



**Titre:** Analyse d'une unité de déparaffinage au solvant à l'aide des  
Title: méthodes multivariées PCA et PLS

**Auteur:** Dany Lévesque  
Author:

**Date:** 1997

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Lévesque, D. (1997). Analyse d'une unité de déparaffinage au solvant à l'aide des  
Citation: méthodes multivariées PCA et PLS [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie. <https://publications.polymtl.ca/8563/>

 **Document en libre accès dans PolyPublie**  
Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/8563/>  
PolyPublie URL:

**Directeurs de  
recherche:** Michel Perrier  
Advisors:

**Programme:** Non spécifié  
Program:

## **INFORMATION TO USERS**

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**Bell & Howell Information and Learning**  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

**UMI<sup>®</sup>**  
800-521-0600



**UNIVERSITÉ DE MONTRÉAL**

**ANALYSE D'UNE UNITÉ DE DÉPARAFFINAGE AU SOLVANT À L'AIDE DES  
MÉTHODES MULTIVARIABLES PCA ET PLS**

**DANY LÉVESQUE  
DÉPARTEMENT DE GÉNIE CHIMIQUE  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL**

**MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION  
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES  
(GÉNIE CHIMIQUE)**

**JUIN 1997**

**© Dany Lévesque, 1997**



**National Library  
of Canada**

**Acquisitions and  
Bibliographic Services**

**395 Wellington Street  
Ottawa ON K1A 0N4  
Canada**

**Bibliothèque nationale  
du Canada**

**Acquisitions et  
services bibliographiques**

**395, rue Wellington  
Ottawa ON K1A 0N4  
Canada**

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-42914-8

**Canada**

**UNIVERSITÉ DE MONTRÉAL**

**ÉCOLE POLYTECHNIQUE DE MONTRÉAL**

**Ce mémoire intitulé:**

**ANALYSE D'UNE UNITÉ DE DÉPARAFFINAGE AU SOLVANT À L'AIDE DES  
MÉTHODES MULTIVARIABLES PCA ET PLS**

**présenté par: LÉVESQUE Dany**

**en vue de l'obtention du diplôme de: Maîtrise ès science appliquées**

**a été dûment accepté par le jury d'examen constitué de:**

**M. KLVANA Danilo, Ph.D., président**

**M. PERRIER Michel, Ph.D., membre et directeur de recherche**

**M. OULIDI Abder, Ph. D., membre**

## **Remerciements**

**D'abord, je tiens à remercier M. Michel Perrier pour avoir accepté de travailler avec moi. J'ai beaucoup apprécié ses conseils toujours judicieux.**

**Je veux également remercier le professeur John F. MacGregor et le département de génie chimique de l'Université McMaster pour m'avoir permis d'utiliser leur logiciel de modélisation multivariable MACSTAT. Je remercie également M. MacGregor et la compagnie Umetri pour m'avoir permis d'assister au cours intensif sur les méthodes de modélisation multivariable qui m'a donné un bon coup de pouce dans mon projet.**

**Je tiens aussi à souligner les efforts de M. Paul-André Lavoie et M. Marc Cyr qui ont permis la réalisation des tests chromatographiques sur mes échantillons.**

**Finalement, je remercie ma famille pour m'avoir soutenu au cours de mon projet. Je veux tout spécialement remercier mon copain, Bernard, pour son soutien et son encouragement lors de mes bons et mauvais moments.**

## Résumé

L'analyse statistique multivariable d'une unité de déparaffinage au solvant (UDS) est effectuée dans ce projet. L'objectif du projet est d'identifier les paramètres d'opération de l'UDS les plus influents sur le contenu en huile de la cire à l'aide des méthodes de modélisation PCA (Analyse en Composantes Principales) et PLS (Moindres Carrés Partiels) dans le but ultime de faire le contrôle de la qualité de la cire.

L'UDS fait partie de l'usine des huiles lubrifiantes d'une raffinerie. Elle a pour fonction de séparer les paraffines des huiles lubrifiantes afin d'en améliorer les propriétés à froid. Ce procédé est composé d'une section de cristallisation de la cire aidée par l'injection de solvant (mélange de méthyl-éthyl-cétone et de toluène) et d'une section de filtration. Aucun modèle théorique n'existe quant à la cristallisation et la filtration de la cire relativement à son contenu en huile. Ceci s'explique par la complexité des propriétés de cristallisation et de filtration de la cire. De plus, il est important de mentionner que l'UDS procède quatre différents grades d'huile et de cire de façon semi-continue. Chaque grade est procédé tour à tour, à raison de deux à quatre journées de production en continu.

Afin d'atteindre l'objectif fixé, des tests échelons ont été effectués sur plusieurs paramètres d'opération susceptibles d'influencer la qualité de la cire. Deux des quatre grades d'huile et de cire produits à l'UDS ont fait l'objet de l'étude. Neuf paramètres d'opération ont été testés: la concentration du solvant de dilution, les ratios entre les cinq débits de dilution et le débit d'alimentation, les débits de solvant de lavage des filtres, les vitesses de rotation des filtres et les niveaux dans les cuves des filtres. D'autres paramètres d'opération sont considérés importants selon les différentes théories sur l'UDS, mais difficiles à manipuler ou contrôler (températures de cristallisation, pressions d'aspiration des filtres, débit d'alimentation). Comme leur variabilité est importante lors



de l'opération normale du procédé, leur impact sur la qualité de la cire pourra être identifié.

Les tests échelons ont été effectués de façon à créer de la variabilité au niveau des paramètres d'opération importants qui sont très stables lors de l'opération normale du procédé. Ils permettent aussi d'avoir une appréciation de la dynamique du procédé. De plus, ils assurent en partie d'obtenir des relations causales entre les paramètres d'opération et la qualité de la cire.

Le contenu en huile de la paraffine n'est pas mesuré en ligne. Des échantillons ont donc été prélevés aux dix minutes lors des trente journées de tests. La mesure du contenu en huile de la cire est ardue, longue et coûteuse. La mesure de l'indice de réfraction qui est rapide et peu dispendieuse a été identifiée comme alternative. La relation qui existe entre ces deux paramètres est vérifiée pour s'assurer de la validité de l'utilisation de cette mesure. Les deux mesures ont été effectuées sur deux échantillons de cire par journée de tests. Les régressions linéaires obtenues présentent des coefficients de détermination ( $R^2$ ) de 97% et 83% pour les grades 100 et 580 respectivement. Ces résultats appuient donc l'utilisation de l'indice de réfraction pour évaluer la qualité de la cire.

Un processus de prétraitement des données est nécessaire avant d'effectuer la modélisation multivariable des données recueillies. Les méthodes de modélisation PCA et PLS sont efficaces pour modéliser l'opération en régime permanent. Il a donc fallu s'assurer qu'aucune dynamique importante n'est présente dans les bases de données à l'aide du calcul des autocorrélations pour tous les paramètres d'opération, pour chacune des journées de tests. Les délais entre chacun des paramètres d'opération et l'indice de réfraction ont ensuite été identifiés à l'aide des corrélations croisées. Ces délais varient entre 30 et 80 minutes. Il a donc fallu aligner les données dans le temps. Finalement,

chaque paramètre a été centré sur sa moyenne et ramené à une variance de un de façon à ce que tous les paramètres soient sur un pied d'égalité face aux modèles PCA et PLS.

La modélisation PCA et PLS est ensuite entreprise de façon indépendante pour les deux grades. À l'aide de la modélisation PCA, les sources les plus importantes de variabilité au sein des bases de données sont identifiées. L'importance des tests échelons effectués sur le procédé par rapport au bruit est également vérifiée. Les modèles PLS permettent d'identifier les événements et les paramètres d'opération qui ont un effet significatif sur la qualité de la cire.

Les modèles PCA calculés pour les deux grades expliquent environ 80% de la variabilité des cinquante paramètres d'opération à l'aide de dix variables latentes. Ces modèles identifient la combinaison des filtres en opération et les températures du circuit de cristallisation comme les sources de variabilité les plus importantes dans les bases de données des deux grades. La variation du débit d'alimentation est également significative, mais moins importante que les premiers paramètres. Finalement les tests échelons sur les ratios de dilutions sont importants relativement au bruit selon les modèles PCA puisqu'ils sont modélisés par ces derniers.

Le modèle PLS calculé pour le grade 100 explique 87% de la variabilité de l'indice de réfraction à l'aide de sept variables latentes. Celui calculé pour le grade 580 explique 75% de la variabilité de l'indice de réfraction à l'aide de dix variables latentes. Ces modèles identifient le débit de filtrat recyclé au réservoir d'alimentation des filtres comme le paramètre qui a le plus d'effet sur la qualité de la cire. Il doit être maintenu à son maximum pour obtenir un contenu en huile de la cire plus faible. Le deuxième aspect de l'opération retenu par les modèles PLS est le profil de température du circuit de cristallisation. Il a un plus grand impact sur la qualité de la cire de grade 100, mais il est tout de même significatif pour le grade 580. De meilleurs résultats sur la qualité de la cire

sont obtenus lorsque les températures du circuit de cristallisation sont à leur point d'opération le plus faible. Quant aux paramètres d'opération des filtres, ils semblent avoir un plus grand impact sur la qualité de la cire de grade 580 que sur la qualité de la cire de grade 100. Le principal résultat stipule qu'un des filtres est moins efficace lors de la production de la cire de grade 580. Finalement, les tests échelons qui ont eu un effet sur l'indice de réfraction sont ceux qui ont fait varier le débit de filtrat recyclé. Ce qui appuie l'importance de ce débit sur la qualité de la cire. Selon les modèles PLS des deux grades, il est préférable de maintenir les débits de prédilution et de troisième dilution faibles. L'effet sur la qualité de la cire de la plupart des paramètres d'opération considérés importants par les modèles PLS respecte généralement celui identifié par la théorie sur les unités de déparaffinage au solvant.

Deux exemples de cartes de contrôle statistique multivariable sont aussi présentés. La carte de contrôle est basée sur une modélisation PLS des données d'opération qui présentent une bonne qualité de la cire de grade 100. La carte a été testée sur des données contenant des périodes de courts-circuits du réfrigérant aux deux premiers échangeurs. Les courts-circuits sont effectués lors de la production du grade 100 lorsque les échangeurs sont encrassés. La carte de contrôle permet de conclure que le temps requis afin que le procédé revienne à son état initial suite à cette opération varie entre 60 et 80 minutes. Ce qui confirme la nécessité de travailler sur la procédure des courts-circuits de façon à minimiser l'impact de cette opération sur la qualité des produits.

La carte de contrôle statistique multivariable a aussi été testée sur des données d'opération qui présentent un contenu en huile élevé. Les paramètres d'opération identifiés fautifs par cette carte sont en général consistants avec les conclusions tirées à partir du modèle PLS global du grade 100. Le paramètre le plus important est le débit de filtrat recyclé.

Les méthodes de modélisation PCA et PLS s'avèrent donc efficaces pour identifier les sources de variabilité de la qualité de la cire. Leur application est avantageuse puisqu'elle ne nécessite pas une connaissance théorique très poussée de l'unité. De plus, elles apportent beaucoup d'information sur le fonctionnement du procédé. D'ailleurs, les conclusions tirées à partir de l'analyse statistique peuvent facilement servir à ajuster l'opération du procédé puisque les corrections sur les paramètres d'opération suggérées se trouvent à l'intérieur des limites d'opération normale du procédé.

## **Abstract**

The subject of this project is the multivariate statistical analysis of an Oil Dewaxing Unit (ODU). The objective of the project is to identify the most influent operation parameters on the wax oil content using the Principal Components Analysis (PCA) and Partial Least Squares (PLS) multivariate modelling methods in order to control the wax quality.

The ODU is a part of the Lubricating Oil Plant in a refinery. This unit separates the wax from the lubricating oil in order to upgrade the properties of the oil at low temperatures. A crystallisation circuit, helped by the injection of a solvent (MEK-toluene), and a filtration section compose the ODU process. No theoretical models relating the process parameters of the process to the oil content of the wax exists. The complexity of the wax crystallisation and filtration properties explains this situation. It is also important to mention that the ODU processes four grades of oil and wax in a semi-continuous fashion. Each grade is processed one after another, for two to four days of continuous operation.

Several step tests on the production process were done in order to attain the objective of the project. The production of two of the four grades were the object of the study. Nine operation parameters were tested on each grade: the solvent concentration, the flow ratios between the dilutions and the process feed, the cold wash solvent flow and the rotation speed and the vat level of the filters. Other operation parameters (crystallisation temperature profile, filter's suction pressure, feed flow) are considered important in theory, but were hard to manipulate or to control. In general, their variability is important from one day of production to another. Therefore, their impact on the wax quality is likely to be identified by the models.

The step tests were done in order to create variability on the operation parameters that are usually very stable during normal operation. They also give a preview of the process dynamics. They finally assure, in part, causal relationship between the process parameters and the wax quality.

The wax oil content is not measured on line at the ODU. So, wax samples were collected every ten minutes during the thirty days of tests. The wax oil content is an arduous, long and expensive analysis. The measure of the refractive index was identified as a good alternative. The relationship between these two measurements of the wax quality was verified to justify the use of the refractive index. The two measurements were made on two samples for each day of test. Linear regressions were calculated and gave  $R^2$  of 97% and 83 % for grades 100 and 580 respectively. These results justify the use of the wax refractive index to evaluate its quality.

A pretreatment had to be done on the data collected during the tests before calculating the PCA and PLS models. The multivariate modelling methods PCA and PLS are efficient on steady state data. The presence of dynamics was verified with the autocorrelation calculation of each parameter, for every day of tests. No important dynamic periods relative to the steady state periods were identified. The delay between the process parameters and the refractive index was evaluated with the cross-correlation calculation. The delays vary between 30 and 80 minutes. The data table was adjusted in order to align the parameters' data with the refractive index data. Finally, the data was mean centered and autoscaled.

The PCA and PLS models were then calculated. With the PCA models, the most important sources of variability in the data sets were identified. The significance of the step tests relatively to the measurement noise was also verified. The PLS models identified the most influent operation parameters and events on the wax oil content.

The PCA models calculated for the two grades explain around 80% of the process variability with ten latent variables. These models identified the combination of filters in operation and the crystallisation circuit temperatures as the most important sources of variability in the data sets. The variability associated to the feed flow is also important, but with less extent than the previous parameters. The step tests on the flow ratios between the dilutions and the feed are also significant compared to the noise.

The PLS model for the grade 100 explains 87% of the refractive index variability with seven latent variables. The PLS model for the grade 580 explains 75% of the refractive index variability with ten latent variables. These models identified the recycled filtrate flow of the filters feed tank as the most influent parameter on the wax quality. It must be kept at its maximum to obtain a low oil content in the wax. The crystallisation temperature profile is the second most important operation parameter relatively to the wax quality. It has a bigger impact on the grade 100 wax quality, but is still significant to the grade 580. The best results on quality were obtained when the temperatures were at their lowest operation point. The PLS models determined that the filters' operation parameters have a more important influence on the grade 580 wax oil content. The most significant result states that one of the filters should not be used when grade 580 is in production. Finally, the most influent step tests on the wax's quality are those that provoked the recycled filtrate flow to reduce. This enhances the importance of this flow on the wax quality. The two PLS models also conclude that the predilution and third dilution flows should be kept low to improve wax quality. All these conclusions could be explained by the theory and process knowledge on the Oil Dewaxing Unit.

Two examples of multivariate statistical control chart were also presented. The chart was built with a PLS model based on data obtained for good quality grade 100 wax. The chart was tested on data containing coolant by-pass at the two first crystallisers. The by-

pass are done when the crystallisers are clogged. We could conclude with this chart that the time necessary for the process to come back to its initial state after a by-pass is between one and two hours. This procedure should be optimized to reduce its impact on the products quality.

The chart was also tested with data for which bad quality grade 100 wax results were obtained. The parameters that caused this data set to differ from the reference data set are consistent with the ones identified by the PLS models. The most important parameter is the recycled filtrate flow to the feed tank of the filters.

The multivariate statistical analysis methods were very efficient tools to identify the most influent process operation parameters on the wax oil content. Their use is interesting because no extended theoretical knowledge of the process is necessary to use them. They also give a good insight of the process. Furthermore, the recommendations based on these models are realizable since the data used to build the models represent the real process operating range.



**Table des matières**

Remerciements.....	IV
Résumé.....	V
Abstract.....	X
Table des matières.....	XIV
Liste des tableaux.....	XVIII
Liste des figures.....	XIX
Liste des sigles et abréviations.....	XXVIII
Liste des annexes .....	XXXI
1. Introduction.....	1
1.1 Mise en situation.....	1
1.2 Objectif et Méthodologie.....	3
1.2.1 Objectif.....	3
1.2.2 Méthodologie.....	3
1.3 Contenu et organisation .....	3
2. Présentation du procédé.....	5
2.1 Description de l'unité de déparaffinage au solvant .....	6
2.1.1 La cristallisation.....	8
2.1.2 La filtration.....	9
2.1.3 Récupération du solvant.....	10

<b>2.2 Particularités d'opération .....</b>	<b>10</b>
2.2.1 Changement de grade.....	11
2.2.2 Court-circuit .....	11
2.2.3 Disponibilité du solvant.....	11
2.2.4 Capacité de réfrigération .....	12
2.2.5 Perturbations possibles.....	12
<b>3. Présentation des méthodes de modélisation multivariable.....</b>	<b>14</b>
3.1 Revue de littérature.....	15
3.1.1 Quelques applications.....	16
3.2 Théorie .....	19
3.2.1 Théorie sur l'analyse en composantes principales (PCA).....	19
3.2.2 Théorie sur la méthode des moindres carrés partiels (PLS) .....	23
3.3 Élaboration d'un modèle .....	27
3.3.1 Algorithme NIPALS .....	27
3.3.2 Validation croisée .....	30
3.4 Analyse des résidus .....	32
3.5 Particularités des données historiques.....	34
<b>4. Plan d'expérimentation.....</b>	<b>38</b>
4.1 Portée de l'étude multivariable .....	38
4.1.1 Choix des grades étudiés .....	38
4.1.2 Paramètres d'opération mesurés.....	38



5.4.3 Données avec courts-circuits.....	142
5.4.4 Données avec un indice de réfraction élevé.....	144
6. Conclusion et Recommandations.....	146
6.1 Conclusion.....	146
6.2 Recommandations.....	149
Références .....	150
Annexes .....	153

**Liste des tableaux**

<b>Tableau 4.1.1: Liste des variables échantillonnées.....</b>	<b>39</b>
<b>Tableau 5.1.1: Distribution des données pour le grade 100 .....</b>	<b>51</b>
<b>Tableau 5.1.2: Statistiques de l'analyse PCA sur l'espace X du grade 100 .....</b>	<b>52</b>
<b>Tableau 5.1.3: Données statistiques de l'analyse PLS sur l'espace Y et l'espace X, grade 100 .....</b>	<b>92</b>
<b>Tableau 5.1.4: Liste des paramètres d'opération qui influencent la qualité de la cire de grade 100.....</b>	<b>118</b>
<b>Tableau 5.2.1: Distribution des données pour le grade 580 .....</b>	<b>125</b>
<b>Tableau 5.2.2: Statistiques de l'analyse PCA sur l'espace X du grade 580 .....</b>	<b>126</b>
<b>Tableau 5.2.3: Données statistiques de l'analyse PLS sur l'espace Y et l'espace X, grade 580 .....</b>	<b>130</b>
<b>Tableau 5.2.4: Liste des paramètres d'opération qui influencent la qualité de la cire de grade 580.....</b>	<b>133</b>
<b>Tableau 5.4.1: Présentation des données utilisées .....</b>	<b>141</b>

## Liste des figures

Figure 2.1: Schéma de l'usine des huiles lubrifiantes.....	5
Figure 2.1.1: Schéma d'écoulement de l'UDS. ....	7
Figure 3.2.1: Matrice des paramètres d'opération et matrices du modèle PCA. ....	20
Figure 3.2.2: Illustration du principe de projection sur un système de trois dimensions.	21
Figure 3.2.3: Projection des observations sur le plan formé par les deux composantes principales. ....	22
Figure 3.2.4: Matrice $X$ estimée par les matrices $T$ et $P'$ .....	23
Figure 3.2.5: Illustration du principe de projection pour le modèle PLS. ....	25
Figure 3.2.6: Matrices du modèle PLS. ....	26
Figure 3.3.1: Algorithme NIPALS pour le modèle PCA. ....	28
Figure 3.3.2: Algorithme NIPALS pour le modèle PLS. ....	29
Figure 4.3.1: Indice de réfraction de la cire en fonction de son contenu en huile, grade 100.....	44
Figure 4.3.2: Indice de réfraction de la cire en fonction de son contenu en huile, grade 580. ....	44
Figure 4.4.1: Autocorrélation calculée pour la température de la botte du filtre C (haut) et pour l'indice de réfraction (bas) pour le test du 9 juillet 1996 sur le grade 580.....	47
Figure 4.4.2: Corrélation-croisée entre la température de la botte du filtre C et l'indice de réfraction pour le test du 9 juillet 1996 sur le grade 580.....	48

Figure 4.4.3: Délais appliqués sur les paramètres d'opération du grade 100.....	49
Figure 4.4.4: Délais appliqués sur les paramètres d'opération du grade 580.....	49
Figure 5.1.1: Somme des carrés des erreurs de prédiction sur X, modèle PCA, grade 100. .....	54
Figure 5.1.2: Évolution dans le temps du différentiel de température entre la 3 <sup>ème</sup> dilution et l'alimentation.....	54
Figure 5.1.3: Pourcentage de la variabilité des paramètres non expliquée par le modèle PCA.....	55
Figure 5.1.4: Graphique des valeurs réelles normalisées en fonction des valeurs prédites normalisées des paramètres suivants: a) débit de la deuxième dilution, b) différentiel de température entre la troisième dilution et l'alimentation, c) niveau de l'épouseur (section récupération du solvant dans la cire). ....	56
Figure 5.1.5: Exemple de graphiques des variables latentes $t_2/t_1$ et des poids $p_2/p_1$ ...	58
Figure 5.1.6 Graphique $t_2/t_1$ , modèle PCA, grade 100.....	61
Figure 5.1.7: Graphique $p_2/p_1$ , modèle PCA, grade 100.....	61
Figure 5.1.8: $t_1$ en évolution dans le temps, modèle PCA, grade 100.....	62
Figure 5.1.9: $t_2$ en évolution dans le temps, modèle PCA, grade 100.....	63
Figure 5.1.10: $t_3$ en évolution dans le temps, modèle PCA, grade 100.....	65
Figure 5.1.11: Graphique $t_3/t_1$ , modèle PCA, grade 100.....	66
Figure 5.1.12: Graphique $p_3/p_1$ , modèle PCA, grade 100. ....	66
Figure 5.1.13: Graphique $t_3/t_2$ , modèle PCA, grade 100.....	67

Figure 5.1.14: $t_4$ en évolution dans le temps, modèle PCA, grade 100.....	68
Figure 5.1.15: Graphique $t_4/t_1$ , modèle PCA, grade 100.....	70
Figure 5.1.16: Graphique $p_4/p_1$ , modèle PCA, grade 100.....	70
Figure 5.1.17: $t_5$ en évolution dans le temps, modèle PCA, grade 100.....	71
Figure 5.1.18: Graphique $t_5/t_1$ , modèle PCA, grade 100.....	72
Figure 5.1.19: Graphique $p_5/p_1$ , modèle PCA, grade 100.....	73
Figure 5.1.20: $t_6$ en évolution dans le temps, modèle PCA, grade 100.....	75
Figure 5.1.21: Graphique $t_6/t_1$ , modèle PCA, grade.....	75
Figure 5.1.22: Graphique $p_6/p_1$ , modèle PCA, grade 100.....	76
Figure 5.1.23: $t_7$ en évolution dans le temps, modèle PCA, grade 100.....	77
Figure 5.1.24: Graphique $t_7/t_1$ , modèle PCA, grade 100.....	79
Figure 5.1.25: Graphique $p_7/p_1$ , modèle PCA, grade 100.....	80
Figure 5.1.26: $t_8$ en évolution dans le temps, modèle PCA, grade 100.....	81
Figure 5.1.27: Graphique $t_8/t_1$ , modèle PCA, grade 100.....	81
Figure 5.1.28: Graphique $p_8/p_1$ , modèle PCA, grade 100.....	82
Figure 5.1.29: $t_9$ en évolution dans le temps, modèle PCA, grade 100.....	84
Figure 5.1.30: Graphique $t_9/t_1$ , modèle PCA, grade 100.....	85
Figure 5.1.31: Graphique $p_9/p_1$ , modèle PCA, grade 100.....	85
Figure 5.1.32: Graphique $t_{10}/t_1$ , modèle PCA, grade 100.....	86
Figure 5.1.33: Graphique $p_{10}/p_1$ , modèle PCA, grade 100.....	87
Figure 5.1.34: $t_{10}$ en évolution dans le temps, modèle PCA, grade 100.....	87



Figure 5.1.35: Niveau de la botte du filtre A normalisé en évolution dans le temps, grade 100. ....	89
Figure 5.1.36: Température de la botte du filtre A en évolution dans le temps, grade 100. ....	89
Figure 5.1.37: Paramètres d'opération du filtre C, prédiction en fonction de la valeur réelle. ....	91
Figure 5.1.38: Pourcentage de la variabilité des paramètres de X non expliquée par le modèle PLS.....	93
Figure 5.1.39: Somme des carrés des erreurs de prédiction de Y, modèle PLS, grade 100. ....	94
Figure 5.1.40: Graphiques $u_i/t_i$ , modèle PLS, grade 100. ....	96
Figure 5.1.41: $t_1$ en évolution dans le temps, modèle PLS, grade 100. ....	98
Figure 5.1.42: Graphique $t_2/t_1$ , modèle PLS, grade 100. ....	99
Figure 5.1.43: Graphique $w_2/w_1$ , modèle PLS, grade 100. ....	99
Figure 5.1.44: $t_2$ en évolution dans le temps, modèle PLS, grade 100. ....	101
Figure 5.1.45: $t_3$ en évolution dans le temps, modèle PLS, grade 100. ....	103
Figure 5.1.46: Graphique $t_3/t_1$ , modèle PLS, grade 100. ....	103
Figure 5.1.47: Graphique $w_3/w_1$ , modèle PLS, grade 100. ....	104
Figure 5.1.48: Graphique $t_4/t_1$ , modèle PLS, grade 100. ....	105
Figure 5.1.49: Graphique $w_4/w_1$ , modèle PLS, grade 100. ....	106
Figure 5.1.50: $t_4$ en évolution dans le temps, modèle PLS, grade 100. ....	106

