set used to build the latent variable models affects the diagnosis ability of the methods under study, ii) how they behave when new types of failures, not included in the original list of fault candidates, have to be identified and iii) which of them are more suitable for either early or late diagnosis.

2. Materials and methods

2.1. Latent variable-based fault detection

Fault detection in LV-MSPC is carried out by checking whether current measurement data are consistent with past NOC ones. Two are the indices exploited to identify deviations from in-control conditions: Squared Prediction Error (SPE) and Hotelling's T_A^2. They are defined as:

$$\text{SPE} = ||\mathbf{x}^{\mathrm{T}}(\mathbf{I} - \mathbf{P}\mathbf{P}^{\mathrm{T}})||^2 = ||\tilde{\mathbf{x}}||^2 = \sum_{k=1}^{K} \tilde{x}_k^2 \tag{1}$$

$$T_A^2 = \sum_{a=1}^{A} \frac{t_a^2}{s_a^2} \tag{2}$$

where $\mathbf{x}^{\mathrm{T}}$ $(1 \times K)$ is an observation vector containing the value of the K variables registered at a specific time point, $\mathbf{I}$ $(K \times K)$ denotes the identity matrix, $\mathbf{P}$ $(K \times A)$ contains the loadings of the latent variable model built on a group of NOC recordings, $\tilde{\mathbf{x}}$ $(K \times 1)$ corresponds to the residual vector associated to $\mathbf{x}^{\mathrm{T}}$ whose k-th element is connoted by $\tilde{x}_k$, A equals the number of computed latent variables, t_a represents the a^{th} latent variable score of $\mathbf{x}^{\mathrm{T}}$ calculated by projecting it onto the subspace defined by $\mathbf{P}$, and s_a^2 is the variance of the a^{th} latent variable.
Equation 1 applies to PCA. For PLS, it should read:

$$\mathrm{SPE} = \|\mathbf{x}^{\mathrm{T}}(\mathbf{I} - \mathbf{W}(\mathbf{P}^{\mathrm{T}}\mathbf{W})^{-1}\mathbf{P}^{\mathrm{T}})\|^2 = \|\mathbf{x}^{\mathrm{T}}(\mathbf{I} - \mathbf{W}^{*}\mathbf{P}^{\mathrm{T}})\|^2 \tag{3}$$

where $\mathbf{W}$ and $\mathbf{P}$ are the PLS weighting and loadings matrices, respectively.
SPE measures the squared perpendicular distance of $\mathbf{x}^{\mathrm{T}}$ to the NOC model hyperplane while T_A^2 reflects the distance from the origin of the NOC model hyperplane to the projection of $\mathbf{x}^{\mathrm{T}}$ onto it. Thus, process variations which severely break the correlation structure of the original in-control data [24] would lead to abnormally large values of SPE. On the other hand, variations that preserve such a correlation structure (e.g. changes in the operating conditions of the process) would lead to anomalously high values of T_A^2. Specifically, out-of-control signals are generated when SPE and/or T_A^2 are/is found to be beyond their/its individual control limit/limits, derived as in [25–27]
Yue and Qin [15] have proposed a single combined index to be resorted to for simplifying failure detection:

$$\varphi = \frac{\mathrm{SPE}}{\mathrm{SPE}_{\mathrm{limit}}} + \frac{T_A^2}{T_{\mathrm{limit}}^2} \tag{4}$$

where $\mathrm{SPE}_{\mathrm{limit}}$ and T_{limit}^2 symbolise the estimated confidence thresholds of the SPE and Hotelling's T_A^2 statistics, respectively. The distribution of the combined index φ can be approximated as a $g\chi_h^2$ distribution where g and h are derived according to [15].

2.2. Supervised latent variable-based fault diagnosis

In supervised latent variable-based fault diagnosis the identification of a fault is addressed by comparing the current out-of-control behaviour of the process (possibly detected by means of PCA- or PLS-based SPE and T_A^2 indices) with an existing database of reference patterns, each typical of one of say J known failures. The way this comparison is performed depends on the fault diagnosis method as detailed in the following subsections.

2.2.1. SPE-based fault reconstruction (SPE-FR)

In the presence of the j-th fault, $\mathbf{x}^{\mathrm{T}}$ can be represented as the sum of two contributions, related to its in-control ($\mathbf{x}_{\mathrm{NOC}}^{\mathrm{T}}$) and out-of-control ($\mathbf{f}_j^{\mathrm{T}}\boldsymbol{\Xi}_j^{\mathrm{T}}$) variation (see Figure 1 for a graphical sketch):

$$\mathbf{x}^{\mathrm{T}} = \mathbf{x}_{\mathrm{NOC}}^{\mathrm{T}} + \mathbf{f}_j^{\mathrm{T}}\boldsymbol{\Xi}_j^{\mathrm{T}} \tag{5}$$

where $\|\mathbf{f}_j^{\mathrm{T}}\|$ corresponds to the magnitude of the j-th failure ($\mathbf{f}_j$ has dimensions $A_{\mathrm{fault}}\times 1$) and $\boldsymbol{\Xi}_j$ ($K\times A_{\mathrm{fault}}$) is an orthonormal matrix spanning its A_{fault}-dimensional subspace, both empirically derived

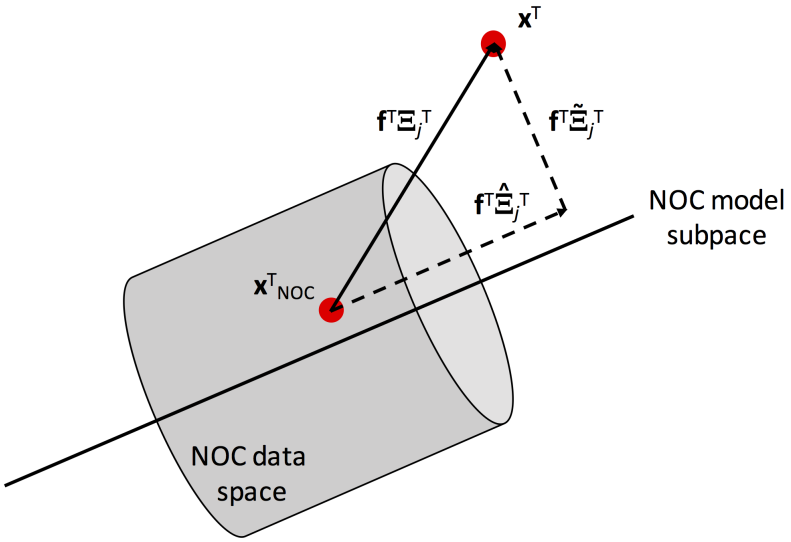

Figure 1 - Graphical representation of $\mathbf{x}^T$ as the sum of two contributions, related to its in-control ($\mathbf{x}^T_{NOC}$) and out-of-control ($\mathbf{f}_j^T\boldsymbol{\Xi}_j^T$) variation. Notice $\boldsymbol{\Xi}_j^T = \hat{\boldsymbol{\Xi}}_j^T + \tilde{\boldsymbol{\Xi}}_j^T$ where $\hat{}$ and $\tilde{}$ denote a matrix projection on the NOC latent variable model subspace and on the residuals subspace, respectively

from available data recorded during its occurrence [15, 28]. The objective of fault reconstruction is to re-estimate $\mathbf{x}^T_{NOC}$ by iteratively eliminating the effect of each one of the J known failures (according to Figure 1 this means bringing $\mathbf{x}^T$ back to $\mathbf{x}^T_{NOC}$ along the direction determined by $\boldsymbol{\Xi}_j$). Thus, that one, whose influence removal yields the best reconstruction of $\mathbf{x}^T_{NOC}$, is identified as the most analogous to the currently observed out-of-control situation.
The quality of the reconstruction can be evaluated by the so-called reconstructed SPE. For the j-th fault, it is defined as:

$$\mathrm{SPE}_j = \|\tilde{\mathbf{x}}_j\|^2 \tag{6}$$

being

$$\tilde{\mathbf{x}}_j^T = \tilde{\mathbf{x}}^T - \mathbf{f}_j^T\tilde{\boldsymbol{\Xi}}_j^T \tag{7}$$

$$\mathbf{f}_j^T = \arg\min \|\tilde{\mathbf{x}}^T - \mathbf{f}_j^T\tilde{\boldsymbol{\Xi}}_j^T\|^2 = \tilde{\mathbf{x}}^T\tilde{\boldsymbol{\Xi}}_j(\tilde{\boldsymbol{\Xi}}_j^T\tilde{\boldsymbol{\Xi}}_j)^{-1} \tag{8}$$

with $\tilde{}$ indicating a vector/matrix projection on the subspace of the residuals of the NOC model. Specifically, the lower SPE_j, the higher the similarity of the j-th failure to the ongoing one.

