GEOMETRIC PROPERTIES OF KERNEL PARTIAL LEAST SQUARES FOR NON-LINEAR PROCESS MONITORING

José L. Godoy[(a)], Germán Bustos[(b)], Alejandro H. Gonzalez[(c)] and Jacinto L. Marchetti[(d)]

[(a,b,c,d)] INTEC (CONICET and Universidad Nacional del Litoral), Güemes 3450 (3000), Santa Fe, Argentina

[(a)]jlgodoy, [(b)]gbustos, [(c)]alejgon, [(d)]jlmarch@santafe-conicet.gov.ar

ABSTRACT

This work proposes a new strategy for monitoring non-linear processes based on Kernel Partial Least Squares (KPLS). When strongly non-linear process are considered, a PLS regression model could not be enough accurate. So, the first stage of the proposed method is to map the input data to a high-dimension space, where a linear regression model can be obtained. Then, an *implicit* linear regression model relating the high-dimension space with output space (output data) is obtained. This model implicitly induces a decomposition of the high-dimension space into the Model subspace and the complementary Residual subspace, being the vectors in the first subspace the effective domain of the linear regression model. Finally, once the space decomposition is understood, new statistics (metrics) for each subspace are proposed to monitor the process and detect possible abnormal behaviours. The effectiveness of the method is tested by means of a synthetic simulation example from the literature.

Keywords: KPLS, space decomposition, multivariate process monitoring, fault detection indexes.

1. INTRODUCTION

KPLS is a promising regression method for deal with nonlinear problems because it can efficiently compute regression coefficients in high-dimensional space by means of the nonlinear kernel function. It is also an efficient method for estimating and predicting quality variables in the strongly nonlinear process by mapping data from the original space into a high-dimensional space. It only requires the use of linear algebra, making it as simple as linear multivariate projection methods, and it can handle a wide range of nonlinearities because of its ability to use different kernel functions. Its application results from a simple example show that the proposed method can effectively capture the nonlinear relationship among variables.

The need for associating input and output data sets obtained by online data login of complex process variables constitutes a problem that requires increasing attention. Lately, KPLS has become a powerful approach to find multivariable non-linear structures, mainly because large non-linear correlations implicit in the data can be adequately overcome.

Fault detection makes use of the so-called fault detection indices based on model. A fault is detected when one of the fault detection indices is beyond its control limit. Once a fault is detected, it is necessary to diagnose its cause.

In this work, the implicit space decomposition (made by KPLS) is obtained and their main geometric properties are used to design a non-linear monitoring strategy using new fault detection indices actuating on each subspace. This allows us to diagnose the fault type. The main contributions of this work are then the introduction of new fault detection indices derived from the KPLS decomposition, and a diagnosis tool based on the statistics that triggers the alarm condition.

2. KERNEL PLS REGRESSION

As usual in data driven method, we first should collect a set of N samples of the predictor vector, $\{\mathbf{x}_i\}_{i=1}^{N}$, with $\mathbf{x}_i \in \Re^m$ for $i,\cdots,N$, and N samples of the response vector, $\{\mathbf{y}_i\}_{i=1}^{N}$, with $\mathbf{y}_i \in \Re^p$ for $i,\cdots,N$, which are called Identification Data Set (IDS). Once these data are collected, the Kernel Partial Least Squares (KPLS) method maps the predictor vectors, $\mathbf{x}_i$, from $\Re^m$ to a high-dimension space $\Re^c$ (with $c \gg m$), where a linear PLS regression model can be created to relate vectors in $\Re^c$ with the response vectors, $\mathbf{y}_i$, in $\Re^p$. In this way, a latent model can be formulated in $\Re^c$, in order to extend linear PLS to non-linear kernel PLS (Rosipal and Trejo, 2001). The non-linear transformation that maps vectors from $\Re^m$ to $\Re^c$ is not made by means of an explicit non-linear function $\boldsymbol{\varphi}$, but by means of a kernel function k selected to compute the following inner products:

$$k\left(\mathbf{x}_i,\mathbf{x}_j\right)=\boldsymbol{\varphi}\left(\mathbf{x}_i\right)'\boldsymbol{\varphi}\left(\mathbf{x}_j\right), \quad \text{for } i=1,\cdots,N, \text{ and } j=1,\cdots,N. \tag{1}$$

Notice that trough the introduction of the kernel trick (Rosipal and Trejo, 2001), $\boldsymbol{\varphi}\left(\mathbf{x}_i\right)'\boldsymbol{\varphi}\left(\mathbf{x}_j\right)=k\left(\mathbf{x}_i,\mathbf{x}_j\right)$, one can avoid both, performing explicit nonlinear mapping and computing dot products in the high-dimensional space. Furthermore, as it is known and will be seen later, the PLS regression method need only dot products to perform the regression. Now, from Eq. (1) it is obtained the *Gram Kernel* matrix, $\mathbf{K} \in \Re^{N\times N}$, defined as follows:

$$\mathbf{K}=\boldsymbol{\Phi}\boldsymbol{\Phi}', \quad \boldsymbol{\Phi}=\left[\boldsymbol{\varphi}\left(\mathbf{x}_1\right)\cdots\boldsymbol{\varphi}\left(\mathbf{x}_N\right)\right]'. \tag{2}$$

The entries of this last matrix are the cross inner products of all mapped predictor vectors, $\{\boldsymbol{\varphi}(\mathbf{x}_i)\}_{i=1}^{N}$.