Figure 5.1.51: $t_3$ en évolution dans le temps, modèle PLS, grade 580. ....	108
Figure 5.1.52: Graphique $t_3/t_1$ , modèle PLS, grade 100. ....	108
Figure 5.1.53: Graphique $w_3/w_1$ , modèle PLS, grade 100. ....	109
Figure 5.1.54: Graphique $t_4/t_1$ , modèle PLS, grade 100. ....	111
Figure 5.1.55: Graphique $w_4/w_1$ , modèle PLS, grade 100. ....	111
Figure 5.1.56: $t_6$ en évolution dans le temps, modèle PLS, grade 100. ....	112
Figure 5.1.57: $t_7$ en évolution dans le temps, modèle PLS, grade 100. ....	113
Figure 5.1.58: Graphique $t_7/t_1$ , modèle PLS, grade 100. ....	114
Figure 5.1.59: Graphique $w_7/w_1$ , modèle PLS, grade 100. ....	114
Figure 5.1.60: Influence des variables sur l'indice de réfraction.....	117
Figure 5.1.61: Débit de filtrat recyclé au réservoir d'alimentation des filtres et indice de réfraction de la cire de grade 100.....	120
Figure 5.1.62: Profil de température du cycle de cristallisation lors de la production de cire à haut indice de réfraction et à bas indice de réfraction. ....	121
Figure 5.1.63: Débit de filtrat recyclé et pression d'aspiration du filtre A. ....	123
Figure 5.2.1: Fraction de la variabilité des paramètres non expliquée par le modèle PCA .....	127
Figure 5.2.2: Somme des carrés des erreurs de prédiction de X, modèle PCA, grade 580 .....	128
Figure 5.2.3: Fraction de la variabilité des paramètres non expliquée par le modèle PLS, grade 580 .....	131

Figure 5.2.4: Somme des carrés des erreurs de prédiction de Y, modèle PLS, grade 580 .....	132
Figure 5.4.1: Somme des carrés des erreurs de prédiction pour les données avec courts-circuits .....	142
Figure 5.4.2: Contribution à l'erreur de prédiction.....	143
Figure 5.4.3: Somme des carrés des erreurs de prédiction pour les données à haut contenu en huile.....	144
Figure I.I: Résultats des analyses chromatographiques sur les différents échantillons d'huiles recueillis lorsque l'indice de réfraction de la cire était: .....	154
Figure I.II: Résultats des analyses de spectrométrie de masse (bombardement au Krypton) sur les échantillons d'huiles recueillis lorsque l'indice de réfraction de la cire était: a) élevé, b) moyen. c) faible.....	155
Figure I.III: Résultats des analyses de spectrométrie de masse (bombardement au Krypton) sur les échantillons de cire recueillis lorsque l'indice de réfraction de la cire était: a) élevé b) faible.....	156
Figure II.I: $t_1$ en évolution dans le temps, modèle PCA, grade 580. ....	158
Figure II.II: Graphique $t_2/t_1$ , modèle PCA, grade 580. ....	159
Figure II.III: Graphique $p_2/p_1$ , modèle PCA, grade 580 .....	159
Figure II.IV: $t_2$ en évolution dans le temps, modèle PCA, grade 580.....	161
Figure II.V: $t_3$ en évolution dans le temps, modèle PCA, grade 580.....	162
Figure II.VI: Graphique $t_3/t_1$ , modèle PCA, grade 580.....	163

Figure II.VII: Graphique $p_3/p_1$ , modèle PCA, grade 580.....	163
Figure II.VIII: Graphique $t_4/t_1$ , modèle PCA, grade 580.....	165
Figure II.IX: Graphique $p_4/p_1$ , modèle PCA, grade 580.....	166
Figure II.X: $t_4$ en évolution dans le temps, modèle PCA, grade 580.....	166
Figure II.XI: $t_5$ en évolution dans le temps, modèle PCA, grade 580.....	168
Figure II.XII: Graphique $t_5/t_1$ , modèle PCA, grade 580.....	169
Figure II.XIII: Graphique $p_5/p_1$ , modèle PCA, grade 580.....	170
Figure II.XIV: $t_6$ en évolution dans le temps, modèle PCA, grade 580.....	171
Figure II.XV: Graphique $t_6/t_1$ , modèle PCA, grade 580.....	172
Figure II.XVI: Graphique $p_6/p_1$ , modèle PCA, grade 580.....	172
Figure II.XVII: $t_7$ en évolution dans le temps, modèle PCA, grade 580.....	173
Figure II.XVIII: Graphique $t_7/t_1$ , modèle PCA, grade 580.....	174
Figure II.XIX: Graphique $p_7/p_1$ , modèle PCA, grade 580.....	174
Figure II.XX: Graphique $t_8/t_1$ , modèle PCA, grade 580.....	176
Figure II.XXI: Graphique $p_8/p_1$ , modèle PCA, grade 580.....	177
Figure II.XXII: $t_8$ en évolution dans le temps, modèle PCA, grade 580.....	177
Figure II.XXIII: $t_9$ en évolution dans le temps, modèle PCA, grade 580.....	178
Figure II.XXIV: Graphique $t_9/t_1$ , modèle PCA, grade 580.....	179
Figure II.XXV: Graphique $p_9/p_1$ , modèle PCA, grade 580.....	179
Figure II.XXVI: Graphique $t_{10}/t_1$ , modèle PCA, grade 580.....	181
Figure II.XXVII: Graphique $p_{10}/p_1$ , modèle PCA, grade 580.....	181

Figure II.XXVIII: $t_{10}$ en évolution dans le temps, modèle PCA, grade 580. ....	182
Figure III.I: Graphiques $u_i/t_i$ , modèle PLS, grade 580. ....	186
Figure III.II: Graphique $t_2/t_1$ , modèle PLS, grade 580. ....	188
Figure III.III: Graphique $w_2/w_1$ , modèle PLS, grade 580. ....	189
Figure III.IV: $t_1$ en évolution dans le temps, modèle PLS, grade 580. ....	189
Figure III.V: $t_2$ en évolution dans le temps, modèle PLS, grade 580. ....	191
Figure III.VI: Graphique $t_3/t_1$ , modèle PLS, grade 580. ....	192
Figure III.VII: Graphique $w_3/w_1$ , modèle PLS, grade 580. ....	193
Figure III.VIII: $t_3$ en évolution dans le temps, modèle PLS, grade 580. ....	193
Figure III.IX: Graphique $t_4/t_1$ , modèle PLS, grade 580. ....	195
Figure III.X: Graphique $w_4/w_1$ , modèle PLS, grade 580. ....	195
Figure III.XI: $t_4$ en évolution dans le temps, modèle PLS, grade 580. ....	196
Figure III.XII: $t_5$ en évolution dans le temps, modèle PLS, grade 580. ....	197
Figure III.XIII: Graphique $t_5/t_1$ , modèle PLS, grade 580. ....	197
Figure III.XIV: Graphique $w_5/w_1$ , modèle PLS, grade 580. ....	198
Figure III.XV: $t_6$ en évolution dans le temps, modèle PLS, grade 580. ....	199
Figure III.XVI: Graphique, $t_6/t_1$ , modèle PLS, grade 580. ....	199
Figure III.XVII: Graphique $w_6/w_1$ , modèle PLS, grade 580. ....	200
Figure III.XVIII: Graphique $t_7/t_1$ , modèle PLS, grade 580. ....	201
Figure III.XIX: Graphique $w_7/w_1$ , modèle PLS, grade 580. ....	201
Figure III.XX: $t_7$ en évolution dans le temps, modèle PLS, grade 580. ....	202

Figure III.XXI: $t_8$ en évolution dans le temps, modèle PLS, grade 580. ....	203
Figure III.XXII: Graphique $t_9/t_1$ , modèle PLS, grade 580. ....	203
Figure III.XXIII: Graphique $w_9/w_1$ , modèle PLS, grade 580. ....	204
Figure III.XXIV: $t_9$ en évolution dans le temps, modèle PLS, grade 580. ....	205
Figure III.XXV: Graphique $t_9/t_1$ , modèle PLS, grade 580. ....	205
Figure III.XXVI: Graphique $w_9/w_1$ , modèle PLS, grade 580. ....	206
Figure III.XXVII: $t_{10}$ en évolution dans le temps, modèle PLS, grade 580. ....	207
Figure III.XXVIII: Graphique $t_{10}/t_1$ , modèle PLS, grade 580. ....	207
Figure III.XXIX: Graphique $w_{10}/w_1$ , modèle PLS, grade 580. ....	208

## Liste des sigles et abréviations

<b>A</b>	Nombre total de dimension des modèles PCA et PLS
<b>a</b>	Numéro de la dimension
<b><math>b_k</math></b>	Coefficient du paramètre $x_k$ dans le modèle de prédiction d'Y
<b>B</b>	Matrice des coefficients $b_k$
<b><math>C_{a1-a2}</math></b>	Contribution à la variation entre deux observations pour une dimension donnée
<b>CSV/SD</b>	Critère d'arrêt du calcul itératif des modèles basé sur la validation croisée
<b>E</b>	Matrice des résidus du modèle sur la matrice X
<b>F</b>	Matrice des résidus du modèle sur la matrice Y
<b>g</b>	Numéro du groupe lors du calcul de la validation croisée
<b>G</b>	Nombre total de groupes utilisés pour faire le calcul de la validation Croisée
<b><math>h_i</math></b>	Résidu sur la relation linéaire entre $t_{ia}$ et $u_{ia}$
<b>i</b>	Numéro de l'observation (ligne) dans les matrices X et Y
<b>j</b>	Compteur dans les algorithmes NIPALS
<b>k</b>	Numéro du paramètres (colonne) dans la matrice X
<b>K</b>	Nombre total de paramètres dans la matrice X
<b>m</b>	Numéro du paramètres (colonne) dans la matrice Y
<b>M</b>	Nombre total de paramètres dans la matrice Y
<b>N</b>	Nombre total d'observations dans les matrices X et Y
<b><math>p'_a</math></b>	Vecteur des poids attribués aux K paramètres dans la $a^{ième}$ dimension
<b><math>p'_{ak}</math></b>	Poids attribués au paramètre k dans la dimension a
<b>P'</b>	Matrice ( $A \times K$ ) des poids attribués aux K paramètres pour les A dimensions

$Press_g$	Somme des carrés de l'erreur de prédiction pour le groupe $g$ lors du calcul de la validation croisée
PRESS	Somme des $Press_g$
$q'_.$	Vecteur des coefficients de corrélation entre $Y$ et $t_.$
$q'_{.m}$	Coefficient attribué à la $m^{ième}$ variable de qualité dans la dimension $a$
$Q'$	Matrice ( $A \times M$ ) des coefficients de corrélation $q'_{.m}$
$SSx_{ordonné}$	Somme des carrés des résidus sur $X$ expliqués par le modèle
$SSy_{ordonné}$	Somme des carrés des résidus sur $Y$ expliqués par le modèle
$SSy_{ordonné}(a)$	Somme des carrés des résidus sur $Y$ expliqués par le modèle pour la dimension $a$ du modèle
$SSx_{valid.croisée}$	Somme des carrés des résidus sur $X$ expliqués par le modèle, validée par la validation croisée
$SSy_{valid.croisée}$	Somme des carrés des résidus sur $Y$ expliqués par le modèle, validée par la validation croisée
$SSx$	Somme des carrés des erreurs de prédiction sur les paramètres de la matrice $X$
$SSy$	Somme des carrés des erreurs de prédiction sur les paramètres de la matrice $Y$
$SPE_x$	Somme des carrés des erreurs de prédiction sur la matrice $X$ pour chaque observation
$SPE_y$	Somme des carrés des erreurs de prédiction sur la matrice $Y$ pour chaque observation
$t_.$	Vecteur des variables latentes $t_{i.}$
$t_{i.}$	Variable latente pour l'observation $i$ dans la $a^{ième}$ dimension pour la matrice $X$
$T$	Matrice ( $N \times A$ ) des $A$ variables latentes $t_{i.}$
$u_.$	Vecteur des variables latentes $t_{.a}$



$u_{ia}$	Variable latente pour l'observation $i$ dans la $a^{ième}$ dimension pour la matrice $Y$
$U$	Matrice ( $N \times A$ ) des $A$ variables latentes $u_{ia}$
$w'_a$	Vecteur des coefficients de corrélation entre les paramètres de $X$ et $u_a$
$w'_{ak}$	Coefficient de corrélation entre le paramètre $k$ et la variable latente $u_a$
$W'$	Matrice ( $A \times K$ ) des coefficients $w'_{ak}$
$x_k$	Paramètres $k$ de la matrice $X$
$x_{ik}$	Paramètres $k$ de la matrice $X$ pour la $i^{ième}$ dimension
$X$	Matrice ( $N \times K$ ) des paramètres d'opération du procédé
$x_g$	Paramètre d'opération faisant partie du groupe $g$ lors du calcul du $Press_g$ pour le modèle PCA
$\hat{x}_g$	Prédiction des paramètres d'opération faisant partie du groupe $g$ à l'aide du modèle réduit
$y_m$	$m^{ième}$ variable de qualité de la matrice $Y$
$y_{im}$	Variable de qualité $m$ de la matrice $Y$ pour la $i^{ième}$ observation
$Y$	Matrice ( $N \times M$ ) des variables de qualité
$y_g$	Paramètre d'opération faisant partie du groupe $g$ lors du calcul du $Press_g$ pour le modèle PLS
$\hat{y}_g$	prédiction des paramètres d'opération faisant partie du groupe $g$ à l'aide du modèle réduit

## Liste des Annexes

**Annexe I: Résultats des analyses chromatographiques effectuées sur la cire et l'huile.....**

.....153

**Annexe II: Analyse détaillée des variables latentes du modèle PCA pour le grade 580 ....**

.....157

**Annexe III: Analyse détaillée des variables latentes du modèle PLS pour le grade 580 ....**

..... 183

## **1. Introduction**

### **1.1 Mise en situation**

De nos jours, de puissants ordinateurs permettent d'effectuer facilement l'acquisition d'une multitude de données d'opération à une fréquence d'échantillonnage élevée. Le premier objectif de l'acquisition de données est de recueillir autant d'information que possible sur le procédé dans le but d'améliorer son opération et sa performance. Or, une si grande quantité de données ne peut être analysée aisément. Ainsi, plusieurs paramètres mesurés sont carrément ignorés par le personnel des usines. Il en résulte une perte d'information importante.

Les méthodes statistiques multivariées PCA (analyse en composantes principales) et PLS (moindres carrés partiels) permettent d'extraire l'information importante sur le procédé à partir de tous les paramètres d'opération simultanément. Ces méthodes réduisent la dimensionnalité du problème d'analyse de données en résumant les données en un nombre restreint d'événements indépendants les uns des autres. L'utilisation de ces méthodes est de plus en plus populaire. Elles ont été utilisées pour différentes applications, et ont fait preuve d'une efficacité remarquable.

Les méthodes de modélisation multivariable PCA et PLS sont utilisées dans ce projet pour effectuer l'analyse d'une unité de déparaffinage au solvant (UDS). L'UDS fait partie de l'usine des huiles lubrifiantes des raffineries. Les huiles lubrifiantes ont longtemps été le seul produit de grande valeur des usines d'huiles lubrifiantes. Au cours des dernières années, la valeur marchande des paraffines de bonne qualité a augmentée. L'intérêt de pouvoir produire des paraffines à faible contenu en huile est donc grandissant: en plus de procurer plus d'huiles lubrifiantes, la diminution du contenu en huile de la cire augmentera sa valeur de vente. La stabilisation de la qualité des paraffines

permettra aussi de s'assurer de plusieurs clients de ce produit. En effet, selon chacun, la qualité requise du produit diffère.

L'absence de modèle théorique de la cristallisation de la cire relativement à son contenu en huile motive l'utilisation des méthodes de modélisation statistique pour mieux connaître l'UDS. La complexité des propriétés de cristallisation de la cire rend sa modélisation très ardue. Seules des modélisations empiriques ont déjà été effectuées (Tyumkin, Shabalina, Zanozina et Diskina, 1989). Elles ne sont d'ailleurs applicables qu'aux procédés qui opèrent dans les mêmes conditions.

La modélisation de la filtration de la cire est aussi difficile. Le gâteau de cire étant significativement compressible, la complexité de la modélisation de la filtration augmente substantiellement. Plusieurs paramètres de modélisation de la filtration de gâteaux compressibles sont très variables et impossibles ou très difficiles à mesurer. C'est pourquoi, la modélisation théorique de l'UDS est un exercice fastidieux.

L'UDS est donc un procédé mal connu au point de vu du contenu en huile de la cire. Par contre, il est clair que plusieurs paramètres d'opération ont un impact sur la qualité de la cire. De plus, ils sont pour la plupart disponibles pour être manipulés afin de contrôler le contenu en huile de la cire. Les méthodes multivariées peuvent apporter l'information nécessaire quant à l'impact de tous ces paramètres sur la qualité de la cire.

## **1.2 Objectif et Méthodologie**

### **1.2.1 Objectif**

L'objectif du projet est d'identifier les paramètres d'opération de l'unité de déparaffinage au solvant qui influencent le plus le contenu en huile de la cire à l'aide des méthodes de modélisation PCA (analyse des composants principaux) et PLS (moindres carrés partiels), dans le but final de pouvoir contrôler la qualité de la cire.

### **1.2.2 Méthodologie**

L'atteinte de l'objectif nécessite l'accomplissement des étapes suivantes:

- Étudier le procédé de déparaffinage au solvant;
- Identifier les paramètres d'opération susceptibles d'avoir un impact sur le contenu en huile de la cire;
- Vérifier la relation entre le contenu en huile de la cire et l'indice de réfraction;
- Réaliser des tests échelons sur les paramètres d'opération choisis;
- Analyser les données à l'aide des méthodes multivariées PCA (analyse des composants principaux) et PLS (moindres carrés partiels).

## **1.3 Contenu et organisation**

En premier lieu, le procédé de déparaffinage au solvant est présenté. Le schéma du procédé est illustré et le procédé est décrit. Quelques particularités de l'UDS sont ensuite énoncées.

La théorie relative aux méthodes de modélisation statistique multivariable utilisées au cours de ce projet est présentée au chapitre 3. Les algorithmes de calcul, les principales considérations analytiques et une revue de littérature sur les applications de ces méthodes sont présentés.

**Le chapitre 4 expose le plan d'expérience suivi lors du projet. La portée de l'étude est définie et les tests échelons réalisés à l'UDS sont présentés. De plus, quelques aspects importants quant au matériel et aux méthodes d'expérimentation et d'analyse sont énoncés.**

**Les résultats de l'analyse statistique multivariable sont présentés et analysés au chapitre 5. Ils sont ensuite discutés et expliqués en relation avec les théories connues sur l'UDS. Finalement, un exemple de carte de contrôle de la qualité multivariable qui utilise un modèle PLS comme base de comparaison est finalement présenté au chapitre 6.**

## 2. Présentation du procédé

L'usine des huiles lubrifiantes constitue un secteur de production important de la raffinerie Shell de Montréal-Est. Elle a pour fonction de produire quatre huiles de base. La production de ces huiles s'effectue à travers quatre unités. Un schéma de l'usine est présenté à la figure suivante.

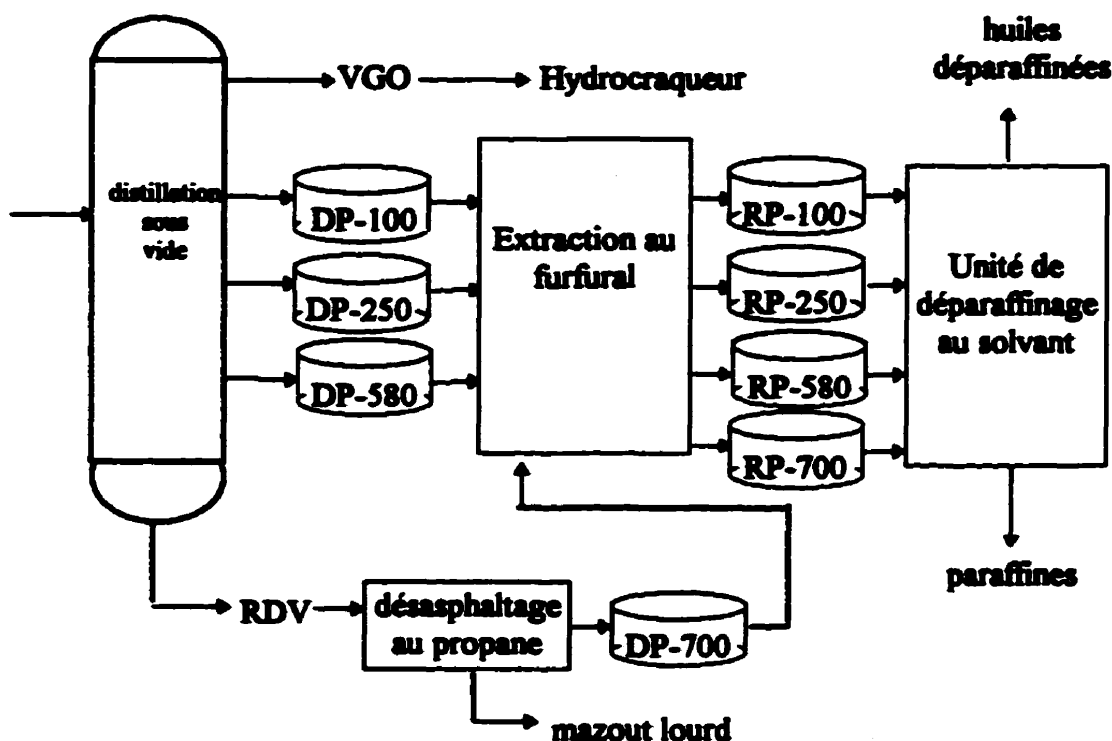


Figure 2.1: Schéma de l'usine des huiles lubrifiantes.

L'unité de distillation sous vide fractionne les produits lourds du brut paraffinique en cinq composants: VGO, DP-100, DP-250, DP-580 et RDV. Le VGO, fraction plus légère,

est acheminée à l'hydrocraqueur situé dans une autre usine de la raffinerie. Le produit de bas de colonne, le RDV, est acheminé à l'unité de désasphaltage au propane pour en extraire les asphaltènes. De l'huile lubrifiante lourde (DP-700) et du mazout lourd y sont produits. L'unité d'extraction au furfural extrait les composés aromatiques des distillats paraffineux (DP-100, DP-250, DP-580 et DP-700). Ils deviennent alors le raffinat paraffineux (RP-100, RP-250, RP-580 et RP-700). Finalement, l'unité de déparaffinage au solvant (UDS) sépare les paraffines (SW-100, SW-250, SW-580 et SW-700) des huiles lubrifiantes. Les paraffines et les huiles (HVI-100, HVI-250, HVI-580 et HVI-700) constituent les deux principaux produits de l'usine.

L'unité d'extraction au furfural et l'unité de déparaffinage au solvant procèdent chacun des grades tour à tour, à raison de deux à quatre jours de production continue du même grade. Les distillats et les raffinats paraffineux sont entreposés dans des réservoirs distincts en attendant leur traitement à ces unités.

### **2.1 Description de l'unité de déparaffinage au solvant**

L'unité de déparaffinage au solvant (UDS) sépare les paraffines des huiles lubrifiantes afin d'améliorer les propriétés des huiles à basse température. En effet, les paraffines solidifient à une température très élevée. Il est donc nécessaire de les extraire des huiles lubrifiantes afin d'en abaisser le point d'écoulement. Un faible point d'écoulement des huiles permet le bon fonctionnement des moteurs à basse température. Ce critère de qualité est très important dans les pays hivernaux, comme le Canada.

L'UDS se divise en trois sous-sections:

- la cristallisation;
- la filtration, et;
- la récupération du solvant.



La figure suivante présente le schéma d'écoulement de l'UDS.

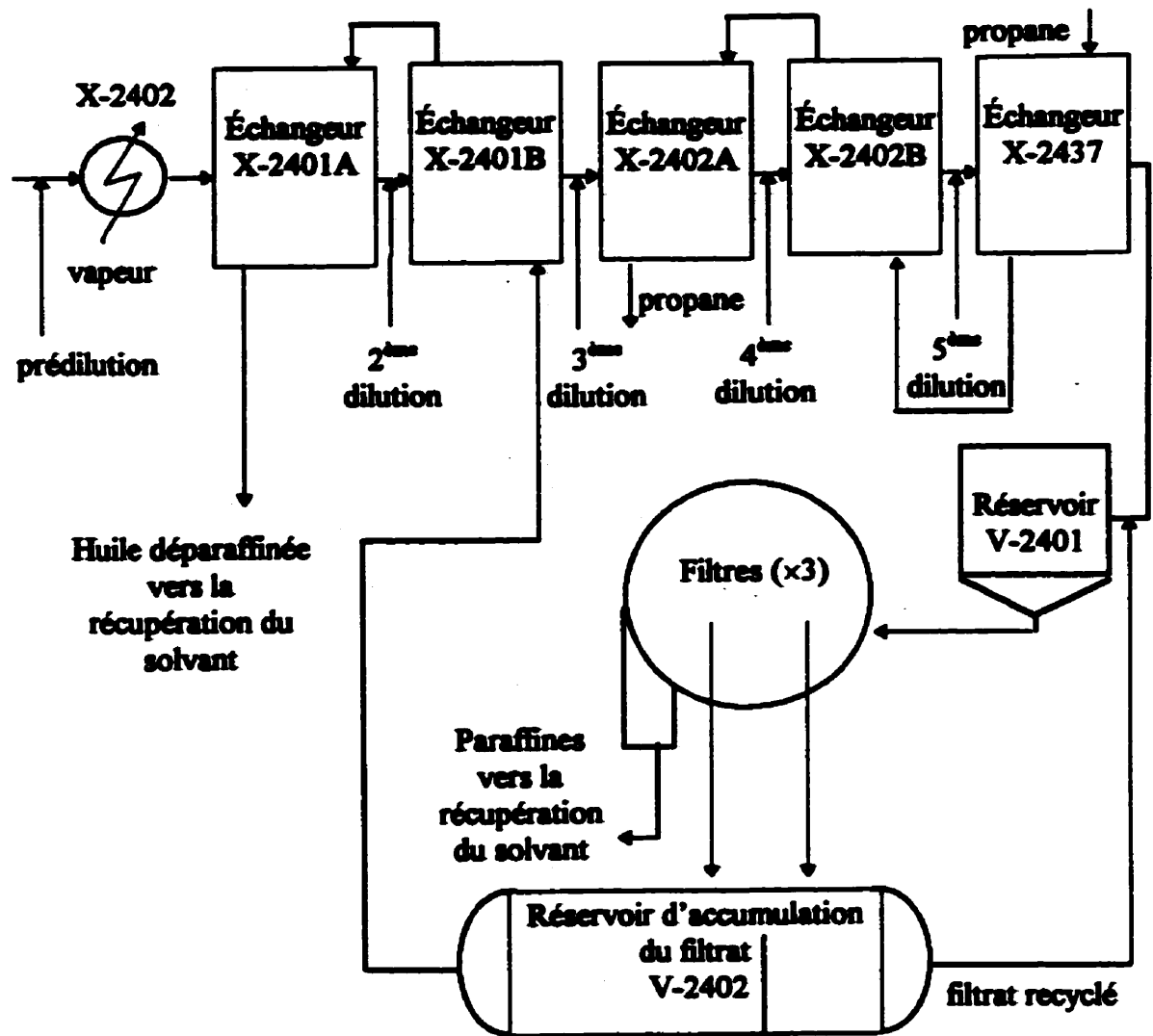


Figure 2.1.1: Schéma d'écoulement de l'UDS.