2.2.2. Combined index-based fault reconstruction (CI-FR)

The fault reconstruction approach, described in the previous subsection, has been extended by Yue and Qin [15] so that both SPE and Hotelling's T_A^2 are taken into account when diagnosing a failure. Here the combined index φ_j substitutes SPE_j for the fault identification:

$$\varphi_j = \frac{\mathrm{SPE}_j}{\mathrm{SPE}_{\mathrm{limit}}} + \frac{T^2_{A,j}}{T^2_{\mathrm{limit}}} \tag{9}$$

where $T^2_{A,j} = \|(\mathbf{x}^{\mathrm{T}} - \mathbf{f}_j^{\mathrm{T}}\boldsymbol{\Xi}_j^{\mathrm{T}})\mathbf{P}\mathbf{S}^{-\frac{1}{2}}\|^2$ and $\mathbf{S}$ is the scores covariance matrix resulting from the NOC latent variable model. Similarly to SPE-FR, the lower φ_j, the higher the similarity of the j-th failure to the currently observed deviation from NOC.

2.2.3. Fault signature (FS)

A fault signature consists of the direction vectors describing how the process moves in both the latent variable model and residual subspace immediately after the fault detection [16]. Fault diagnosis is then enabled by comparing such vectors calculated based on the outlying observations with those yielded by known failures.

Mathematically, rewriting Equation (5), the sample vector $\mathbf{x}^{\mathrm{T}}$, in the presence of the j-th fault, can be expressed as:

$$\mathbf{x}^{\mathrm{T}} = \mathbf{x}^{*\mathrm{T}} + \mathbf{z}_j^{\mathrm{T}} \tag{10}$$

where $\mathbf{x}^{*\mathrm{T}}$ symbolises the in-control observation immediately preceding it in time. $\mathbf{z}_j^{\mathrm{T}}$ can be further decomposed using the NOC latent variable model into two contributions, one lying on the model space ($\hat{\mathbf{z}}_j^{\mathrm{T}}$) and the other on the residual space ($\tilde{\mathbf{z}}_j^{\mathrm{T}}$):

$$\mathbf{z}_j^{\mathrm{T}} = \hat{\mathbf{z}}_j^{\mathrm{T}} + \tilde{\mathbf{z}}_j^{\mathrm{T}} = \mathbf{z}_j^{\mathrm{T}}\mathbf{P}\mathbf{P}^{\mathrm{T}} + \mathbf{z}_j^{\mathrm{T}}(\mathbf{I} - \mathbf{P}\mathbf{P}^{\mathrm{T}}) \tag{11}$$

$\hat{\mathbf{z}}_j^{\mathrm{T}}$ and $\tilde{\mathbf{z}}_j^{\mathrm{T}}$ are then normalised to be insensitive to their magnitudes as:

$$\hat{\mathbf{z}}_{j,\mathrm{norm}}^{\mathrm{T}} = \frac{\hat{\mathbf{z}}_j^{\mathrm{T}}}{\|\hat{\mathbf{z}}_j^{\mathrm{T}}\|} \tag{12}$$

$$\tilde{\mathbf{z}}_{j,\mathrm{norm}}^{\mathrm{T}} = \frac{\tilde{\mathbf{z}}_j^{\mathrm{T}}}{\|\tilde{\mathbf{z}}_j^{\mathrm{T}}\|} \tag{13}$$

$\hat{\mathbf{z}}_{j,\mathrm{norm}}^{\mathrm{T}}$ and $\tilde{\mathbf{z}}_{j,\mathrm{norm}}^{\mathrm{T}}$ denote the signatures of the j-th fault and the whole sets of $\hat{\mathbf{z}}_{j,\mathrm{norm}}^{\mathrm{T}}$ and $\tilde{\mathbf{z}}_{j,\mathrm{norm}}^{\mathrm{T}}$ vectors for $j = 1, \ldots, J$ constitute the reference fault signature libraries:

$$\hat{\mathbf{Z}} = \{\hat{\mathbf{z}}_{1,\mathrm{norm}}^{\mathrm{T}},\ \hat{\mathbf{z}}_{2,\mathrm{norm}}^{\mathrm{T}},\ \ldots,\ \hat{\mathbf{z}}_{J,\mathrm{norm}}^{\mathrm{T}}\} \quad \tilde{\mathbf{Z}} = \{\tilde{\mathbf{z}}_{1,\mathrm{norm}}^{\mathrm{T}},\ \tilde{\mathbf{z}}_{2,\mathrm{norm}}^{\mathrm{T}},\ \ldots,\ \tilde{\mathbf{z}}_{J,\mathrm{norm}}^{\mathrm{T}}\} \tag{14}$$

$\hat{\mathbf{Z}}$ and $\tilde{\mathbf{Z}}$ contain all the available information (in both the model and the residual subspaces) about the J known faults. New fault signatures can be included in $\hat{\mathbf{Z}}$ and $\tilde{\mathbf{Z}}$ after a new type of fault has been detected.

As a new observation, $\mathbf{x}_{\mathrm{new}}^{\mathrm{T}}$, is detected as an outlier, its signature is compared with those in the fault library. To this end, $\mathbf{x}_{\mathrm{new}}^{\mathrm{T}}$ is first subjected to the same decomposition as in Equations (10) and (11):

$$\mathbf{x}_{\mathrm{new}}^{\mathrm{T}} = \mathbf{x}_{\mathrm{new}}^{*\mathrm{T}} + \mathbf{z}_{\mathrm{new}}^{\mathrm{T}} \tag{15}$$

$$\mathbf{z}_{\mathrm{new}}^{\mathrm{T}} = \hat{\mathbf{z}}_{\mathrm{new}}^{\mathrm{T}} + \tilde{\mathbf{z}}_{\mathrm{new}}^{\mathrm{T}} = \mathbf{z}_{\mathrm{new}}^{\mathrm{T}}\mathbf{P}\mathbf{P}^{\mathrm{T}} + \mathbf{z}_{\mathrm{new}}^{\mathrm{T}}(\mathbf{I} - \mathbf{P}\mathbf{P}^{\mathrm{T}}) \tag{16}$$

$\hat{\mathbf{z}}_{\mathrm{new,norm}}^{\mathrm{T}}$ and $\tilde{\mathbf{z}}_{\mathrm{new,norm}}^{\mathrm{T}}$ are then normalised:

$$\hat{\mathbf{z}}_{\mathrm{new,norm}}^{\mathrm{T}} = \frac{\hat{\mathbf{z}}_{\mathrm{new}}^{\mathrm{T}}}{\|\hat{\mathbf{z}}_{\mathrm{new}}^{\mathrm{T}}\|} \tag{17}$$

$$\tilde{\mathbf{z}}^{\mathrm{T}}_{\mathrm{new,norm}} = \frac{\tilde{\mathbf{z}}^{\mathrm{T}}_{\mathrm{new}}}{||\tilde{\mathbf{z}}^{\mathrm{T}}_{\mathrm{new}}||} \tag{18}$$

and the cosines of the angles between $\hat{\mathbf{z}}^{\mathrm{T}}_{\mathrm{new,norm}}$ and $\tilde{\mathbf{z}}^{\mathrm{T}}_{\mathrm{new,norm}}$ and the known fault signatures (namely $\alpha_{\hat{\mathbf{z}}^{\mathrm{T}}_{\mathrm{new,norm}}}$ and $\alpha_{\tilde{\mathbf{z}}^{\mathrm{T}}_{\mathrm{new,norm}}}$) are then computed: the closer these values to one, the higher the similarity between the two compared signatures (see Figure 2 for a schematic illustration of the fault signatures associated to the same type of fault occurred in two different process operating regions). Notice that SPE-FR and FS are strictly related, but the way fault directions/subspaces are estimated in these two methodologies is clearly distinct and this will obviously have an impact on the diagnosis outcomes.

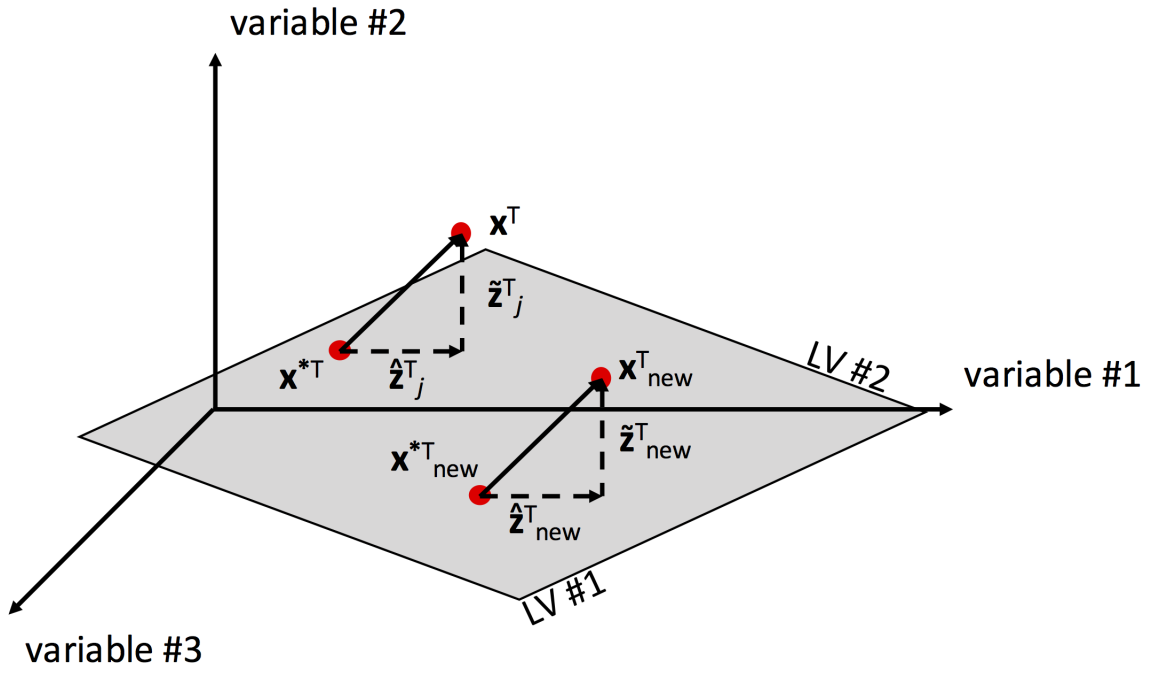

Figure 2 - Schematic representation of the fault signatures associated to the same type of fault occurred at two different process operating regions. LV stands for Latent Variable