As in the PLS method, it is assumed here a non-linear KPLS model with zero mean. To center the mapped data in the high-dimensional space, $\Re^c$, matrix $\mathbf{K}$ should be substituted by the centered matrix $\bar{\mathbf{K}}$, given by:

$$\bar{\mathbf{K}} = \bar{\boldsymbol{\Phi}}\bar{\boldsymbol{\Phi}}' = \mathbf{K} - \mathbf{KE} - \mathbf{EK} + \mathbf{EKE}, \quad (3)$$

where $\mathbf{E}$ is a matrix with unitary entries, and $\bar{\mathbf{K}}$ and $\bar{\boldsymbol{\Phi}}$ are the centered versions of $\mathbf{K}$ and $\boldsymbol{\Phi}$.

Now, from matrices $\bar{\mathbf{K}}$ and $\mathbf{Y}=[\mathbf{y}_1,\cdots,\mathbf{y}_N]' \in \Re^{N\times p}$, the training KPLS algorithm is derived as a sequence of modified NIPALS steps, as follows (Rosipal and Trejo, 2001):

First, set the index a=1, $\bar{\mathbf{K}}_1 = \bar{\mathbf{K}}$ and $\mathbf{Y}_1 = \mathbf{Y}$. Then:

1. Set the vector of $\mathbf{Y}_a$-scores $\mathbf{u}_a \in \Re^N$ as the maximum-variance column of $\mathbf{Y}_a$.
2. Calculate the vector of $\bar{\boldsymbol{\Phi}}_a$-scores $\mathbf{t}_a \in \Re^N$ as: $\mathbf{t}_a = \bar{\mathbf{K}}_a\mathbf{u}_a, \quad \mathbf{t}_a \leftarrow \mathbf{t}_a / \|\mathbf{t}_a\|$
3. Regress columns of $\mathbf{Y}_a$ on $\mathbf{t}_a$: $\mathbf{c}_a = \mathbf{Y}_a'\mathbf{t}_a$, where $\mathbf{c}_a$ is a weight vector.
4. Calculate the new score vector $\mathbf{u}_a$ for $\mathbf{Y}_a$ as: $\mathbf{u}_a = \mathbf{Y}_a\mathbf{c}_a, \quad \mathbf{u}_a \leftarrow \mathbf{u}_a / \|\mathbf{u}_a\|$
5. Repeat steps 2-3 until convergence of $\mathbf{t}_a$.
6. Deflate the matrices: $\bar{\mathbf{K}}_{a+1} \leftarrow (\mathbf{I} - \mathbf{t}_a\mathbf{t}_a')\bar{\mathbf{K}}_a(\mathbf{I} - \mathbf{t}_a\mathbf{t}_a'),$ $\mathbf{Y}_{a+1} \leftarrow \mathbf{Y}_a - \mathbf{t}_a\mathbf{t}_a'\mathbf{Y}_a,$
7. Set a=a+1 and return to step 1. Stop when a>A, where A is the number of latent variables selected in the high-dimensional space, $\Re^c$.

Notice that the selection of A is determined by supervising the deflation of $\mathbf{Y}_a$. Again, as in linear PLS, the prediction on training data has the following from (Rosipal and Trejo, 2001):

$$\begin{aligned}\hat{\mathbf{Y}} = \bar{\boldsymbol{\Phi}}\mathbf{B}_{PLS} &= \bar{\boldsymbol{\Phi}}\bar{\boldsymbol{\Phi}}'\mathbf{U}\left(\mathbf{T}'\bar{\mathbf{K}}\mathbf{U}\right)^{-1}\mathbf{T}'\mathbf{Y} \\ &= \bar{\mathbf{K}}\mathbf{U}\left(\mathbf{T}'\bar{\mathbf{K}}\mathbf{U}\right)^{-1}\mathbf{T}'\mathbf{Y} = \mathbf{T}\mathbf{T}'\mathbf{Y} = \mathbf{T}\mathbf{C}',\end{aligned} \quad (4)$$

where the matrices $\mathbf{T}=[\mathbf{t}_1...\mathbf{t}_A]$, $\mathbf{U}=[\mathbf{u}_1...\mathbf{u}_A]$ and $\mathbf{C}=[\mathbf{c}_1...\mathbf{c}_A]$ are orthogonal by columns. Notice that the regression coefficient matrix $\mathbf{B}_{PLS}$ exists but is never computed by the KPLS algorithm, since the kernel substitution avoids the necessity of an explicit computation.

Equality (4) shows that the output response can be obtained from the inner products of the mapped predictor vectors. So, for a new observation $\mathbf{x}$ of the predictor vector, the response vector will be given by:

$$\begin{aligned}\hat{\mathbf{y}}' = \bar{\boldsymbol{\varphi}}(\mathbf{x})'\mathbf{B}_{PLS} &= \bar{\mathbf{k}}'\mathbf{U}\left(\mathbf{T}'\bar{\mathbf{K}}\mathbf{U}\right)^{-1}\mathbf{T}'\mathbf{Y} \\ &= \bar{\mathbf{k}}'\mathbf{R}\mathbf{C}' = \bar{\mathbf{k}}'\mathbf{M},\end{aligned} \quad (5)$$

where $\bar{\mathbf{k}}$ is the vector of centered kernel functions evaluated in the pairs ($\mathbf{x}$, $\mathbf{x}_j$) with j=1,…,N. From Eq. (3), this vector is given by:

