

Informative PLS score-loading plots for process understanding and monitoring**

ROLF ERGON*

Keywords: *PLS, score-loading correspondence, biplot, process understanding and monitoring*

Principal component regression (PCR) based on principal component analysis (PCA) and partial least squares regression (PLSR) are well known projection methods for analysis of multivariate data. They result in scores and loadings that may be visualized in a score-loading plot (biplot) and used for process monitoring. The difficulty with this is that often more than two principal or PLS components have to be used, resulting in a need to monitor more than one such plot. However, it has recently been shown that for a scalar response variable all PLSR/PCR models can be compressed into equivalent PLSR models with two components only. After a summary of the underlying theory, the present paper shows how such two-component PLS (2PLS) models can be utilized in informative score-loading biplots for process understanding and monitoring. The possible utilization of known projection model monitoring statistics and variable contribution plots is also discussed, and a new method for visualization of contributions directly in the biplot is presented. An industrial data example is included.

1. Introduction

Partial least squares regression (PLSR) and principal component regression (PCR) are well known methods for prediction of e.g. a scalar response variable y from multivariate regressor variables \mathbf{z}^T according to $\hat{y}_{\text{new}} = \mathbf{z}_{\text{new}}^T \hat{\mathbf{b}}$ (Martens & Næs, 1989). In these methods the estimator $\hat{\mathbf{b}}$ is found from modeling data collected in a regressor matrix $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_p] = [\mathbf{z}_1 \ \mathbf{z}_2 \ \dots \ \mathbf{z}_N]^T$ and a response vector $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_N]^T$, where the basic idea in PLSR is to maximize the covariance between \mathbf{X} and \mathbf{y} , while PCR is based on principal component analysis (PCA) of \mathbf{X} (Martens & Næs, 1989). The reason for use of PLSR, PCR or some other regularization method is that an ordinary least squares (LS) problem becomes ill-posed due to a large p (many variables) relative to N (few samples in the modeling set) or/and strongly collinear variables in \mathbf{X} . An important result of PLSR/PCR is also a compression of \mathbf{X} into a few components, as in the PLS factorization

$$\mathbf{X} = \hat{\mathbf{t}}_1 \hat{\mathbf{w}}_1^T + \hat{\mathbf{t}}_2 \hat{\mathbf{w}}_2^T + \dots + \hat{\mathbf{t}}_A \hat{\mathbf{w}}_A^T + \hat{\mathbf{E}} = \hat{\mathbf{T}} \hat{\mathbf{W}}^T + \hat{\mathbf{E}}, \quad (1)$$

where $A \ll p$, and the interpretation of these components is a central part of PLSR/PCR modeling. The underlying assumption is here that the \mathbf{X} and \mathbf{y} data are generated by a few latent variables (plus noise), which has been shown to be true in a vast number of practical applications.

In a process monitoring application the response y is typically a product quality that cannot be measured on-line, while \mathbf{z}^T is made up by easily accessible process measurements (MacGregor & Kourti, 1995), and possibly also by manipulated inputs. An

*Telemark University College, Porsgrunn, Norway.

**Reprinted from JOURNAL OF PROCESS CONTROL, VOL. 14, NO. 8, Rolf Ergon, "Informative PLS score-loading plots for process understanding and monitoring", pp. 889–897, Copyright (2004), with permission from Elsevier Science.

important part of the PLSR/PCR modeling is visualization by use of score and loading plots, and the interpretation of these plots for e.g. product quality estimation, and fault detection and diagnosis (Chiang *et al.*, 2001; Kourti & MacGregor, 1996). Most multivariate monitoring techniques also use squared prediction error (SPE) plots (also known as Q plots) and Hotelling's T^2 plots that summarize the information, give a statistical basis for the interpretation and preserve the time sequence of the data (Chiang *et al.*, 2001; Kourti & MacGregor, 1996). As part of trend and fault diagnoses, specific process samples may also be analyzed by use of variable contribution plots (Chiang *et al.*, 2001; Kourti & MacGregor, 1996). The overall monitoring problem is, however, increasingly difficult with more than two PLSR/PCR components, as is often necessary also in the scalar response case. For example will the use of four components make it necessary to monitor either a T^2 plot (and lose some details) or three score plots (e.g. \hat{t}_2 , \hat{t}_3 and \hat{t}_4 vs. \hat{t}_1). And as shown in an industrial data example in Section 6, some important process understanding that can be gained from a two-component model, is very much lost for models with more than two components.

In a recent work (Ergon, 2003) it has been shown that for a scalar response variable, all PLSR/PCR models can be reduced to an equivalent model with two components only (2PLS), while at the same time the estimator $\hat{\mathbf{b}}$ is unchanged. The 2PLS factorization is

$$\mathbf{X} = \hat{\mathbf{t}}_1 \hat{\mathbf{w}}_1^T + \tilde{\mathbf{t}}_2 \tilde{\mathbf{w}}_2^T + \tilde{\mathbf{E}} = \tilde{\mathbf{T}} \tilde{\mathbf{W}}^T + \tilde{\mathbf{E}}, \quad (2)$$

i.e. all y -relevant variation in \mathbf{X} is captured in one score plot (\tilde{t}_2 vs. \hat{t}_1) and one loading plot (\tilde{w}_2 vs. \hat{w}_1). An attractive feature with this factorization is also that \tilde{t}_2 is orthogonal to both y and $\hat{y} = \mathbf{X}\hat{\mathbf{b}}$. Note that the first component in (2) is the same as in (1). Also note that the residual $\tilde{\mathbf{E}}$ in (2) generally is different from $\hat{\mathbf{E}}$ in (1), because more y -orthogonal variation in \mathbf{X} may be captured in $\tilde{\mathbf{E}}$.

The main focus of the present paper is on the utilization of 2PLS realizations in process monitoring applications based on correspondence between scores and loadings, and use of combined score-loading plots (biplots) (Skagerberg & Sundin, 1993; Ergon 2002). Such plots will make it possible for a process operator monitoring a single plot to detect a deviation from normal plant operation, to evaluate the importance of this with regard to the response y , and to get some indications on which process variables that may cause the deviation. The standard techniques for use of SPE, T^2 and contribution plots are applicable also for a 2PLS model, although they may to some extent be less necessary. With two components only, a statistical T^2 upper limit may be shown as a confidence ellipse in the score plot, while the contributions from the different variables may be shown directly in the score-loading plot (see Section 5 for theory and Section 6 for an industrial data example).

Some background in PCR and PLSR modeling is given in Section 2, followed by a short presentation of the 2PLS algorithm in Section 3. The score-loading correspondence is discussed in Section 4, followed by a summary of multivariate monitoring statistics and variable contribution analyses in Section 5, the industrial data example in Section 6, and conclusions in Section 7.