2.2.4. Partial Least Squares Discriminant Analysis (PLS-DA)

Supervised fault diagnosis can be thought of as a discrimination problem. In fact, since historical data associated to e.g. J different classes (different kinds of fault) are available, one can build a classification model to be used in order to predict the belonging category of new out-of-control observations. PLS-DA is one of the simplest way of doing that: the process data coming from the failure database, $\mathbf{X}$ ($N \times K$), are regressed via Partial Least Squares (PLS [29, 30]) on a dummy binary-coded response matrix, $\mathbf{Y}$ ($N \times J$), made up by a set of row vectors, constructed so that, if their corresponding measurement records are members of the j-th class, they have a 1-value in their j-th entry and 0-values in all the other ones. The PLS-DA solution has been found to be equivalent to that resulting from Linear Discriminant Analysis (LDA [31]), but, due to its latent

variable-based nature, it permits to overcome the aforementioned drawbacks which classical statistical tools typically suffer from [17]. Whenever new faulty observations are registered, their $\mathbf{y}^{\mathrm{T}}$ $(1 \times J)$ vectors can be predicted and the final assignment accordingly addressed.

2.3. Datasets

To evaluate and compare the performance of the four supervised fault diagnosis methods under study, two different datasets were analysed: the first was generated by a Simulink-based model of a binary distillation column, while the second relates to a pasteurisation process performed in a laboratory-scale plant.

2.3.1. Dataset #1 - Distillation process

A Simulink-based model of a binary distillation column [32, 33] was exploited to simulate data resulting from a continuous process of fractionation of a methanol/ethanol mixture. A total number of 48 variables contaminated by homoscedastic measurement noise was generated (see Table 1). As detailed in Table 2, five data blocks ($\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{F}_3$, $\mathbf{F}_4$ and $\mathbf{F}_5$), related to disturbances

Table 1 - Distillation process dataset: list of the 48 measured variables

Variable ID	Variable description
T_{F}	Feed temperature
z_{F}	Feed composition
F_{V}	Feed volumetric flow rate
D	Top volumetric flow rate
V	Boil-up volumetric flow rate
B	Bottom volumetric flow rate
L	Reflux volumetric flow rate
T_{1-41}	1-41 column tray temperature

of specific nature, were obtained together with data simulated under NOC. $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$ are

Table 2 - Distillation process dataset: features of the 5 simulated data blocks

Block ID	Status of the process	Block dimensions
$\mathbf{F}_1$	Fault #1 - Disturbance in z_{F}	3984×48
$\mathbf{F}_2$	Fault #2 - Disturbance in T_{F}	3984×48
$\mathbf{F}_3$	Fault #3 - Disturbance in F_{V}	3984×48
$\mathbf{F}_4$	Fault #4 - PI control loop 1 failure	447×48
$\mathbf{F}_5$	Fault #5 - PI control loop 2 failure	268×48

constituted by 8 sub-matrices, respectively, each one containing 498 time samples, during which the corresponding failure was induced at different magnitudes (fault magnitude is tuned manually). As such a parameter could not be tuned for faults #4 and #5, both $\mathbf{F}_4$ and $\mathbf{F}_5$ include only 2 sub-arrays, yielded by two repeated simulations of the same type of interference (the dimensions of these sub-arrays are 223×48 and 224×48 for $\mathbf{F}_4$, and 139×48 and 129×48 for $\mathbf{F}_5$).

2.3.2. Dataset #2 - Pasteurisation process

Pasteurisation data were collected by an Armfield PCT23 MKII process plant trainer, which permits to monitor the 12 variables listed in Table 3. The instrumentation is equipped with an

Table 3 - Pasteurisation process dataset: list of the 12 measured variables

Variable ID	Variable description
$Level_{\mathrm{T}}$	Feed tank level
T_1	Liquid temperature in the pasteurisation tube
T_2	Heating water temperature
T_3	Final product temperature
T_4	Liquid temperature when preheating fresh feed
T_5	Fresh feed temperature after preheating
F	Liquid flow rate
P_1	Water heating power measure 1
P_2	Water heating power measure 2
P_3	Water heating power measure 3
$Pump_1$	Feed liquid peristaltic pump opening percentage
$Pump_2$	Heating water peristaltic pump opening percentage

electrical console for fault simulation. Here, as indicated in Table 4, 11 different data blocks were generated together with data recorded under NOC. They are all constituted by sub-matrices of

Table 4 - Pasteurisation process dataset: features of the 11 data blocks

Block ID	Status of the process	Block dimensions
$\mathbf{F}_1$	Fault #1 - T_1 set-point change	75×12
$\mathbf{F}_2$	Fault #2 - Heating water peristaltic pump failure	75×12
$\mathbf{F}_3$	Fault #3 - Liquid flow rate sensor failure	75×12
$\mathbf{F}_4$	Fault #4 - T_1 sensor failure (overestimation of T_1)	75×12
$\mathbf{F}_5$	Fault #5 - T_1 sensor failure (underestimation of T_1)	75×12
$\mathbf{F}_6$	Fault #6 - T_4 sensor failure (overestimation of T_4)	60×12
$\mathbf{F}_7$	Fault #7 - T_5 sensor failure (underestimation of T_5)	75×12
$\mathbf{F}_8$	Fault #8 - T_5 sensor failure (overestimation of T_5)	75×12
$\mathbf{F}_9$	Fault #9 - Pasteurisation tube outlet valve failure	60×12
$\mathbf{F}_{10}$	Fault #10 - Liquid flow rate set-point negative change	75×12
$\mathbf{F}_{11}$	Fault #11 - Liquid flow rate set-point positive change	75×12

dimension 15×12, yielded by repeated inductions of the same type of disturbance.

3. Comparative study

Both the original datasets were split into training and test sets so that entire replicates of all the faults under study were included in both of them. The former were employed as reference failure database, while the latter permitted to assess the performance of the single considered diagnostic strategies. Notice that every test set does not contain exact replicates of the faults of the corresponding training set, but replicates of different magnitude and/or with different noise contributions of the various concerned failures. In other words, in all the replicates of a single fault the evolution of the measured variables over time is affected by the failure in approximately the same way and what varies is just the severity of the breakage of NOC and/or the fluctuations of such a time evolution due to e.g. instrumental response variability. This represents a rather realistic scenario for the application of supervised fault diagnosis methodologies. Moreover, in order to avoid biased results, the fault replicates of the highest and lowest magnitude were always added to the respective training set.

3.1. Diagnosis performance indices

The accuracy degree of the diagnosis of the j-th failure can be estimated by the sensitivity and specificity indices, defined as:

$$\text{sensitivity}_j = \frac{\text{TP}_j}{\text{TP}_j + \text{FN}_j} \times 100 \tag{19}$$

$$\text{specificity}_j = \frac{\text{TN}_j}{\text{FP}_j + \text{TN}_j} \times 100 \tag{20}$$

where TP_j, FN_j, TN_j and FP_j stand for True Positives (the number of observations correctly identified as affected by the j-th fault), False Negatives (the number of observations mistakenly identified as not affected by the j-th fault), True Negatives (the number of observations correctly identified as not affected by the j-th fault) and False Positives (the number of observations mistakenly identified as affected by the j-th fault), respectively. The sensitivity index measures the proportion of observations affected by the j-th fault which were correctly pointed out as such by the diagnosis method. The specificity index reflects the proportion of observations not affected by the j-th fault which were correctly pointed out as such by the diagnosis method.
For a fair and more general comparison of the investigated approaches, the average sensitivity and the average specificity over all the considered types of faults were calculated:

$$\text{average sensitivity} = \frac{\sum_{j=1}^{J} \text{sensitivity}_j}{J} \tag{21}$$

$$\text{average specificity} = \frac{\sum_{j=1}^{J} \text{specificity}_j}{J} \tag{22}$$

3.2. Fault diagnosis criteria

An important point to stress regards how to decide whether a particular observation is affected by a specific type of fault and subsequently determine TP_j, FN_j, TN_j and FP_j. Here two criteria were applied:

- C_1: an observation is signalled as affected by the most suspected fault. C_1 might be useful when short intervention time and fast troubleshooting are needed as process engineers will have to focus only on a single type of deviation from NOC;
- C_2: an observation is signalled as affected by all the types of faults for which a certain condition is found to be fulfilled. C_2 might be beneficial when intervention time is not critical and process engineers may want to investigate all the possible reasons behind a particular out-of-control signal. Notice that this criterion is particularly interesting when new types of faults have to be identified.