$$\bar{\mathbf{k}} = \left[\bar{k}_1(\mathbf{x})\cdots\bar{k}_N(\mathbf{x})\right]' = \mathbf{k} - \mathbf{Ke} - \mathbf{Ek} + \mathbf{EKe}, \quad (6)$$

where $\mathbf{e}$ is an unitary vector and $\mathbf{k} = \left[k_1(\mathbf{x})\cdots k_N(\mathbf{x})\right]' \in \Re^{N\times N}$ is the kernel functions vector, where the element j is given by the kernel function evaluated at ($\mathbf{x}$, $\mathbf{x}_j$), i.e. $k_j(\mathbf{x}) = k(\mathbf{x}_j, \mathbf{x})$. Each of the elements of vector $\bar{\mathbf{k}}$ is computed as follows:

$$\begin{aligned}\bar{k}_j(\mathbf{x}) = \bar{k}(\mathbf{x}_j, \mathbf{x}) = k(\mathbf{x}, \mathbf{x}) &- \frac{2}{N}\sum_{j=1}^{N}k(\mathbf{x}_j, \mathbf{x}) \\ &+ \frac{1}{N^2}\sum_{j=1}^{N}\sum_{n=1}^{N}k(\mathbf{x}_j, \mathbf{x}_n)\end{aligned} \quad (7)$$

Now, we focus our attention on the latent vectors $\mathbf{t} \in \Re^A$. From Eq. (4), it follows that for a new observation $\mathbf{x}$, this vector will be given by:

$$\mathbf{t}' = [t_1\cdots t_A]' = \bar{\mathbf{k}}'\mathbf{R}, \quad (8)$$

where $t_a = \bar{\mathbf{k}}'\mathbf{r}_a$ a=1,....A, with $\mathbf{R} = [\mathbf{r}_1\cdots\mathbf{r}_A]$.

Therefore, for a new observation $\mathbf{x}$, the prediction can be computed from $\mathbf{t}$, as (Eq. (5)):

$$\hat{\mathbf{y}}' = \mathbf{t}'\mathbf{C}'. \quad (9)$$

Now, given a new measurement predictor vector $\mathbf{x}$ in original units, the response vector (also in original units) is predicted by:

$$\hat{\mathbf{y}} = \mathbf{D}_{\mathbf{y}}\mathbf{M}'\bar{\mathbf{k}}\left(\mathbf{D}_{\mathbf{x}}^{-1}(\mathbf{x} - \bar{\mathbf{x}})\right) + \bar{\mathbf{y}}, \quad (10)$$

where the sample standard deviations $\mathbf{D}_{\mathbf{x}} = diag(\hat{\sigma}_{x_1}...\hat{\sigma}_{x_m})$, $\mathbf{D}_{\mathbf{y}} = diag(\hat{\sigma}_{y_1}...\hat{\sigma}_{y_p})$, and the means $\bar{\mathbf{x}}$, $\bar{\mathbf{y}}$, are determined in the training procedure.

3. KPLS-BASED MODELING FOR PROCESS MONITORING

The Kernel PLS algorithm induces an external model which decomposes $\bar{\boldsymbol{\Phi}}$ and $\mathbf{Y}$ in score vectors $\mathbf{t}_a$, weight vectors ($\mathbf{p}_a$ and $\mathbf{c}_a$), and residual error matrices ($\tilde{\bar{\boldsymbol{\Phi}}}$ and $\tilde{\mathbf{Y}}$), as follows:

$$\bar{\boldsymbol{\Phi}} = \mathbf{T}\mathbf{P}' + \tilde{\bar{\boldsymbol{\Phi}}}, \quad (11)$$

$$\mathbf{Y} = \mathbf{T}\mathbf{C}' + \tilde{\mathbf{Y}}, \quad (12)$$

In agreement with the standard linear PLS model (Godoy et al, 2011) it is assumed that the score variables t_a are good predictor of $\mathbf{Y}$. Furthermore, it is also assumed a linear inner relation between the scores of $\mathbf{t}_a$ and $\mathbf{u}_a$, that is,

$$\mathbf{U} = \mathbf{T}\mathbf{B} + \mathbf{H}, \quad (13)$$

where **B** is the $A \times A$ diagonal matrix and **H** denotes the matrix of residuals. Equivalently, we have $\mathbf{U}\mathbf{B}^{-1} - \mathbf{H}\mathbf{B}^{-1} = \mathbf{T}$, where $\mathbf{C}' = \mathbf{B}\mathbf{Q}'$ with **Q** orthonormal by columns. Then, the following regression kernel-based model is obtained:

$$\begin{aligned} \mathbf{Y} &= \bar{\mathbf{K}}\mathbf{R}\mathbf{C}' + \mathbf{H}\mathbf{B}^{-1}\mathbf{C}' + \tilde{\mathbf{Y}}_2 = \hat{\mathbf{Y}} + \tilde{\mathbf{Y}}_1 + \tilde{\mathbf{Y}}_2, \\ \mathbf{Y} &= \mathbf{T}\mathbf{C}' + \tilde{\mathbf{Y}}_1 + \tilde{\mathbf{Y}}_2 \end{aligned} \tag{14}$$