### **2.1.1 La cristallisation**

Le circuit de cristallisation du procédé est constitué de cinq échangeurs de chaleur. L'huile déparaffinée est utilisée comme liquide réfrigérant aux deux premiers échangeurs. Le propane sert de réfrigérant aux trois derniers.

Du solvant de dilution est injecté à l'entrée de chaque échangeur de façon à favoriser la cristallisation de la cire: le solvant a pour effet d'augmenter la température de cristallisation de cette dernière. Ainsi, la température de refroidissement requise pour cristalliser assez de paraffines afin d'obtenir le point d'écoulement désiré pour l'huile est moindre que sans solvant. Le solvant permet également de diminuer la viscosité du mélange d'huiles et de paraffines au cours du refroidissement, ce qui facilite l'écoulement du mélange à travers le procédé. La quantité de solvant injectée à l'entrée de chaque échangeur est fixée en ratio par rapport au débit d'alimentation à l'unité. Les ratios diffèrent d'un grade à l'autre.

Le solvant de dilution idéal doit répondre à plusieurs caractéristiques:

- pouvoir solvant très élevé pour les huiles;
- solubilité des paraffines très faible;
- température de cristallisation très faible.

Il est très difficile de trouver un solvant répondant à toutes ces caractéristiques. Un mélange est donc utilisé: méthyl-éthyl-cétone (MEK) et toluène. Le toluène a un pouvoir solvant très élevé pour les hydrocarbures. De façon à compenser l'effet solvant du toluène pour les paraffines, un antisolvant est ajouté, le MEK. La composition optimale du solvant est obtenue suite à plusieurs expérimentations en laboratoire. La quantité de MEK est fixée aussi haute que possible pour augmenter la température de cristallisation de la cire. La limite supérieure de concentration de MEK est définie par la concentration à laquelle les huiles deviennent immiscibles dans le mélange MEK-toluène.

Comme plusieurs grades d'huiles et de paraffines sont produits à l'unité, la composition optimale du solvant pour le grade le plus lourd est utilisée pour tous les grades. D'autres informations plus spécifiques sur le solvant de dilution sont disponibles dans McKetta (1989).

Suite à la cristallisation, le mélange solvant-raffinat paraffineux est acheminé dans le réservoir d'alimentation des filtres (V-2401). Le mélange est filtré presque immédiatement de façon à éviter qu'il se réchauffe et que les paraffines se solubilisent.

### **2.1.2 La filtration**

La filtration est effectuée à l'aide de trois filtres rotatifs en parallèle. Quatre filtres sont disponibles, mais seulement trois sont nécessaires à la production. Le mélange paraffines cristallisées-huiles-solvant est alimenté par gravité dans les cuves des trois filtres. Le gâteau est lavé avec le même mélange de solvant (MEK-toluène) que celui utilisé pour les dilutions. Le solvant de lavage doit être à la même température que la solution filtrée de façon à ne pas solubiliser la paraffine ou cristalliser l'huile. La quantité de solvant de lavage utilisée devrait représenter environ trois fois le volume des pores du gâteau (McCabe, Smith et Harriott (1985)).

Le gâteau de cire est décollé du baril du filtre rotatif par une lame (doctor blade) aidée d'un souffle de gaz inerte. Il est ensuite réchauffé et acheminé à la section de récupération pour séparer les paraffines du solvant de dilution.

Le filtrat riche en huile se retrouve dans le réservoir d'accumulation (V-2402). Vue sa très basse température, il servira incessamment comme source de réfrigération aux deux premiers échangeurs. Le filtrat riche en solvant est recyclé au réservoir d'alimentation des filtres (V-2401). Ce débit de recirculation est présentement manipulé pour contrôler

le niveau du réservoir. Selon McKetta (1989), il devrait plutôt être utilisé pour ajuster le ratio liquide-solide du mélange alimenté aux filtres.

Lorsque la toile de filtration est encrassée, le lavage à chaud des filtres est effectué. Les filtres sont lavés tour à tour à l'aide de solvant chaud de même composition que le solvant froid. La fréquence des lavages dépend du grade produit à l'unité. Plus le grade est lourd, plus la fréquence de lavage est élevée. Le réservoir d'alimentation V-2401 sert de réservoir tampon lors des lavages.

Les différents paramètres d'opération des filtres sont fixés selon le grade produit. La capacité de filtration représente souvent le facteur limitant de la production, surtout pour les grades les plus lourds. Les principaux paramètres de filtration sont la vitesse de rotation, la pression d'aspiration, le niveau dans les cuves, le débit de solvant de lavage, et le niveau et la température de la botte des filtres.

### **2.1.3 Récupération du solvant**

La section de récupération du solvant consiste à séparer l'huile et la cire du solvant de dilution, et à purifier ce dernier autant que possible. La récupération du solvant n'a pas d'impact sur le contenu en huile de la cire. Les résultats d'analyses de concentration du solvant effectuées par la raffinerie démontrent que le contenu en huile du solvant sec et humide est presque nul. Ce qui appuie l'hypothèse que toute l'huile contenue dans le gâteau à la sortie des filtres reste dans la cire plutôt que de s'évaporer avec le solvant lors de la séparation. Cette section de l'unité n'est donc pas analysée dans ce projet.

### **2.2 Particularités d'opération**

L'opération de l'UDS présente plusieurs particularités qui peuvent avoir un impact sur la qualité du produit. Cette section explique les principales particularités de l'opération:

changement de grade, court-circuit, disponibilité du solvant, et capacité de refroidissement. Ensuite, quelques autres sources de perturbation sont énoncées.

### **2.2.1 Changement de grade**

L'UDS est opérée de façon semi-continue. C'est-à-dire que les différents grades d'huile sont traités un à un à l'unité. Le même grade est traité pour des périodes qui varient entre deux et quatre jours. À chaque changement de grade, les paramètres d'opération de l'unité sont ajustés pour le nouveau grade. L'atteinte du régime permanent lors du changement de grade varie beaucoup d'une fois à l'autre. Ce qui constitue une source de variation inévitable de la qualité du produit final dans le réservoir de produit fini.

### **2.2.2 Court-circuit**

Lors de la production du grade 100, les températures du circuit de cristallisation sont beaucoup plus froides que pour les autres grades. Les échangeurs s'encrassent donc plus rapidement. L'encrassement est dû surtout à la présence d'eau dans le solvant humide. Cette eau cristallise et colle aux parois des échangeurs. Lorsque la perte de charge dans les échangeurs devient trop élevée, le court-circuitage du réfrigérant aux deux premiers échangeurs est nécessaire. Le court-circuit fait augmenter les températures dans le circuit de cristallisation de façon très considérable. La glace fond, mais la qualité des produits est très affectée.

### **2.2.3 Disponibilité du solvant**

Le solvant de dilution qui est séparé de la cire et de l'huile dans la section de récupération du solvant est réutilisé dans le procédé. Deux types de solvant existent à l'UDS: le solvant humide et le solvant sec. Selon la performance de la section de récupération, une quantité plus ou moins grande de solvant humide est produite. La quantité de solvant sec

correspondante est obtenue. Pour le solvant sec, il existe un réservoir tampon d'où du solvant est retiré lorsqu'il en manque dans le procédé. Le surplus de solvant sec y est aussi acheminé le cas échéant.

Il n'existe pas de réservoir tampon pour le solvant humide. Or, les opérateurs sont contraints à utiliser la quantité de solvant humide disponible. Il arrive donc qu'ils doivent alimenter plus ou moins de prédilution et de deuxième dilution pour ajuster le niveau du réservoir de solvant humide.

#### **2.2.4 Capacité de réfrigération**

L'unité de déparaffinage au solvant est très limitée par sa capacité de refroidissement. Le contrôle des différentiels de températures entre le débit d'alimentation et les dilutions et le contrôle du profil de température du circuit de cristallisation sont donc limités.

#### **2.2.5 Perturbations possibles**

La température de l'alimentation au procédé n'est pas contrôlée. Ceci constitue une source de perturbation importante puisque le profil de température de cristallisation s'en trouve modifié. Ce paramètre est très important quant à la qualité de la cire produite. La température extérieure a aussi un impact important sur le procédé, notamment sur le pouvoir réfrigérant de l'unité.

La configuration du procédé en elle-même constitue une source de propagation des perturbations. Il existe plusieurs interactions entre les paramètres d'opération du procédé. Entre autres, la réutilisation en circuit fermé du solvant a un impact sur la stabilité du procédé. Par exemple, une modification de la température du solvant froid aura un effet sur le profil de température par l'entremise des injections de solvant dans le

circuit de cristallisation. Il y aura aussi un impact sur la filtration puisque ce solvant sert également au lavage du gâteau.

L'utilisation de l'huile déparaffinée comme liquide réfrigérant aux deux premiers échangeurs est une autre source de propagation des perturbations. Elle transmet les variations de la température de filtration et de la température du solvant de lavage au circuit de cristallisation. Le profil de température de cristallisation est ainsi perturbé.

À cause de ces interactions, plusieurs paramètres d'opération de l'unité varient ensemble. L'identification de la source de variation initiale est donc difficile. Les méthodes d'analyse statistiques multivariées présentées au prochain chapitre aideront à identifier les paramètres d'opération de l'UDS qui varient ensemble ou non.

### **3. Présentation des méthodes de modélisation multivariable**

Aujourd'hui, les systèmes informatiques permettent d'effectuer l'acquisition de centaines de données d'opération à la fois, et ce à une fréquence d'échantillonnage très élevée. Toutes ces données représentent une source importante d'information sur le procédé. Malheureusement, plusieurs des paramètres mesurés et historisés sont ignorés.

En effet, pour interpréter rapidement cette quantité énorme de données de façon simultanée, des outils adéquats sont requis. Or, les méthodes de modélisation multivariable PCA (Analyse en composantes principales) et PLS (Moindres carrés partiels) permettent d'utiliser l'ensemble des données recueillies et d'en tirer le maximum d'information.

Les méthodes de modélisation multivariable PCA et PLS sont d'autant plus intéressantes qu'elles sont structurées de façon à pouvoir utiliser des bases de données plus ou moins bien conditionnées (voir la liste ci-bas), difficiles ou impossibles à traiter avec les méthodes de modélisation multivariable traditionnelles. Elles peuvent:

- modéliser une très grande quantité de paramètres d'opération;
- modéliser une base de données comportant plus de paramètres d'opération que d'observations, ou vice versa;
- modéliser une base de données dont les paramètres sont fortement colinéaires;
- modéliser une base de données comportant des données manquantes;
- modéliser des données bruitées (erreur de mesure, erreur d'analyse, ratio signal/bruit faible), et ;
- aident à identifier les données aberrantes.



La représentation graphique des résultats s'ajoute aux particularités de ces méthodes de modélisation. L'interprétation des résultats est très visuelle.

Ces méthodes sont utilisées dans plusieurs buts. Entre autres:

- effectuer le suivi d'un procédé et le diagnostic des perturbations;
- effectuer le contrôle de la qualité des produits;
- identifier et comprendre les relations entre les paramètres d'opération (PCA);
- identifier et comprendre les relations entre les variables de qualité (PCA);
- identifier et comprendre les relations entre les paramètres d'opération et les variables de qualité (PLS), et;
- optimiser le procédé.

Ce chapitre présente d'abord une brève revue de la littérature illustrant quelques applications de ces méthodes. Les aspects théoriques en relation avec le présent projet sont ensuite exposés. L'algorithme de calcul utilisé est présenté par la suite. Enfin, quelques aspects pratiques quant à l'utilisation de données historiques sont soulignés.

### **3.1 Revue de littérature**

Jolliffe mentionne dans son ouvrage que l'analyse des composantes principales est probablement la technique d'analyse multivariable la plus vieille et la mieux connue. Le premier article répertorié au sujet des composantes principales date de 1901. Il a été rédigé par Karl Pearson. La technique telle qu'elle est utilisée aujourd'hui a été élaborée par Harold Hotelling qui publia son premier article sur ce sujet en 1933. Le développement de l'analyse en composantes principales a été très actif vers la fin des années 1930 et au début des années 1940. Son évolution a ensuite été ralentie jusqu'à l'avènement d'ordinateurs permettant d'appliquer ces techniques à des problèmes de dimension raisonnable (Jackson 1991). En effet, le calcul manuel des composantes

principales devient fastidieux lorsque le nombre de paramètres à modéliser est supérieur à quatre. Or, l'analyse en composantes principales devient pratique lorsque le nombre de paramètres à modéliser est très grand. C'est pourquoi l'apparition d'ordinateurs puissants a promu le développement de telles méthodes.

La méthode de modélisation PCA a fait ses preuves dans plusieurs domaines des sciences sociales, économiques et naturelles. À titre d'exemple, elle a été appliquée aux études de démographie, en psychologie, en agriculture, en biologie, en chimie, en climatologie, en écologie, en recherche en alimentation, en géologie, et en météorologie.

La méthode des moindres carrés partiels a été développée entre la fin des années 1960 et le début des années 1980. Herman Wold a élaborer cette méthode pour effectuer la modélisation de systèmes complexes en sciences sociales. Son fils, Svante Wold, a procédé à quelques raffinements de la méthodes vers la fin des années 1970. La méthode des moindres carrés partiels s'est avérée très efficace et même supérieure aux méthodes de régression multiple et de régression par étapes (stepwise regression) dans le domaine des sciences naturelles (Wold 1989). Les méthodes de modélisation PLS ont fait leur preuve dans le domaine de la chimie, de la biologie, du génie chimique, de la métallurgie, et plusieurs autres. Les prochains paragraphes présentent quelques exemples d'application de ces méthodes.

### **3.1.1 Quelques applications**

Les méthodes de modélisation multivariable sont utilisées dans plusieurs buts et sur plusieurs types de procédés. Slama (1991) utilise les modèles PCA et PLS pour modéliser un craqueur catalytique fluidisé. Elle a fait l'identification des différentes sources de variabilité présentes lors de l'opération du procédé, et elle a construit un modèle prédictif des variables de qualité à partir des paramètres d'opération.

Kresta, MacGregor et Marlin (1991) utilisent les modèles PCA et PLS pour effectuer le suivi de la performance d'un réacteur à lit fluidisé et d'une colonne à distiller extractive (séparation du mélange azéotropique acétone-méthanol-eau). Il s'agit d'une modélisation statique des procédés, faite à partir d'une base de données générée par simulations. Elle est intéressante en ce qui a trait à la capacité des modèles PLS et PCA à modéliser des systèmes non linéaires.

Hodouin, MacGregor, Hou et Franklin (1993) ont quant à eux testé les méthodes multivariées PCA et PLS sur des données provenant d'une usine de broyage et de flottation. Ils ont vérifié l'efficacité de ces méthodes pour effectuer l'analyse préliminaire des données, classer les différents régimes opératoires, contrôler la qualité des produits et finalement, réaliser une modélisation empirique de l'opération.

La méthode des moindres carrés partiels a été comparée à la modélisation par réseaux neuronaux par Dayal, MacGregor, Taylor, Kildaw et Marcikic (1994). Ils ont appliqué ces méthodes à un lessiveur Kamyr dans une usine de pâte et papier. L'objectif était d'obtenir un modèle prédictif de l'indice Kappa (contenu en lignine de la pâte à la sortie du lessiveur). La méthode des moindres carrés s'est avérée plus intéressante que les réseaux neuronaux parce qu'elle permet entre autres d'obtenir de l'information concrète sur le procédé, et permet d'éliminer efficacement les paramètres moins significatifs pour obtenir un modèle inférentiel plus parcimonieux.

Quelques développements intéressants ont été apportés aux modèles PCA et PLS. Notamment la possibilité de former des blocs avec les paramètres lorsque le procédé se subdivise en plusieurs sections bien définies. Le suivi de la performance du procédé peut donc être fait indépendamment pour chacune des sous-sections du procédé, en parallèle ou non avec le suivi global. Cette configuration permet normalement d'identifier la

provenance des perturbations efficacement. MacGregor, Jaeckle, Kaparissides et Koutoudi (1994) ont appliqué les modèles PCA et PLS multiblocs à un réacteur tubulaire servant à produire du polyéthylène basse densité.

Quelques procédés en cuvée ont fait l'objet de la modélisation PLS. Pour ce faire, les modèles ont été légèrement modifiés: PCA et PLS multi-directions (multiway). Dans ce type de modélisation, une troisième dimension est ajoutée à la base de données analysée par les modèles PCA et PLS. L'acquisition des données est effectuée pour plusieurs cuvées et le modèle PLS est calculé à partir de toutes les données sur toutes les cuvées. Kourti, Nomikos et MacGregor (1995), Nomikos et MacGregor (1994) et Kosanovich, Dahl et Piovoso (1996) présentent des applications de ces méthodes multi-directions pour des procédés de polymérisation.

Finalement, Piovoso et Kosanovich (1994) présentent de nouvelles façons d'utiliser les méthodes de modélisation multivariable PCA et PLS. Ils utilisent la méthode PLS pour développer un modèle servant à estimer et contrôler la concentration du distillat d'une colonne à distiller binaire. De plus, ils ont développé un contrôleur par rétroaction basé sur un modèle statique PCA/PCR. Ce contrôleur est implanté sur la colonne à distiller. Les performances de ce contrôleur sont comparées à celle d'un contrôleur PI accompagné de l'estimateur PLS. En résumé, le contrôleur PCA/PCR est plus agressif que le contrôleur PI. Aussi, moins de bruit sur la variable manipulée et un temps de réponse plus rapide résultent de son utilisation.

Les possibilités des modèles PCA et PLS sont très diversifiées. Plusieurs autres applications ne sont pas mentionnées ici. Ces méthodes sont encore sujettes à évoluer énormément.

### **3.2 Théorie**

La théorie sur les méthodes multivariées PCA et PLS présentée dans ce chapitre ne touche que les aspects en relation avec le présent projet. Plusieurs documents sont disponibles pour obtenir des informations plus approfondies sur le sujet. Entre autres, les publications de Jackson (1991) et de Jolliffe (1986) font une bonne présentation de l'analyse en composantes principales. La méthode des moindres carrés partiels est très bien présentée par Martens et Naes (1991), Wold (1989) et Geladi et Kowalski (1986). Kourti et MacGregor (1995) font une bonne revue de la méthode des moindres carrés partiels et de ses possibilités d'applications.

#### **3.2.1 Théorie sur l'analyse en composantes principales (PCA)**

L'analyse en composantes principales consiste à estimer la structure de corrélation qui existe entre les paramètres d'opération. Elle identifie les paramètres d'opération qui varient ensemble. Elle résume donc les événements qui se produisent en cours d'opération.

Les  $K$  paramètres échantillonnés  $N$  fois sont placés dans une matrice nommée  $X$  de dimension  $N \times K$ . L'analyse en composantes principales approxime la matrice  $X$  par deux petites matrices  $T$  ( $N \times A$ ) et  $P'$  ( $A \times K$ ). Il s'agit d'une projection des données de  $X$  (espace de  $K$  dimensions) dans un espace de  $A$  dimensions, où  $A$  est inférieur ou égal à  $K$  et à  $N$ .

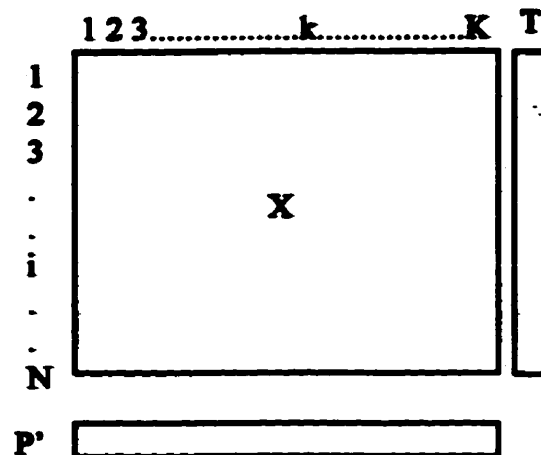
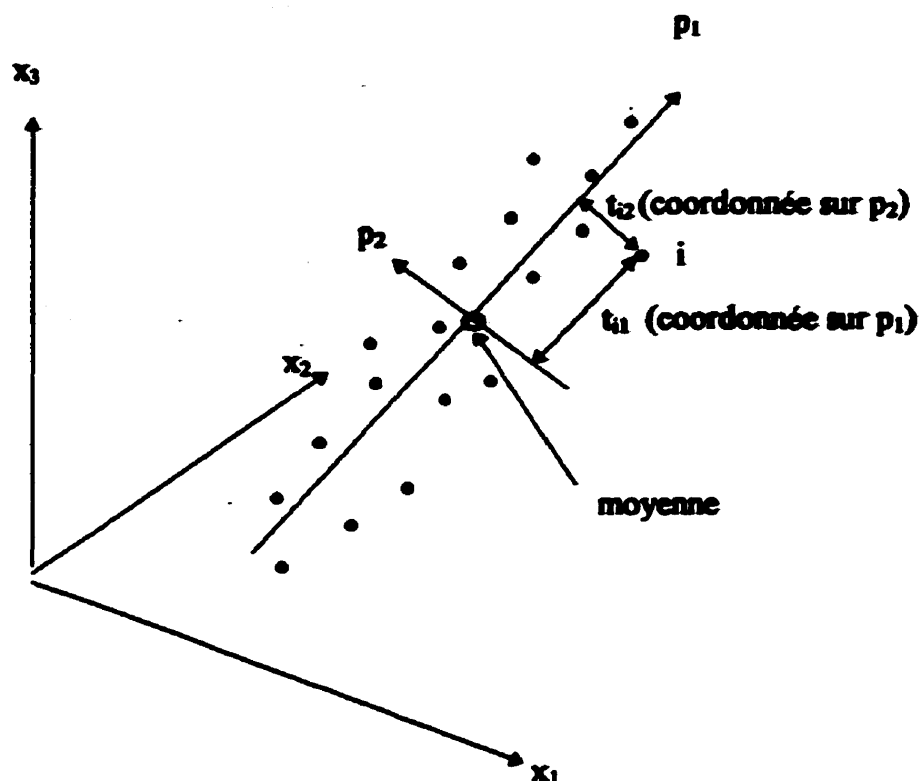


Figure 3.2.1: Matrice des paramètres d'opération et matrices du modèle PCA.

Ce principe de projection peut être facilement visualisé pour les systèmes contenant au plus trois paramètres d'opération ( $K \leq 3$ ). La figure 3.2.2 représente une base de données composée de trois paramètres et  $N$  observations. Chaque point représente une observation de la matrice  $X$  (une ligne). Chaque axe représente un paramètre de  $X$  (une colonne). L'ensemble des points représente donc l'évolution des paramètres d'opération à travers les  $N$  échantillons. L'origine des axes est placée au centre de gravité de l'ensemble des points, soit sur la moyenne de chacun des paramètres. Ce n'est pas montré sur la figure afin de conserver la clarté du dessin.

La première composante principale est représentée par une droite. Cette droite passe à travers l'origine (moyenne) et approxime le mieux possible la matrice  $X$  selon la méthode des moindres carrés. La direction de la droite est déterminée par le vecteur  $p_1$ . Il est orienté selon la variance maximale de  $X$ . Les points ( $i$ ) sont projetés sur cette droite et leurs coordonnées sont les variables  $t_{i1}$ .



**Figure 3.2.2: Illustration du principe de projection sur un système de trois dimensions.**

**La deuxième composante principale est orthogonale à la première et passe également par la moyenne. Elle améliore l'approximation de  $X$  autant que possible.**

**Les deux composantes principales forment un plan sur lequel sont projetés les observations. La distance entre les observations et leur projection sur le plan constitue le résidu du modèle. Ce résidu servira à calculer les intervalles de confiance du modèle. (voir figure 3.2.3.)**

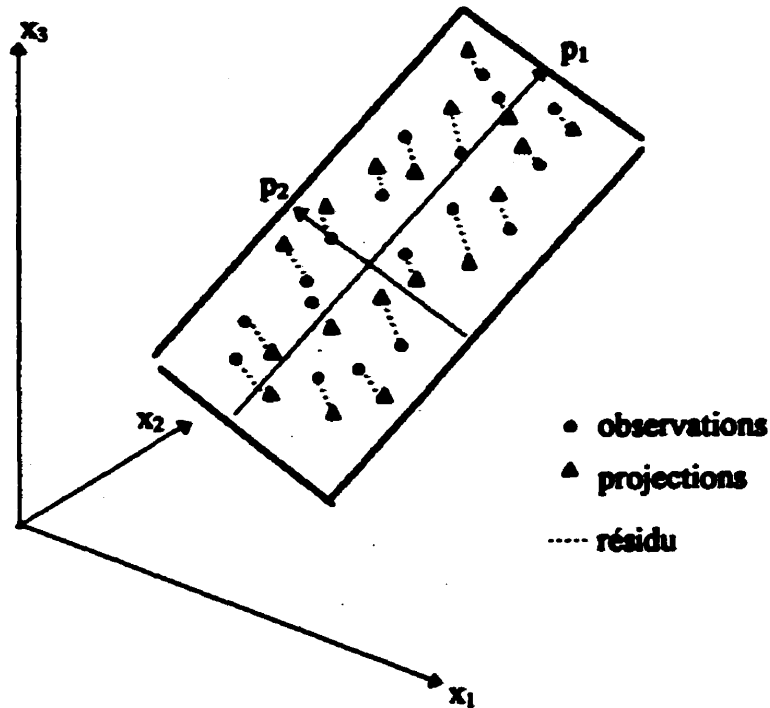


Figure 3.2.3: Projection des observations sur le plan formé par les deux composantes principales.

Ces propriétés géométriques sont conservées dans la  $K^{\text{ème}}$  dimension (Wold et al. 1987). Ainsi, le modèle PCA obtenu est de la forme suivante:

$$X = \bar{x} + t_1 \cdot p'_1 + t_2 \cdot p'_2 + t_3 \cdot p'_3 + \dots + E \quad (3.1)$$

$$X = \bar{x} + T \cdot P' + E$$

La figure 3.2.4 présente le système matriciel qui résulte de cette modélisation. Les variables latentes  $t_k$  sont de nouvelles variables qui résument les vrais paramètres d'opération. Chacune décrit un événement distinct, indépendant des autres à cause de la propriété d'orthogonalité des modèles PCA (Slama 1991). Les variables  $t_k$  qui constituent le vecteur  $t$ , sont la combinaison linéaire des paramètres  $x_k$  avec les poids  $p_{ik}$



comme coefficients. Ces derniers forment le vecteur  $p'_a$ . Les paramètres d'opération attribués d'un poids  $p_{ak}$  élevé sont ceux qui ont causé l'événement que représente la  $a^{\text{ième}}$  dimension du modèle.