For the various techniques dealt with:

a) SPE-FR:
   - C_1: an observation is identified as affected by the fault yielding the lowest reconstructed SPE;
   - C_2: an observation is identified as affected by the faults yielding reconstructed SPE values below an empirical significance threshold.

b) CI-FR:
   - C_1: an observation is identified as affected by the fault yielding the lowest φ value;
   - C_2: an observation is identified as affected by the faults yielding φ values below an empirical significance threshold.

c) FS:
   - C_1: an observation is identified as affected by the fault yielding the $\alpha_{\hat{\mathbf{x}}^{\mathrm{T}}_{\mathrm{new,norm}}}$ and $\alpha_{\tilde{\mathbf{x}}^{\mathrm{T}}_{\mathrm{new,norm}}}$ values closest to 1;
   - C_2: an observation is identified as affected by the faults yielding $\alpha_{\hat{\mathbf{x}}^{\mathrm{T}}_{\mathrm{new,norm}}}$ and $\alpha_{\tilde{\mathbf{x}}^{\mathrm{T}}_{\mathrm{new,norm}}}$ values beyond an empirical significance threshold.

d) PLS-DA:
   - C_1: an observation is assigned to the class for which it showed its highest y-predicted value;
   - C_2: an observation is assigned to the classes for which it showed y-predicted values higher than an empirical significance threshold [34].

3.3. Diagnosis and model windows

To evaluate the effect of the size of the training set on the performance of the diagnosis techniques, 1, 30 and 120 observations per fault replicate for the distillation process, and 1, 6 and 12 observations per fault replicate for the pasteurisation process, respectively, were used to build the supervised latent variable models. The selection of the dimension of such a model window W_{m} rested on the diverse dynamics of the two processes, on data availability and on the nature of the simulated induced failures.
Similarly, to study the temporal evolution of the diagnosis performance from the first time instant at which each fault is detected and determine whether a specific approach was more suitable for either early or late diagnosis, 1 to 120 observations per failure for the distillation process and 1 to 12 observations per failure for the pasteurisation process were included in the test set, respectively. This way, a range of various diagnosis windows W_{d} was taken into account.

3.4. Pre-adjustment of the empirical significance thresholds for C_2

For enabling a fair comparison based on C_2, the aforementioned empirical thresholds were estimated in order to achieve a similar average specificity rate (for the training set and, thus, at least for $W_{\mathrm{d}} = W_{\mathrm{m}}$) for all the diagnosis strategies. This guarantees that the average sensitivity values resulting from the application of the different approaches are commensurable.
On the contrary, when the assignation criterion is C_1 there is no need of such a pre-calibration: in this case the average sensitivity and the average specificity are strictly inter-correlated and only diverge for a scaling factor.

4. Results

4.1. Average sensitivity and specificity

Figures 3-6 show the trend of the average sensitivity and average specificity yielded by each methodology under study as the size of W_{d} increases and for every considered W_{m} size Figure 3 displays the outcomes of the analysis of the distillation process dataset based on the C_1 assignation criterion; Figure 4 displays the outcomes of the analysis of the distillation process dataset based on the C_2 assignation criterion; Figure 5 displays the outcomes of the analysis of the pasteurisation process dataset based on the C_1 assignation criterion; Figure 6 displays the outcomes of the analysis of the pasteurisation process dataset based on the C_2 assignation criterion. From them it follows:

a) SPE-FR: for the distillation process dataset SPE-FR exhibited the best performance when the model window was not too small ($W_{\mathrm{m}} = 30$ or $W_{\mathrm{m}} = 120$). In particular, it led to satisfactorily good results for small diagnosis windows (early diagnosis) when the assignation criterion was C_1 (see Figures 3c, 3d, 3e and 3f). Concerning C_2, SPE-FR also accomplished a rather correct early fault diagnosis being only outperformed by CI-FR (see Figures 4c, 4d, 4e and 4f). However, as the diagnosis window size increased (late diagnosis), its performance worsened.
   For the pasteurisation process dataset SPE-FR exhibited satisfactory results for small W_d regardless of the assignation criterion (see Figures 5 and 6);

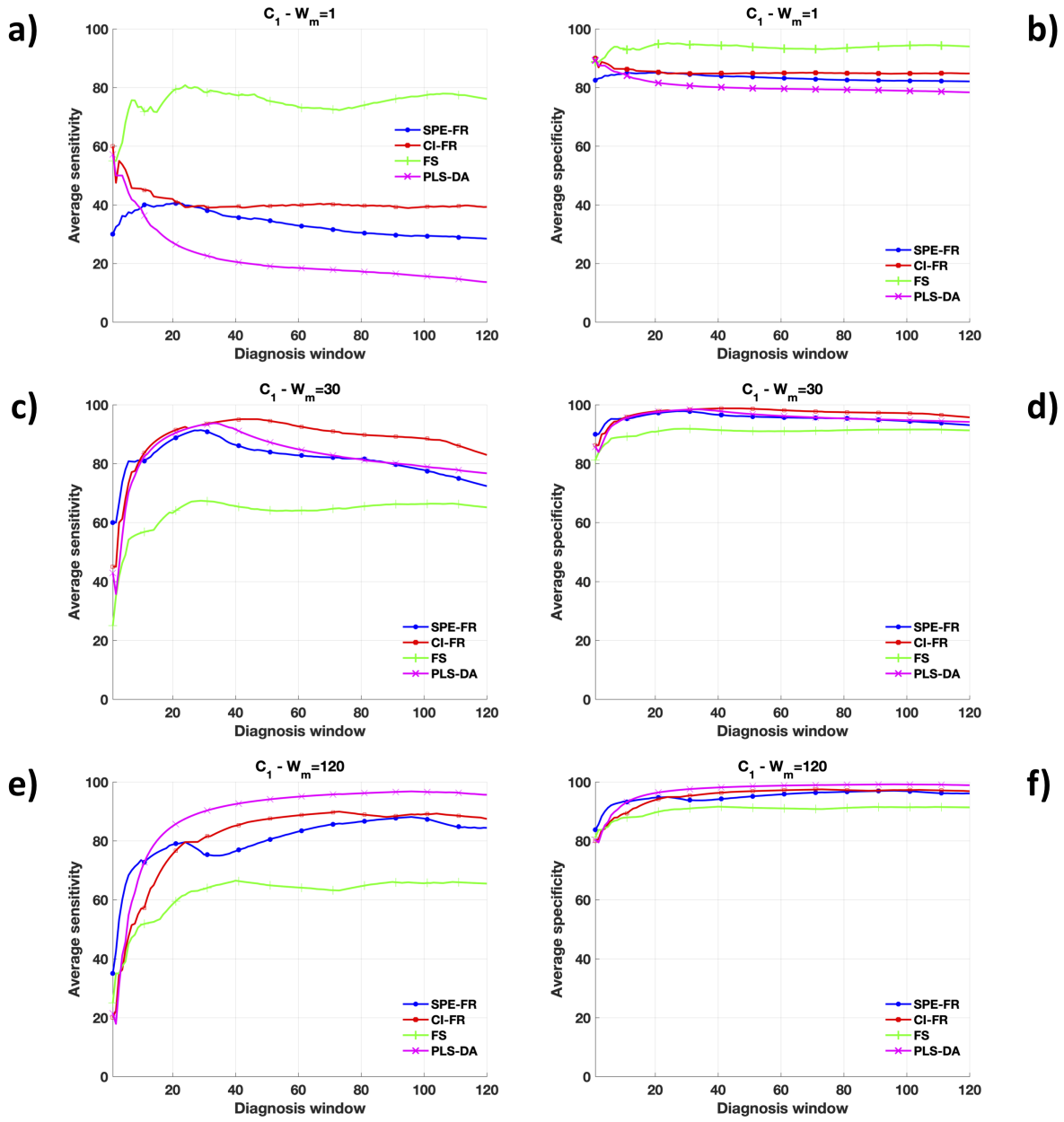

Figure 3 - Distillation process dataset: C_1 assignation criterion. a) Average sensitivity trend for $W_m = 1$; b) Average specificity trend for $W_m = 1$; c) Average sensitivity trend for $W_m = 30$; d) Average specificity trend for $W_m = 30$; e) Average sensitivity trend for $W_m = 120$; f) Average specificity trend for $W_m = 120$. To ease visualisation, symbols are represented only for a subset of equidistant time points