where $\tilde{\mathbf{Y}} = \tilde{\mathbf{Y}}_1 + \tilde{\mathbf{Y}}_2$. We call **V** the pseudo-inverse of $\mathbf{P}'$ ($\mathbf{V}'\mathbf{P} = \mathbf{P}'\mathbf{V} = \mathbf{I}$); then, the prediction of **T** is directly obtained from Eq. (11), as:

$$\mathbf{T} = \bar{\mathbf{\Phi}}\mathbf{V}, \tag{15}$$

because the row space of $\tilde{\bar{\mathbf{\Phi}}}$ belongs to the null-space of the linear transformation $\mathbf{P}'$, i.e. $\tilde{\bar{\mathbf{\Phi}}}\mathbf{V} = \mathbf{0}$ (Godoy et al, 2011). If $\mathbf{P}' = \mathbf{W}_A\mathbf{\Sigma}_A\mathbf{V}_A'$ is the compact singular value decomposition (SVD) of $\mathbf{P}'$, then $\mathbf{V} \triangleq (\mathbf{P}')^- = \mathbf{V}_A\mathbf{\Sigma}_A^{-1}\mathbf{W}_A'$ where $^{-1}$ denote a generalized inverse (Meyer, 2000). Equivalently, from Eq. (14), **D** is the pseudo-inverse of $\mathbf{C}'$ ($\mathbf{C}'\mathbf{D} = \mathbf{D}'\mathbf{C} = \mathbf{I}$) and, since $\tilde{\mathbf{Y}}_2\mathbf{D} = 0$, then: $\mathbf{T} + \mathbf{H}\mathbf{B}^{-1} = \mathbf{Y}\mathbf{D}$.

An interesting point of the propose procedure is that the matrices **V**, **P**, **B** and **Q** will never be estimated (otherwise it would be impractical). They are only defined to develop the proof that follows, in order to find metrics based on kernel substitution trick.

3.1. Underlying KPLS decomposition of the input mapped and output spaces

After synthesizing an *in-control* KPLS model, the measurement vectors $\bar{\boldsymbol{\varphi}}(\mathbf{x}) \in \Re^c$ and $\mathbf{y} \in \Re^p$ are (underlying) decomposed as described below.

First the latent vector is computes as $\bar{\mathbf{k}}'\mathbf{R} = \bar{\boldsymbol{\varphi}}(\mathbf{x})'\mathbf{V} = \mathbf{t}'$ (see Eq. 15), where $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_A]$ is the matrix of PLS components, given by $\mathbf{v}_a = \sum_{j=1}^{N} \alpha_j \bar{\boldsymbol{\varphi}}(\mathbf{x}_j)$ with $\alpha_j \in \Re$. Then, a new mapped vector $\bar{\boldsymbol{\varphi}}(\mathbf{x})$ (associated to the measurements **x**) can be decomposed as:

$$\begin{aligned} \bar{\boldsymbol{\varphi}}(\mathbf{x}) &= \hat{\bar{\boldsymbol{\varphi}}}(\mathbf{x}) + \tilde{\bar{\boldsymbol{\varphi}}}(\mathbf{x}) \quad \in \Re^c \\ \hat{\bar{\boldsymbol{\varphi}}}(\mathbf{x}) &= \mathbf{P}\mathbf{V}'\bar{\boldsymbol{\varphi}}(\mathbf{x}) = \mathbf{P}\mathbf{t} \in W_M \subseteq \Re^c \\ \tilde{\bar{\boldsymbol{\varphi}}}(\mathbf{x}) &= (\mathbf{I} - \mathbf{P}\mathbf{V}')\bar{\boldsymbol{\varphi}}(\mathbf{x}) \in W_R \subseteq \Re^c \end{aligned} \tag{16}$$

where $\mathbf{P} = \bar{\mathbf{\Phi}}'\mathbf{T} = \bar{\mathbf{\Phi}}'\bar{\mathbf{K}}\mathbf{R}$, with **P** and **V** orthogonal by columns. The oblique projections that appear in Eq. (16), decompose the high-dimensional space, $\Re^c$, into two complementary subspaces W_M and W_R, i.e. $W_M \oplus W_R \equiv \Re^c$ (Godoy et al, 2011). Matrix $\mathbf{P}\mathbf{V}'$ is the projector onto the model subspace $W_M \equiv Span\{\mathbf{P}\} \subseteq \Re^c$, along the residual subspace $W_R \equiv Span\{\mathbf{V}\}^{\perp} \subseteq \Re^c$ direction (Godoy et al, 2011), where $^{\perp}$ denotes the orthogonal complement of the subspace (notice that range of **P** is W_M and the null-space is W_R). Hence, mapped predictor space is decomposed in an underlying form, by KPLS, into complementary oblique subspaces.

In addition, if we generalize the results of Godoy et al (2011) for the case KPLS, response space is also decomposed (by KPLS) in complementary oblique subspaces, which is related with the predictor modeled subspace according to:

$$\begin{aligned} \mathbf{y} &= \hat{\mathbf{y}}^* + \tilde{\mathbf{y}}_2 \in \mathbb{R}^p, \\ \hat{\mathbf{y}}^* &= \mathbf{C}\mathbf{D}'\mathbf{y} \in S_{MY} \equiv Span\{\mathbf{C}\}, \\ \tilde{\mathbf{y}}_2 &= (\mathbf{I} - \mathbf{C}\mathbf{D}')\mathbf{y} \in S_{RY} \equiv Span\{\mathbf{D}\}^{\perp} \end{aligned} \tag{17}$$

$$\begin{aligned} \hat{\mathbf{y}}^* &= \hat{\mathbf{y}} + \tilde{\mathbf{y}}_1, \\ \hat{\mathbf{y}} &= \mathbf{C}\mathbf{R}'\bar{\mathbf{k}}(\mathbf{x}) \in S_{MY}, \\ \tilde{\mathbf{y}}_1 &= \hat{\mathbf{y}}^* - \hat{\mathbf{y}} = \mathbf{C}\mathbf{D}'\mathbf{y} - \mathbf{C}\mathbf{R}'\bar{\mathbf{k}}(\mathbf{x}) \in S_{MY} \end{aligned} \tag{18}$$

Figure 1 illustrates part of underlying space decomposition developed in this work. Each mapped measurement vector is decomposed and their projections are compared with their control limits.