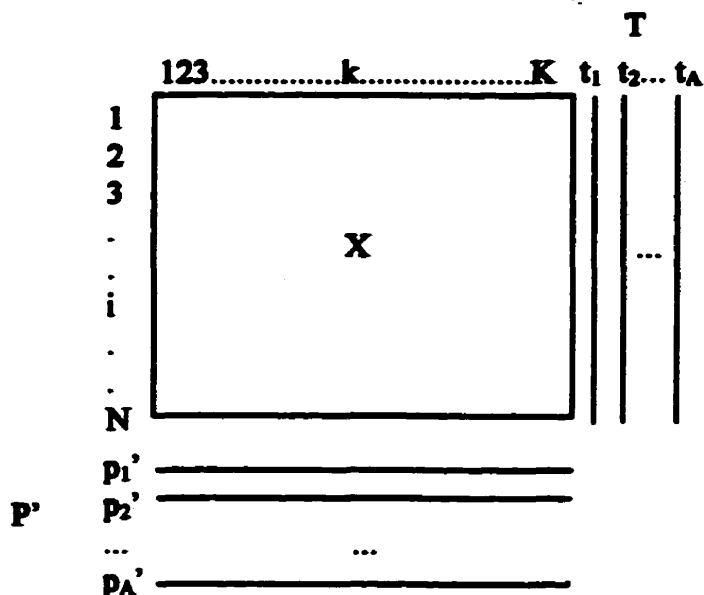


Figure 3.2.4: Matrice  $X$  estimée par les matrices  $T$  et  $P$ .

L'algorithme NIPALS (Nonlinear Iterative Partial Least Squares) est utilisé pour calculer les matrices  $T$  et  $P'$ . Cet algorithme est présenté dans la section 3.2. Le nombre optimal de variables latentes nécessaires pour modéliser la matrice  $X$  est déterminé à l'aide de la validation croisée, présentée également dans la section 3.3.

### 3.2.2 Théorie sur la méthode des moindres carrés partiels (PLS)

La méthode des moindres carrés partiels modélise les variables de qualité du procédé à l'aide des paramètres d'opération. La matrice des paramètres d'opération  $X$  ( $N \times K$ ) est la même que celle utilisée pour l'analyse en composantes principales. La matrice des variables de qualité est construite de la même façon que la matrice  $X$ : les  $M$  variables de

qualité forment les colonnes de la matrice Y tandis que les N observations constituent les lignes de la matrice. À chaque ligne  $i$  de la matrice Y est associée la ligne  $i$  de la matrice X.

La méthode PLS repose sur les mêmes théories que l'analyse en composantes principales. Elle diffère un peu puisque le modèle obtenu doit à la fois modéliser les espaces X et Y et expliquer de façon optimale l'espace Y. Pour arriver à cette fin, des concessions sont faites au point de vue de la modélisation de l'espace X. C'est-à-dire, que l'approximation de X ne sera pas optimale, mais plutôt ajustée de façon à mieux expliquer Y. Une légère rotation de l'axe  $w$  par rapport à la direction de variance maximale en résulte (Geladi et Kowalski 1986). Par exemple, la figure 3.2.5 présente les deux espaces X et Y.

La ligne pointillée dans l'espace X indique la direction de variance maximale de la matrice X. Les observations sont centrées sur leur moyenne à priori, de façon à situer l'origine des axes au centre de gravité de l'ensemble de points. La première composante PLS est une droite dans l'espace X ( $w_1$ ) et une droite dans l'espace Y ( $q_1$ ). Ces droites sont celles qui approximent le mieux les données de X et Y tout en faisant en sorte que  $t_1$  et  $u_1$  soient le mieux corrélées:

$$u_{i1} = b_1 t_{i1} + h_i \quad \text{où } h_i \text{ est le résidu} \quad (3.2)$$

La seconde composante PLS est représentée par une autre droite ( $w_2$  et  $q_2$ ) dans chacun des espaces X et Y qui sont orthogonales aux droites de la première composante. Cette composante améliore autant que possible la corrélation entre X et Y.

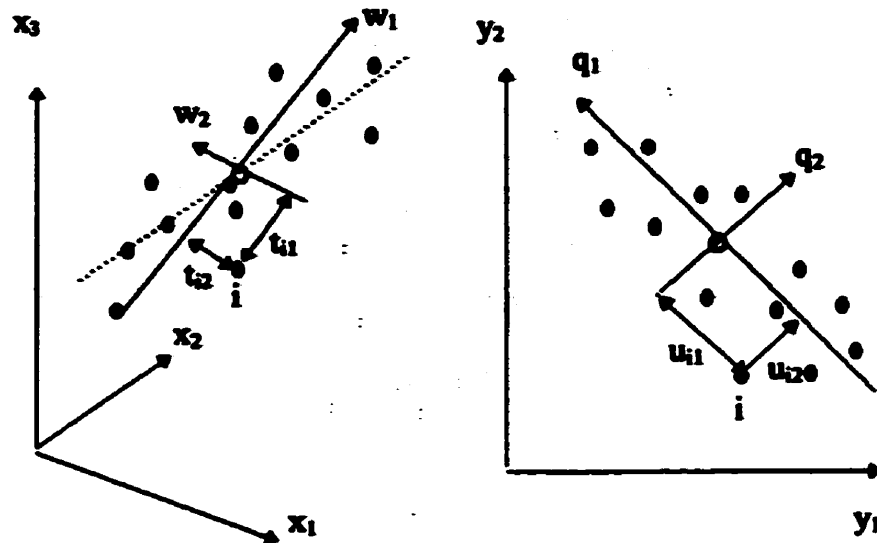


Figure 3.2.5: Illustration du principe de projection pour le modèle PLS.

Les deux composantes forment des plans dans chacun des espaces. Le modèle PLS consiste ainsi à projeter les observations sur ces plans. La distance entre les observations et le plan dans l'espace  $X$  sert à calculer l'intervalle de tolérance du modèle quant aux données qui pourront être utilisées pour faire de la prédiction. En effet, la prédiction ne peut être faite que pour des données similaires aux données de référence. La distance entre le plan de l'espace  $Y$  et les observations sert à calculer l'intervalle de confiance de la corrélation entre  $X$  et  $Y$ .

Le modèle obtenu est présenté schématiquement à la figure suivante. Les vecteurs  $p'_A$  sont calculés à partir des variables latentes  $t_A$  et de la matrice  $X$ :  $p'_A = \frac{t'_A \cdot X}{t'_A \cdot t_A}$ . Ils sont orthogonaux et  $t \cdot p'$  représente la meilleure approximation de  $X$ . Les variables latentes  $t_A$  résument les paramètres de  $X$  qui sont corrélés avec  $Y$ . Les poids  $w'_A$  sont les

coefficients de corrélation entre les paramètres  $x_k$  et les variables latentes  $u_i$ . Les colonnes de  $X$  fortement corrélées avec  $Y$  sont donc attribuées de poids  $w'_k$  élevés.

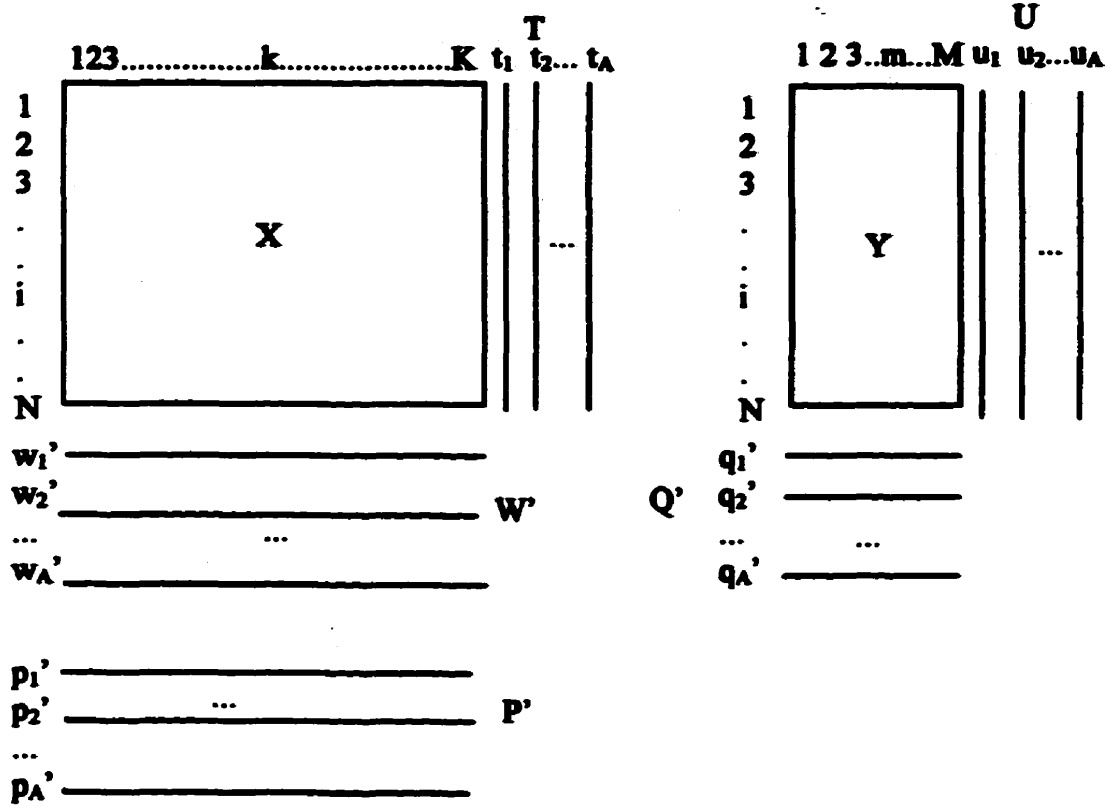


Figure 3.2.6: Matrices du modèle PLS.

Le modèle mathématique obtenu (Wold, Kettaneh-Wold et MacGregor 1996) est donc:

$$\begin{aligned}
 X &= \bar{x} + T \cdot P + E \\
 Y &= \bar{y} + U \cdot Q + F \\
 Y &= \bar{y} + X \cdot B + F \\
 B &= W \cdot (P'W)^{-1} \cdot Q
 \end{aligned}
 \tag{3.3}$$

E et F sont les matrices de résidus sur X et Y. L'algorithme NIPALS quelque peu modifié est utilisé pour calculer le modèle PLS. Le nombre de dimensions optimal pour décrire Y est aussi déterminé par la validation croisée.

La prochaine section présente les algorithmes de calcul NIPALS pour les deux méthodes multivariées discutées ici. La validation croisée y est également présentée.

### **3.3 Élaboration d'un modèle**

#### **3.3.1 Algorithme NIPALS**

Il existe plusieurs façons de calculer les composantes principales et les composantes PLS. L'algorithme NIPALS est probablement le plus utilisé. Il est présenté pour le calcul des modèles PCA et PLS.

Le terme NIPALS signifie «Nonlinear Iterative Partial Least Squares». Cet algorithme calcule les composantes principales une à une. La première dimension est d'abord calculée à partir de la matrice X pour le modèle PCA et à partir de X et Y pour le modèle PLS. Les dimensions subséquentes sont calculées à partir des matrices des résidus sur X et sur Y suite à leur modélisation à l'aide des dimensions précédentes.

##### **3.3.1.1 Algorithme NIPALS pour le modèle PCA**

L'algorithme NIPALS pour le modèle PCA calcule de façon itérative les composantes principales qui décrivent X. La dimension finale du modèle PCA notée A est déterminée par la validation croisée. Le schéma présenté à la figures 3.2.1 illustrent cet algorithme.

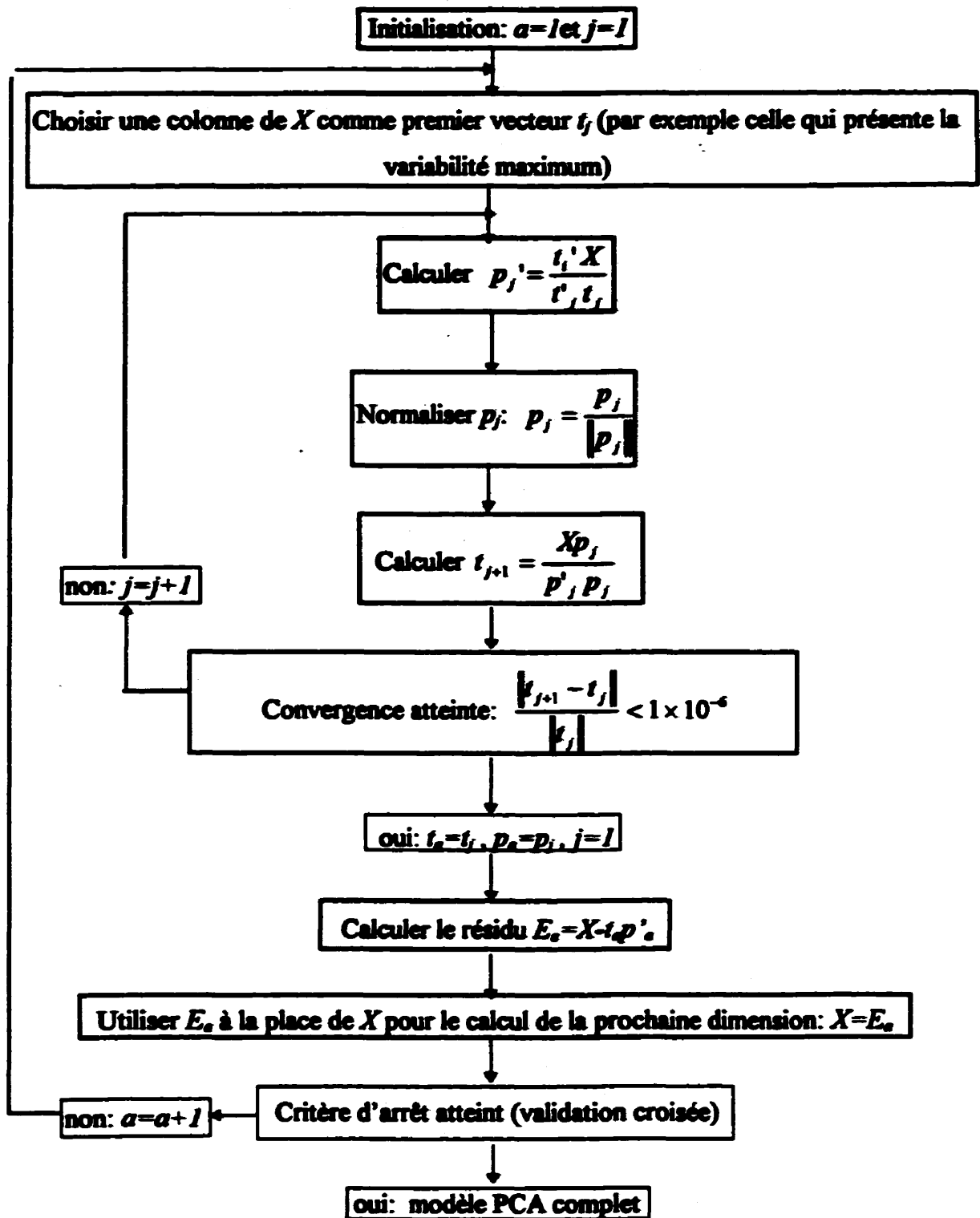


Figure 3.3.1: Algorithme NIPALS pour le modèle PCA.

### 3.3.1.2 Algorithme NIPALS pour le modèle PLS

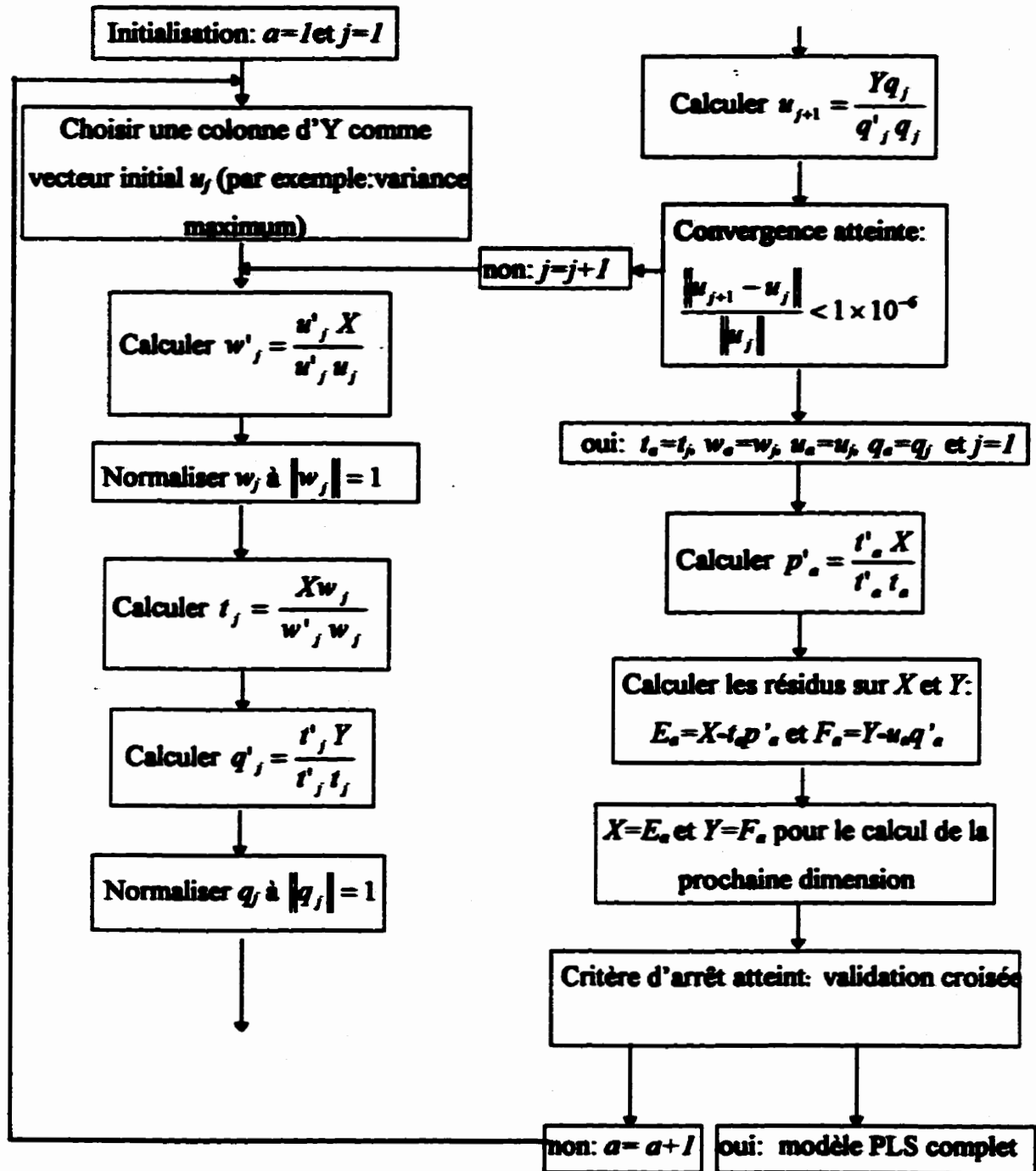


Figure 3.3.2: Algorithme NIPALS pour le modèle PLS.

L'algorithme NIPALS pour le modèle PLS calcule les variables latentes  $t$  et  $u$  simultanément. Il est conçu de façon à ce qu'il y ait échange d'information entre l'espace  $X$  et l'espace  $Y$ . Une légère rotation des composantes  $p$  du modèle PCA en résulte pour obtenir une meilleure approximation de  $Y$ . Pour plus d'informations sur l'algorithme, l'article de Geladi et Kowalski (1986) est une bonne référence.

### **3.3.2 Validation croisée**

La détermination de la dimension du modèle est un aspect important de la modélisation multivariable. L'objectif de la modélisation est d'obtenir un modèle qui a un bon pouvoir prédictif, tout en étant parcimonieux quant au nombre de dimensions utilisées.

La validation croisée est communément utilisée pour déterminer la dimensionnalité des modèles PCA et PLS. Elle permet d'évaluer quelle fraction des données représente le «signal» et quelle fraction représente le «bruit» (Wold 1978).

#### **3.3.2.1 Validation croisée pour le modèle PCA**

La validation croisée consiste d'abord à former de façon aléatoire  $G$  groupes de données dans la matrice  $X$ . Typiquement,  $G$  est compris entre 4 et 7. Dans chaque groupe, les données doivent être dispersées à travers la matrice de façon à ce qu'aucune des lignes ou colonnes n'ait la majorité de ses données sélectionnées. Chacune des données de la matrice ne peut faire partie que d'un seul groupe. Ensuite, un des groupes de données est retiré de la matrice  $X$ . Les nouveaux vecteurs  $p_{A+1}$  et  $t_{A+1}$  sont calculées à partir de la matrice réduite ( $A$  étant la dimension actuelle du modèle). Ensuite, les données manquantes sont prédites à partir du modèle de  $(A+1)$  dimensions et la somme des carrés des erreurs de prédiction du modèle pour les données extraites est calculée:



$$Press_g = \sum \sum (x_{ig} - \hat{x}_{ig})^2 \quad (3.4)$$

où  $x_{ig}$ : donnée extraite

$\hat{x}_{ig}$ : donnée extraite prédite par le modèle

Press: Prediction Error Sum of Squares

La matrice X est reconstituée à son état initial, et la procédure est reprise pour le groupe de données suivant, jusqu'à ce tous les G groupes de données aient été retirés. L'erreur de prédiction totale est alors calculée:

$$PRESS = \sum_{g=1}^G Press_g \quad (3.5)$$

Le ratio  $\frac{CSV}{SD} = \sqrt{\frac{PRESS}{SSx}}$  est utilisé de façon intuitive comme critère d'arrêt de l'algorithme NIPALS (Slama 1991).  $SSx$  est l'erreur de prédiction du modèle de A dimensions. Lorsque ce ratio est inférieur à 1, la dimension (A+1) est conservée dans le modèle: elle réduit la somme des carrés des erreurs de prédiction de façon significative. Si non, le processus itératif de l'algorithme est arrêté. Le modèle PCA est complet à A dimensions.

### 3.3.2.2 Validation croisée pour le modèle PLS

La validation croisée pour le modèle PLS consiste à extraire G groupes d'observations (lignes) de la matrice X de façon aléatoire, avec les lignes correspondantes dans la matrice Y. Les nouveaux paramètres du modèle  $t_{A+1}$ ,  $u_{A+1}$ ,  $w_{A+1}$ ,  $q_{A+1}$  et  $p_{A+1}$  sont alors calculés à l'aide des observations restantes. Les variables y extraites sont ensuite prédites

à l'aide des observations de  $X$  extraites et du modèle de  $A+1$  dimensions. Le carré de l'erreur de prédiction est ensuite calculé:

$$Press_g = \sum \sum (y_g - \hat{y}_g)^2 \quad (3.6)$$

où  $y_g$ : variable de la matrice  $Y$  extraite

$\hat{y}_g$ : variable prédite à l'aide du modèle

La statistique  $Press_g$  est calculée pour les  $G$  groupes d'observations. Une fois que tous les  $G$  groupes ont été retirés une fois, le carré de l'erreur de prédiction total est calculé:

$$PRESS = \sum_{g=1}^G Press_g \quad (3.7)$$

Le critère d'arrêt est ensuite calculé de la même façon que pour le modèle PCA. La première dimension pour laquelle le critère CSV/SD est supérieur ou égale à 1 est rejetée, et le modèle final est composé de l'ensemble des  $A$  premières dimensions.

### 3.4 Analyse des résidus

Les résidus contiennent beaucoup d'information sur les données modélisées. Une appréciation du modèle est obtenue par l'analyse des ces résidus. Trois types d'erreurs de prédiction peuvent être calculées:

- la somme des carrés des erreurs de prédiction sur la matrice:  $SSx_{\text{matrice}}$  et  $SSy_{\text{matrice}}$ ;
- la somme des carrés des erreurs de prédiction sur les paramètres:  $SSx$  et  $SSy$ , et;
- la somme des carrés des erreurs de prédiction sur les observations:  $SPEx$  et  $SPEy$ .

La somme des carrés de l'erreur de prédiction sur la matrice X dans le cas du modèle PCA et sur la matrice Y dans le cas du modèle PLS est un indicateur de la capacité du modèle à prédire cette matrice. La somme des carrés des erreurs de prédiction sur la matrice X dans le cas du modèle PLS indique la fraction de la matrice X qui est utilisée pour prédire la matrice Y. Les résidus sur la matrice X servent aussi à calculer l'intervalle de tolérance à l'intérieur duquel de nouvelles données doivent se situer pour être considérées similaires à celles de la base de données de référence. En effet, les modèles PCA et PLS ne sont valables comme modèles de prédiction que pour des données similaires à la base de données de référence. Les résidus sur les observations de Y servent quant à eux à calculer l'intervalle de confiance du modèle PLS.

La somme des carrés des erreurs de prédiction sur les paramètres permet d'identifier les paramètres les moins bien modélisés dans le cas du modèle PCA. Elle indique quels paramètres participent à l'explication de Y dans le cas du modèle PLS.

La somme des carrés des erreurs de prédiction sur les observations identifie les observations les moins bien modélisées par les modèles. Les observations pour lesquelles les résidus sont trop élevés seront identifiées comme aberrantes. Elles pourront être extraites de la base de données ou corrigées, selon le cas.

Une fois que le modèle est satisfaisant, c'est-à-dire que son pouvoir prédictif est optimisé et que les données aberrantes sont enlevées si nécessaire, l'analyse des résultats de modélisation constitue la prochaine étape du travail. En effet, les modèles PCA et PLS ne sont pas une fin en soi mais bien un moyen de retirer de l'information du procédé. La capacité d'analyser ces données et de distinguer les liens causaux des relations de corrélation est cruciale. L'utilisation de données historisées par un système d'acquisition de données industriel pour construire un modèle prédictif présente plusieurs particularités qui sont énoncées dans la prochaine section.

### **3.5 Particularités des données historiques**

Les banques de données historisées par les systèmes d'acquisition de données industriels sont une source abondante d'informations sur le procédé. L'utilisation de ces données pour construire des modèles prédictifs fait appel à plusieurs considérations importantes.

#### **3.5.1.1 Dynamique du procédé**

Le premier aspect important à considérer pour les modèles PCA et PLS est la présence de périodes dynamiques dans la base de données. La dynamique procure de l'information très intéressante quant au comportement du procédé, mais complique sa modélisation. Elle introduit souvent des relations non linéaires entre les paramètres d'opération. Or, les méthodes PCA et PLS sont des modèles linéaires, efficaces pour modéliser l'opération en régime permanent. La présence de périodes dynamiques dans la base de données nécessite donc des dimensions supplémentaires pour réaliser la modélisation. De plus, si la dynamique change selon la zone d'opération du procédé, ou qu'elle est très différente pour chacune des variables, la structure de corrélation sera perturbée, ce qui aura des conséquences plus ou moins graves sur le pouvoir prédictif du modèle (Slama 1991).

Le calcul des autocorrélations a été effectué de façon à s'assurer que la proportion de périodes dynamiques n'est pas trop importante. Cet aspect du travail est discuté dans le prochain chapitre.

#### **3.5.1.2 Les délais**

La présence de délais importants entre les paramètres d'opération du procédé peut avoir un impact significatif sur le modèle. En effet, si les périodes de délai sont grandes par rapport aux périodes d'opération en régime permanent, les modèles PCA et PLS

représenteront moins bien le comportement du procédé. Dans cette situation, il est nécessaire d'aligner les données. Les corrélations croisées ont été utilisées pour vérifier la présence des délais et les évaluer.