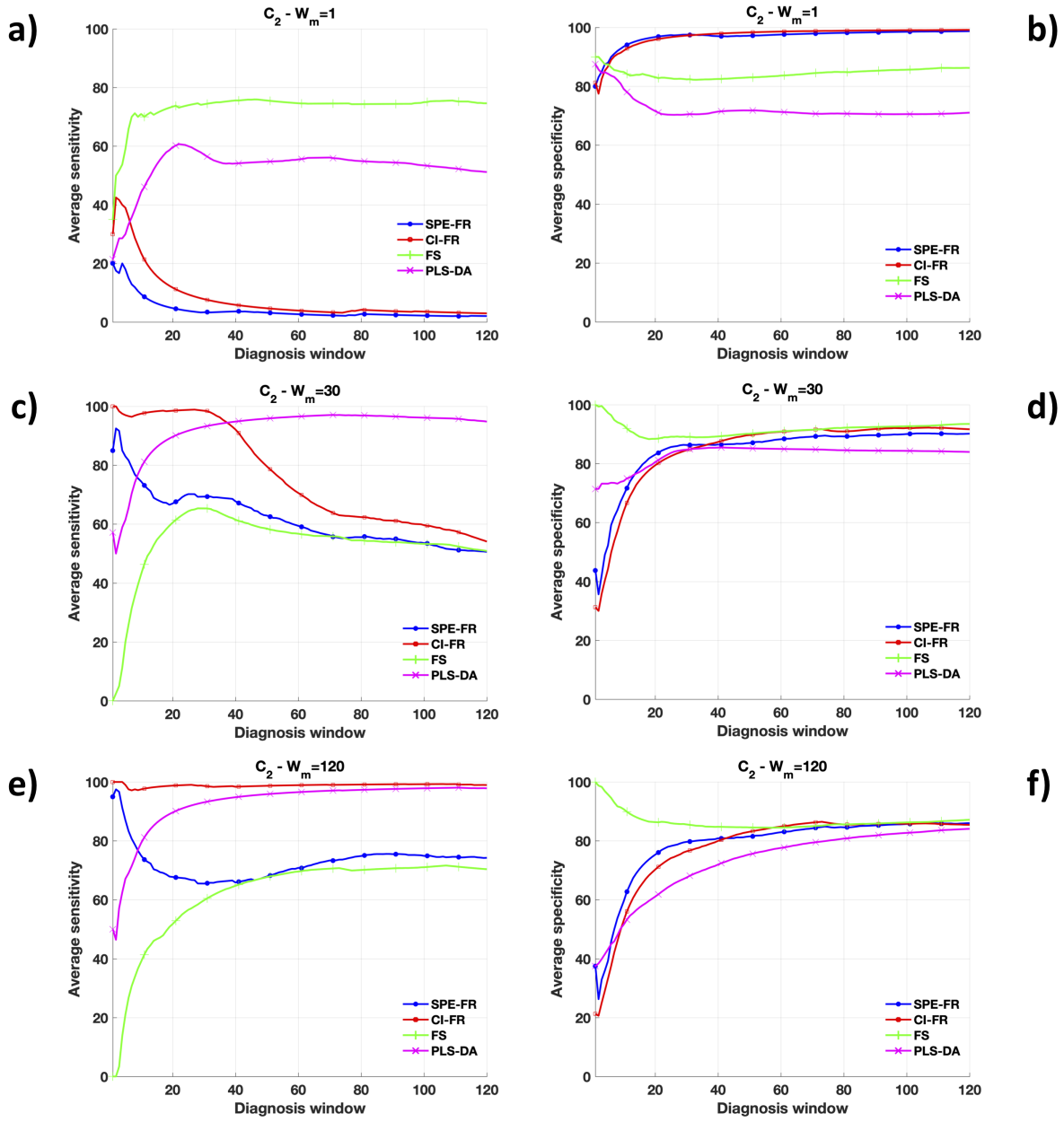

Figure 4 - Distillation process dataset: C_2 assignation criterion. a) Average sensitivity trend for $W_m = 1$; b) Average specificity trend for $W_m = 1$; c) Average sensitivity trend for $W_m = 30$; d) Average specificity trend for $W_m = 30$; e) Average sensitivity trend for $W_m = 120$; f) Average specificity trend for $W_m = 120$. To ease visualisation, symbols are represented only for a subset of equidistant time points

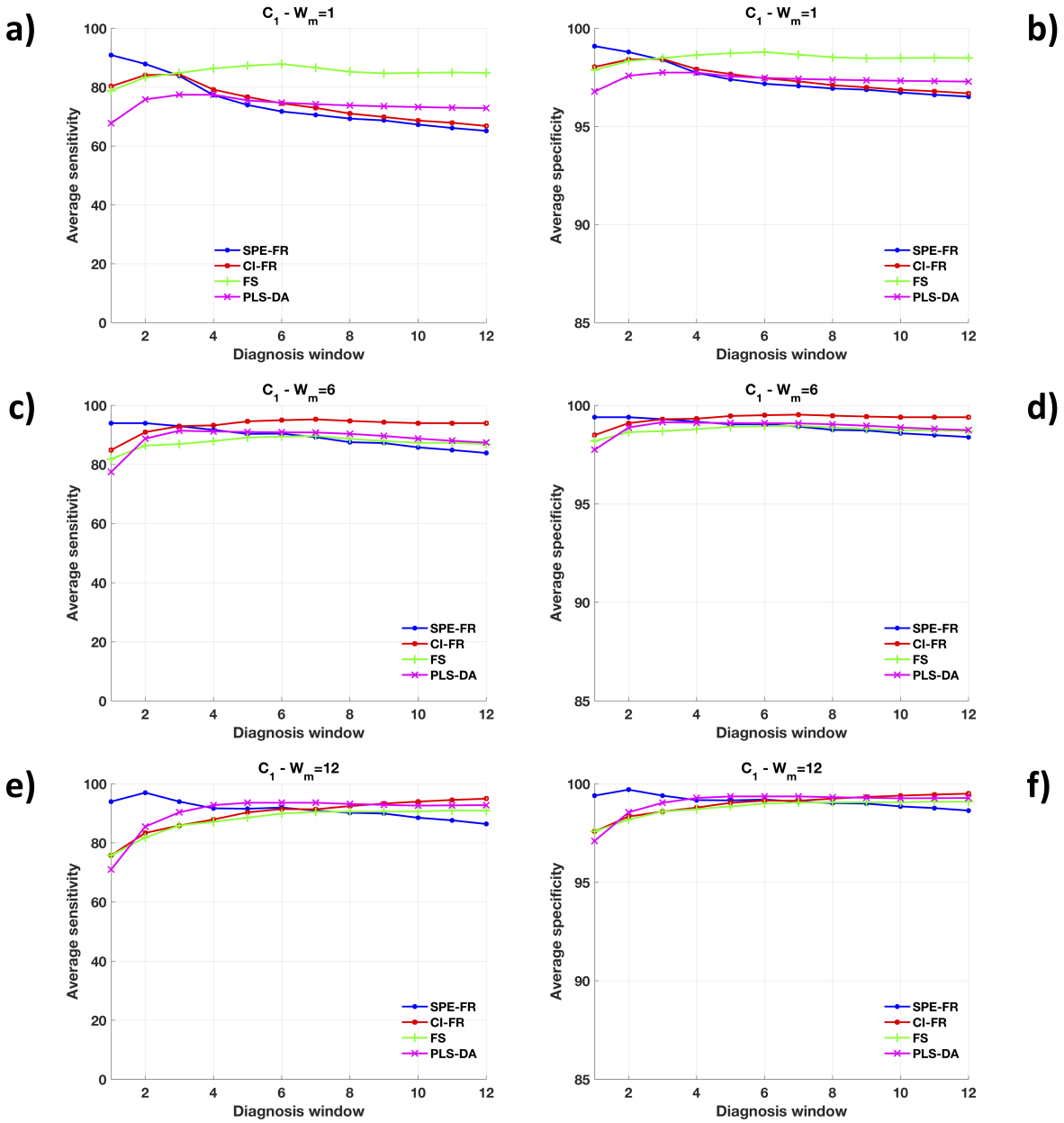

Figure 5 - Pasteurisation process dataset: C_1 assignation criterion. a) Average sensitivity trend for $W_m = 1$; b) Average specificity trend for $W_m = 1$; c) Average sensitivity trend for $W_m = 6$; d) Average specificity trend for $W_m = 6$; e) Average sensitivity trend for $W_m = 12$; f) Average specificity trend for $W_m = 12$