4. PROCESS MONITORING BASED ON KPLS

4.1. Fault detection indexes

The multivariate process monitoring strategy uses statistical indexes associated to different subspaces for fault detection purposes. Based on the *in-control* KPLS model it can be analyzed every future behavior of the process by mapping the new observations **x** to the modeled subspace and to the (complementary) residual subspace, $\mathbf{x} \mapsto \hat{\bar{\boldsymbol{\varphi}}}(\mathbf{x}) + \tilde{\bar{\boldsymbol{\varphi}}}(\mathbf{x})$. In each of these subspaces it is possible to measure distances, independently. However, since no explicit formulas are available for the projections of Eq. (16), it must be found new statistics using the kernel substitution to obtain (estimated) norms for the projected vectors.

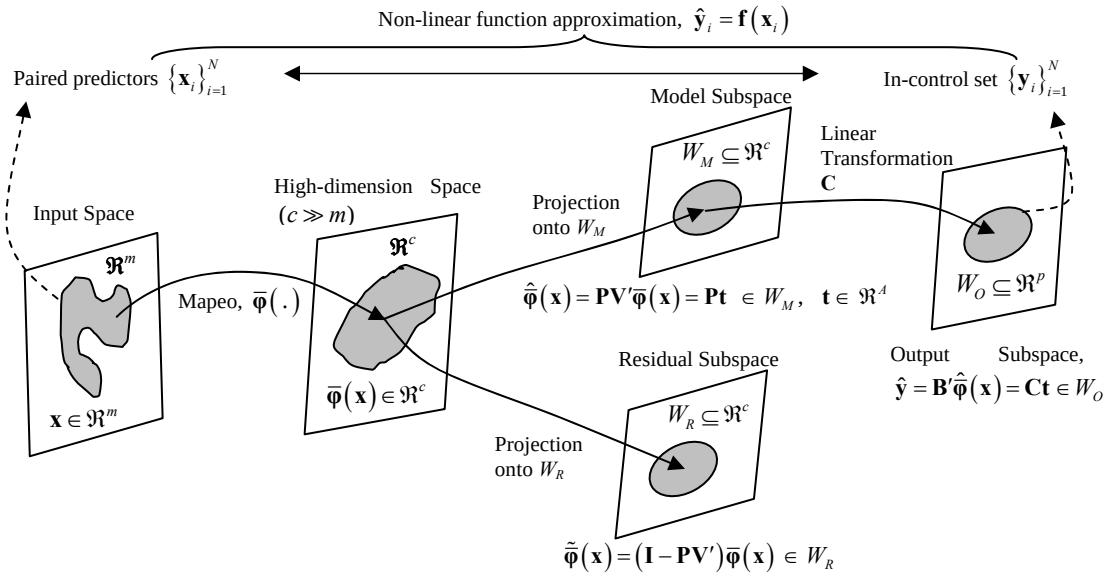

Figure 1: KPLS intrinsic architecture showing the input underlying decomposition with their relations and control limits.

For instance, to detect a significant change in W_M, the following Hotelling's T^2 statistic for $\mathbf{t}$ is defined:

$$T_t^2(\mathbf{x}) = \mathbf{t}'\boldsymbol{\Lambda}^{-1}\mathbf{t} = (N-1)\bar{\mathbf{k}}'\mathbf{R}\mathbf{R}'\bar{\mathbf{k}}, \tag{19}$$

where $\boldsymbol{\Lambda} = (N-1)^{-1}\mathbf{T}'\mathbf{T} = (N-1)^{-1}\mathbf{I}$.

When a new special event (originally not considered by the in-control KPLS model) occurs, the new mapped observation $\bar{\boldsymbol{\varphi}}(\mathbf{x})$ will move out from W_M, into W_R. The squared prediction error of $\bar{\boldsymbol{\varphi}}(\mathbf{x})$ (SPE_X), or distance to the $\bar{\boldsymbol{\varphi}}(\mathbf{x})$-model, is defined as:

$$\begin{aligned} SPE_X(\mathbf{x}) &= \left\|\tilde{\bar{\boldsymbol{\varphi}}}(\mathbf{x})\right\|^2 = \left\|\bar{\boldsymbol{\varphi}}(\mathbf{x}) - \mathbf{PV}'\bar{\boldsymbol{\varphi}}(\mathbf{x})\right\|^2 \\ &= \bar{\boldsymbol{\varphi}}(\mathbf{x})'\bar{\boldsymbol{\varphi}}(\mathbf{x}) - 2\bar{\boldsymbol{\varphi}}(\mathbf{x})'\hat{\bar{\boldsymbol{\varphi}}}(\mathbf{x}) + \hat{\bar{\boldsymbol{\varphi}}}(\mathbf{x})'\hat{\bar{\boldsymbol{\varphi}}}(\mathbf{x}) \\ &= \bar{k}(\mathbf{x},\mathbf{x}) - 2\bar{\mathbf{k}}'\bar{\mathbf{K}}\mathbf{R}\mathbf{t} + \mathbf{t}'\mathbf{t} \end{aligned} \tag{20}$$