2. PCR and PLSR modeling

Multivariate calibration using PCR/PLSR directly or implicitly assumes a latent variables (LV) model

$$\mathbf{y} = \mathbf{T}\mathbf{Q}^T + \mathbf{f} \quad (3)$$

$$\mathbf{X} = \mathbf{T}\mathbf{L}^T + \mathbf{E}, \quad (4)$$

which assuming $\mathbf{L}^T\mathbf{L} = \mathbf{I}$ and using the LS solutions $\hat{\mathbf{Q}}^T = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{y}$ and $\hat{\mathbf{T}} = \mathbf{X}\mathbf{L}$ results in $\hat{\mathbf{y}} = \hat{\mathbf{T}}(\hat{\mathbf{T}}^T\hat{\mathbf{T}})^{-1}\hat{\mathbf{T}}^T\mathbf{y}$ and thus the estimator

$$\hat{\mathbf{b}}_{LV} = \mathbf{L}(\mathbf{L}^T\mathbf{X}^T\mathbf{X}\mathbf{L})^{-1}\mathbf{L}^T\mathbf{X}^T\mathbf{y}. \quad (5)$$

The PCR and PLSR algorithms use different factorizations of \mathbf{X} as summarized below, and thus also different factorizations of \mathbf{y} . PCR is based on the PCA factorization or singular value decomposition

$$\begin{aligned} \mathbf{X} = \mathbf{USV}^T &= [\mathbf{U}_1 \quad \mathbf{U}_2] \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^T + \hat{\mathbf{E}}_{PCA} \\ &= \hat{\mathbf{T}}_{PCA}\hat{\mathbf{P}}_{PCA}^T + \hat{\mathbf{E}}_{PCA} = \hat{\mathbf{t}}_1^{PCA}(\hat{\mathbf{p}}_1^{PCA})^T + \hat{\mathbf{t}}_2^{PCA}(\hat{\mathbf{p}}_2^{PCA})^T + \dots + \hat{\mathbf{t}}_A^{PCA}(\hat{\mathbf{p}}_A^{PCA})^T + \hat{\mathbf{E}}_{PCA}, \end{aligned} \quad (6)$$

where $\hat{\mathbf{T}}_{PCA} = \mathbf{U}_1\mathbf{S}_1 \in \mathbb{R}^{N \times A}$ and $\hat{\mathbf{P}}_{PCA} = \mathbf{V}_1 \in \mathbb{R}^{p \times A}$, and where both $\hat{\mathbf{T}}_{PCA}$ and $\hat{\mathbf{P}}_{PCA}$ are orthogonal, with $\hat{\mathbf{P}}_{PCA}$ also orthonormal. The estimator is determined by (5), with \mathbf{L} replaced by $\hat{\mathbf{P}}_{PCA}$. The number of components to include is based on either cross-validation or independent test set validation (Martens & Næs, 1989). In the established terminology the columns of $\hat{\mathbf{T}}_{PCA}$ are called score vectors, where the elements are scores, while the columns of $\hat{\mathbf{P}}_{PCA}$ are called loading vectors, where the elements are loadings.

The basic idea in PLSR is that the covariance between \mathbf{X} and \mathbf{y} should be maximized. The original so-called orthogonalized PLSR algorithm of Wold is based on a factorization with orthogonal score vectors and non-orthogonal loading vectors (Martens & Næs, 1989; Ergon, 2003). In the present context it is, however, more relevant to refer to the Martens factorization (Martens & Næs, 1989)

$$\mathbf{X} = \hat{\mathbf{t}}_1\hat{\mathbf{w}}_1^T + \hat{\mathbf{t}}_2\hat{\mathbf{w}}_2^T + \dots + \hat{\mathbf{t}}_A\hat{\mathbf{w}}_A^T + \hat{\mathbf{E}} = \hat{\mathbf{T}}\hat{\mathbf{W}}^T + \hat{\mathbf{E}}, \quad (7)$$

where the loading weight matrix $\hat{\mathbf{W}} = [\hat{\mathbf{w}}_1 \quad \hat{\mathbf{w}}_2 \quad \dots \quad \hat{\mathbf{w}}_A]$ is orthonormal, while the score matrix $\hat{\mathbf{T}} = \mathbf{X}\hat{\mathbf{W}}$ is non-orthogonal. The common estimator in the Wold and Martens PLSR algorithms is (Helland, 1988)

$$\hat{\mathbf{b}} = \hat{\mathbf{W}}(\hat{\mathbf{W}}^T\mathbf{X}^T\mathbf{X}\hat{\mathbf{W}})^{-1}\hat{\mathbf{W}}^T\mathbf{X}^T\mathbf{y}. \quad (8)$$

3. Compression into two PLS components

3.1. Basic insight

The basic insight behind the 2PLS algorithm is illustrated in Figure 1. The estimator $\hat{\mathbf{b}}$ is found in the space spanned by the loading weight vectors in $\hat{\mathbf{W}} = [\hat{\mathbf{w}}_1 \quad \hat{\mathbf{w}}_2 \quad \dots \quad \hat{\mathbf{w}}_A]$, i.e. it is a linear combination of these vectors. It is, however, also found in the plane defined by $\hat{\mathbf{w}}_1$ and a vector $\hat{\mathbf{w}}_2$ orthogonal to $\hat{\mathbf{w}}_1$, which is a linear combination of the vectors $\hat{\mathbf{w}}_2, \hat{\mathbf{w}}_3, \dots, \hat{\mathbf{w}}_A$. The matrix $\hat{\mathbf{W}} = [\hat{\mathbf{w}}_1 \quad \hat{\mathbf{w}}_2]$ is thus the loading weight matrix in a two-component PLS solution (2PLS) giving exactly the same estimator $\hat{\mathbf{b}}$ as the original solution using any number A components (Ergon, 2003).

What matters in the original PLS model is not the matrix $\hat{\mathbf{W}}$ as such, but the space spanned by $\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \dots, \hat{\mathbf{w}}_A$ (Helland, 1988), and in the 2PLS model it is the plane spanned by $\hat{\mathbf{w}}_1$ and $\hat{\mathbf{w}}_2$ that is essential. The reason for keeping $\hat{\mathbf{w}}_1$ is discussed in Ergon (2003). Note that all samples in \mathbf{X} (row vectors) in the original PLS model are projected onto the space spanned by $\hat{\mathbf{w}}_2, \hat{\mathbf{w}}_3, \dots, \hat{\mathbf{w}}_A$. They may thus be further projected onto the plane spanned by $\hat{\mathbf{w}}_1$ and $\hat{\mathbf{w}}_2$, and form a single score plot containing all \mathbf{y} -relevant information. If for some reason e.g. $\hat{\mathbf{w}}_2$ is more informative than $\hat{\mathbf{w}}_1$, a plane through $\hat{\mathbf{w}}_2$ and $\hat{\mathbf{b}}$ may

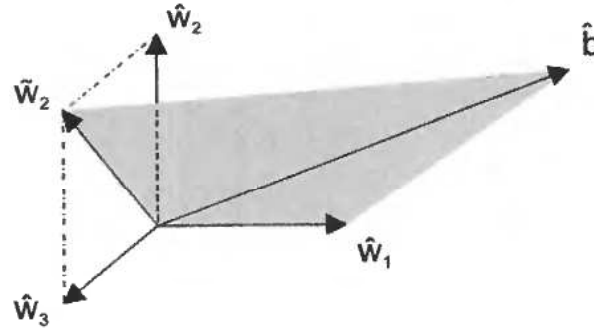


Figure 1. Illustration of basic insight behind the 2PLS factorization, assuming $A = 3$ original components. The PLSR estimator $\hat{\mathbf{b}}$ is found in the space spanned by $\hat{\mathbf{w}}_1$, $\hat{\mathbf{w}}_2$ and $\hat{\mathbf{w}}_3$, but also in the shadowed plane spanned by $\hat{\mathbf{w}}_1$ and $\tilde{\mathbf{w}}_2$.

be a better alternative. It will in any case result in a 2PLS model that gives the estimator $\hat{\mathbf{b}}$, as will in fact all planes through $\hat{\mathbf{b}}$ that are at the same time subspaces of the column space of $\hat{\mathbf{W}}$.