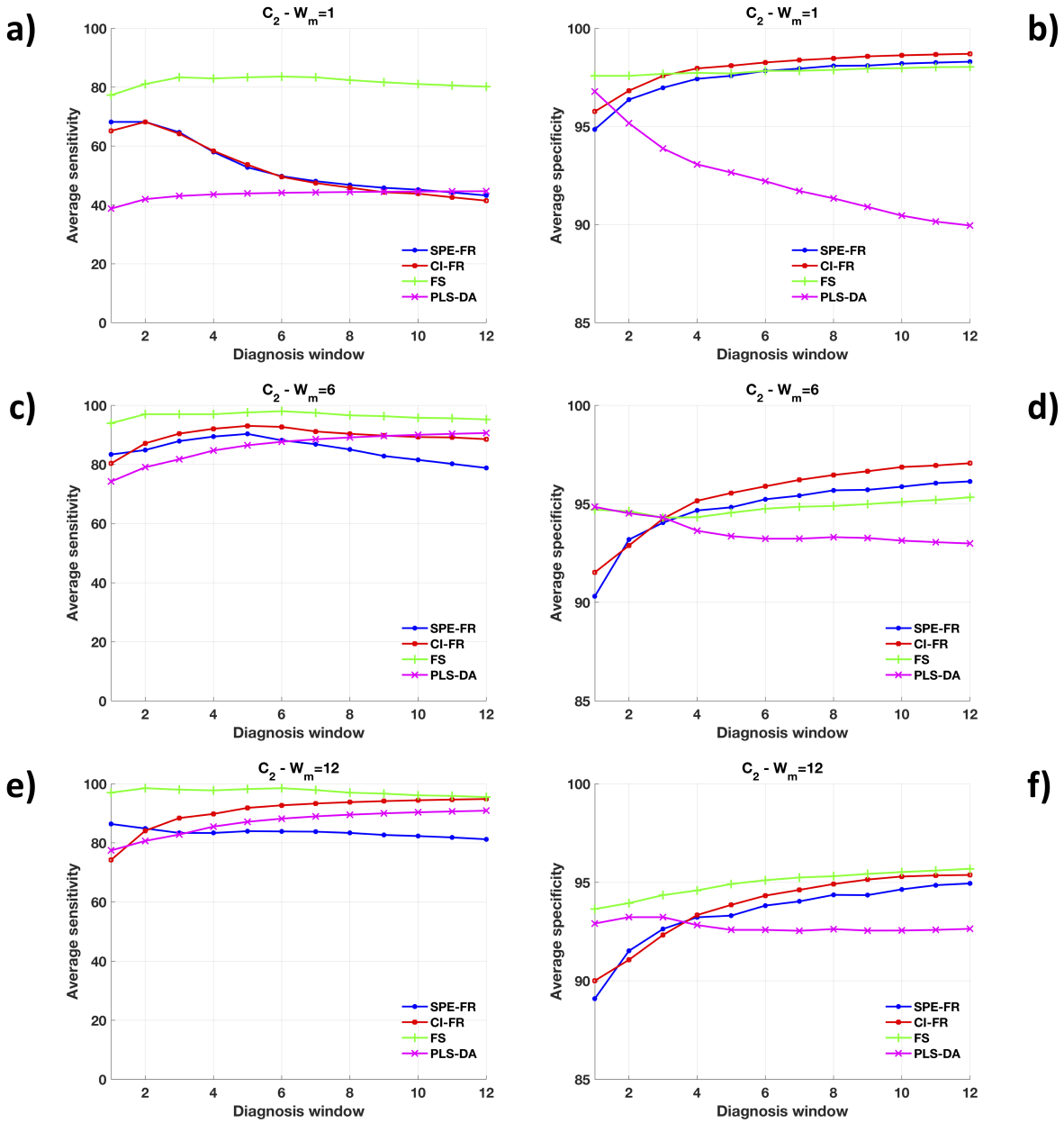

Figure 6 - Pasteurisation process dataset: C_2 assignation criterion. a) Average sensitivity trend for $W_m = 1$; b) Average specificity trend for $W_m = 1$; c) Average sensitivity trend for $W_m = 6$; d) Average specificity trend for $W_m = 6$; e) Average sensitivity trend for $W_m = 12$; f) Average specificity trend for $W_m = 12$

b) CI-FR: CI-FR returned the most satisfactory outcomes for medium-length diagnosis windows. For the distillation process dataset it clearly outperformed the other techniques in terms of sensitivity in early diagnosis for $W_m = 30$ and $W_m = 120$ when C_2 was chosen as assignation criterion (see Figures 4c, 4e). Nevertheless, such a high sensitivity is counterbalanced by low specificity values, which represents a clear symptom of fault identification issues (see Figures 4d, 4f). As the size of the diagnosis window increased, also its performance generally worsened but more gradually than for SPE-FR, highlighting CI-FR may guarantee a better late diagnosis than SPE-FR.

   Regarding the pasteurisation process dataset CI-FR exhibited satisfactory late diagnosis results for the C_2 assignation criterion when $W_m = 6$ and $W_m = 12$ (see Figures 6c, 6d, 6e and 6f);

c) FS: FS clearly outmatched the others considered approaches when very small model windows were concerned ($W_m = 1$, see Figures 3a and 3b, 4a and 4b, 5a and 5b, 6a and 6b). However, this does not mean that the best early fault diagnosis was achieved by this method. In fact, the other strategies guaranteed on average better early diagnosis results when wider model windows were resorted to. Anyway, it should be taken into account that FS was originally developed for cases in which $W_m = 1$ while here, for properly addressing the comparative study, it was adapted for dealing also with larger model windows.

   For the pasteurisation process dataset, FS exhibited the best results among all the compared diagnosis strategies for the C_2 assignation criterion regardless of W_m (see Figure 6). This is a consequence of the smaller W_m values set in this specific circumstance which make the model window size influencing less the diagnosis performance.

d) PLS-DA: PLS-DA generally showed one of the best late diagnosis performance among the compared approaches. It was found to provide a highly correct fault identification (especially for the distillation process dataset) when large model and diagnosis windows were dealt with (see e.g Figures 3e and 3f, 4c and 4d, 4e and 4f). Nevertheless, PLS-DA proved to suffer from severe limitations when handling small model windows (Building a classification/discrimination model with very few members per category is, in fact, not a reliable practice) (see Figures 3a and 3b, 4a and 4b, 5a and 5b, 6a and 6b).

4.2. Diagnosis of new types of faults

Another important aspect of supervised diagnosis techniques is their ability of identifying types of faults not included in the original reference library. For its evaluation, each data block associated to a specific fault was iteratively left out of the calibration set and the nature of the corresponding test observations subsequently assessed by applying the C_2 assignation criterion (C_1, which forces the test observations to be identified as affected by a unique reference fault, is not appropriate for this assessment). For each one of such observations the diagnosis was assumed to be successful if a) none of the reference failures led to a reconstructed SPE value below the respective significance threshold (for SPE-FR), b) none of the reference failures led to a combined index value below the respective significance threshold (for CI-FR), c) none of the reference failures led to cosine values of $\alpha_{\hat{\mathbf{x}}^T_{new,norm}}$ and $\alpha_{\tilde{\mathbf{x}}^T_{new,norm}}$ beyond the respective significance threshold (for FS) or d) its SPE value was found to be beyond the 99% confidence limit of this statistic (for PLS-DA). Also here, to get an idea of the effect of the model window size and of the early and late fault diagnosis performance of the compared methodologies the study was conducted on the distillation process

dataset for $W_m = 1$, $W_m = 30$ and $W_m = 120$ and for $W_d = 30$ (early diagnosis) and $W_d = 120$ (late diagnosis) as well as on the pasteurisation process dataset for $W_m = 1$, $W_m = 6$ and $W_m = 12$ and for $W_d = 6$ (early diagnosis) and $W_d = 12$ (late diagnosis).

Table 6 - Distillation process dataset: diagnosis of new types of faults. Average correct fault diagnosis rates for $W_d = 30,\ 120$ and $W_m = 1,\ 30,\ 120$

	$W_d = 30$ (early diagnosis)			$W_d = 120$ (late diagnosis)		
	$W_m = 1$	$W_m = 30$	$W_m = 120$	$W_m = 1$	$W_m = 30$	$W_m = 120$
SPE-FR	94.8%	69.7%	52.7%	97.8%	74.3%	61.9%
CI-FR	93.7%	63.5%	45.5%	97.9%	77.7%	60.6%
FS-FR	31.3%	80.7%	75.3%	34.2%	92.2%	82.7%
PLS-DA	42.2%	72.2%	100.0%	56.7%	84.5%	99.8%

Table 7 - Pasteurisation process dataset: diagnosis of new types of faults. Average correct fault diagnosis rates for $W_d = 6,\ 12$ and $W_m = 1,\ 6,\ 12$

	$W_d = 6$ (early diagnosis)			$W_d = 12$ (late diagnosis)		
	$W_m = 1$	$W_m = 6$	$W_m = 12$	$W_m = 1$	$W_m = 6$	$W_m = 12$
SPE-FR	80.3%	62.5%	55.6%	84.3%	67.2%	61.2%
CI-FR	84.6%	69.5%	60.4%	88.0%	75.9%	64.0%
FS-FR	87.9%	75.5%	64.9%	87.4%	75.1%	66.4%
PLS-DA	53.5%	44.9%	44.9%	54.0%	45.4%	46.0%

Tables 6 and 7 list the average correct fault diagnosis rates accomplished by SPE-FR, CI-FR, FS and PLS-DA, respectively.