where $\bar{\boldsymbol{\varphi}}(\mathbf{x})'\hat{\bar{\boldsymbol{\varphi}}}(\mathbf{x}) = \bar{\boldsymbol{\varphi}}(\mathbf{x})'\mathbf{Pt} = \bar{\boldsymbol{\varphi}}(\mathbf{x})'\bar{\boldsymbol{\Phi}}'\bar{\mathbf{K}}\mathbf{R}\mathbf{t} = \bar{\mathbf{k}}'\bar{\mathbf{K}}\mathbf{R}\mathbf{t}$.

Then, SPE_X can be used for detecting a change in W_R. When the process is in-control, the SPE_X index represents the fluctuations that can not be explained by the KPLS model. The distance from the regression model in S_{MY} is defined as:

$$SPE_{Y1} = \left\|\tilde{\mathbf{y}}_1\right\|^2 = \left\| \begin{bmatrix} \mathbf{CD}' & -\mathbf{CR}' \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{k}}(\mathbf{x}) \end{bmatrix} \right\|^2 \tag{21}$$

and the distance from the $\mathbf{y}$-model in S_{RY} is defined as:

$$SPE_{Y2} = \left\|\tilde{\mathbf{y}}_2\right\|^2 = \left\|(\mathbf{I} - \mathbf{CD}')\mathbf{y}\right\|^2 \tag{22}$$

Frequently, $\mathbf{R}_{\hat{\bar{\boldsymbol{\varphi}}}}$ and $\mathbf{R}_{\hat{\mathbf{y}}}$ are singular. Then, the generalized Mahalanobis' distance for $\hat{\bar{\boldsymbol{\varphi}}}$ and $\hat{\mathbf{y}}$ are:

$$D_{\hat{\bar{\boldsymbol{\varphi}}}} = \hat{\bar{\boldsymbol{\varphi}}}'\mathbf{R}_{\hat{\bar{\boldsymbol{\varphi}}}}^{-}\hat{\bar{\boldsymbol{\varphi}}}, \tag{23}$$

$$D_{\hat{\mathbf{y}}} = \hat{\mathbf{y}}'\mathbf{R}_{\hat{\mathbf{y}}}^{-}\hat{\mathbf{y}}. \tag{24}$$

where the correlation matrices are given by:

$$\begin{aligned} \mathbf{R}_{\hat{\mathbf{y}}} &= (N-1)^{-1}\hat{\mathbf{Y}}'\hat{\mathbf{Y}} \\ &= \mathbf{C}\left[(N-1)^{-1}\mathbf{T}'\mathbf{T}\right]\mathbf{C}' = (N-1)^{-1}\mathbf{CC}' \end{aligned} \tag{25}$$

$$\begin{aligned} \mathbf{R}_{\hat{\bar{\boldsymbol{\varphi}}}} &= (N-1)^{-1}\hat{\bar{\boldsymbol{\Phi}}}'\hat{\bar{\boldsymbol{\Phi}}} \\ &= (N-1)^{-1}\mathbf{PT}'\mathbf{TP}' = (N-1)^{-1}\mathbf{PP}' \end{aligned} \tag{26}$$

Since $\mathbf{C}$ is orthogonal by columns, the property of the generalized inverse of a SVD yields: $\mathbf{R}_{\hat{\mathbf{y}}}^{-} = \left(\mathbf{C}\left[(N-1)^{-1}\mathbf{I}\right]\mathbf{C}'\right)^{-} = (N-1)\mathbf{CC}'$. Then, replacing this and Eq. (9) into the Eq. (24), result the following equality:

$$D_{\hat{\mathbf{y}}} = \hat{\mathbf{y}}'\mathbf{R}_{\hat{\mathbf{y}}}^{-}\hat{\mathbf{y}} = \mathbf{t}'\mathbf{C}'((N-1)\mathbf{CC}')\mathbf{Ct} = (N-1)\mathbf{t}'\mathbf{t} = T_t^2. \tag{27}$$

Similarly, $\mathbf{R}_{\hat{\bar{\boldsymbol{\varphi}}}}^{-} = (N-1)\mathbf{PP}'$ (see Eq. 26). Then, result:

$$D_{\hat{\bar{\boldsymbol{\varphi}}}} = \hat{\bar{\boldsymbol{\varphi}}}'\left[(N-1)\mathbf{PP}'\right]\hat{\bar{\boldsymbol{\varphi}}} = (N-1)\mathbf{t}'\mathbf{t} = T_t^2 = D_{\hat{\mathbf{y}}}. \tag{28}$$

and consequently, the metrics on $\hat{\bar{\boldsymbol{\varphi}}}(\mathbf{x})$, $\mathbf{t}$ or $\hat{\mathbf{y}}$ are equivalents.

Hence, the process output (quality variables) can be monitored through a PLSR-based input statistic. Then, we propose monitoring using four non-overlapped metrics (SPE_X, T_t^2, SPE_{Y1}, and SPE_{Y2}) which completely

cover both measurement spaces, each one on a different subspace.