3.2. Algorithm

As illustrated in Figure 1, the central problem is to find a second vector that together with $\hat{\mathbf{w}}_1$ spans the shadowed plane that includes $\hat{\mathbf{b}}$. One way of doing this follows from the estimator formulation (8), in that $(\hat{\mathbf{W}}^T \mathbf{X}^T \mathbf{X} \hat{\mathbf{W}})^{-1} \hat{\mathbf{W}}^T \mathbf{X}^T \mathbf{y} = [\hat{a}_1 \ \hat{a}_2 \ \dots \ \hat{a}_A]^T$, and thus

$$\hat{\mathbf{b}} = \hat{a}_1 \hat{\mathbf{w}}_1 + [\hat{\mathbf{w}}_2 \ \hat{\mathbf{w}}_3 \ \dots \ \hat{\mathbf{w}}_A][\hat{a}_2 \ \hat{a}_3 \ \dots \ \hat{a}_A]^T = \hat{a}_1 \hat{\mathbf{w}}_1 + \tilde{a}_2 \tilde{\mathbf{w}}_2. \quad (9)$$

We summarize the 2PLS compression and its properties in Theorem 1 below (see Ergon (2003) for proof and Matlab code). The second vector spanning the shadowed plane in Figure 1 is not necessarily $\tilde{\mathbf{w}}_2$, but it is convenient to make use of the orthogonal vectors $\hat{\mathbf{w}}_1$ and $\tilde{\mathbf{w}}_2$ in the theorem and its proof.

Theorem 1 The original PLSR estimator (8) can be written as

$$\hat{\mathbf{b}} = \tilde{\mathbf{W}}(\tilde{\mathbf{W}}^T \mathbf{X}^T \mathbf{X} \tilde{\mathbf{W}})^{-1} \tilde{\mathbf{W}}^T \mathbf{X}^T \mathbf{y}, \quad (10)$$

where $\tilde{\mathbf{W}} = [\hat{\mathbf{w}}_1 \ \tilde{\mathbf{w}}_2]$ is the new orthonormal loading (weight) matrix. The corresponding factorization of \mathbf{X} is

$$\mathbf{X} = \tilde{\mathbf{T}} \tilde{\mathbf{W}}^T + \tilde{\mathbf{E}} = \hat{\mathbf{t}}_1 \hat{\mathbf{w}}_1^T + \tilde{\mathbf{t}}_2 \tilde{\mathbf{w}}_2^T + \tilde{\mathbf{E}}, \quad (11)$$

and $\tilde{\mathbf{w}}_2$ is

$$\tilde{\mathbf{w}}_2 = \frac{[\hat{\mathbf{w}}_2 \ \hat{\mathbf{w}}_3 \ \dots \ \hat{\mathbf{w}}_A][\hat{a}_2 \ \hat{a}_3 \ \dots \ \hat{a}_A]^T}{\|[\hat{\mathbf{w}}_2 \ \hat{\mathbf{w}}_3 \ \dots \ \hat{\mathbf{w}}_A][\hat{a}_2 \ \hat{a}_3 \ \dots \ \hat{a}_A]^T\|}. \quad (12)$$

Here $[\hat{a}_2 \ \hat{a}_3 \ \dots \ \hat{a}_A]^T$ is extracted from $[\hat{a}_1 \ \hat{a}_2 \ \dots \ \hat{a}_A]^T = (\hat{\mathbf{W}}^T \mathbf{X}^T \mathbf{X} \hat{\mathbf{W}})^{-1} \hat{\mathbf{W}}^T \mathbf{X}^T \mathbf{y}$, while $\hat{\mathbf{t}}_1 = \mathbf{X} \hat{\mathbf{w}}_1$ is the same as in the factorization (7), and $\tilde{\mathbf{t}}_2 = \mathbf{X} \tilde{\mathbf{w}}_2$. Furthermore, $\tilde{\mathbf{t}}_2$ is orthogonal to both \mathbf{y} and $\hat{\mathbf{y}} = \mathbf{X} \hat{\mathbf{b}}$, i.e. $\mathbf{y}^T \tilde{\mathbf{t}}_2 = 0$ and $\hat{\mathbf{y}}^T \tilde{\mathbf{t}}_2 = 0$. ■

Note that this theorem is valid for any number A of original PLS components. Also note that the residual $\tilde{\mathbf{E}}$ may be different from the original residual $\hat{\mathbf{E}}$, i.e. some extra \mathbf{y} -orthogonal structured variation in \mathbf{X} may be captured in $\tilde{\mathbf{E}}$ (see theoretical example in Subsection 3.3 below).

Remark 1 Since $\hat{\mathbf{W}}$ in the estimator (8) may be replaced by $\hat{\mathbf{W}}\mathbf{M}$, where \mathbf{M} is any invertible transformation matrix, it follows from Theorem 1 that any plane containing \mathbf{b} that is also a subspace of the column space of $\hat{\mathbf{W}}$ may be used instead of the shadowed plane in Figure 1. The theorem may thus be given a more general formulation.

3.3. Theoretical example

Assume an ordinary non-orthogonalized PLS factorization according to (7) with three components, i.e.

$$\mathbf{X} = \hat{\mathbf{t}}_1 \hat{\mathbf{w}}_1^T + \hat{\mathbf{t}}_2 \hat{\mathbf{w}}_2^T + \hat{\mathbf{t}}_3 \hat{\mathbf{w}}_3^T + \hat{\mathbf{E}}, \quad (13)$$

and the corresponding score matrix $\hat{\mathbf{T}} = [\hat{\mathbf{t}}_1 \ \hat{\mathbf{t}}_2 \ \hat{\mathbf{t}}_3] = \mathbf{X}[\hat{\mathbf{w}}_1 \ \hat{\mathbf{w}}_2 \ \hat{\mathbf{w}}_3]$. Also assume a new sample $\mathbf{z}_{\text{new}}^T$ such that $\hat{\mathbf{t}}_{\text{new},1} = \mathbf{z}_{\text{new}}^T \hat{\mathbf{w}}_1 = 0$. The predicted response is then

$$\begin{aligned} \hat{y} &= \mathbf{z}_{\text{new}}^T \hat{\mathbf{W}} (\hat{\mathbf{W}}^T \mathbf{X}^T \mathbf{X} \hat{\mathbf{W}})^{-1} \hat{\mathbf{W}}^T \mathbf{X}^T \mathbf{y} = \mathbf{z}_{\text{new}}^T [\hat{\mathbf{w}}_1 \ \hat{\mathbf{w}}_2 \ \hat{\mathbf{w}}_3] \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{bmatrix} \\ &= \hat{a}_2 \mathbf{z}_{\text{new}}^T \hat{\mathbf{w}}_2 + \hat{a}_3 \mathbf{z}_{\text{new}}^T \hat{\mathbf{w}}_3 = \hat{a}_2 \hat{t}_{\text{new},2} + \hat{a}_3 \hat{t}_{\text{new},3}. \end{aligned} \quad (14)$$

From this follows that it is possible to have scores $\hat{t}_{\text{new},2} \neq 0$ and $\hat{t}_{\text{new},3} \neq 0$, and at the same time have $\hat{y} = 0$.

After the compression into two components according to Theorem 1, the factorization is

$$\mathbf{X} = \tilde{\mathbf{t}}_1 \tilde{\mathbf{w}}_1^T + \tilde{\mathbf{t}}_2 \tilde{\mathbf{w}}_2^T + \tilde{\mathbf{E}}, \quad (15)$$

and the corresponding score matrix $\tilde{\mathbf{T}} = [\tilde{\mathbf{t}}_1 \ \tilde{\mathbf{t}}_2] = \mathbf{X}[\tilde{\mathbf{w}}_1 \ \tilde{\mathbf{w}}_2]$, while the predicted response with $\mathbf{z}_{\text{new}}^T \tilde{\mathbf{w}}_1 = 0$ as above is

$$\hat{y} = \mathbf{z}_{\text{new}}^T [\tilde{\mathbf{w}}_1 \ \tilde{\mathbf{w}}_2] \begin{bmatrix} \tilde{a}_1 \\ \tilde{a}_2 \end{bmatrix} = \tilde{a}_2 \mathbf{z}_{\text{new}}^T \tilde{\mathbf{w}}_2 = \tilde{a}_2 \tilde{t}_{\text{new},2}. \quad (16)$$

If now $\hat{y} = 0$ we must also have $\tilde{t}_{\text{new},2} = 0$, i.e. we will not see any deviation from the origin in the $\tilde{\mathbf{t}}_2$ vs. $\tilde{\mathbf{t}}_1$ score plot. This also illustrates how some \mathbf{y} -orthogonal variation in \mathbf{X} may be captured in $\tilde{\mathbf{E}}$.