The statistical significance of the observed differences among them was assessed by an ANalysis Of VAriance (ANOVA) carried out by taking into account three factors (i.e. fault diagnosis method, model window and diagnosis window) and all their possible interactions. Figures 7 and 8 display the 95% Least Significant Difference (LSD) intervals for all of these factors and interactions whose effect was actually detected as statistically significant ($p-$value $\ll 0.05$). They mainly corroborate all the conclusions drawn in Section 4.1. Overall, concerning the distillation process dataset, SPE-FR, CI-FR and PLS-DA outmatched FS (see Figure 7a), while, regarding the pasteurisation process dataset, SPE-FR, CI-FR and FS clearly performed better than PLS-DA probably due to the lower number of considered training observations per fault (see Figure 8a). Large model windows enabled the best fault diagnosis in the former case-study (see Figure 7b), while for the latter small model windows returned the best outcomes (see Figure 8b). Large diagnosis windows were found to be optimal in both the explored scenarios (see Figure 7c and 8c). Here in addition, it is interesting to notice how the model window size affected the ability of the compared approaches of correctly identifying new types of faults (see Figures 7d and 8d). SPE-FR and CI-FR were found to behave rather similarly, being their early and late diagnosis performance higher when small

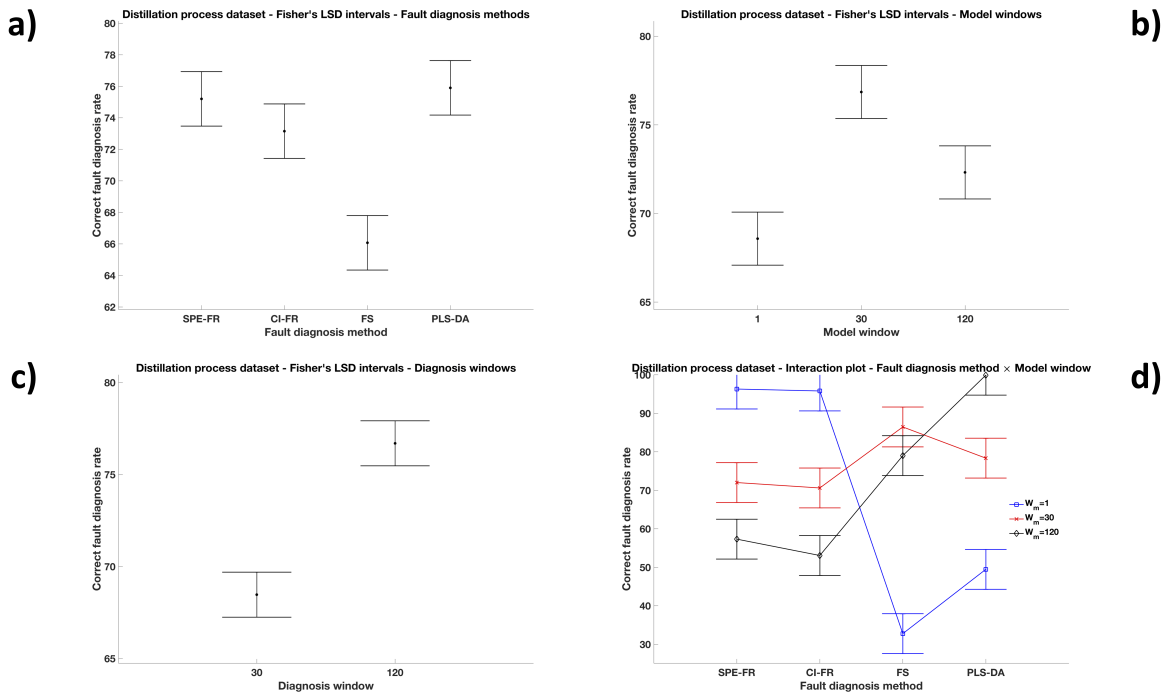

Figure 7 - Distillation process dataset: ANOVA 95% Least Significant Difference (LSD) intervals for mean correct fault diagnosis rate a) fault diagnosis method, b) model window, c) diagnosis window, and d) interaction fault diagnosis method/model window. In d), the theoretical LSD intervals are cut at a value of correct fault diagnosis rate equal to 100%

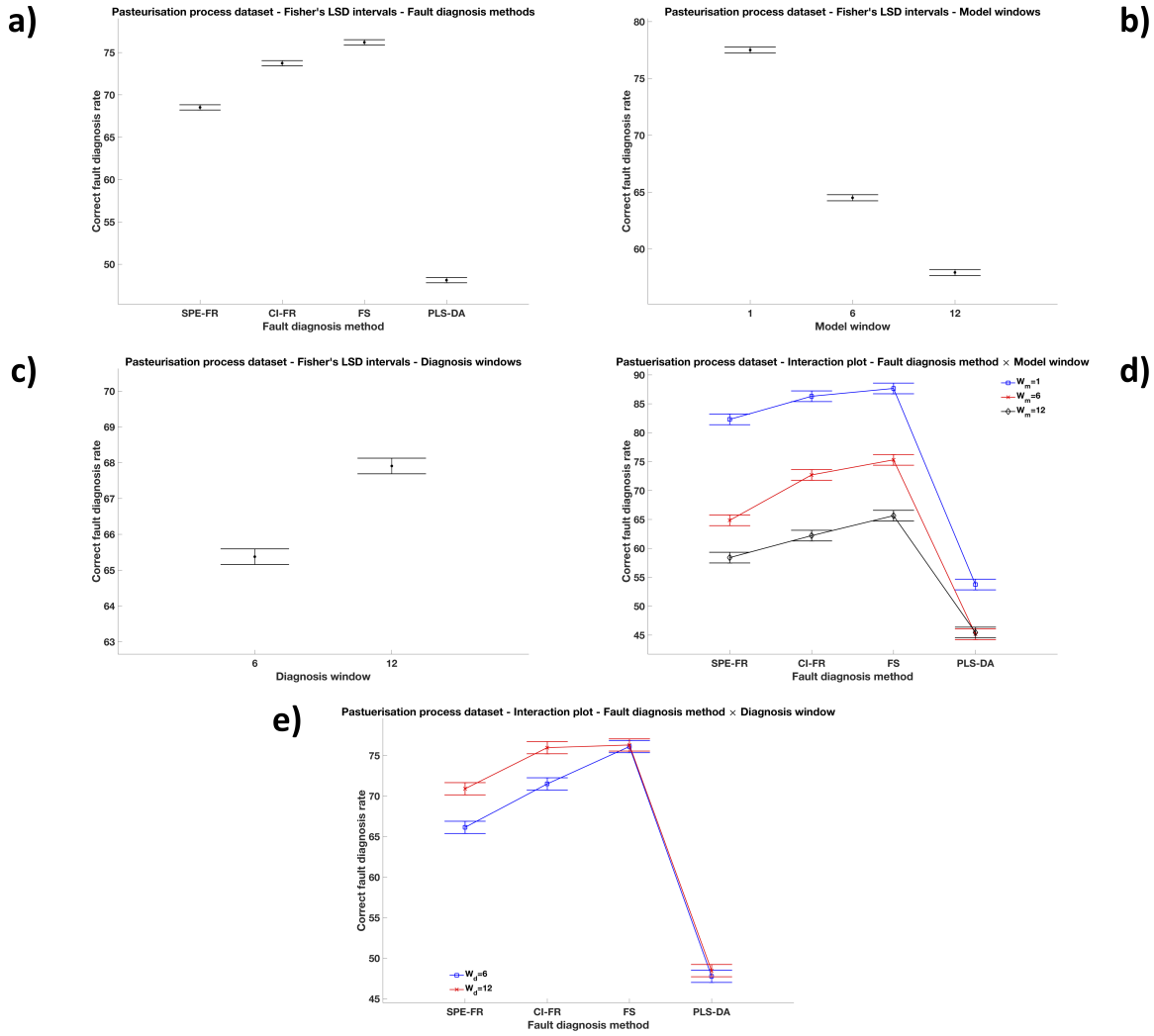

Figure 8 - Pasteurisation process dataset: ANOVA 95% Least Significant Difference (LSD) intervals for mean correct fault diagnosis rate a) fault diagnosis method, b) model window, c) diagnosis window, d) interaction fault diagnosis method/model window, and e) interaction fault diagnosis method/diagnosis window

model windows were concerned. As expected, for the pasteurisation process dataset, FS returned the best results when $W_m = 1$. This is not the case for the distillation process dataset, probably due to the different dynamics of the simulated failures (they reach their steady state after several time instants). Again for the distillation process dataset, PLS-DA yielded its best correct fault diagnosis rate when the largest model window ($W_{\mathrm{m}} = 120$) was dealt with. On the other hand, less notable changes in the PLS-DA performance were observed for the pasteurisation process dataset as W_{m} varied, probably owing to the smaller differences in the size of the various model windows.

5. Conclusions

SPE-FR, CI-FR, FS and PLS-DA showed substantially different performance depending on the size of the model window and on whether they were applied for early or late fault diagnosis. Notice that class-modelling approaches like SIMCA were a priori excluded from the comparison because their use is not recommended in purely discrimination scenarios (i.e. when the C_1 assignation criterion is exploited) [35]. FS generally yielded better results when W_{m} was small. SPE-FR, CI-FR and PLS-DA required instead larger sets of training observations to enable a correct fault diagnosis. However, as pointed out in Section 4.1, the use of small/large model windows not necessarily guarantees a better early/late diagnosis performance. All these outcomes suggest that a hybrid strategy based on the employment of a method more suitable for early diagnosis combined with another more suitable for late diagnosis can represent a feasible option in complex case-studies.