#### **3.5.1.3 Instrumentation**

L'instrumentation joue également un rôle important. En effet, au cours de la période d'échantillonnage, l'ajustement des instruments de mesure se détériore, ou encore, ils sont recalibrés. Ces changements seront classifiés par les modèles PCA et PLS comme des événements. Certains d'entre eux seront identifiables lors de l'analyse, mais la plupart ne le seront pas. Ceci s'ajoute au problème d'analyse des résultats.

#### **3.5.1.4 Boucles de contrôle**

Les variables contrôlées d'un procédé ne varient pas beaucoup en général, à moins qu'un changement majeur désiré par les opérateurs ne survienne. Ces variables ont habituellement un effet important sur les variables de qualité. Or, les boucles de contrôle restreignent leur variabilité à l'intérieur d'une zone qui n'aura aucun impact sur ces dernières. Elles ne seront donc pas considérées significatives par le modèle PLS. Donc, c'est l'impact des variables non contrôlées sur la qualité des produits qui est modélisé.

De plus, certaines boucles de contrôle peuvent introduire des relations inverses aux relations réelles entre les variables de qualité et les paramètres d'opération. Cette situation subvient, par exemple, pour une colonne à distiller dont la concentration du distillat est contrôlée à l'aide du débit de reflux. La relation réelle bien connue dit que la pureté du distillat augmente lorsque le débit de reflux augmente. Or, à cause de la boucle de contrôle, on observera le contraire: afin de maintenir le niveau de pureté à son point de consigne, lorsque la pureté diminue, le contrôleur augmentera le débit de reflux.

Lorsque la pureté augmente, le contrôleur diminuera le débit de reflux. Dans ce cas, c'est l'effet de rétroaction qui est modélisé, pas le procédé.

### **3.5.1.5 Colinéarité des paramètres**

La colinéarité entre les paramètres d'opération permet souvent d'obtenir un modèle prédictif très efficace vu la grande source d'informations que chaque paramètre procure. Par contre, l'analyse d'un tel modèle peut être très ardue. En effet, lorsqu'un événement se produit, et que tous les paramètres colinéaires varient ensemble, les modèles PCA et PLS distribuent le poids de cet événement également entre tous les paramètres. Plus il y a de paramètres colinéaires, plus leur poids respectif est petit. Ainsi, d'importants événements peuvent être réduits en importance, ou passer complètement inaperçus à cause de cet effet de colinéarité. Donc, si un paramètre est attribué d'un poids élevé, c'est qu'il est relié à un événement considérable. Par contre, rien ne peut être conclu quant aux paramètres attribués de poids plus petits.

### **3.5.1.6 Relation de cause à effet**

Lors de l'analyse d'un modèle multivariable PLS construit à partir d'une base de données historisée, il faut faire attention aux relations identifiées entre les paramètres d'opération et les variables de qualité. En effet, certains paramètres considérés importants par le modèle ne peuvent être en fait que le reflet d'un événement caché. Il se peut qu'aucune relation causale n'existe entre ce paramètre et les variables de qualité. Ce type de modèle ne peut donc pas servir directement pour optimiser le procédé. Afin de s'assurer de liens causaux entre les variables de qualité et les paramètres d'opération, il faut utiliser des plans d'expérience qui consistent à faire varier plusieurs paramètres à la fois, et garder aussi constants que possible tous les paramètres qui ne sont pas testés.

Lorsque l'utilisation de tels plans est impossible, ce qui est fréquent sur les unités de production industrielles, le modèle PLS peut être utilisé pour se diriger dans la bonne voie. Toutes les connaissances du procédé doivent être mises à contribution dans cette éventualité.

Une fois que ces aspects sur les données historisées sont bien connus, l'analyse des résultats de modélisation sera plus éclairée. La prochaine section fait un bref historique des méthodes de modélisation multivariables PCA et PLS et fait une revue littéraire de quelques applications retrouvées dans la littérature.

## **4. Plan d'expérimentation**

### **4.1 Portée de l'étude multivariable**

La portée de l'analyse a été déterminée suite à l'étude du procédé de déparaffinage au solvant. Deux grades de production sont étudiés, et cinquante paramètres d'opération sont échantillonnés. Des tests échelons sont effectués sur les neuf paramètres jugés les plus importants et relativement faciles à manipuler.

#### **4.1.1 Choix des grades étudiés**

Sur les quatre grades produits à l'unité de déparaffinage au solvant, les grades 100 et 580 sont choisis comme sujets de l'étude. Le grade 100 est le plus léger et le grade 580 le plus lourd de la même catégorie de cire (cire cristalline: cristaux  $> 4$  microns). Il existe un grade plus lourd, SW-700, qui contient des paraffines microcristallines (cristaux  $< 4$  microns). Les résultats obtenus pour ce type de produit seraient difficilement comparables aux résultats obtenus pour le grade 100. De plus, en étudiant les deux grades choisis, la plage d'opération des cires cristallines est couverte.

#### **4.1.2 Paramètres d'opération mesurés**

L'unité de déparaffinage au solvant (UDS) se divise en trois sections: la cristallisation, la filtration et la récupération du solvant. Les deux premières sections sont étudiées. En effet, selon la théorie et l'expérience de la raffinerie, les deux premières sections sont les étapes cruciales quant au contenu final en huile de la cire. Ainsi, une grande partie des paramètres d'opération de l'unité sont éliminés (environ 100).

Au sein des deux sections choisies, plusieurs paramètres d'opération sont écartés parce qu'ils sont redondants. Cette situation se retrouve surtout au niveau des températures du



circuit de cristallisation. Cinquante-trois variables sont donc retenues sur les 75 paramètres d'opération disponibles. L'une de ces variables est le niveau de l'épouseur dans la section de récupération du solvant de la cire. Ce choix est motivé par les quelques débordements de l'épouseur qui ont occasionné une augmentation subite de l'indice de réfraction de la cire. Les modèles pourront ainsi y référer.

Trois paramètres ont été retirés par la suite. Les instruments de mesure ne fonctionnaient pas adéquatement. Il s'agit de la pression à la sortie de l'échangeur X-2437, du différentiel de température entre la deuxième dilution et l'alimentation et du différentiel de température entre la quatrième dilution et l'alimentation. Cinquante variables sont finalement retenues pour l'analyse. La liste des variables échantillonnées est présentée dans le tableau suivant:

**Tableau 4.1.1: Liste des variables échantillonnées**

<b>Variable</b>	<b>Paramètre représenté</b>
1	Alimentation à l'unité UDS
2	Vapeur 40 lb/eau recyclée à l'échangeur X-2402
3	Débit de prédilution vers X-2404
4	Température de la prédilution
5	Pression à l'entrée de l'échangeur X-2401A
6	Température à l'entrée de l'échangeur X-2401A
7	Température à la 3 <sup>e</sup> passe de l'échangeur X-2401A
8	Débit de la deuxième dilution
9	Pression à l'entrée de l'échangeur X-2401B
10	Température à la sortie de l'échangeur X-2401A
11	Débit de la troisième dilution
12	Différentiel de température 3 <sup>e</sup> dilution/alimentation
13	Pression à l'entrée de l'échangeur X-2402A
14	Température à la sortie de l'échangeur X-2401B
15	Débit de la quatrième dilution
16	Température à la sortie de l'échangeur X-2402A
17	Pression à l'entrée de l'échangeur X-2402B
18	Débit de la cinquième dilution
19	Différentiel de température 5 <sup>e</sup> dilution/alimentation

Tableau 4.1.1: Liste des variables échantillonnées (suite)

Variable	Paramètre représenté
20	Pression à l'entrée de l'échangeur X-2437
21	Température à la sortie de l'échangeur X-2402B
22	Température à la sortie de l'échangeur X-2437
23	Niveau du réservoir d'alimentation aux filtres
24	Débit de filtrat recyclé
25	Température du solvant froid
26	Niveau de la cuve du filtre A
27	Solvant froid du filtre A
28	Aspiration basse du filtre A
29	Vitesse de rotation du filtre A
30	Niveau botte du filtre A
31	Température de la botte du filtre A
32	Niveau cuve du filtre B
33	Solvant froid du filtre B
34	Aspiration basse du filtre B
35	Vitesse de rotation du filtre B
36	Niveau de la botte du filtre B
37	Température de la botte du filtre B
38	Niveau de la cuve du filtre C
39	Solvant froid du filtre C
40	Aspiration basse du filtre C
41	Vitesse de rotation du filtre C
42	Niveau de la botte du filtre C
43	Température de la botte du filtre C
44	Niveau de la cuve du filtre D
45	Solvant froid du filtre D
46	Aspiration basse du filtre D
47	Vitesse de rotation du filtre D
48	Niveau de la botte du filtre D
49	Température de la botte du filtre D
50	Niveau de l'épandeur de cire

## **4.2 Tests échelons**

Les tests échelons sont effectués dans le but d'obtenir un modèle représentatif de toute la plage d'opération du procédé. Ces tests échelons se répercutent sur tout le procédé et créent une variabilité significative des paramètres d'opération par rapport au bruit de mesure. Neufs variables sont testées une à une. Les tests ont été effectués sur les paramètres d'opération suivants:

- Concentration du solvant en méthyl-éthyl-cétone;
- Ratio entre le débit de prédilution et le débit d'alimentation au procédé;
- Ratio entre le débit de deuxième dilution et le débit d'alimentation au procédé;
- Ratio entre le débit de troisième dilution et le débit d'alimentation au procédé;
- Ratio entre le débit de quatrième dilution et le débit d'alimentation au procédé;
- Ratio entre le débit de cinquième dilution et le débit d'alimentation au procédé;
- Niveau des cuves des filtres;
- Solvant froid aux filtres;
- Vitesse de rotation des filtres;

Le test échelon sur la concentration du solvant n'est pas inclus dans la modélisation multivariable. L'intervalle de temps entre l'échantillonnage des données pour ce test est différent des autres tests: quinze minutes sont requises pour effectuer l'échantillonnage et la mesure de la concentration du solvant. Les paramètres d'opération sont échantillonnés aux dix minutes lors des autres tests.

L'intervalle d'échantillonnage de dix minutes est fixé par le temps requis pour échantillonner la cire et mesurer de son indice de réfraction au laboratoire de l'unité. Dans la prochaine section, la démarche utilisée pour effectuer la mesure de l'indice de réfraction est décrite.

### **4.3 Matériel et méthodes**

#### **4.3.1 Contenu en huile de la cire**

La première étape du projet consiste à justifier l'utilisation de l'indice de réfraction de la cire pour approximer son contenu en huile. Il s'agit d'identifier la relation qui existe entre les deux paramètres. Ainsi, des échantillons de cire recueillis lors des journées de tests sont analysés au laboratoire de la raffinerie selon les méthodes d'analyse ASTM suivies normalement pour évaluer le contenu en huile de la cire (ASTM D 721 pour le grade 100 et ASTM 3235 pour le grade 580). Ces analyses sont toujours effectuées par le même technicien de laboratoire dans le cadre de ce projet.

En résumé, les méthodes ASTM consistent à dissoudre un échantillon de cire dans un solvant, refroidir la solution à  $-32^{\circ}\text{C}$  pour précipiter la cire, et filtrer le mélange. Ensuite, le solvant est évaporé du filtrat et le résidu (l'huile) est pesé. La précision du test, selon les méthodes ASTM, est d'environ 1 % pour le grade 100 et 2 % pour le grade 580.

Les méthodes ASTM sont identiques pour les deux grades. La seule différence réside dans la composition du solvant utilisé. La méthode ASTM D 721 (grade 100) utilise du méthyle-éthyle-cétone pur tandis que la méthode ASTM 3235 (GRADE 580) utilise un mélange égal de méthyle-éthyle-cétone et de toluène.

#### **4.3.2 Indice de réfraction**

L'indice de réfraction de la cire est mesuré au laboratoire des huiles lubrifiantes, toujours par la même personne. Pour ce faire, un réfractomètre à plaque chauffante est utilisé. La plaque chauffante sert à contrôler très précisément la température à laquelle l'indice de réfraction est mesuré. La température est réglée à  $70.0^{\circ}\text{C}$ . L'indice de réfraction est très sensible à température de mesure. C'est pourquoi le réfractomètre est calibré chaque

**semaine par une personne désignée par la raffinerie. La procédure suivie pour effectuer la mesure est très simple:**

- **échantillonner la cire liquide;**
- **déposer une goutte de cire sur le prisme du réfractomètre;**
- **laisser réchauffer la goutte de cire deux minutes;**
- **placer la ligne de démarcation au centre de la croix du réfractomètre;**
- **lire l'indice de réfraction.**

**L'erreur de mesure est d'environ  $\pm 0,0001$  pour le grade 100 et de  $\pm 0,00025$  pour le grade 580. L'indice de réfraction de la cire de grade 580 est plus difficile à mesurer. La ligne de démarcation qu'il faut aligner sur la croix du réfractomètre n'est pas aussi claire que celle du grade 100.**

**Les corrélations obtenues entre l'indice de réfraction et le contenu en huile de la cire sont présentées aux figures suivantes. Les données ont été normalisées afin de respecter la confidentialité de celles-ci. Le coefficient de détermination pour la corrélation du grade 100 est de 97 %, tandis que le coefficient de détermination obtenu pour le grade 580 est de 83 %. Les limites de confiance à 95 % de la corrélation sont illustrées sur les graphiques.**

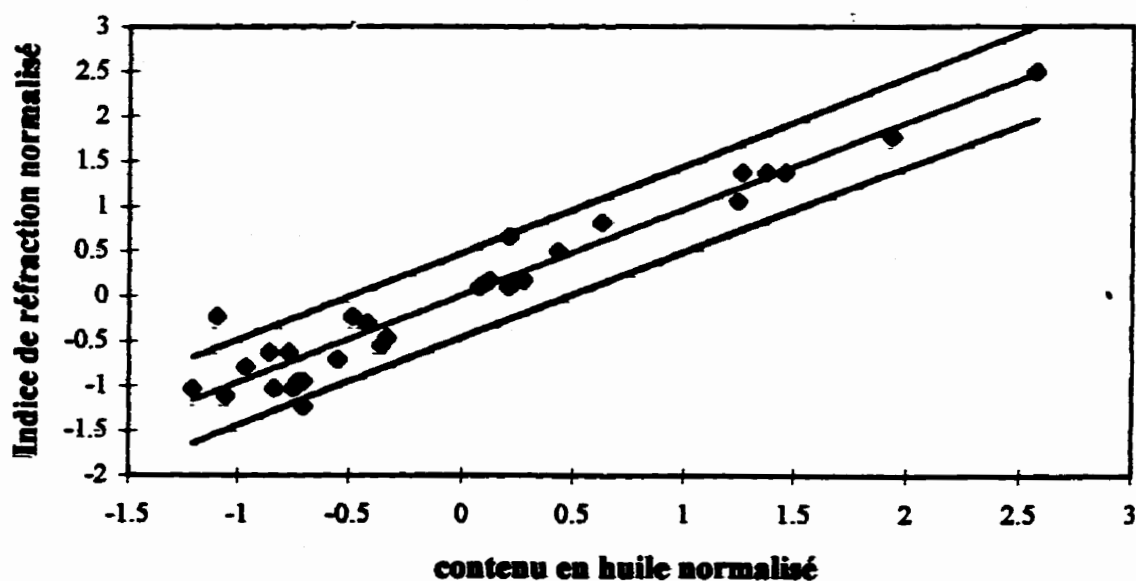


Figure 4.3.1: Indice de réfraction de la cire en fonction de son contenu en huile, grade 100.

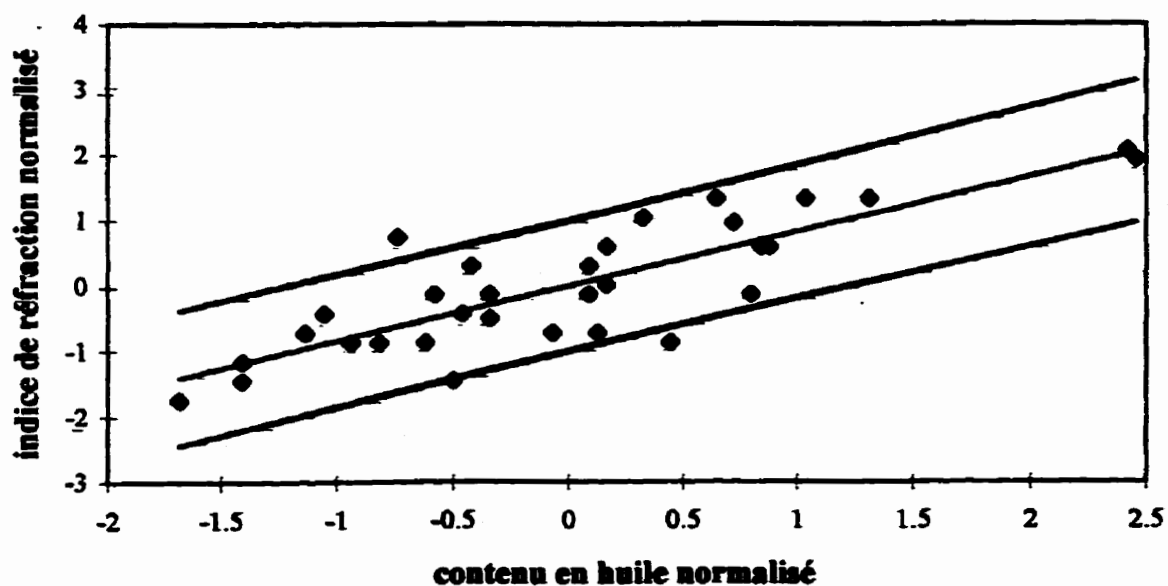


Figure 4.3.2: Indice de réfraction de la cire en fonction de son contenu en huile, grade 580.

Les corrélations indiquent que le contenu en huile dans la cire est bien relié de façon linéaire à son indice de réfraction. Ainsi, l'utilisation de l'indice de réfraction comme indicateur du contenu en huile de la cire est justifiée. De plus, la mesure de l'indice de réfraction est beaucoup plus simple, plus rapide et moins coûteuse que la mesure du contenu en huile dans la cire.

Le coefficient de détermination de la corrélation obtenu pour le grade 580 est moins bon que celui obtenu pour le grade 100. En effet, les mesures de l'indice de réfraction et du contenu en huile dans la cire sont moins précises pour ce grade, ce qui se reflète dans la corrélation.

#### **4.3.3 Analyses chromatographiques de l'huile contenue dans la cire**

La composition de l'alimentation à l'UDS est fonction du produit brut alimenté à l'unité ainsi que des étapes de transformation qui précèdent cette unité: la distillation sous vide et l'extraction au fufural. L'hypothèse que le contenu en huile de la cire soit relié à la composition de l'huile semble plausible. Des tests de chromatographie gazeuse ont été effectués sur des échantillons d'huile recueillis lors de la production de cire à haut et bas contenu en huile. Des tests de spectrométrie de masse ont été effectués également sur les échantillons de cire et d'huile.

Les résultats des analyses chromatographiques sont présentés en annexe I. Ces résultats infirment l'hypothèse posée plus haut. Les huiles semblent composées de constituants chimiques très semblables. Les pics sont situés au mêmes endroits sur les graphiques. Ils sont d'amplitudes différentes parce que la quantité d'huile injectée dans le chromatographe n'était pas mesurée exactement. Les tests de spectrométrie de masse sur l'huile ont été effectués à l'aide de deux types de bombardements: par atomes de krypton et par atomes de xénon. Les résultats dans tous les cas indiquent qu'aucune différence ne

peut être détectée entre les différents échantillons d'huile au point de vue des composés ionisables.

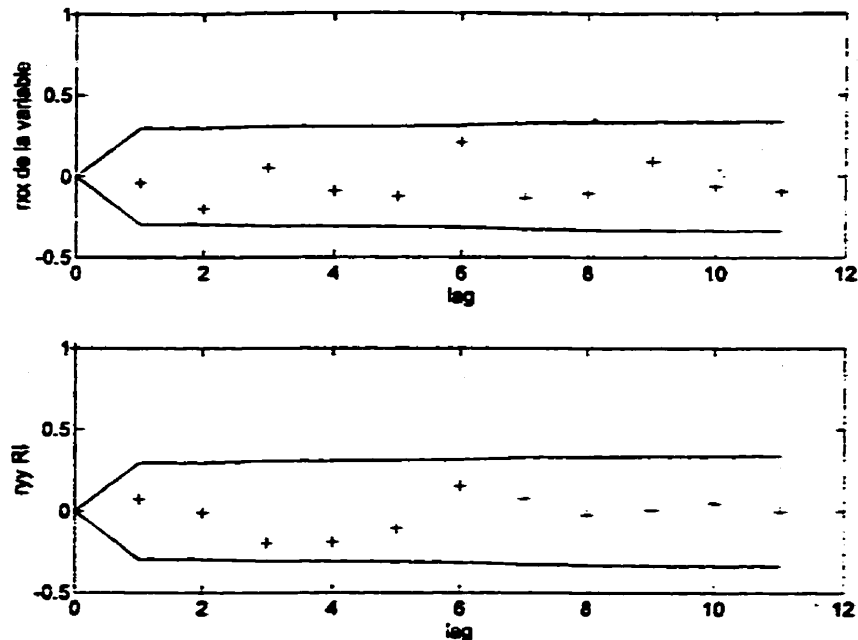
Les tests de spectrométrie de masse effectués sur des échantillons de cire concluent également que celles-ci ont les mêmes compositions chimiques également. Les bombardements magnétiques effectués à l'aide d'atome d'azote, de krypton n'ont pu détecter de différence entre les échantillons de cire quant aux composés ionisables.

Suite à ces analyses, on peut conclure que la variabilité observée sur l'indice de réfraction lors des tests échelons est vraisemblablement due à l'opération du procédé. Une variation évidente de la composition chimique de l'alimentation n'est pas en cause selon les tests effectués.

#### **4.4 Traitement préliminaire des données**

Les données utilisées pour effectuer la modélisation multivariable ont été traitées préalablement à la modélisation. Les modèles PCA et PLS sont des modèles linéaires, valables pour l'opération en régime permanent. Les périodes dynamiques ne doivent pas être trop importantes par rapport aux périodes de régime permanent pour que les méthodes soient efficaces. Le calcul des autocorrélations pour chacun des paramètres étudiés, pour chacune des journées de tests a donc été effectué pour chacun des grades afin de vérifier l'importance des périodes dynamiques dans l'opération. La différenciation a été appliquée lorsque les données n'étaient pas stationnaires. Un exemple du résultat du calcul des autocorrélations est présenté à la figure 4.4.1.





**Figure 4.4.1: Autocorrélation calculée pour la température de la botte du filtre C (haut) et pour l'indice de réfraction (bas) pour le test du 9 juillet 1996 sur le grade 580.**

Les autocorrélations ont permis de constater que les paramètres d'opération ainsi que l'indice de réfraction ne présentent pas de dynamique importante. Une corrélation significative à la première période de décalage est observée pour certains paramètres, mais pas de façon constante sur toutes les journées d'opération. Ainsi, il n'est pas jugé nécessaire de juxtaposer au bloc X un bloc supplémentaire de données décalées d'un intervalle de temps. La juxtaposition de ce bloc de données aurait pour effet d'augmenter considérablement la dimension du problème. De plus l'objectif de l'étude n'est pas de produire un modèle aussi prédictif que possible, mais bien d'identifier les variables les plus influentes sur l'indice de réfraction de la cire.

Vu cet objectif, il est important d'aligner les paramètres et l'indice de réfraction dans le temps. C'est-à-dire de retirer les périodes de délai sur chacun des paramètres. Afin de déterminer les délais entre les paramètres d'opération et l'indice de réfraction, les

corrélations-croisées ont été calculées. Cet exercice a été fait pour chacune des journées d'opération, pour les deux grades. À l'aide des résultats obtenus pour les paramètres les mieux corrélés avec l'indice de réfraction, les délais ont été déterminés. Un exemple du résultat du calcul des corrélation-croisées est présenté à la figure 4.4.2. Les graphiques aux figures 4.4.3 et 4.4.4 présentent les délais moyens pour chacun des paramètres, pour chacun des grades.

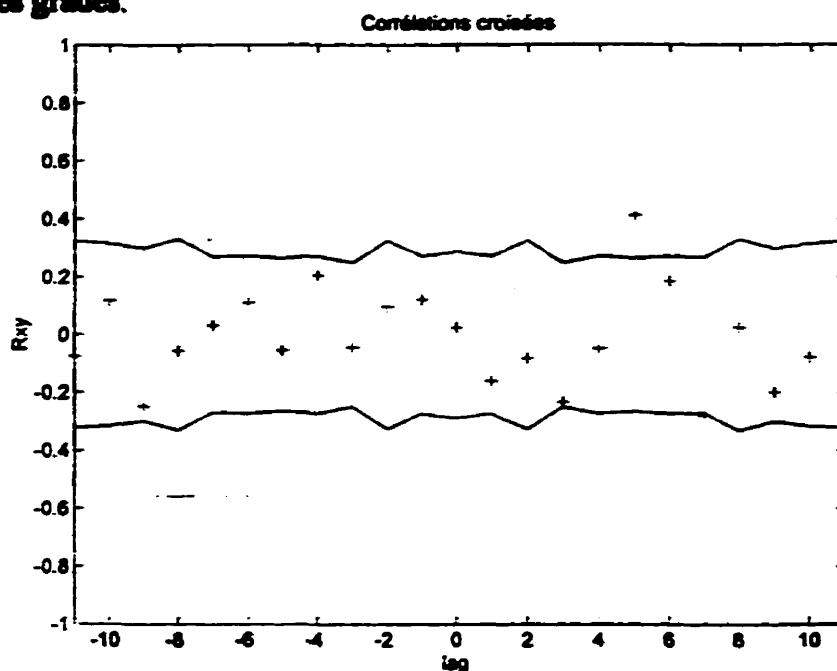


Figure 4.4.2: Corrélation-croisée entre la température de la botte du filtre C et l'indice de réfraction pour le test du 9 juillet 1996 sur le grade 580.