Note, however, that it is generally possible to have deviations from the origin in the $\tilde{\mathbf{t}}_2$ vs. $\tilde{\mathbf{t}}_1$ score plot also when $\hat{y} = 0$. The general expression for \hat{y} is

$$\hat{y} = \mathbf{z}_{\text{new}}^T [\tilde{\mathbf{w}}_1 \ \tilde{\mathbf{w}}_2] \begin{bmatrix} \tilde{a}_1 \\ \tilde{a}_2 \end{bmatrix} = \hat{a}_1 \mathbf{z}_{\text{new}}^T \tilde{\mathbf{w}}_1 + \tilde{a}_2 \mathbf{z}_{\text{new}}^T \tilde{\mathbf{w}}_2 = \hat{a}_1 \hat{t}_{\text{new},1} + \tilde{a}_2 \tilde{t}_{\text{new},2}, \quad (17)$$

which may give $\hat{y} = 0$ also when $\hat{t}_{\text{new},1} \neq 0$ and $\tilde{t}_{\text{new},2} \neq 0$. In such cases, however, it will be possible to see that $\hat{y} = 0$ directly in the score plot, as demonstrated in the industrial data example in Section 6.

4. Score-loading correspondence

Score-loading correspondence for process monitoring is discussed in Ergon (2002), and a short summary only is given here. A convenient starting point is then the PCA/SVD (6), where the number of components to include is both data and application

dependent. The essential fact in the present context is that both $\hat{\mathbf{T}}_{\text{PCA}}$ and $\hat{\mathbf{P}}_{\text{PCA}}$ are orthogonal, with $\hat{\mathbf{P}}_{\text{PCA}}$ also orthonormal, resulting in the LS estimate

$$\hat{\mathbf{T}}_{\text{PCA}} = [\hat{\mathbf{t}}_1^{\text{PCA}} \quad \hat{\mathbf{t}}_2^{\text{PCA}} \quad \dots \quad \hat{\mathbf{t}}_A^{\text{PCA}}] = [\hat{\mathbf{t}}_1^{\text{PCA}} \quad \hat{\mathbf{t}}_2^{\text{PCA}} \quad \dots \quad \hat{\mathbf{t}}_N^{\text{PCA}}]^T = \mathbf{X} \hat{\mathbf{P}}_{\text{PCA}}. \quad (18)$$

Assuming the same covariance matrix as for the row vectors in the modeling data matrix \mathbf{X} , a new row vector $\mathbf{z}_{\text{new}}^T$ thus results in

$$(\hat{\mathbf{t}}_{\text{new}}^{\text{PCA}})^T = \mathbf{z}_{\text{new}}^T \hat{\mathbf{P}}_{\text{PCA}}. \quad (19)$$

If in addition $\mathbf{z}_{\text{new}}^T = [0 \dots 0 \quad x_{\text{new},j} \quad 0 \dots 0]$ we obtain

$$(\hat{\mathbf{t}}_{\text{new}}^{\text{PCA}})^T = [0 \dots 0 \quad x_{\text{new},j} \quad 0 \dots 0] \begin{bmatrix} (\hat{\mathbf{p}}_1^{\text{PCA}})^T \\ \vdots \\ (\hat{\mathbf{p}}_j^{\text{PCA}})^T \\ \vdots \\ (\hat{\mathbf{p}}_p^{\text{PCA}})^T \end{bmatrix} = x_{\text{new},j} (\hat{\mathbf{p}}_j^{\text{PCA}})^T. \quad (20)$$

This means that the scores $\hat{\mathbf{t}}_{\text{new}}^{\text{PCA}} = [\hat{t}_{\text{new},1} \quad \hat{t}_{\text{new},2} \quad \dots \quad \hat{t}_{\text{new},A}]^T$ correspond to the loadings $\hat{\mathbf{p}}_j^{\text{PCA}} = [\hat{p}_{j,1} \quad \hat{p}_{j,2} \quad \dots \quad \hat{p}_{j,A}]^T$ of the variable j in \mathbf{X} that causes the score to deviate from zero, with $x_{\text{new},j}$ as a scaling coefficient. When several variables in combination causes $\hat{\mathbf{t}}_{\text{new}}^{\text{PCA}}$, the scores correspond to weighted sums of loadings.

For the 2PLS factorization (11) it follows that $\tilde{\mathbf{T}} = \mathbf{X} \tilde{\mathbf{W}}$, where $\tilde{\mathbf{W}}$ just as $\hat{\mathbf{P}}$ in PCA is orthonormal, and the score-loading correspondence requirements are thus fulfilled. When one variable only in $\mathbf{z}_{\text{new}}^T$ is different from zero, the new score in the single 2PLS score plot will thus be found in exactly the same direction as the corresponding loading in the single 2PLS loading plot (see industrial data example in Section 6). Note that $\hat{\mathbf{b}} = \tilde{\mathbf{W}}[\hat{a}_1 \quad \hat{a}_2]^T$ just as $\mathbf{z}_{\text{new}} = \tilde{\mathbf{W}}[\hat{t}_{\text{new},1} \quad \hat{t}_{\text{new},2}]^T$ can be plotted in the score plot, while the loadings $[\hat{a}_1 \quad \hat{a}_2]$ in $\hat{\mathbf{y}} = \tilde{\mathbf{T}}[\hat{a}_1 \quad \hat{a}_2]^T$ just as the loadings $[\hat{w}_{j,1} \quad \hat{w}_{j,2}]$ in a variable vector $\mathbf{x}_j = \tilde{\mathbf{T}}[\hat{w}_{j,1} \quad \hat{w}_{j,2}]^T$ can be plotted in the loading plot. From this follows that $\hat{\mathbf{b}}$ and $\hat{\mathbf{y}}$ (based on the modeling data) in a score-loading biplot will have identical positions.

Finally note that the PLS factorization according to Wold, with orthogonal score vectors, does not fulfill the score-loading correspondence requirements (although the deviation may in some cases be small) (Ergon, 2002).

5. Monitoring statistics and contribution plots for projection methods

5.1. Squared prediction error

Process and product monitoring based on projection methods like PLS and 2PLS, rely on an in-control model developed by use of a reference or modeling data set, which defines an A -dimensional projection space. It is therefore first of all necessary to check if the perpendicular distance of a new sample $\mathbf{z}_{\text{new}}^T$ from this projection space is within acceptable limits. As a measure of this distance it is common to use the squared prediction error

$$\text{SPE}_{\mathbf{z}_{\text{new}}} = \sum_{j=1}^p (x_{\text{new},j} - \hat{x}_{\text{new},j})^2. \quad (21)$$

The SPE will detect the occurrence of a totally new type of process development. This is also applicable when using the 2PLS data compression, the only special thing being that the projection space is the plane illustrated in Figure 1. If statistical limits for such a detection of special events are of interest, $\hat{x}_{\text{new},j}$ in (21) may be computed by use of a

PCA model. Upper control limits can then be computed using approximate results for the distribution of quadratic forms, often referred to as Q -statistic (MacGregor & Kourti, 1995; Chiang *et al.*, 2001; Kourti & MacGregor, 1996).