An interesting aspect is that process dynamics seems to be correlated to W_{m}. In fact, the highest fault diagnosis sensitivity and specificity rate were on average achieved for low W_{m} when the pasteurisation dataset was handled and for large W_{m} in the distillation case-study. And this is not surprising if it is considered that the dynamics of the pasteurisation process is much faster than the dynamics of the distillation one, i.e. the fault steady-state is reached in a shorter time for the first system than for the second. In other words, it can be said that process dynamics can be taken into account in an indirect way by properly tuning the W_{m} value according to the specific scenario at hand.

Regarding the ability of the four techniques of diagnosing new types of failures not included in the original reference library, SPE-FR, CI-FR and FS led to better diagnosis results for relatively small model windows. Conversely, PLS-DA exhibited its best performance when W_{m} was large (at least for the distillation process dataset).

The intrinsic nature of the different types of faults was also found to have a considerable effect on the diagnosis power of the compared approaches, even though reporting average trends constitutes an easy and direct way to give a global idea of how such approaches may behave under various conditions.

The presented study additionally highlighted sensitivity and specificity may constitute efficient indices to assess the potential of fault diagnosis methodologies when lots of distinct failures need to be discriminated.

6. References

[1] T. Kourti, J. MacGregor, Process analysis, monitoring and diagnosis, using multivariate projection methods, Chemometr. Intell. Lab. 28 (1995) 3–21.

[2] T. Kourti, Comprehensive Chemometrics, 1st Edition, Vol. 4, Elsevier B.V., Oxford, UK, 2009, Ch. Multivariate statistical process control and process control, using latent variables, pp. 21–54.

[3] A. Ferrer, Comprehensive Chemometrics, 1st Edition, Vol. 1, Elsevier B.V., Oxford, UK, 2009, Ch. Statistical control of measures and processes, pp. 97–126.

[4] A. Ferrer, Multivariate Statistical Process Control based on Principal Component Analysis (MSPC-PCA): some reflections and a case study in an autobody assembly process, Qual. Eng. 19 (2007) 311–325.

[5] P. Van den Kerkhof, G. Gins, J. Vanlaer, J. Van Impe, Dynamic model-based fault diagnosis for (bio)chemical batch processes, Comput. Chem. Eng. 40 (2012) 12–21.

[6] P. Van den Kerkhof, J. Vanlaer, G. Gins, J. Van Impe, Analysis of smearing-out in contribution plot based fault isolation for Statistical Process Control, Chem. Eng. Sci. 104 (2013) 285–293.

[7] M. Bauer, N. Thornhill, A practical method for identifying the propagation path of plant-wide disturbances, J. Process Contr. 18 (2008) 707–719.

[8] Y. Shu, J. Zhao, Data-driven causal inference based on a modified transfer entropy, Comput. Chem. Eng. 57 (2013) 173–180.

[9] T. Rato, M. Reis, On-line process monitoring using local measures of association. Part II: design issues and fault diagnosis, Chemometr. Intell. Lab. 142 (2015) 265–275.

[10] T. Rato, M. Reis, Markovian and non-Markovian sensitivity enhancing transformations for process monitoring, Chem. Eng. Sci. 163 (2017) 223–233.

[11] T. Kourti, J. MacGregor, Multivariate SPC methods for process and product monitoring, J. Qual. Technol. 28 (1996) 409–428.

[12] P. Miller, R. Swanson, C. Heckler, Contribution plots: a missing link in multivariate quality control, Appl. Math. Comput. Sci. 8 (1998) 775–792.

[13] C. Alcala, S. Qin, Reconstruction-based contribution for process monitoring, Automatica 45 (2009) 1593–1600.

[14] R. Dunia, S. Qin, Subspace approach to multidimensional fault identification and reconstruction, AIChE J. 44 (1998) 1813–1831.

[15] H. Yue, S. Qin, Reconstruction-based fault identification using a combined index, Ind. Eng. Chem. Res. 40 (2001) 4403–4414.

[16] S. Yoon, J. MacGregor, Fault diagnosis with multivariate statistical models part I: using steady state fault signatures, J. Process Contr. 11 (2001) 387–400.

[17] S. Wold, C. Albano, W. Dunn III, K. Esbensen, E. Hellberg, E. Johansson, M. Sjöström, Food research and data analysis, 1st Edition, Elsevier Applied Science, London, UK, 1983, Ch. Pattern recognition: finding and using regularities in multivariate data, pp. 147–188.

[18] M. Barker, W. Rayens, Partial Least Squares for discrimination, J. Chemometr. 17 (2003) 166–173.

[19] S. Wold, Pattern recognition by means of disjoint principal components models, Pattern Recogn. 8 (1976) 127–139.

[20] S. Wold, M. Sjöström, SIMCA: a method for analyzing chemical data in terms of similarity and analogy, in: Chemometrics: Theory and Application First Edition, Vol. 52, American Chemical Society, Washington D.C., USA, 1977, pp. 243–282.

[21] S. Qin, Survey on data-driven industrial process monitoring and diagnosis, Annu. Rev. Control 36 (2012) 220–234.

[22] E. Russell, L. Chiang, R. Braatz, Data-driven methods for fault detection and diagnosis in chemical processes, 1st Edition, Springer-Verlag London Ltd, London, UK, 2012.

[23] J. MacGregor, A. Cinar, Monitoring, fault diagnosis, fault-tolerant control and optimization: data driven methods, Comput. Chem. Eng. 47 (2012) 111–120.

[24] T. Rato, M. Reis, Non-causal data-driven monitoring of the process correlation structure: a comparison study with new methods, Comput. Chem. Eng. 71 (2014) 307–322.

[25] J. Jackson, G. Mudholkar, Control procedures for residuals associated with Principal Component Analysis, Technometrics 21 (1979) 341–349.

[26] G. Box, Some theorems on quadratic forms applied in the study of analysis of variance problems: effect of inequality of variance in one-way classification, Ann. Math. Stat. 25 (1954) 290–302.
[27] N. Tracy, J. Young, R. Mason, Multivariate control charts for individual observations, J. Qual. Technol. 24 (1992) 88–95.
[28] S. Valle-Cervantes, S. Qin, M. Piovoso, M. Bachmann, N. Mandakoro, Extracting fault subspaces for fault identification of a polyester film process, in: Proc. ACC, IEEE, Arlington, VA, 2001, pp. 4466–4471.
[29] P. Geladi, B. Kowalski, Partial Least Squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[30] H. Martens, T. Næs, Multivariate Calibration, 1st Edition, John Wiley & Sons Ltd., 1989.
[31] R. Fisher, The use of multiple measurements in taxonomic problems, Ann. Eugenic. 7 (1936) 179–188.
[32] S. Skogestad, M. Morari, Understanding the dynamic behavior of distillation columns, Ind. Eng. Chem. Res. 27 (1988) 1848–1862.
[33] P. Villalba, J. Sanchis, A. Ferrer, A graphical user interface for PCA-based MSPC: a benchmark software for multivariate statistical process control in MATLAB, Chemometr. Intell. Lab. 185 (2019) 135–152.
[34] N. Pérez, J. Ferré, R. Boqué, Calculation of the reliability of classification in discriminant Partial Least-Squares classification, Chemometr. Intell. Lab. 95 (2009) 122–128.
[35] O. Rodionova, P. Oliveri, A. Pomerantsev, Rigorous and compliant approaches to one-class classification, Chemometr. Intell. Lab. 159 (2016) 89–96.

Data-driven supervised fault diagnosis methods based on latent variable models: a comparative study

Supporting Material

The following two sections provide additional details about the processes considered in this comparative study.

Distillation process

The Simulink dynamic non-linear model used to generate the simulated dataset is based on [1] and relates to a methanol/ethanol distillation column, comprehensively described in [2]. This model encompasses four controlled variables (distillate composition, composition at the bottom of the column and liquid holdups in the reflux drum and reboiler), four manipulated variables (volumetric flow rates at the top and at the bottom of the column and internal - boil-up and reflux - flow rates at the top and at the bottom of the column) and three disturbance sources (feed composition, feed volumetric flow rate and feed temperature). The distillation column is stabilised by closing two decentralised (SISO) control loops: the first regulates the distillate composition by the top internal flow, while the second regulates the bottom composition by the bottom internal flow. The model features 41 process stages associated to the column reboiler, the column condenser and 39 column trays.

Pasteurisation process

The pasteurisation process is run in an Armfield PCT23 MKII bench mounted plant trainer characterised by i) a miniaturised three-stage plate heat exchanger including a preheating unit and fed by a hot water circulator (maximum heating power of 1.92 kW), ii) two independent feed tanks, iii) a holding tube coated with thermal insulated material and featuring a divert valve, and iv) two variable-speed peristaltic pumps. The equipment incorporates an electrical fault simulation and control console.
The instrumentation is able to keep the product stream, flowing through the holding tube, at high temperature (controlled by the hot water circulation) for a certain time period for bacteriological purposes. The process encompasses three control loops, regulating the water temperature, the temperature at the outlet of the holding tube and the product flow.

References

[1] S. Skogestad, M. Morari, Understanding the dynamic behavior of distillation columns, Ind. Eng. Chem. Res. 27 (1988) 1848–1862.

[2] P. Villalba, J. Sanchis, A. Ferrer, A graphical user interface for PCA-based MSPC: a benchmark software for multivariate statistical process control in MATLAB, Chemometr. Intell. Lab. 185 (2019) 135–152.