4.2. Fault diagnosis by means of alarmed subspaces

An anomaly is a change in the measurements following or not the correlation structure captured by the PLSR model. If the change produces an out-of-control point, the anomaly source can be classified according to a) an excessively large operation change of the normal operation; b) a significant increase of variability; c) the alteration of cross-correlations, and d) sensor faults. Cases a) and b) involve changes in the measurement vector following the modeled correlation structure; while cases c) and d) involve changes in some variables altering the correlation pattern with the others. In fact, an abnormal process behavior involves a deviation of the modeled correlations, thus increasing the value of the proper metric being used. In order to classify the abnormalities, we analyze the effect on each subspace (see Table 1). Rows 1), 2), 3) and 6) feature complex process changes; while rows 4) and 5) represent localized sensor faults. By analyzing the contributions to an alarmed index (Alcala and Qin, 2011), such as SPE_X (or SPE_Y), it could be possible to discriminate changes in the **X**- / **Y**-outer part against sensor fault in **x** / **y** (see this ambiguity in Table 1).

In summary, the proposed monitoring strategy is based on an input and output space PLS decomposition, which classifies the type of process fault or anomaly according to the statistic that triggers the alarm condition.

Table 1. Fault diagnosis based on alarmed index.

Fault/Anomaly in	SPE_X	T_t^2	SPE_{Y1}	SPE_{Y2}
1- Inner part, *d***B**	–	–	×	–
2- **X** outer part, *d***P**	×	–	o	–
3- **Y** outer part, *d***Q**	–	–	o	×
4- **x** sensor	×	–	o	–
5- **y** sensor	–	–	o	×
6- latent space upset	–	×	–	–

×: high value. –: negligible value. o: high/low value.

5. APPLICATION STUDY

5.1. A Non-linear Numerical Simulation Example

A simulated case study is used to evaluate the performance of the proposed monitoring technique for a variety of fault scenarios. To understand the implicit decompositions and statistics as monitoring tools, we simulated different faults in a synthetic system. We use the nonlinear multidimensional simulation example devised in Refs. Zhang et al (2008) and Zhao et al (2006). It is defined as follows:

$$x_1 = t^2 - t + 1 + \varepsilon_1,$$
$$x_2 = \sin(t) + \varepsilon_2,$$
$$x_3 = t^3 + t + \varepsilon_3,$$
$$y = x_1^2 + x_1 x_2 + 3\cos(x_3) + \varepsilon_4$$

where t is uniformly distributed over [-1, 1] and ε_i, i=1,2,3,4 are noise components uniformly distributed over [-0.1, 0.1]. The generated data of 300 samples are segmented into training and testing data sets. They are illustrated in the Fig. 2. The first 200 samples are selected for training, and the subsequent 100 samples are used as testing data set. It is apparent that the input variables are driven by one latent variable t only in this case. From Fig. 2, we can easily see that the response variable is nonlinearly correlated with the input variables. In this work, we used the radial basis kernel, $k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2 / h\right)$, in our implementation. When using this kernel function the value of parameter $h=2\sigma^2=0.06$ (σ=0.1732) has a significant influence on the KPLS prediction performance (Zhang et al, 2008).

The mean squared error (MSE) is used to evaluate the estimator. Figure 3a show the $\%RMSE = 100\sqrt{MSE}$ (standard error) for each a component using training data or testing data (or external validation data). Notice that $MSE(a)$ refers to MSE when the first a latent variables are used. The selection of an adequate number A of latent variables to be included in the KPLS model is crucial; if more than necessary variables are used, an undesirable over-fitting might reduce the predictive ability. We use the Wold's $R_{0.9}$ rule, which can be written as $R(a+1) = MSE(a+1))/MSE(a)$. For successive a values, the sequence is stopped when $R(a+1)>0.9$; and hence $A=a$. Conceptually, this criterion states that an additional latent variable will not be included in the KPLS model unless it provides a meaningful prediction improvement, and consequently it gives the maximum number on components to be included in the model. Figure 3b show the above indicator vs. the number of components for training data and testing data. The ratio for training data is larger than 0.9 in $a+1=12$, hence the most parsimonious model corresponds to $A=11$ (Zhang et al 2008). The same ratio with testing data is for monitor overtraining when is selecting a reliable A value. Figures 4 show the prediction results of the training and testing data using the Eq. 10. The upper part of Figure 4 shows the actual and predicted values, and the lower part shows the errors between both values.

Six faults were simulated, which are described in the Table 3. Table 1 shows the diagnosis expected in each simulated fault/anomaly. Figure 5 shows the time evolution of each statistic normalized by its control limit, which clearly exhibit the six simulated abnormalities, where it is possible to detect and classify each type of simulated fault, on the basis of the information given in Table 1. Since $y \in \Re^1$ (is not multivariable), then $SPE_{Y2} \equiv 0$ and $SPE_{Y1} = \left\|y - \hat{y}\right\|^2$. Hence, the y-sensor fault will occur in SPE_{Y1} (see Table 1). The simulation results show that the developed strategy is able to identify abnormalities attributed to sensor faults, process changes, process upsets and disturbances. The model supporting this monitoring approach is based on normal operating data. Process

data recollection constitutes a critical step when developing empirical models for monitoring.