5.2. Hotelling's T^2 statistic

T^2 plots for many PLS components Assuming that the SPE for a new sample $\mathbf{z}_{\text{new}}^T$ is acceptable, a meaningful comparison with the reference data set is possible. Further assuming normal distributions, this may be done using the Hotelling's T^2 statistic (Johnson & Wichern, 1992) based on the estimated score covariance matrix, which with centered data and use of the PLS factorization (7) is

$$\mathbf{S} = \frac{1}{N-1} \hat{\mathbf{T}}^T \hat{\mathbf{T}}. \quad (22)$$

With centered data the T^2 statistic for a new observation $\mathbf{z}_{\text{new}}^T$ is

$$T^2 = \hat{\mathbf{t}}_{\text{new}}^T \mathbf{S}^{-1} \hat{\mathbf{t}}_{\text{new}}, \quad (23)$$

where $\hat{\mathbf{t}}_{\text{new}}^T = \mathbf{z}_{\text{new}}^T \hat{\mathbf{W}}$, i.e. $[\hat{t}_{\text{new},1} \ \hat{t}_{\text{new},2} \ \dots \ \hat{t}_{\text{new},A}] = [x_{\text{new},1} \ x_{\text{new},2} \ \dots \ x_{\text{new},p}] \hat{\mathbf{W}}$. The upper control limit for T^2 based on N past multivariate data and A PLS components is

$$T_{\text{UCL}}^2 = \frac{(N^2 - 1)A}{N(N - A)} F_{\alpha}(A, N - A), \quad (24)$$

where $F_{\alpha}(A, N - A)$ is the upper 100α % critical point of the F distribution with $(A, N - A)$ degrees of freedom (Chiang *et al.*, 2001; Kourti & MacGregor, 1996). For a sequence of new samples, the T^2 value may thus be plotted and compared with T_{UCL}^2 , and a fault alarm signal given according to some more or less conservative rule.

Confidence ellipse for two PLS components Since the 2PLS algorithm results in two components only, the T_{UCL}^2 limit may be replaced by a confidence ellipse in a score plot, based on the eigenvalues (direction of axes) and eigenvectors (length of axes) of \mathbf{S} and using the same F distribution limit as in (24) (Johnson & Wichern, 1992). Such a confidence ellipse may also be included in the score-loading biplot discussed in Section 4, as shown in the industrial data example in Section 6. In addition to the information of a violated upper limit, such a biplot also gives information on which variables or group of variables that are involved, as discussed below.

5.3. Contribution plots

SPE contribution plots When the SPE or T^2 plots indicate that the process is operating outside the normal region, it is of interest to see which variables that contribute the most to this. For a specific sample of interest this can be done by plotting variable contribution plots. The SPE contribution plot is simply a plot (often a bar plot) showing how the different variables contribute to the sum (21) for a specific observation.

Contribution plots on the scores The contribution plots on a score $\hat{t}_{\text{new},a}$ could in the same way show the variable contributions to

$$\hat{t}_{\text{new},a} = \sum_{j=1}^p x_{\text{new},j} \hat{w}_{j,a}. \quad (25)$$

Since the different scores have different influence on T^2 according to (23), and in order to minimize ambiguity, it is common practice (Chiang *et al.*, 2001; Kourti & MacGregor, 1996) to use the following procedure:

1. Compute

$$cont_{j,a}^{new} = \frac{\hat{t}_{new,a}}{s_a^2} x_{new,j} \hat{w}_{j,a}, \quad (26)$$

where s_a^2 is the score vector variance based on the modeling data.

2. Set $cont_{j,a}^{new} = 0$ if it is negative, i.e. if the sign of the contribution $x_{new,j} \hat{w}_{j,a}$ is opposite to the sign of the score $\hat{t}_{new,a}$.
3. Compute and plot

$$CONT_j^{new} = \sum_{a=1}^A cont_{j,a}^{new} \quad (27)$$

for $j = 1, 2, \dots, p$.

Alternative 2PLS contribution plot on the scores With one score plot only, an alternative approach is to plot the weighted loadings in

$$\tilde{\mathbf{W}}^{weighted} = \begin{bmatrix} x_{new,1} \hat{w}_{1,1} & x_{new,1} \tilde{w}_{1,2} \\ x_{new,2} \hat{w}_{2,1} & x_{new,2} \tilde{w}_{2,2} \\ \vdots & \vdots \\ x_{new,p} \hat{w}_{p,1} & x_{new,p} \tilde{w}_{p,2} \end{bmatrix} = [\hat{\mathbf{w}}_1^{weighted} \quad \tilde{\mathbf{w}}_2^{weighted}] \quad (28)$$

directly in the score-loading biplot. The original loadings $[\hat{w}_{j,1} \quad \tilde{w}_{j,2}]$ will after multiplication by $x_{new,j}$ be moved radially, indicating the strength by which they attract (positive $x_{new,j}$) or repel (negative $x_{new,j}$) the scores. These weighted loadings will thus indicate how the contributions from different variables vary with time, as illustrated in the industrial data example in Section 6.

6. Industrial data example

The following example uses multivariate regression data from a mineral processing plant (Höskuldsson, 1996) (the 'cleaner' data, originally published in Hodouin *et al.*, (1993)). The problem considered here is to predict a given quality $y_{new,4}$ from twelve known process variables $\mathbf{z}_{new}^T = [x_{new,1} \quad x_{new,2} \quad \dots \quad x_{new,12}]$. For the purpose of finding an initial PLS factorization, samples 1 to 120 in the data sets *xce* and *yce* (Höskuldsson, 1996) were used for modeling, while samples 151 to 240 in the same data sets were used for validation. The data were centered and scaled to unit variance (autoscaled), and the result was a PLSR model with $A = 6$ components, resulting in a validation root mean square error $RMSEP = 0.15$ (as compared with $RMSEP = 1$ for $A = 0$). The PLS factorization of the autoscaled data matrix $\mathbf{X}^{modeling}$ was finally compressed into a 2PLS factorization (11), resulting in the loading matrix $\tilde{\mathbf{W}}$.

6.1. Score-loading plot for process understanding

To illustrate the use of 2PLS for process understanding, new \mathbf{X} data were introduced as $\mathbf{X}^{test} = \mathbf{I}_{12}$, i.e.

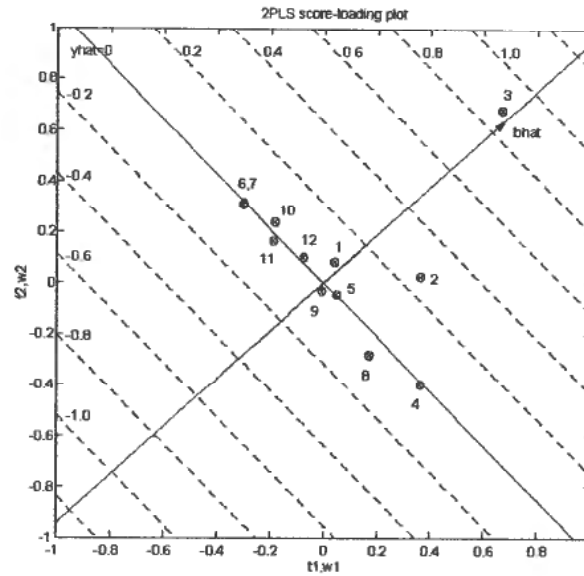


Figure 2. Score-loading correspondence plot with loadings from loading matrix $\tilde{\mathbf{W}}$ (o-marked) and scores from test matrix \mathbf{X}^{test} according to (29) (x-marked), together with the estimator $\hat{\mathbf{b}}$ (marked with arrow). Parallel dashed lines indicate constant predictions \hat{y} . Note the total score-loading correspondence, and the nearly y -orthogonal deviations for scores no. 4, 5 and 6 (see comparison with ordinary PLS score plots below).

$$\mathbf{X}^{\text{test}} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}. \quad (29)$$

The 2PLS loadings in $\tilde{\mathbf{W}}$, $[\hat{w}_{j,1} \ \hat{w}_{j,2}]_{j=1,2,\dots,12}$, and scores $[\hat{t}_{i,1} \ \hat{t}_{i,2}]_{i=1,2,\dots,12}^{\text{test}} = (\mathbf{z}_i^{\text{test}})^T \tilde{\mathbf{W}}$ were then plotted together with the estimator $\hat{\mathbf{b}}$ according to (10) (see Figure 2).