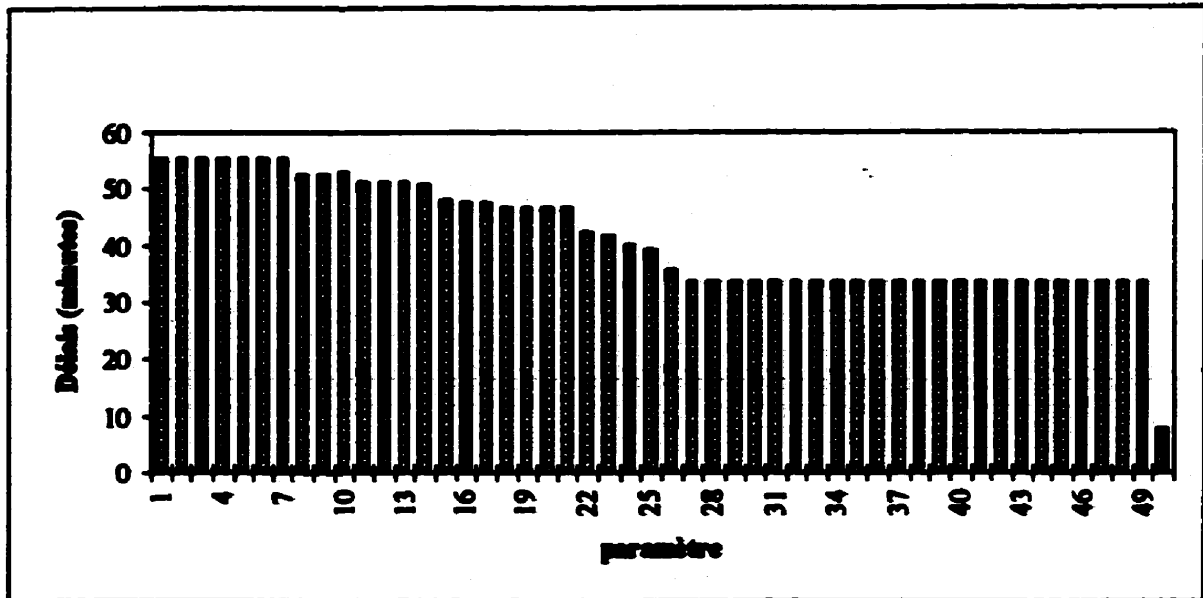


Figure 4.4.3: Délais appliqués sur les paramètres d'opération du grade 100.

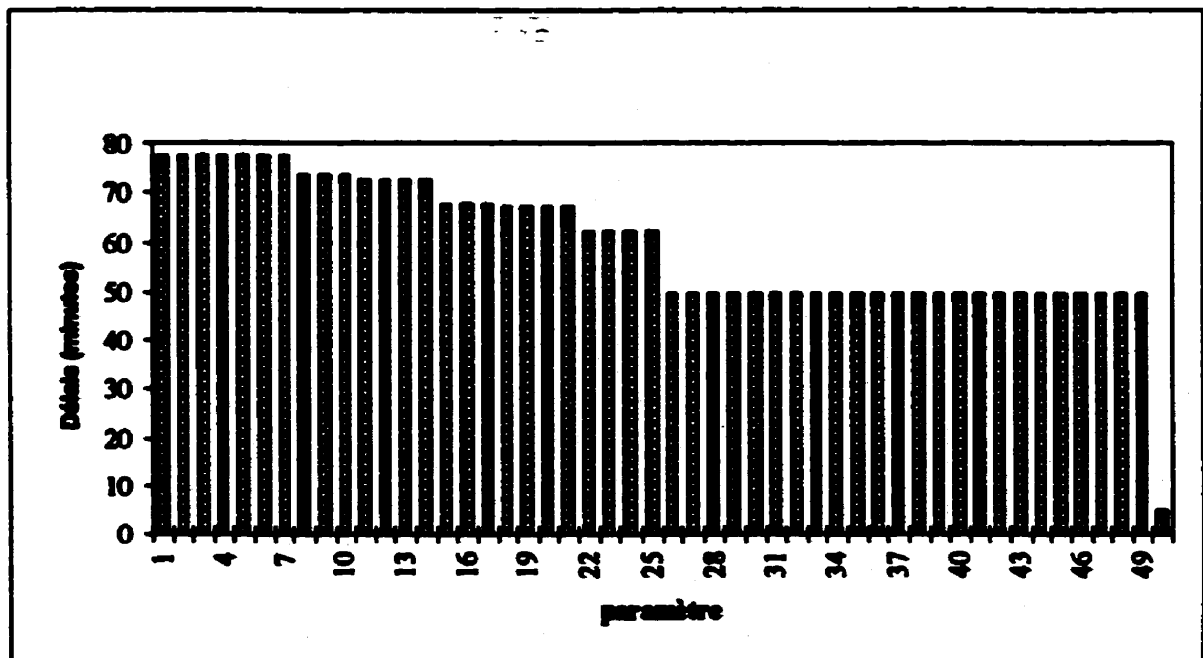


Figure 4.4.4: Délais appliqués sur les paramètres d'opération du grade 580.

## **5. Présentation et analyse des résultats**

Les résultats de modélisation multivariable de l'unité de déparaffinage au solvant est présentée dans ce chapitre. D'abord, l'analyse multivariable pour le grade 100 est présentée. L'analyse en composantes principales est présentée en détails. Une discussion des résultats est ensuite présentée. L'analyse des moindres carrés partiels est ensuite effectuée de façon détaillée. Le calcul de l'influence des paramètres sur l'indice de réfraction de la cire est ensuite présenté. Les résultats obtenus sont finalement discutés. L'analyse multivariable du grade 580 est résumée par la suite. La discussion des résultats de modélisation PCA et PLS obtenus pour ce grade sont présentées.

### **5.1 Analyse multivariable sur les données du grade 100**

#### **5.1.1 Distribution des données pour le grade 100**

Les analyses multivariées PCA et PLS s'effectuent à partir des données recueillies au cours des treize journées de tests réalisés lors de la production du grade 100. Les périodes de courts-circuits du réfrigérant aux deux premiers échangeurs (voir chapitre 2) sont excluent de la base de données analysée. En effet, la variabilité qu'ils créent est disproportionnée par rapport à la variabilité due à l'opération normale du procédé: les températures du circuit de cristallisation, les débits de prédilution et de deuxième dilution ainsi que l'indice de réfraction de la cire varient dans une proportion jamais vue lors de l'opération normale du procédé et sont très corrélés entre eux. Ainsi, les méthodes de modélisation utilisées n'accordent d'importance qu'à cet événement qui présente une structure de corrélation très forte, et tendent à négliger la variabilité naturelle du procédé. Or, il y a beaucoup plus d'intérêt à modéliser cette variabilité dite naturelle que celle due aux courts-circuits, ces derniers étant bien connus et inévitables. C'est pourquoi ils ont

été retirés. La base de données X contient donc 50 paramètres d'opération et 437 observations. Les données utilisées sont distribuées de la façon suivante:

**Tableau 5.1.1: Distribution des données pour le grade 100**

<b>Dates</b>	<b>Test échelon</b>	<b>Plage d'observations</b>
5 juillet 1996	Solvant froid au filtre A	1 à 15
5 juillet 1996		16 à 36
15 juillet 1996	Vitesse de rotation des filtres	37 à 64
16 juillet 1996	Ratio 5 <sup>d</sup> dilution	65 à 86
5 août 1996	Ratio 4 <sup>d</sup> dilution	87 à 113
5 août 1996		114 à 127
6 août 1996	Ratio 4 <sup>d</sup> dilution	128 à 156
12 août 1996	Ratio prédilution	157 à 198
13 août 1996	Ratio 3 <sup>d</sup> dilution	199 à 243
21 août 1996	Ratio 2 <sup>d</sup> dilution	244 à 286
6 septembre 1996	Ratio 5 <sup>d</sup> dilution	287 à 326
14 septembre 1996	Régime permanent	327 à 351
1 <sup>er</sup> octobre 1996	Niveau dans les curves	352 à 363
2 octobre 1996	Solvant aux filtres	364 à 388
10 octobre 1996	Contrôle de niveau du réservoir d'alimentation aux filtres	389 à 437

### **5.1.2 Analyse en composantes principales pour le grade 100**

L'analyse PCA permet d'obtenir une vue d'ensemble de l'état du procédé: détecter certains changements d'opération, identifier des tendances ou des regroupements de données et identifier des événements spéciaux. Les données statistiques de l'analyse des

composants principaux sont présentées au tableau 5.2. La somme des carrés de l'erreur initiale de la matrice  $X$  est de 21 800.

Tableau 5.1.2: Statistiques de l'analyse PCA sur l'espace  $X$  du grade 100

t	$SSx_{\text{expliquée}}$ %	$SSx_{\text{expliquée}}$ cumulée %	$SSx_{\text{résiduelle}}$ cumulée %	Différence $SSx_{\text{expliquée}} - SSx_{\text{résiduelle}}$	CSV/SD
1	27,78	27,78	26,53	1,25	0,857
2	15,63	43,41	41,37	2,04	0,901
3	10,26	53,67	51,63	2,04	0,925
4	7,55	61,22	59,08	2,14	0,940
5	5,40	66,61	64,47	2,14	0,957
6	4,45	71,06	69,46	1,60	0,957
7	3,71	74,77	73,02	1,75	0,966
8	3,07	77,84	75,82	2,02	0,979
9	2,37	80,19	78,75	1,44	0,979
10	2,10	82,30	80,67	1,63	0,988
11	1,86	84,15	81,91	2,24	1,011

La première étape de l'analyse consiste à déterminer la dimension du modèle. Pour ce faire, la validation croisée est utilisée. Le critère d'arrêt stipule qu'une variable latente est significative si le rapport CSV/SD est inférieur à 1. La dimension du modèle est donc de 10 variables latentes pour l'analyse des composants principaux (voir le tableau 5.2). Or, il existe exactement 10 périodes d'opération qui ne sont pas séparées par plusieurs journées. La modélisation de la variabilité due aux changements de période d'opération sera probablement très importante.

Toutes les variables latentes retenues semblent avoir un bon pouvoir prédictif puisque la différence entre la somme des carrés des erreurs expliquées par le modèle ( $SSx_{\text{expliquée}}$ ) et

la somme des carrés des erreurs expliquées par le modèle et validée par la validation croisée ( $SS_{X_{\text{validation croisée}}}$ ) sont très semblables. Le modèle explique donc des événements bien réels, et non pas du bruit ou des erreurs de mesure. Pour vérifier la validité du modèle, l'analyse des résidus est effectuée. Elle permet de vérifier la proportion des observations et des paramètres qui sont plus ou moins bien expliqués par le modèle. À la lumière de ces analyses, la décision de conserver le modèle pour fins d'analyse sera plus éclairée.

### Résidus sur les observations

La figure 5.1.1 présente le résultat du calcul de la somme des carrés des erreurs de prédiction sur les paramètres pour chacune des 437 observations (SPEx). Selon le résultat du calcul, le modèle PCA semble prédire le comportement du procédé de façon adéquate. Seules quatre observations se trouvent à l'extérieur des limites de confiance du modèle. Les limites de confiance sont déterminées à partir de la table de probabilité Chi carré avec  $\alpha = 0,99$ . Pour plus d'information, consultez Jackson (1991) au sujet de la statistique  $Q^2$ .

Un analyse plus approfondie de ces erreurs de prédiction permet d'en identifier la cause. Dans tous les cas il s'agit de variations subites et ponctuelles du différentiel de température entre la troisième dilution et l'alimentation. Sur le graphique de l'évolution de ce paramètre dans le temps, des impulsions sont observées exactement aux observations spécifiées par le calcul d'erreur (figure 5.1.2). Il s'agit en fait de variations résiduelles issues des courts-circuits. Comme le différentiel de température est le seul paramètre pour lequel de la variabilité attribuable aux courts-circuits est présente dans la base de données, le modèle PCA n'a pu expliquer ces variations à l'aide des autres paramètres d'opération.

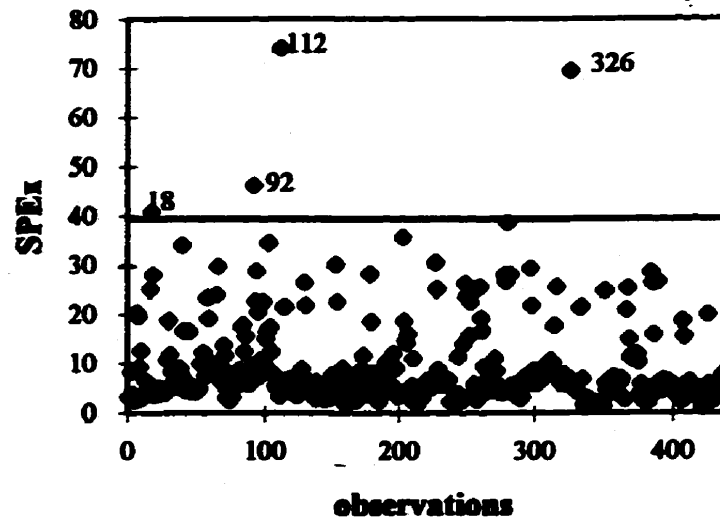


Figure 5.1.1: Somme des carrés des erreurs de prédiction sur X, modèle PCA, grade 100.

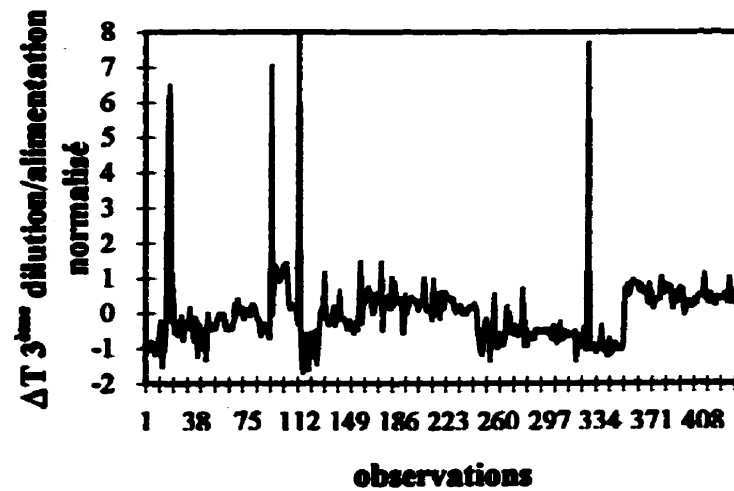


Figure 5.1.2: Évolution dans le temps du différentiel de température entre la 3<sup>ème</sup> dilution et l'alimentation.



### Résidus sur les variables

La figure 5.1.3 illustre la fraction de la variabilité de chacun des paramètres qui n'est pas expliquée par le modèle PCA. Tous les paramètres sont bien expliqués, à part trois: le débit de la deuxième dilution, le différentiel de température entre la troisième dilution et l'alimentation et le niveau de l'épouseur dans la section de récupération du solvant dans la cire. Une grande proportion de la variabilité des deux derniers paramètres reste inexpliquée à cause de quelques observations vraiment différentes des autres qui fausse la structure relationnelle entre ces paramètres et les autres. Dans le cas du débit de la deuxième dilution, il s'agit de l'ensemble des données qui sont mal expliquées par le modèle (voir figure 5.1.4).

La prochaine étape de l'analyse consiste à identifier quels sont les événements que décrivent les variables latentes. En théorie, chacune des variables latentes représente un événement distinct qui s'est produit lors de l'opération. En réalité, la présence de relations non linéaires entre les différents paramètres du procédé ou de périodes transitoires lors de l'opération nécessite des variables latentes supplémentaires pour expliquer un même événement.

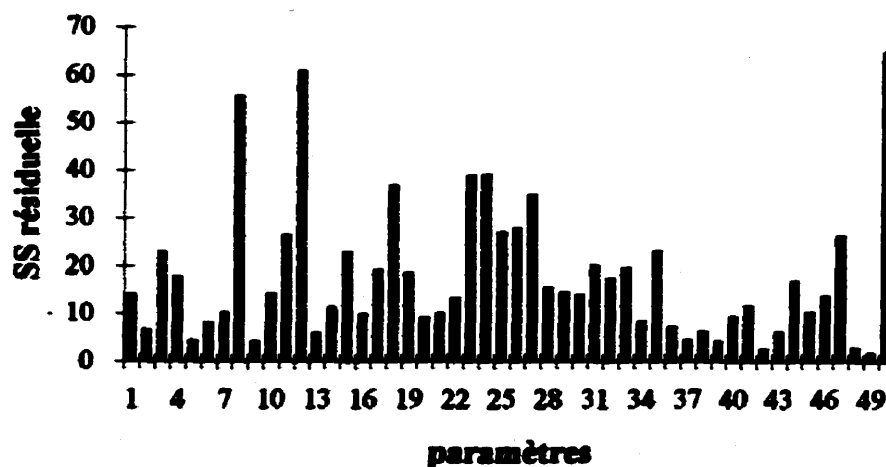
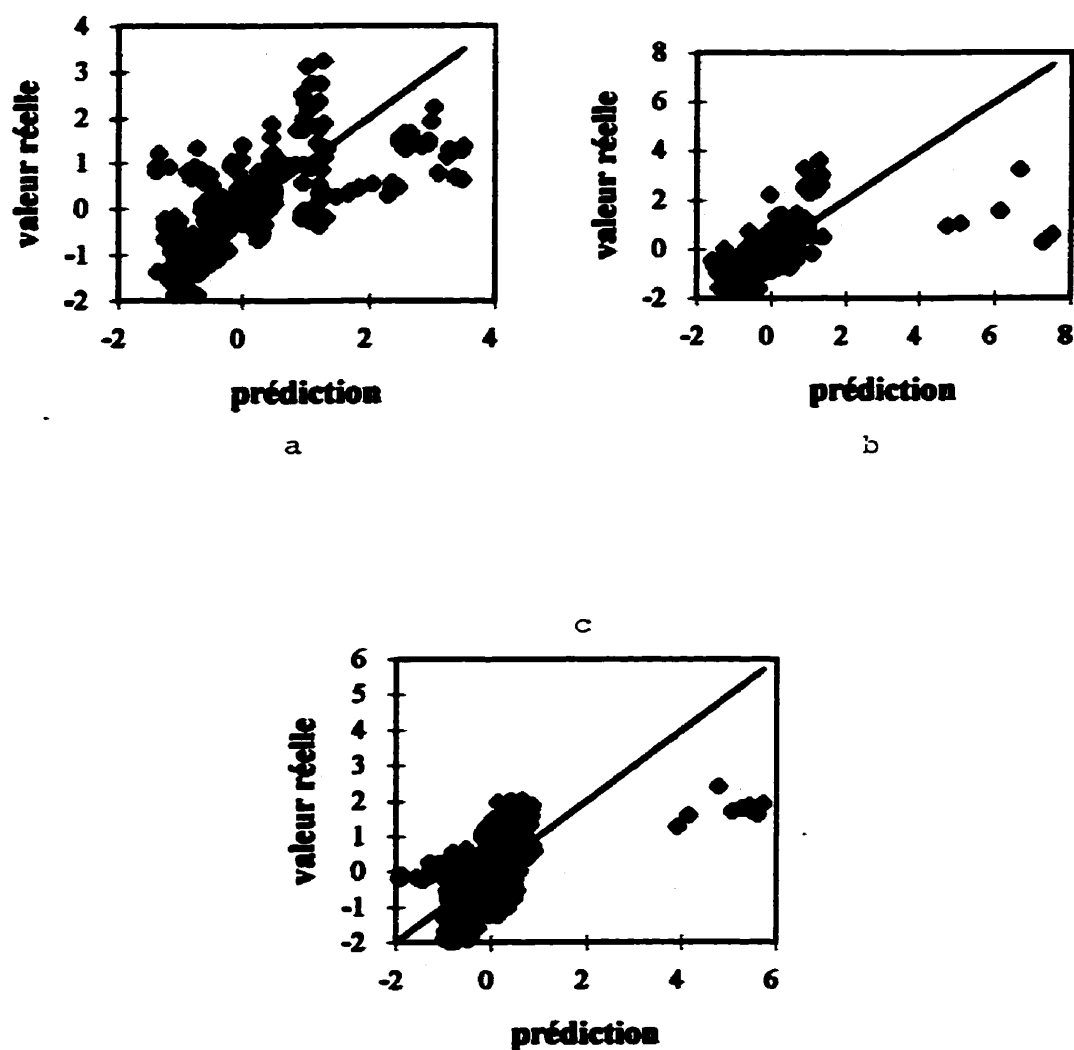


Figure 5.1.3: Pourcentage de la variabilité des paramètres non expliquée par le modèle PCA.



**Figure 5.1.4: Graphique des valeurs réelles normalisées en fonction des valeurs prédites normalisées des paramètres suivants: a) débit de la deuxième dilution, b) différentiel de température entre la troisième dilution et l'alimentation, c) niveau de l'épouseur (section récupération du solvant dans la cire).**

### 5.1.2.1 Analyses des variables latentes

L'analyse des variables latentes s'effectue efficacement en les traçant les unes en fonction des autres. Par exemple, le graphique de la variable latente  $t_2$  en fonction de  $t_1$  (figures 5.1.6 et 5.1.7) permet de visualiser le procédé en deux dimensions, dans le plan  $t_1/t_2$ , plutôt que d'effectuer l'analyse en observant tous les paramètres d'opération à la fois (50 dimensions). En fait, ce plan  $t_2/t_1$  peut être comparé, par analogie, à une fenêtre à travers laquelle on regarde à l'extérieur. Selon la position de la fenêtre, différents événements sont observés, ou sont vus d'un angle différent. La position de la fenêtre sur le procédé est modifiée en choisissant différentes variables latentes pour former le plan. Si l'observateur regarde à travers plusieurs fenêtres, il capte l'ensemble de l'action.

Dans cette étude, l'analyse PCA est effectuée à partir de neuf plans. Chacune des variables latentes est tracée en fonction de la variable latente  $t_1$ . Le choix des variables latentes qui forment les plans est arbitraire. Ce choix est fait ici parce que  $t_1$  est la variable latente qui explique le plus la variabilité du procédé. Les mêmes résultats d'analyse seraient obtenus à partir d'autres plans.

Les graphiques des variables latentes en fonction de l'évolution du procédé dans le temps permettent aussi de retirer de l'information précieuse. Les changements abrupts d'opération y sont facilement identifiables.

Les variables latentes  $t_i$  sont une combinaison linéaire normalisée de tous les paramètres d'opération avec  $p_i$  comme coefficients. Les poids  $p_i$  attribués aux paramètres d'opération dans la dimension  $i$  permettent d'identifier l'événement que représente la variable latente  $t_i$ . Les paramètres associées à un grand poids définissent l'événement en cause. L'analyse des variables latentes est donc effectuée en observant en parallèle les

graphiques  $t_{2}/t_{1}$  et les graphiques  $p_{2}/p_{1}$  pour chacune des dimensions du modèle. Par exemple, des graphiques fictifs sont présentés à la figure 5.1.5.

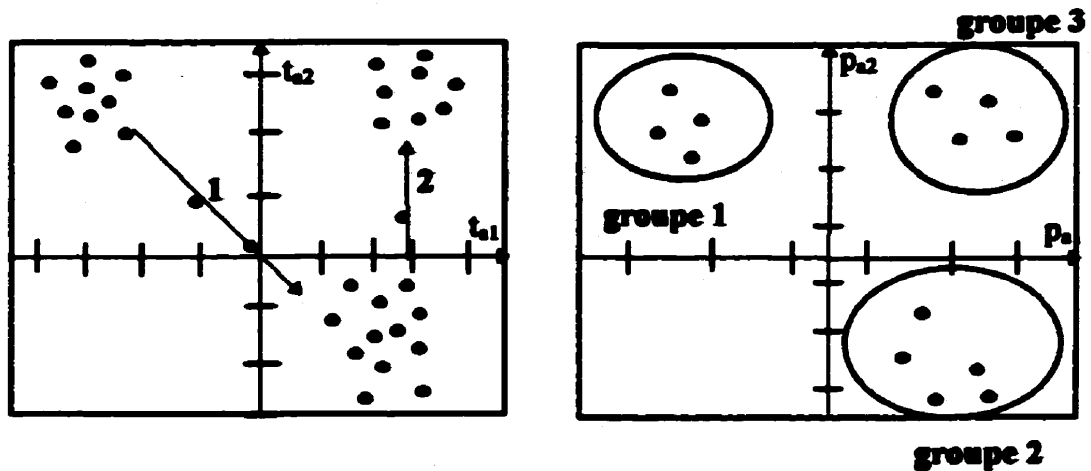


Figure 5.1.5: Exemple de graphiques des variables latentes  $t_2/t_1$  et des poids  $p_2/p_1$ .

Les paramètres impliqués dans le déplacement des observations sur le graphique  $t_2/t_1$  sont identifiées sur le graphique  $p_2/p_1$ . Le déplacement représenté par la flèche 1 sur le graphique  $t_2/t_1$  est dû à l'augmentation de la valeur des paramètres du groupe 2 et à la diminution de la valeur des paramètres du groupe 1 sur le graphique  $p_2/p_1$ . De la même façon, l'augmentation de la valeur des paramètres du groupe 3 et la diminution de la valeur des paramètres du groupe 2 sur le graphique  $p_2/p_1$  entraîne le déplacement représenté par la flèche 2 sur le graphique  $t_2/t_1$ .

Les paramètres hautement corrélés qui expliquent un événement se distribuent le poids qui les lie à cet événement. Ainsi, plus il y a de variables hautement corrélées, plus les poids sont petits, et moins ces variables se démarquent du reste des paramètres. Il devient alors plus difficile de bien identifier l'événement en cause. Il faut donc utiliser son

bon jugement. Par exemple, les conclusions tirées de ces graphiques peuvent souvent être vérifiées sur les graphiques des paramètres en question en évolution dans le temps.

Un autre outil aide à identifier les événements que modélise une dimension du modèle. Il s'agit de la contribution à la variation de la variable latente. Pour chacun des paramètres du procédé, pour une variation d'une variable latente  $t$ , entre deux observations  $o_1$  et  $o_2$  la contribution est calculée pour chacun des paramètres d'opération:

$$C_{o1-o2} = p'_a (X_{o2} - X_{o1}) \quad (5.1)$$

Les paramètres pour lesquels la contribution est élevée sont ceux qui ont varié entre les deux points d'observation  $o_1$  et  $o_2$ .

Les prochains paragraphes présentent l'analyse des variables latentes du modèle PCA pour le grade 100.

### **Dimension 1**

#### **Variable latente $t_1$**

La variable latente  $t_1$  divise les données en quatre sous-groupes. On observe ces regroupements de données autant sur le graphique  $t_2/t_1$ , figures 5.1.6 et 5.1.7 que sur le graphique de  $t_1$  en évolution dans le temps, figure 5.1.8:

- groupe I: opération du 5 juillet 1996. Les filtres A,B et C sont en service.
- groupe II: opération du 15 juillet au 6 septembre 1996. Les filtres A,B et D sont en service.
- groupe III: opération du 14 septembre 1996. Les filtres A,C et D sont en service.
- groupe IV: opération du 1<sup>er</sup> octobre au 10 octobre 1996. Les filtres A,B,C sont en service.

Ainsi, si le filtre C fonctionne,  $t_1$  diminue et si le filtre D fonctionne  $t_1$  augmente.

Les groupes I et IV sont distincts malgré que la combinaison des filtres en fonction soit la même. Ils se distinguent par leur profil de refroidissement qui est plus graduel au premier échangeur pour le groupe IV. La pression à l'entrée du cristalliseur X-2437 est également beaucoup plus faible pour l'opération du groupe IV.

La première dimension retient donc les paramètres pour lesquelles des variations importantes sont observées lors de l'opération.