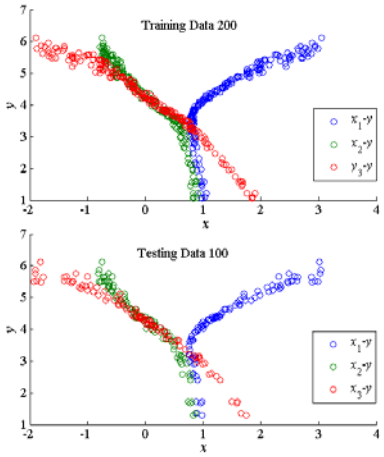

Figure 2. Relationship between inputs and the response.

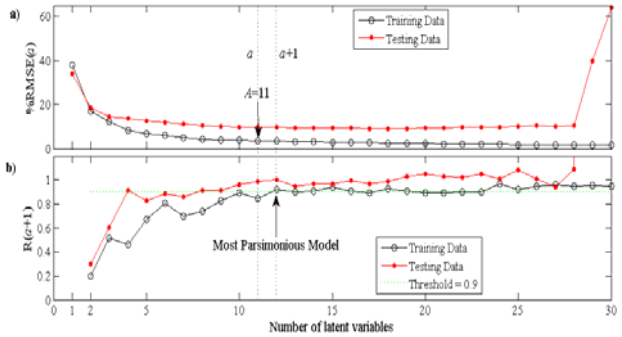

Figure 3. KPLS Model order determination. a) Residual error variance vs. number of latent variables. b) Ratio of residual error variances successive.

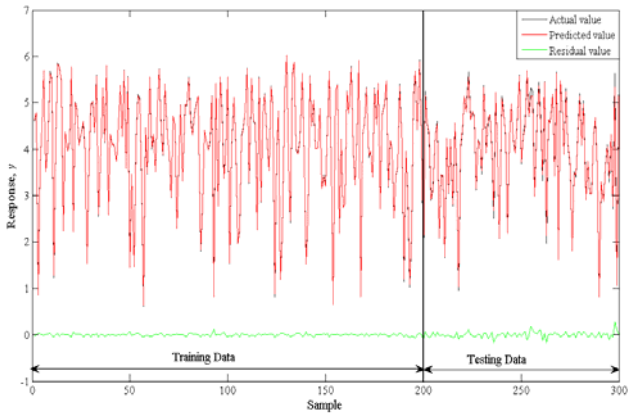

Figure 4. Prediction results via KPLS method.

Table 3: Simulated fault scenarios.

Type	Samples	Magnitude	Diagnosis
4	20..30	dx_1 = -1.5	2/4
4	50..60	dx_2 = 1.5	2/4
4	80..90	dx_3 = 1.5	2/4
2	110..120	$x_1 = 3t^3 - t + 1 + \varepsilon_1$	2/4
6	140..150	t = 1.01 (fixed)	6
5	170..180	dy = 1.5	5

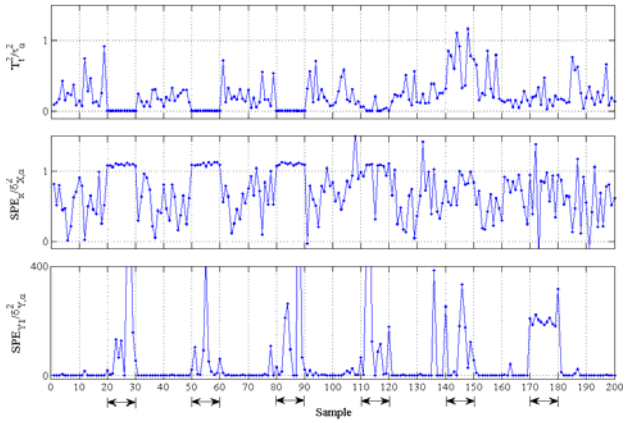

Figure 5: Time evolution of the combined index and each normalized KPLS-statistic.

6. CONCLUSIONS AND FUTURE WORKS

Many multivariate process monitoring systems could be based on a non-linear KPLS model that represents 'in-control' conditions. As in more traditional, a meaningful deviation of the variables from their expected trajectories serves for the detection and diagnosis of abnormal process behaviors. The results of a non-linear simulation example illustrates that the proposed strategy is efficient and accurate.

However, these results are preliminary, more realistic applications are necessary for a reliable validating of the method and to learn more about the proposed non-linear strategy.

ACKNOWLEDGMENTS

The authors are grateful for financial support received from CONICET, MinCyT, Universidad Nacional del Litoral, and Universidad Tecnológica Nacional (Argentina).

REFERENCES

Godoy J.L., Vega J.R., Marchetti J.L., 2011. Geometric Properties of Partial Least Squares Regression Applied to Process Monitoring. *III MACI 2011*. Bahia Blanca (Argentina), 9 to 11 may.

Meyer D., 2000. *Matrix analysis and applied linear algebra*. SIAM, USA.

Zhao, S. J.; Zhang, J.; Xu, Y. M.; Xiong, Z. H., 2006. Nonlinear Projection to Latent Structures Method and its Applications. *Ind. Eng. Chem. Res.* 45, 3843-3852.

Zhang X., Yan W., and Shao H., 2008. Nonlinear multivariate quality estimation and prediction based on kernel partial least squares. *Ind. Eng. Chem. Res.*, 47 (4), 1120–1131.

Rosipal R. and Trejo L.J., 2001. Kernel Partial Least Squares Regression in Reproduciing Kernel Hilbert Space. J. of Machine Learning Research, 2, 97-123.