Figure 2 also shows lines for constant \hat{y} based on (using (10) with $\mathbf{X}\tilde{\mathbf{W}} = \tilde{\mathbf{T}}$)

$$\hat{y}_i = [\hat{t}_{i,1} \ \hat{t}_{i,2}] (\tilde{\mathbf{T}}^T \tilde{\mathbf{T}})^{-1} \tilde{\mathbf{T}}^T \mathbf{y} = [\hat{t}_{i,1} \ \hat{t}_{i,2}] \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = \hat{a}_1 \hat{t}_{i,1} + \hat{a}_2 \hat{t}_{i,2}, \quad (30)$$

and an axis for \hat{y} and $\hat{\mathbf{b}}$ perpendicular to those lines. Note that $\hat{\mathbf{y}}$ (based on the modeling data) in the loading plot and $\hat{\mathbf{b}}$ in the score plot have the same position in the biplot. For an operator support application, the entire plot can easily be rotated such that the \hat{y} axis becomes horizontal pointing to the right.

The plot in Figure 2 can be used to gain process understanding. As can be seen the estimator $\hat{\mathbf{b}}$ and thus the predicted response \hat{y} is strongly correlated with variable 3 and to some extent also with variables 1 and 2, while the other variables have little to do with \hat{y} (as indicated by the o-marked loadings close to the line for $\hat{y} = 0$). This can also be seen by inspection of

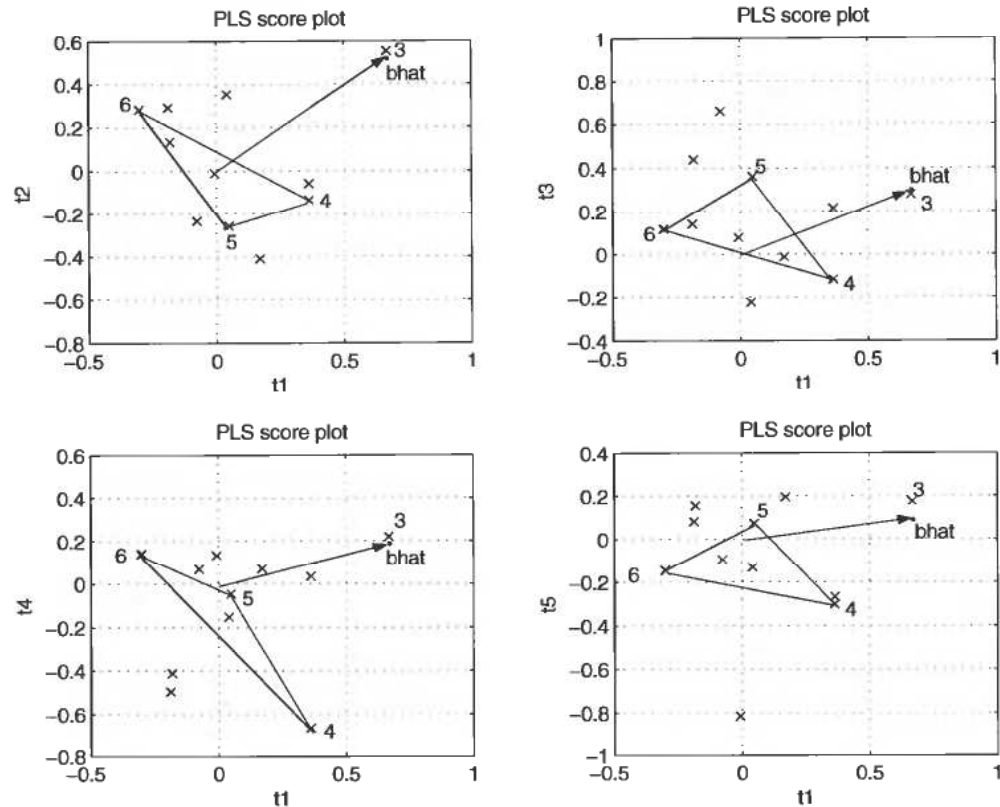


Figure 3. Score plots for test matrix \mathbf{X}^{test} and projected estimator $\hat{\mathbf{b}}$ using ordinary PLS factorization. Note that although the scores 4, 5 and 6 are nearly y -orthogonal, i.e. located in a plane perpendicular to $\hat{\mathbf{b}}$, this cannot be seen from these projections.

$$\hat{\mathbf{b}} = \begin{bmatrix} 0.0805 \\ 0.2588 \\ 0.8772 \\ -0.0092 \\ 0.0033 \\ -0.0030 \\ -0.0055 \\ -0.0636 \\ -0.0257 \\ 0.0286 \\ -0.0226 \\ 0.0120 \end{bmatrix}, \quad (31)$$

but Figure 2 also gives information on y -orthogonal properties (variable 4 gives larger y -orthogonal score movements than variable 5 etc.). Given sufficient process knowledge also the y -orthogonal scores may be informative.

For a comparison score plots for the ordinary PLS factorization are shown in Figure 3, where projections of $\hat{\mathbf{b}}$ are also included. Note that e.g. score no. 5 shows considerable

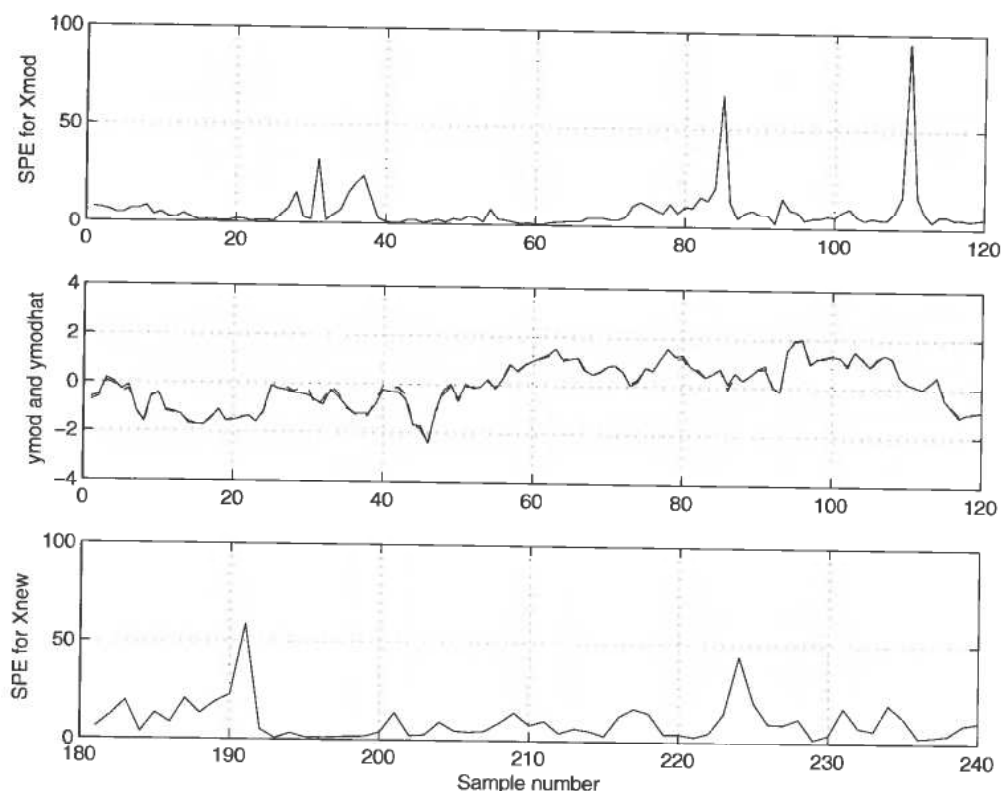


Figure 4. Squared prediction error (SPE) for modeling data (top), the corresponding values for \hat{y} and y (middle), and SPE for the new validation data (bottom).

deviations from the origin, although the deviation for this score in Figure 2 is very small. This illustrates that some y -orthogonal variations in X that in the ordinary PLS factorization are part of \hat{X} , in the 2PLS factorization are captured in \tilde{E} . Also note that it cannot be seen in Figure 3 that the scores no. 4, 5 and 6 are nearly y -orthogonal, i.e. that the estimator \hat{b} is perpendicular to a plane through these scores. Finally note that the $\hat{t}_{i,1}$ scores are the same in both Figure 2 and Figure 3, as follows from the 2PLS factorization (11) as compared with the PLS factorization (7).