#### Poids $p_1$

Les variables qui se distinguent des autres par leur poids pour la dimension 1 sont en premier lieu les paramètres d'opération des filtres C et D:

- niveau de la cuve des filtres C et D;
- débit de solvant froid des filtres C et D;
- aspiration basse des filtres C et D;
- vitesse de rotation des filtres C et D;
- niveau et température des bottes des filtres C et D;

Les températures du début du cycle de réfrigération qui influencent la variable  $t_1$  sont:

- la température de la vapeur 40lb/eau recyclée X-2402;
- la température de l'alimentation à l'échangeur X-2401A;
- la température de la 3<sup>e</sup> passe de l'échangeur X-2401A;
- la température à la sortie de l'échangeur X-2401A;
- la température à la sortie de l'échangeur X-2401B.

La pression à l'entrée de l'échangeur X-2437 est aussi une variable importante pour  $t_1$ .

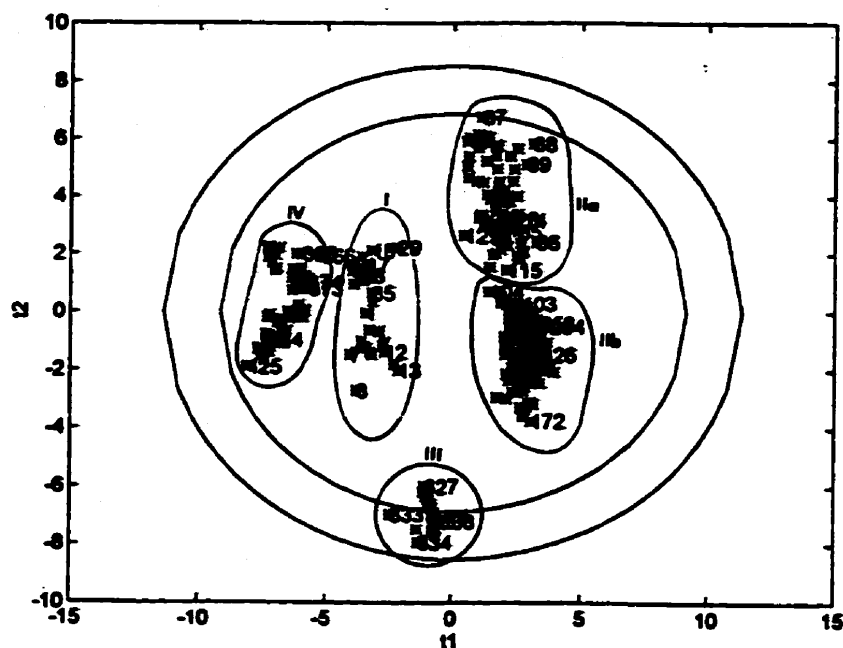


Figure 5.1.6 Graphique  $t_2/t_1$ , modèle PCA, grade 100.

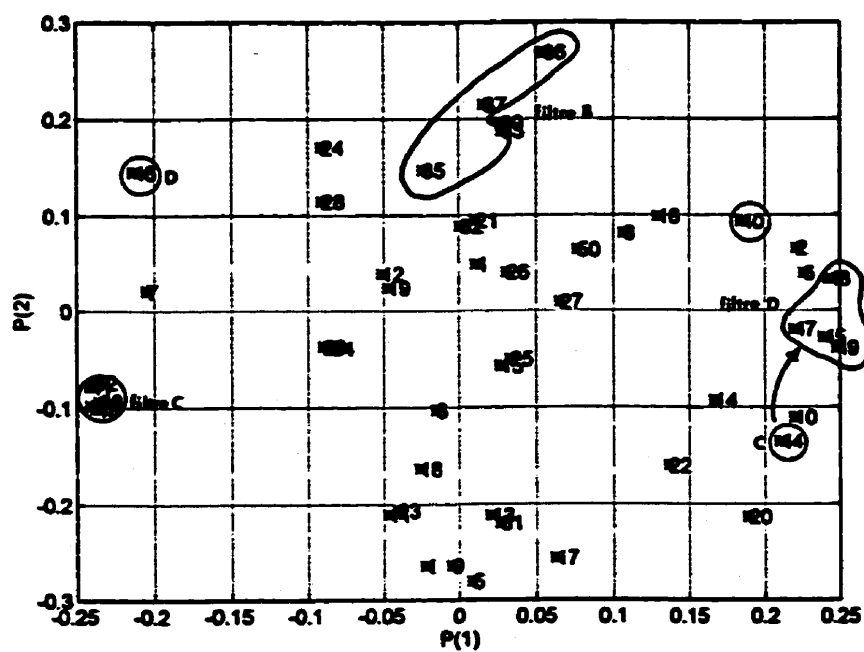


Figure 5.1.7: Graphique  $p_2/p_1$ , modèle PCA, grade 100.

Les chiffres attribués aux points sur le dernier graphique représentent la position de la variable échantillonnée dans le tableau 4.1.1.

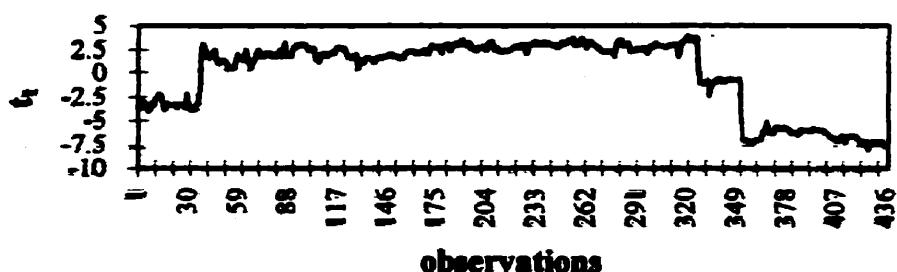


Figure 5.1.8:  $t_1$  en évolution dans le temps, modèle PCA, grade 100.

## Dimension 2

### Variable latente $t_2$

Plusieurs aspects différents de l'opération sont regroupés dans cette dimension. Ces aspects sont le reflet des différentes façons dont le procédé est opéré. De façon plus importante, ce composant décrit la variation du débit d'alimentation et des pressions:

- • groupe I: débit d'alimentation moyen et pression moyenne;
- groupe IIa: débit d'alimentation et pressions plus faibles;
- groupe IIb: débit d'alimentation et pressions plus élevés;
- groupe III: débit d'alimentation élevé, pressions moyennes;
- groupe IV: débit d'alimentation élevé et pressions faibles;

L'opération à faibles débit et pressions est représentée par le plan positif de  $t_2$ . Les observations 1 à 15 et 16 à 36, qui correspondent respectivement à l'opération avant et après le court-circuit du réfrigérant aux deux premiers échangeurs sont séparées à cause de la chute de pression qui en résulte.



La variable latente  $t_2$  explique aussi la variation de la température et du niveau de la botte du filtre A. Lorsque le niveau de la botte du filtre est élevé et la température est faible, les observations se déplacent vers le plan positif de  $t_2$ . Le niveau dans la botte du filtre A est contrôlé par la température de la botte: Lorsque la température est plus faible, la cire s'écoule plus difficilement de la botte, donc le niveau est plus haut. Ces variables sont importantes dans cette dimension puisqu'elles varient beaucoup d'une période d'opération à l'autre.

Certains paramètres du filtre B sont modélisés par cette dimension. Lorsque le filtre B est arrêté (groupe III), les données se déplacent vers la partie négative de  $t_2$ .

$t_2$  modélise aussi l'augmentation du niveau du réservoir V-2401 entre les observations 80 et 115. Il s'agit ici d'un résultat du test échelon sur la quatrième dilution du 5 août 1996.

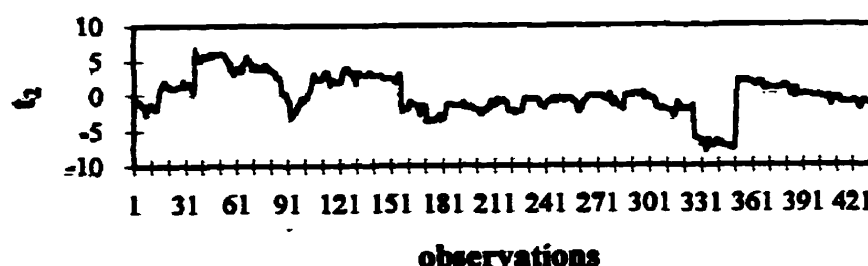


Figure 5.1.9:  $t_2$  en évolution dans le temps, modèle PCA, grade 100.

#### Poids $p_2$

Les variables qui décrivent les variations du débit d'alimentation et des pressions sont des variables les plus importantes pour  $t_2$ :

- débit d'alimentation;
- pression entrée X-2401A;

- pression entrée X-2401B;
- pression entrée X-2402A;
- pression entrée X-2402B;
- pression entrée X-2437;
- niveau du réservoir d'alimentation aux filtres;

Les paramètres d'opération du filtre B modélisés par la deuxième dimension sont les suivants. Ces variables sont toutes à zéro durant l'arrêt du filtre:

- niveau de la botte du filtre B;
- température de la botte du filtre B;
- solvant froid alimenté au filtre B;

La température et le niveau de la botte du filtre A sont également importants, ainsi que la température à la sortie de l'échangeur X-2401B. Cette dimension vient compléter la modélisation de la première dimension pour ces paramètres.

### **Dimension 3**

#### **Variable latente $t_3$**

Sur les graphiques de  $t_3/t_1$  et  $t_3/t_2$ , seul le groupe III est clairement séparé des autres groupes. Les données du groupe III se démarquent des autres par la combinaison des filtres qui sont en fonction durant cette période (ACD).

L'allure de la courbe  $t_3$  en évolution dans le temps est très semblable à celle de  $t_2$  (comparer les figures 5.1.9 et 5.1.10). Le graphique de  $t_3/t_2$  confirme cette observation. On s'attend à ce que  $t_3$  modélise à peu de chose près le même événement que  $t_2$ :

- La variable  $t_2$  modélise le filtre B en partie, et certaines variations d'opération au niveau du débit d'alimentation et des pressions.

- La variable  $t_3$  modélise le filtre B et différencie les périodes d'opération en tenant compte de la température du solvant froid, la différence de température entre l'alimentation et la cinquième dilution, ainsi que de la température à la sortie de l'échangeur X-2402A.

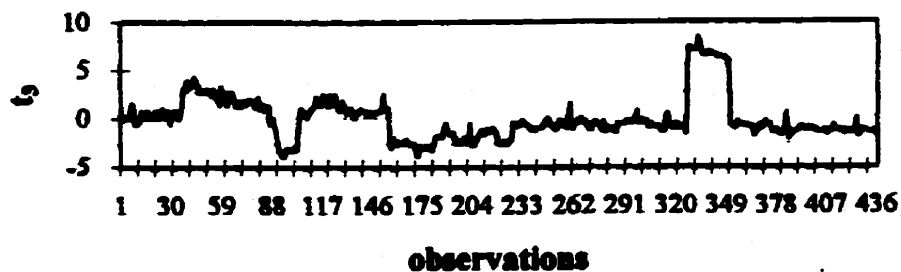


Figure 5.1.10:  $t_3$  en évolution dans le temps, modèle PCA, grade 100.

Les deux variables latentes modélisent donc les mêmes changements de périodes d'opération au niveau de variables différentes:

- changement de période d'opération du 5 juillet au 15 juillet;
- changement de période d'opération du 16 juillet au 5 août;
- changement de période d'opération du 6 août au 21 août;

Le test échelon du 5 août sur la quatrième dilution est aussi modélisé par cette dimension parce que les températures du circuit de cristallisation qui lui sont associées sont importantes dans cette dimension.

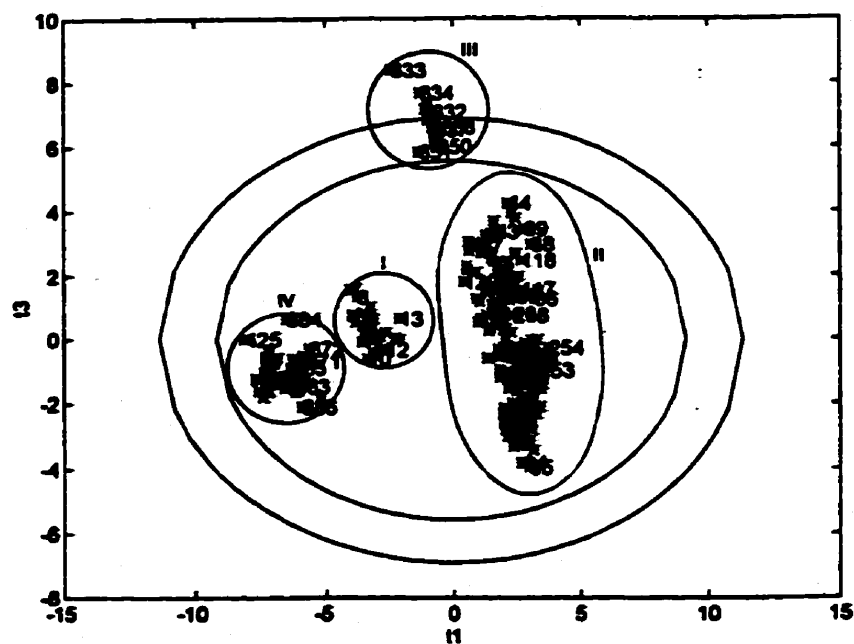


Figure 5.1.11: Graphique  $t_2/t_1$ , modèle PCA, grade 100.

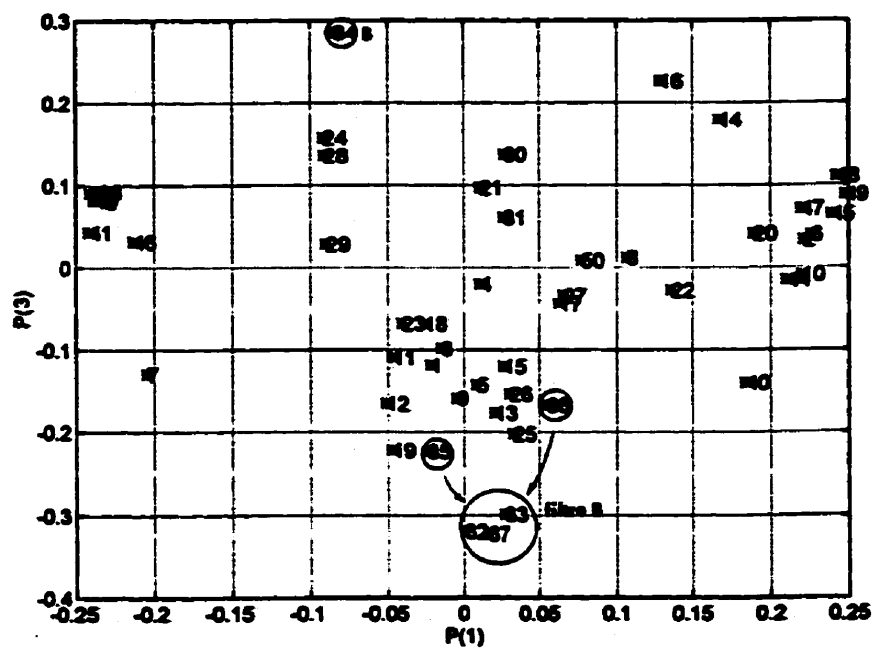


Figure 5.1.12: Graphique  $p_2/p_1$ , modèle PCA, grade 100.

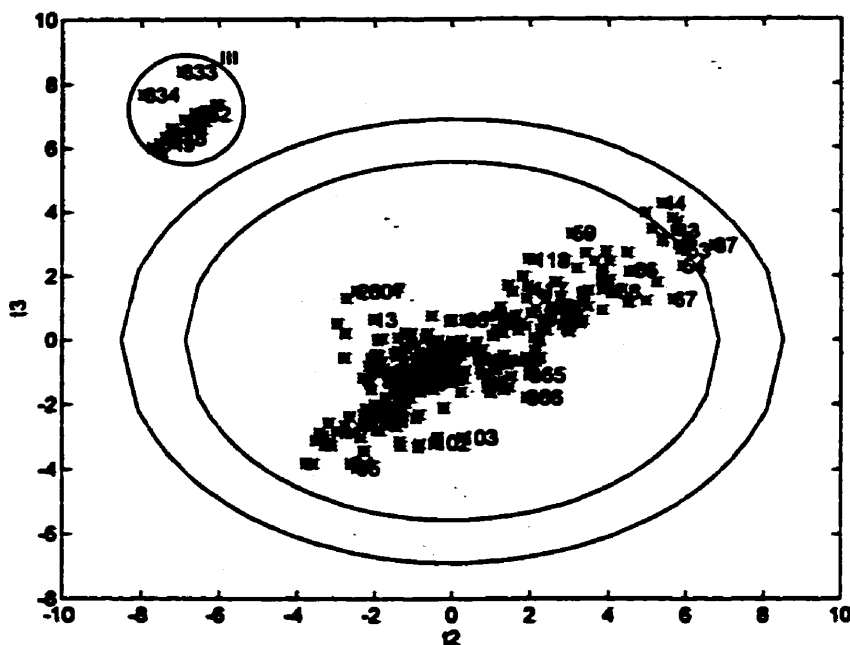


Figure 5.1.13: Graphique  $t_3/t_2$ , modèle PCA, grade 100.

### Poids $p_1$

Les paramètres d'opération du filtre B qui ont le plus de poids sont:

- le niveau de la curve du filtre B;
- le débit de solvant froid du filtre B;
- la vitesse de rotation du filtre B;
- le niveau et température de la botte du filtre B;
- la pression d'aspiration basse du filtre B.

En moindre importance, les variables du circuit de cristallisation qui modélisent les changements de période d'opération et le test échelon du 5 août sont:

- la température à la sortie X-2402A;
- la température à la sortie X-2401B;
- la température du solvant froid;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation;
- le  $\Delta T$  3<sup>e</sup> dilution/alimentation;
- les pressions à l'entrée des échangeurs;
- le débit de filtrat recyclé.

#### Dimension 4

##### Variable latente $t_4$

La quatrième dimension du modèle explique la variabilité des températures de la fin du cycle de cristallisation. C'est pourquoi le groupe de données I, pour lequel les températures en question sont plus élevées, se démarque des autres groupes. Le tests échelons sur la quatrième dilution du 5 août 1996 (observations 87 à 127) a aussi un impact sur ces températures. Cet impact est observé notamment sur le graphique de  $t_4$  en évolution dans le temps.

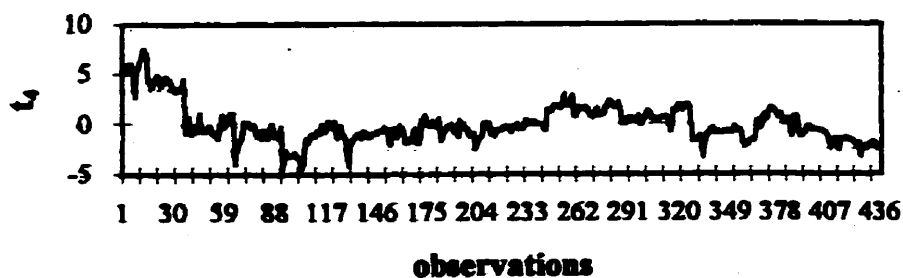


Figure 5.1.14:  $t_4$  en évolution dans le temps, modèle PCA, grade 100.

La variable latente  $t_4$  modélise également le comportement du filtre A. Les lavages chauds du filtre sont représentés par l'excursion des observations vers la partie négative de  $t_4$ . Les observations reviennent à leur position initiale lorsque le lavage se termine:

- observations 7 à 8, 128 à 129, 424 à 425: le lavage du filtre A commence;
- observations 8 à 10, 129 à 130, 425 à 426: le lavage est fini.

#### Poids $p_4$

Les variables qui se démarquent le plus dans cette dimension sont:

- le  $\Delta T$  5<sup>e</sup> dilution/alimentation;
- la température à la sortie de X-2402B;
- la température à la sortie de X-2401B.

Cette dimension semble compléter les variations du profil de refroidissement modélisé en  $t_1$ .

Quelques autres variables sont importantes:

- le débit de la deuxième dilution: relié à la variation de la température à la sortie du deuxième échangeur (X-2401B);
- le débit de la quatrième dilution: relié à la variation de la température à la sortie du quatrième échangeur (X-2402B)

Les paramètres du filtre A modélisés par la quatrième dimension sont:

- la température de la botte du filtre A;
- le niveau de la cuve du filtre A;

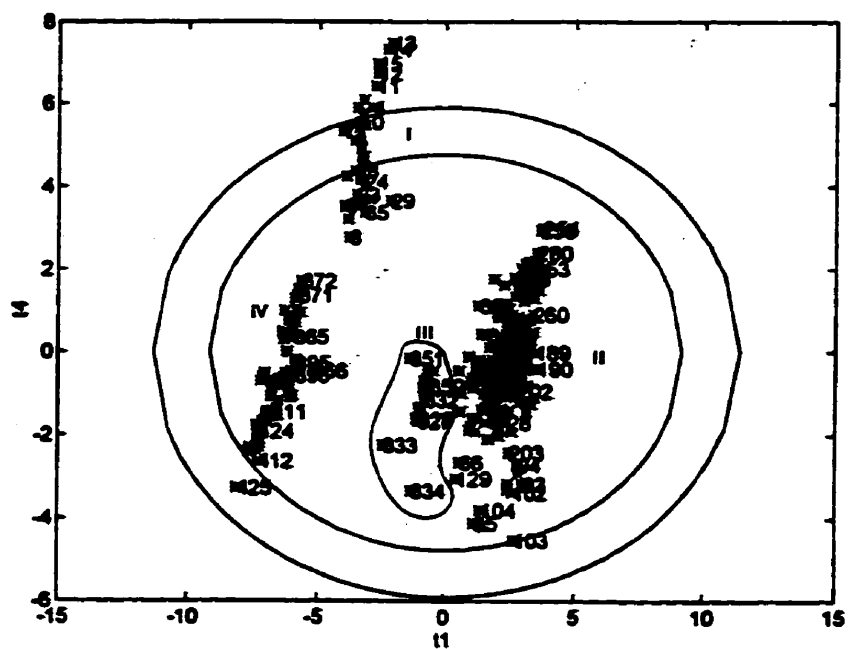


Figure 5.1.15: Graphique  $t_2/t_1$ , modèle PCA, grade 100

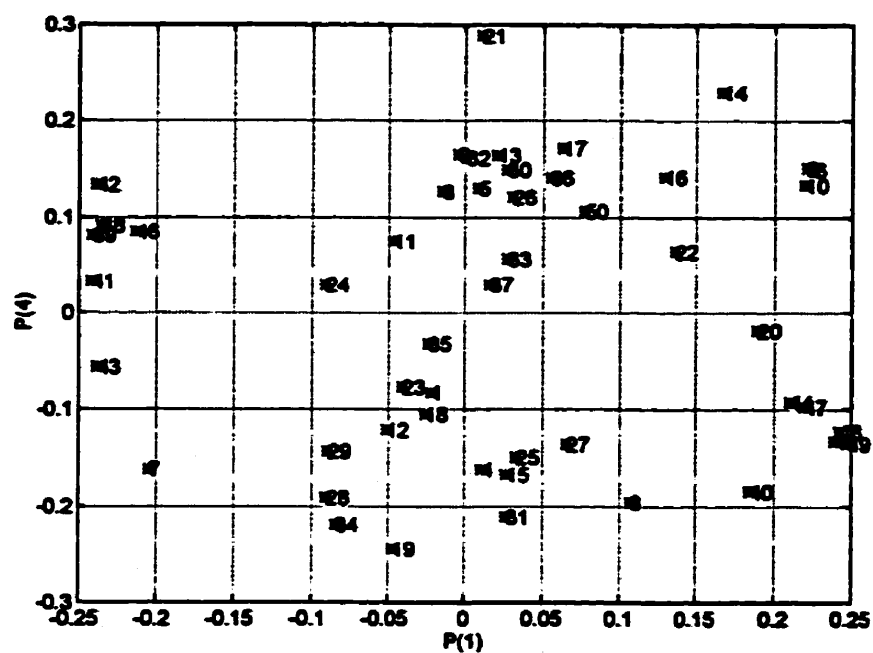


Figure 5.1.16: Graphique  $p_4/p_1$ , modèle PCA, grade 100.



## Dimension 5

### Variable latente $t_5$

Cette dimension semble modéliser les comportements du procédé à l'intérieur d'une même journée d'opération, plutôt que des changements de période d'opération.

La variable latente  $t_5$  modélise en premier lieu le comportement du filtre A. D'abord, sur le graphique de  $t_5$  en fonction du temps, on observe les lavages (pics). Sur le graphique  $t_5/t_1$ , lorsqu'un lavage a lieu, les observations se déplacent vers le côté positif de  $t_5$  pour revenir ensuite à leur position initiale. Par exemple:

- observations 7 à 8;
- observations 128 à 129;
- observations 314 à 315;
- observations 424 à 425.

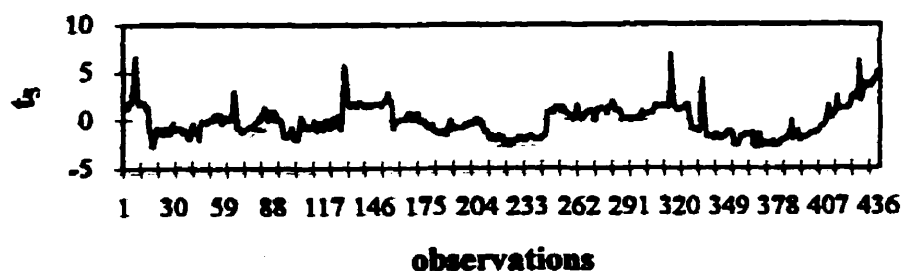


Figure 5.1.17:  $t_5$  en évolution dans le temps, modèle PCA, grade 100.

Le test sur le contrôle du niveau V-2401 à l'aide des vitesses de rotation et des pressions d'aspiration est aussi modélisé par cette variable latente (observations 389 à 437). Durant ce test, les vitesses de rotation des filtres sont augmentées tout au long de la journée, ce qui fait augmenter  $t_5$  (voir figure 5.1.17).

**t, modélise aussi les pressions à l'entrée des échangeurs qui varient d'une journée d'opération à l'autre, mais aussi à l'intérieur d'une même journée: elles augmentent jusqu'au court-circuit du réfrigérant. Par exemple:**

- **observations 400 à 437: augmentation des pressions;**
- **observations 15 à 16: chute des pressions suite au court-circuit.**

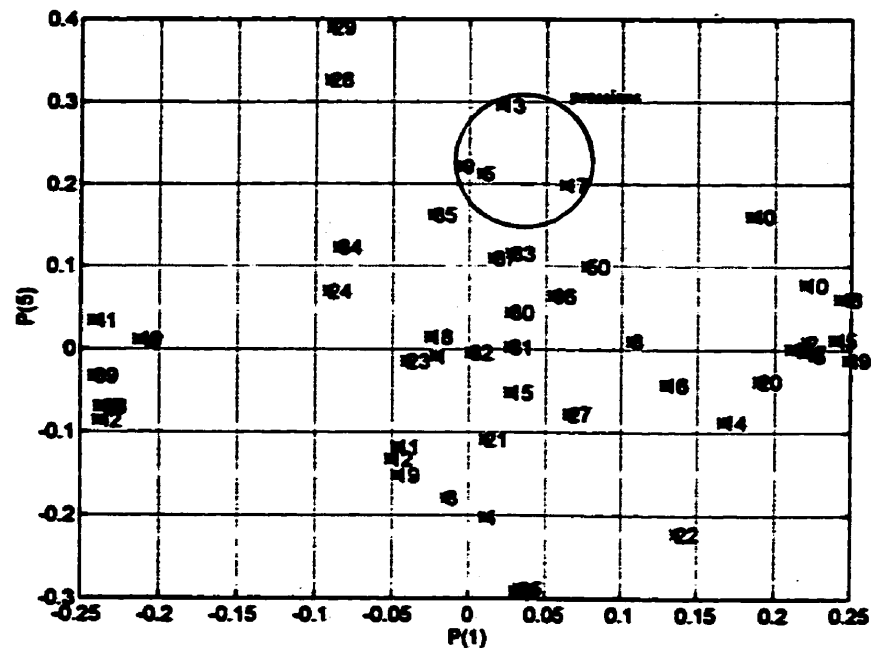


Figure 5.1.19: Graphique  $p_5/p_1$ , modèle PCA, grade 100.