6.2. Plots for process monitoring

As part of a process monitoring system, it is necessary to evaluate the SPE value according to (21), based on $\hat{X} = \tilde{T}\tilde{W}^T$ according to (11). As a basis for comparison, we may first analyze the modeling data by plotting SPE (Figure 4, top) and both y and \hat{y} (Figure 4, middle). The predictions are quite good also for samples with an SPE peak, i.e. the peaks do not indicate very special events. We then plot SPE for the new validation data (Figure 4, bottom), and conclude that these values are in no way extreme, i.e. not located further away from the 2PLS projection plane than for the modeling data. We can thus draw meaningful conclusions regarding the new samples from the score-loading plot.

After appropriate scaling and definition of the normal operating region based on historical data, the plot in Figure 2 may also be used for process monitoring purposes.

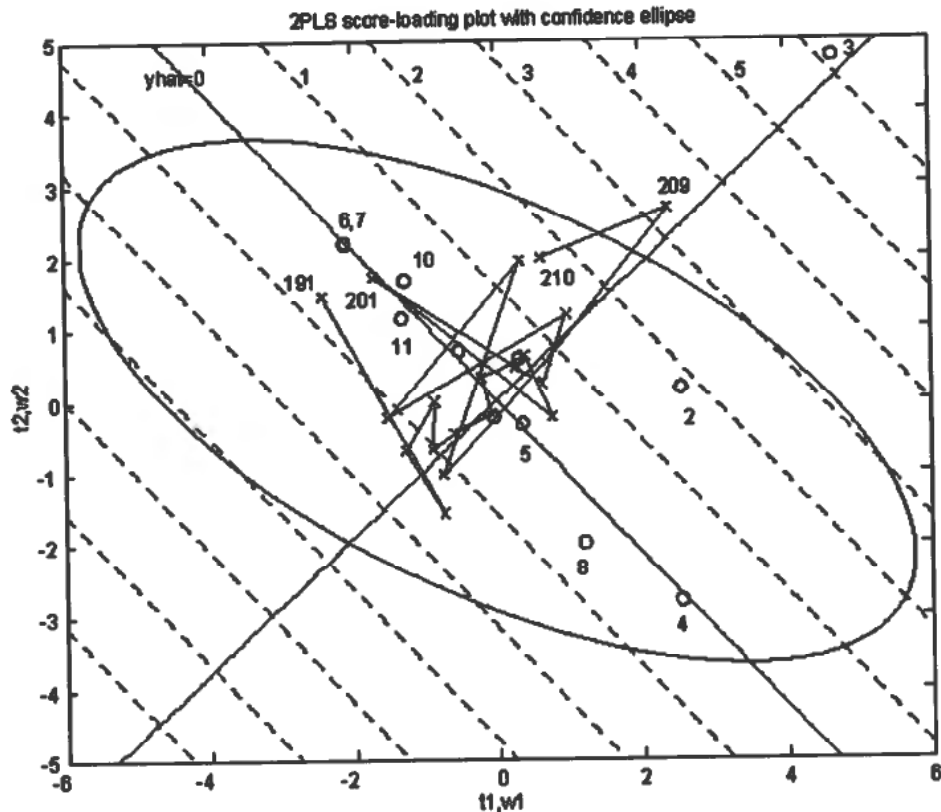


Figure 5. Score-loading plot with new scores (\times -marked) and fixed loadings from the modeling data with scaling constant 7 (o -marked). The ellipse shows T_{UCL}^2 according to (24) for $\alpha = 0.01$.

The last few of the new scores at each time instant may then be shown, indicating how the process conditions are developing, while older scores must fade away in order to give room for new information. A confidence ellipse based on the modeling data and corresponding to the T_{UCL}^2 limit (24) may also be included, as shown in Figure 5.

In order to give as clear a picture as possible the o -marked loadings based on the modeling data are here plotted using a common scaling constant 7, as compared to Figure 2. The loadings will in any case indicate the direction and relative strength of the variable attractions on the scores. A large deviation from the origin thus not only signals a special plant operation situation, but the direction also indicates which regressor variable or variables that cause the deviation. It is for example rather obvious that observation 209 outside of the normal operating region is very much influenced by an especially high positive value of variable 3. The deviations of samples 191, 201 and 210 may more clearly be caused by several variables. All of this will be clarified by use of a weighted loading plot, as shown in Figure 6 below. Note that sample 191 has the maximal SPE value in Figure 4, but that the 2PLS projection is well within the T_{UCL}^2 ellipse.

In order to clarify the information contained in the score-loading plot in Figure 5, the alternative 2PLS contribution plot according to (28) utilizing weighted loadings may be used, as shown in Figure 6 for four samples. This shows that

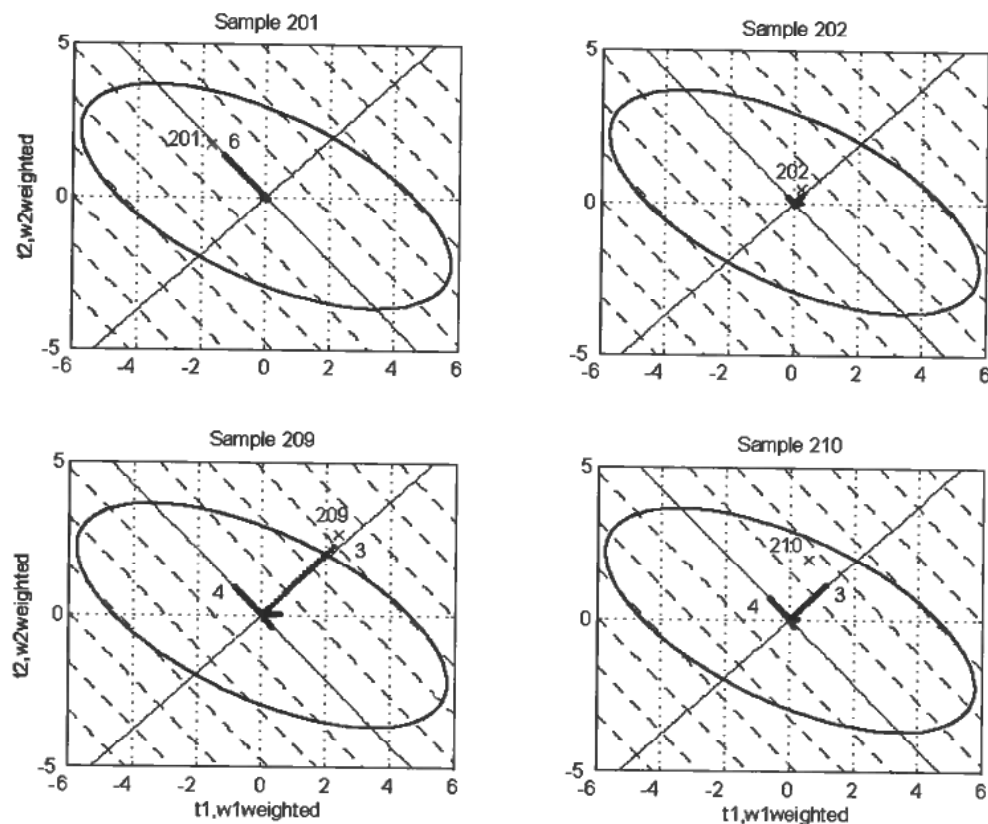


Figure 6. Score-loading plots with scores (\times -marked) and weighted loadings (solid vector lines) according to (28). Note that the score at each sample is the exact sum of the weighted loadings (the most influential ones are numbered).

- the deviation of sample 201 perpendicular to the \hat{y} axis (i.e. $\hat{y} \approx 0$) is mainly caused by a high positive value of variable 6, and not a high positive value of variable 7 or a high negative value of variable 4, as may be concluded from Figure 5
- for sample 202 all variables have small values, keeping the score close to the origin
- the deviation of sample 209 in the direction of the \hat{y} axis (i.e. a high positive value of \hat{y}) is mainly caused by a high positive value of variable 3, although variable 4 also has a high negative value, and some other variables are also involved
- the deviation of sample 210 is clearly caused by a high positive value of variable 3 and a high negative value of variable 4.

From a single plot with process scores and weighted loadings it will thus be possible

- to see a deviation from normal process operation, and the eventual violation of the upper control limit T_{UCL}^2
- to see whether and to which degree the deviation gives a change in predicted response \hat{y}
- to get some information on which variables are causing the deviation, as well as sign and magnitude.

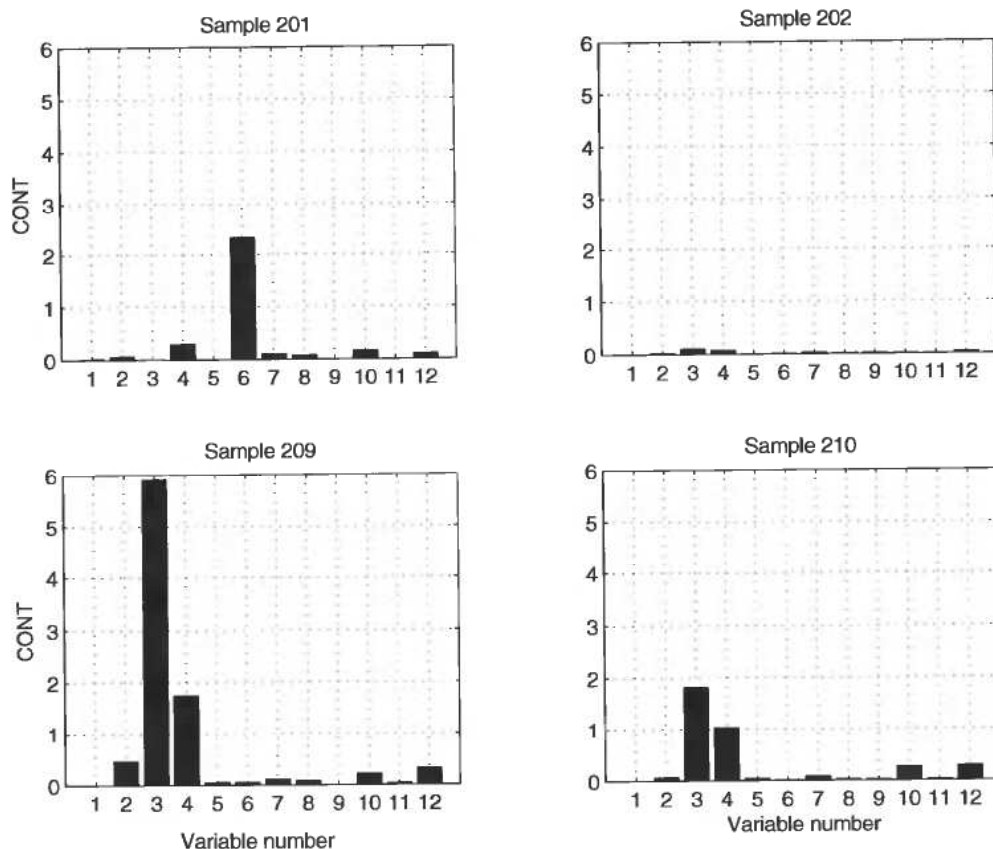


Figure 7. Contribution plots according to (27).

The corresponding traditional contribution plots according to (27) are shown in Figure 7. These plots give no new information, at least not when the number of variables is as low as in this example, and it is in fact obscured that variable 4 influences the samples 209 and 210 by having a high negative value. This is caused by both the scaling factor

$$\frac{\hat{t}_{\text{new},a}}{s_a^2}$$

in (26), and that negative contributions are set equal to zero.

7. Conclusions

Background theory on PLS modeling, compression into two-component PLS (2PLS) realizations, score-loading correspondence, monitoring statistics and contribution plots is presented. The 2PLS data compression makes it possible to construct a single dynamic and informative score-loading biplot, utilizing score-loading correspondence, as illustrated in an example using mineral processing plant data (Figure 6).

One potential use of this is to gain understanding of how and to which extent different process variables affect a specific response variable. When used in process monitoring the biplot will make it possible to see a deviation from normal process

operation and when the T^2_{UCL} upper control limit is violated, to see whether and to which degree such a deviation gives a change in the predicted response, and to judge which process variables are contributing to the deviation. In the same biplot it is also possible to see the sign and magnitude of the variables that contribute to a given deviation. Another result of the 2PLS data compression is that more of the response-orthogonal variations in the process data are captured in the residuals, and thus a more response-relevant model is obtained.

Further research will investigate fault diagnosis based on fault signatures in the score-loading biplot. It might be that some specific process faults will give scores that result in a specific trace or pattern, such that the underlying problem can be revealed at a glance.

References

- CHIANG, L.H., RUSSEL, E.L. & BRAATZ, R.D. (2001). *Fault Detection and Diagnosis in Industrial Systems*, Springer, London.
- ERGON, R. (2002). PLS score-loading correspondence and a bi-orthogonal factorization, *Journal of Chemometrics*, 16, pp. 368–373.
- ERGON, R. (2003). Compression into two-component PLS factorizations, *Journal of Chemometrics*, 17, pp. 303–312.
- HELLAND, I. S. (1988). On the structure of partial least squares regression, *Communications in statistics*, 17, 581–607.
- HODOUIN, D., MACGREGOR, J.F., HOU, M. & FRANKLIN, M. (1993). Multivariate Statistical Analysis of Mineral Processing Plant Data, *Can. Inst. Mining Bull.* 86, No. 975, 23–34.
- HÖSKULDSSON, A. (1996). *Prediction Methods in Science and Technology, Vol. 1 Basic Theory*, Thor Publishing, Copenhagen.
- JOHNSON, R.A. & WICHERN, D.W. (1992). *Applied Multivariate Statistical Analysis*, Prentice-Hall, New Jersey.
- KOURTI, T. & MACGREGOR, J.F. (1996). Multivariate SPC Methods for Process and Product Monitoring, *Journal of Quality Technology*, 28, pp. 409–428.
- MACGREGOR, J.F. & KOURTI, T. (1995). Statistical Process Control of Multivariate Processes, *Control Eng. Practice* 3, pp. 403–414.
- MARTENS, H. & NÆS, T. (1989). *Multivariate Calibration*, Wiley, New York.
- SKAGERBERG, B. & SUNDIN, L. (1993). Multidimensional monitoring of complex industrial processes, *ABB Review* 4, pp. 31–38.