### Poids $p_3$

Les variables du filtre A modélisées ici sont:

- la pression d'aspiration basse du filtre A;
- la vitesse de rotation du filtre A;
- le niveau de la cuve du filtre A.

Lors des lavages, la vitesse de rotation du filtre augmente, le niveau de la cuve diminue et l'aspiration basse augmente. C'est pourquoi  $t_3$  augmente à cette occasion.

Les pressions à l'entrée des quatre premiers échangeurs sont également importantes.

Les variables suivantes sont reliées aux pressions par les courts-circuits. Le court-circuit fait chuter les pressions et augmenter les températures des solvants. Les pressions se remettent à augmenter durant le reste de la journée et les températures diminuent jusqu'à à leur température initiale:

- température du solvant froid;
- température de la prédilution;

La température à la sortie de l'échangeur X-2437 a aussi de l'importance dans cette dimension. Ce qui vient compléter la modélisation du circuit de cristallisation.

#### **Dimension 6**

##### **Variable latente $t_6$**

La variable latente  $t_6$  modélise le test échelon sur la quatrième dilution du 5 août: observations 87 à 127. Lorsque le débit de la quatrième dilution augmente, les observations sortent de la masse et se placent à l'écart, dans la partie positive de  $t_6$  (voir figure 5.1.20).

$t_6$  modélise aussi certains changements de période d'opération:

- 5 juillet au 15 juillet;
- 6 août au 12 août;
- 13 août au 21 août;
- 2 octobre au 10 octobre.

Ces variations de  $t_c$  sont dus à des changements d'opération divers: température et niveau de la botte du filtre A, débit de prédilution, température de prédilution, paramètres d'opération du quatrième échangeur (X-2402B).

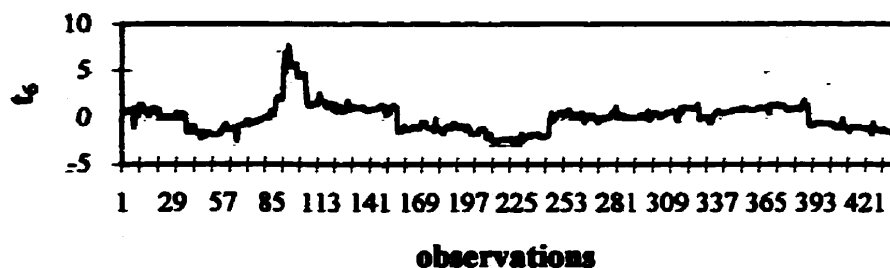


Figure 5.1.20:  $t_c$  en évolution dans le temps, modèle PCA, grade 100.

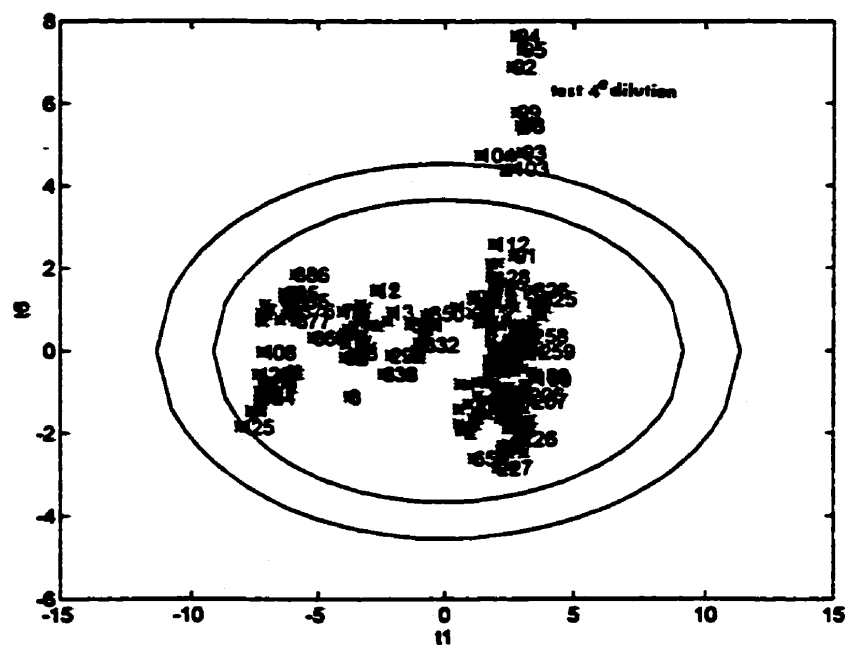


Figure 5.1.21: Graphique  $t_c/t_1$ , modèle PCA, grade 100.

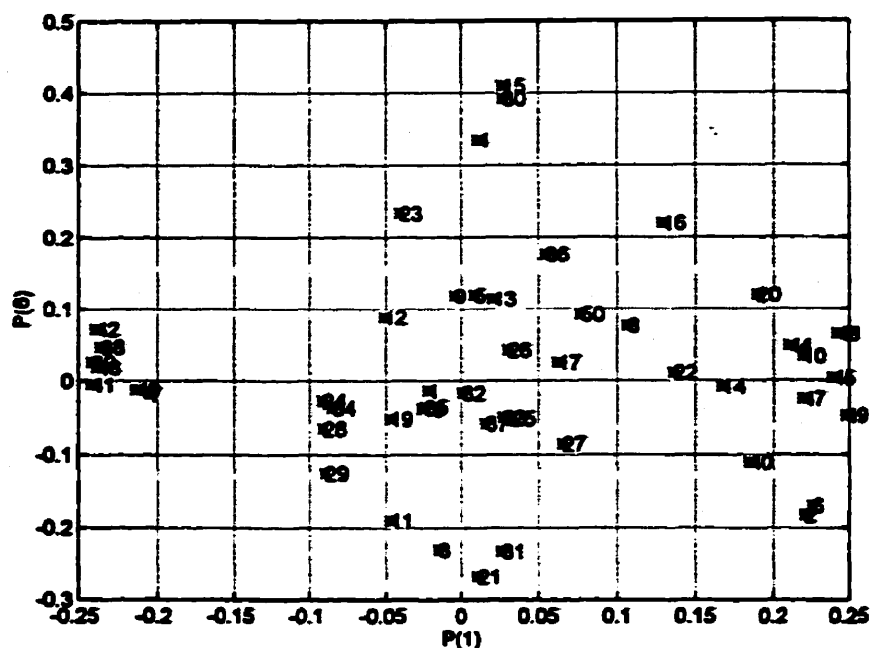


Figure 5.1.22: Graphique  $p_2/p_1$ , modèle PCA, grade 100.

#### Poids $p_i$

Les variables qui modélisent les tests échelons sur la quatrième dilution sont:

- le débit de la quatrième dilution;
- la température à la sortie de l'échangeur X-2402B;
- le débit de la troisième dilution.

L'augmentation du débit de la quatrième dilution a eu pour effet de faire chuter la pression dans le circuit de solvant froid alimenté au cycle de refroidissement. Le point d'injection de la troisième dilution se situe après celle de la quatrième dilution sur le circuit du solvant. C'est pourquoi l'augmentation subite du débit de la quatrième dilution a fait chuter le débit de la troisième dilution.

Les variables suivantes ont également du poids. Elles sont liées aux changements de période d'opération:

- température de la botte du filtre A;
- niveau de la botte du filtre A;
- débit et température de la prédilution.

### Dimension 7

#### Variable latente $t_7$

La variable latente  $t_7$  modélise les lavages du filtre A. Ils sont représentés par les pics sur le graphique de  $t_7$  en évolution dans le temps.

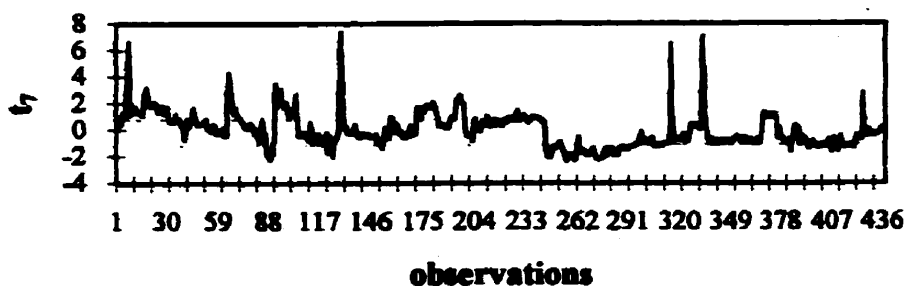


Figure 5.1.20:  $t_7$  en évolution dans le temps, modèle PCA, grade 100.

La septième dimension modélise aussi le test échelon sur la prédilution effectués le 12 août 1996: observations 157 à 198 sur le graphique de  $t_7$  en évolution dans le temps.  $t_7$  explique en grande partie les variations du débit de la prédilution sur toute la période d'étude.

### **Poids $p_7$**

Les paramètres qui modélise les lavages du filtre A pour cette dimension sont:

- le débit de solvant froid au filtre A;
- le niveau de la cuve du filtre A;
- la vitesse de rotation du filtre A;
- le débit de filtrat recyclé;

Le débit de solvant froid étant le plus important de tous au niveau du poids qui lui est attribué.

Les variations du débit de prédilution se reflètent par:

- le débit de prédilution;
- la température du solvant froid;
- la température de la 3<sup>e</sup> passe de l'échangeur X-2401A;
- la température à la sortie de l'échangeur X-2402A;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation.
- le débit de filtrat recyclé;

Le test échelon sur le ratio de la prédilution influence les températures puisque la température du solvant est de beaucoup inférieure à celle de l'alimentation. Le profil de refroidissement est donc modifié lorsqu'on alimente plus ou moins de prédilution.

Le débit d'alimentation est inversement proportionnel au débit de la prédilution selon le graphique de  $p_7/p_1$ . Ceci est contraire à la logique puisque le débit de la prédilution est fixé en ratio par rapport au débit d'alimentation. En fait, il s'agit d'un concours de circonstance. Lorsque le ratio de prédilution a été augmenté lors du test échelon, le débit



d'alimentation a dû être diminué à cause de contraintes d'opération. C'est pourquoi cette relation est observée. Cette relation déterminée par le modèle PCA confirme l'importance du test échelon sur la prédilution. En effet, il s'agit du seul instant de production au cours duquel cette relation est applicable.

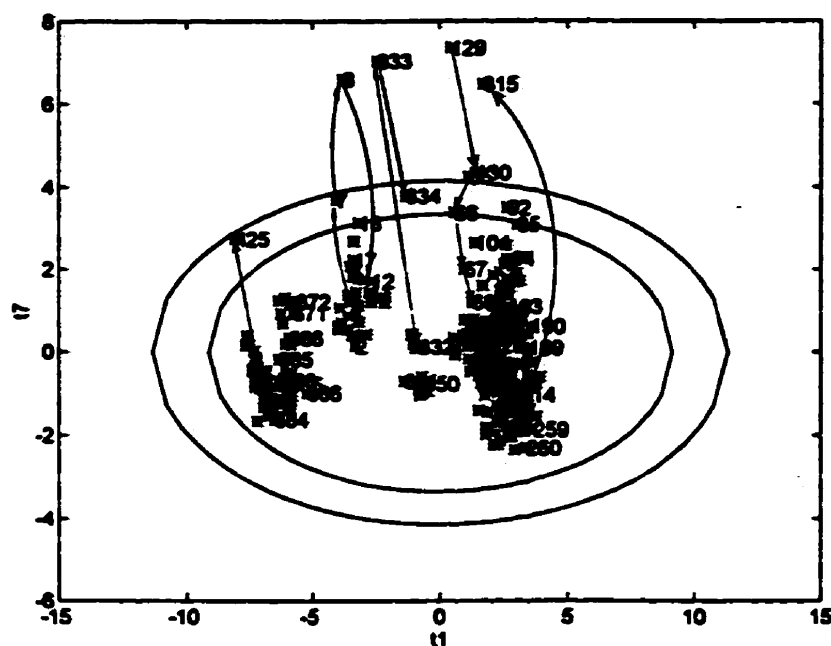


Figure 5.1.21: Graphique  $t_2/t_1$ , modèle PCA, grade 100.

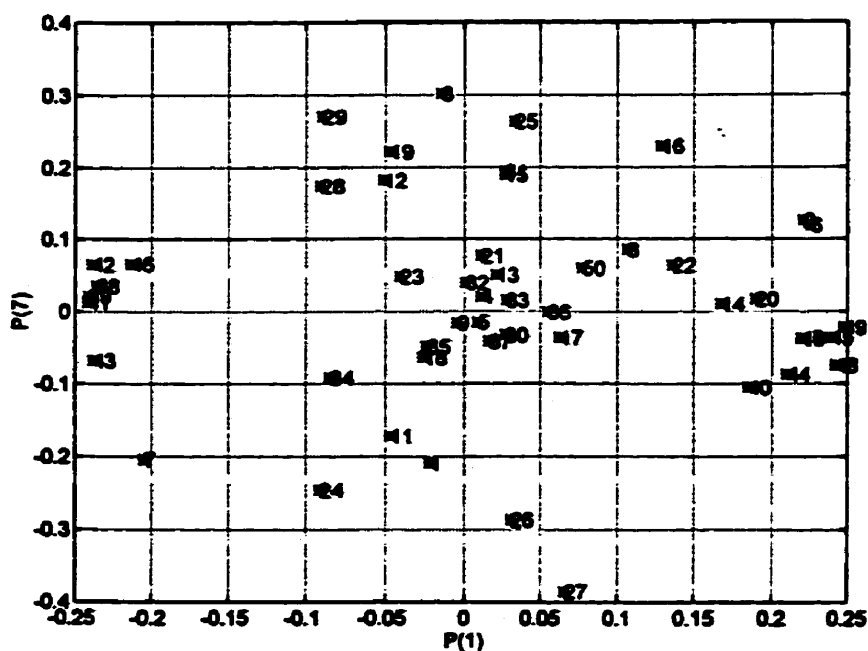


Figure 5.1.22: Graphique  $p_7/p_1$ , modèle PCA, grade 100.

### Dimension 8

#### Variable latente $t_8$

La variable latente  $t_8$  modélise les lavages des filtres A et B. Les lavages sont observés sur le graphique  $t_8/t_1$ : les observations se déplacent vers le bas lors des lavages du filtre A et vers le haut lors des lavages du filtre B.

Aussi, l'augmentation de la température de la prédilution a pour effet de diminuer  $t_8$ :

- observations 113 à 114: données séparées par un court-circuit du réfrigérant aux premiers filtres;
- observations 351 à 352: changement de période d'opération.

De plus, les tests échelons sur la quatrième et la cinquième dilution sont modélisés par cette variable latente.

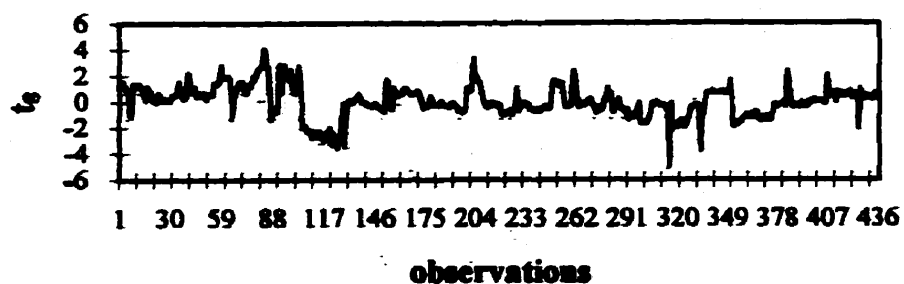


Figure 5.1.23:  $t_2$  en évolution dans le temps, modèle PCA, grade 100.

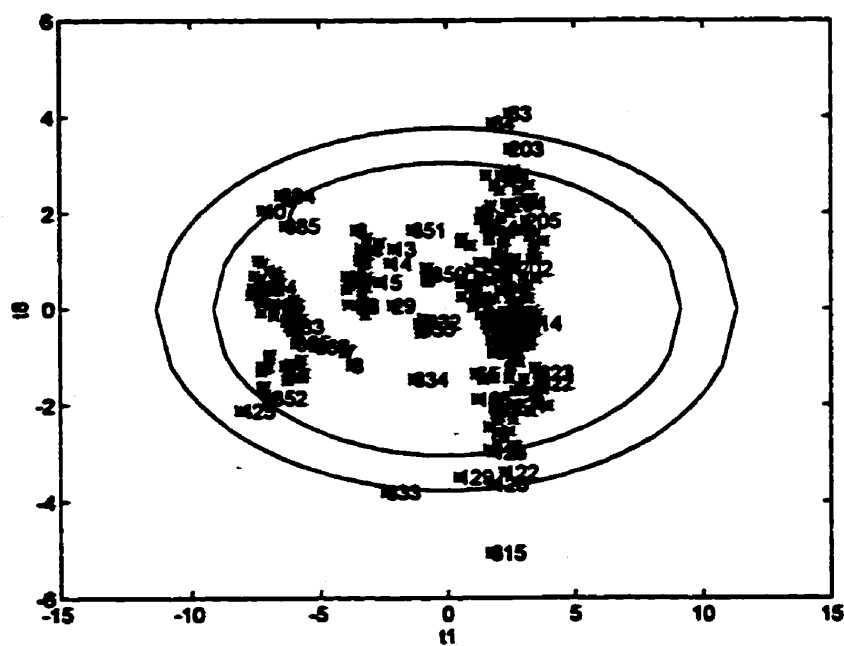


Figure 5.1.24: Graphique  $t_2/t_1$ , modèle PCA, grade 100.

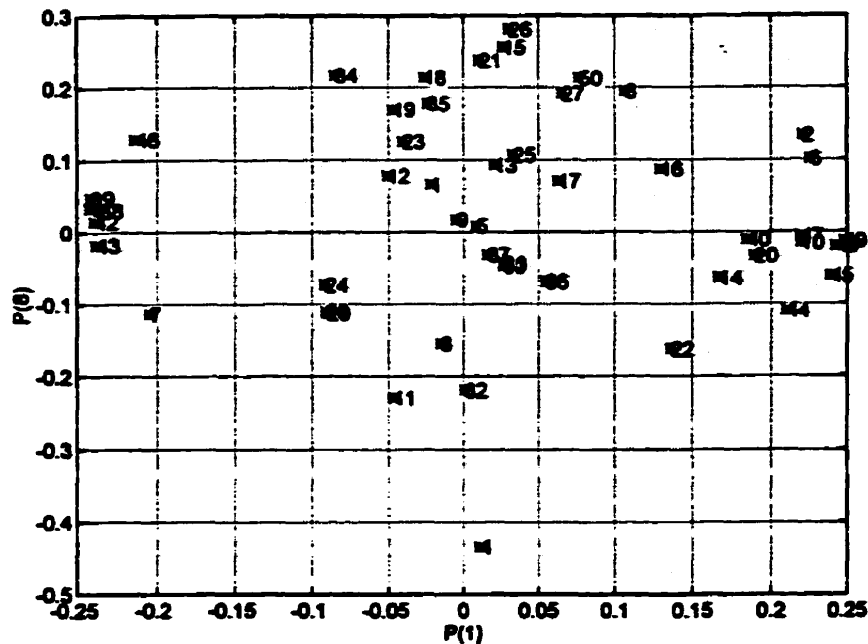


Figure 5.1.25: Graphique  $p_2/p_1$ , modèle PCA, grade 100.

#### Poids $p_2$

Le paramètre d'opération associé au plus haut poids dans cette dimension est la température de la prédilution.

Les paramètres des filtres A et B suivant sont aussi importants:

- niveau de la cuve filtre A;
- solvant froid du filtre A;
- niveau de la cuve du filtre B;
- aspiration basse du filtre B;
- vitesse de rotation du filtre B.

Les débits des deuxième, troisième, quatrième et cinquième dilutions sont aussi significatifs.

### **Dimension 9**

#### **Variable latente $t_2$**

La neuvième dimension du modèle regroupe les tests sur la prédilution et sur la cinquième dilution:

- 16 juillet: test sur la cinquième dilution (observations 65 à 86);
- 6 septembre: test sur la cinquième dilution (observations 287 à 326);
- 12 août: test sur la prédilution (observations 157 à 198);
- observations 16 à 18: retour à sa valeur initiale du débit de la prédilution à la suite du court-circuit;

Lorsque le débit de prédilution augmente,  $t_2$  augmente. Lorsque le débit de cinquième dilution augmente,  $t_2$  diminue.

Le débit de solvant froid au filtre A et sa vitesse de rotation sont modélisés de nouveau par cette dimension: les pics sur le graphique de  $t_2$  en fonction du temps représentent les lavages du filtre A.

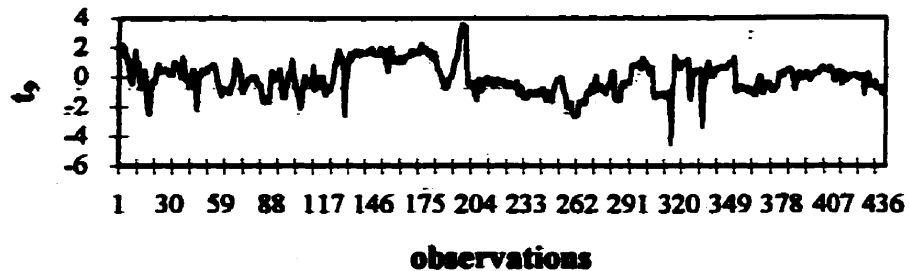


Figure 5.1.26:  $t_p$  en évolution dans le temps, modèle PCA, grade 100.

#### **Poids $p_2$**

Les variables reliées aux tests sur la prédilution sont:

- le débit de prédilution;
- le débit d'alimentation;

Les variables suivantes sont reliées au test sur la cinquième dilution sont:

- le débit de cinquième dilution;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation;
- la température à la sortie de l'échangeur X-2402B;

La vitesse de rotation et le débit de solvant froid du filtre A sont encore importantes dans cette dimension.

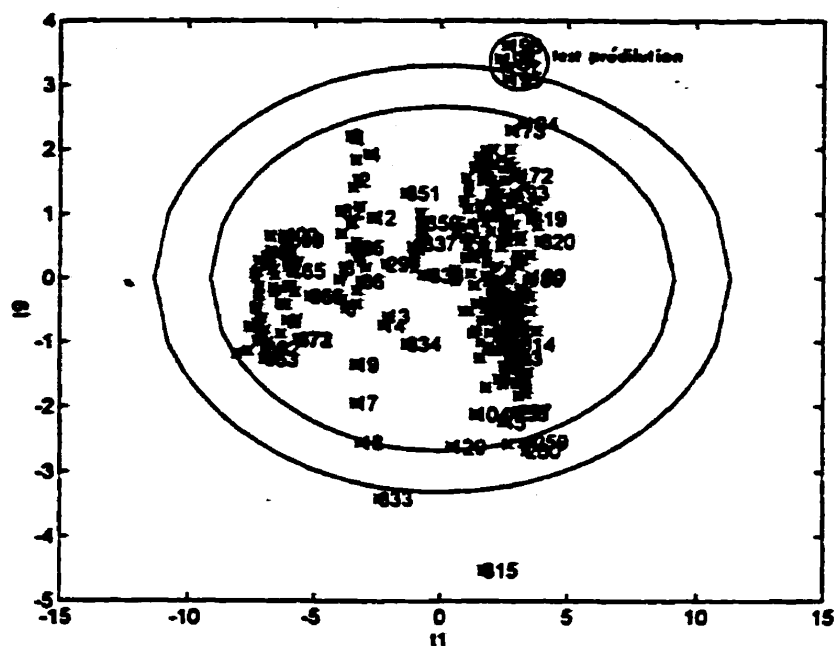


Figure 5.1.27: Graphique  $t_2/t_1$ , modèle PCA, grade 100.

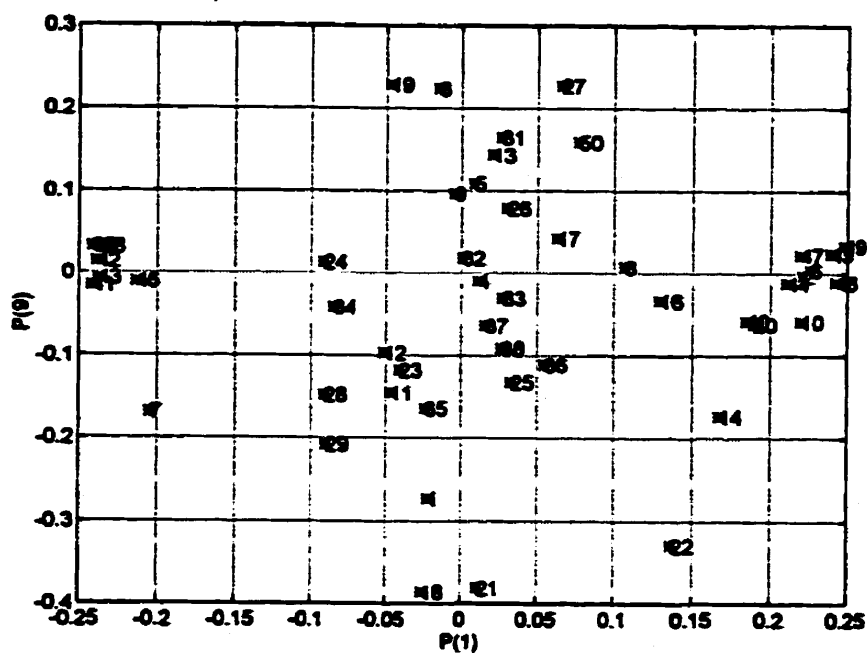


Figure 5.1.28: Graphique  $p_2/p_1$ , modèle PCA, grade 100.

### Dimension 10

#### Variable latente $t_{10}$

La dixième dimension modélise la pression d'aspiration du filtre B et le débit de solvant de lavage froid du filtre A. L'excursion des observations vers des valeurs positives de  $t_{10}$  représente les lavages du filtre A. Les excursions vers des valeurs négatives de  $t_{10}$  représentent les lavages du filtre B.

Les tests échelons sur la prédilution et la cinquième dilution sont également importants dans cette dimension.

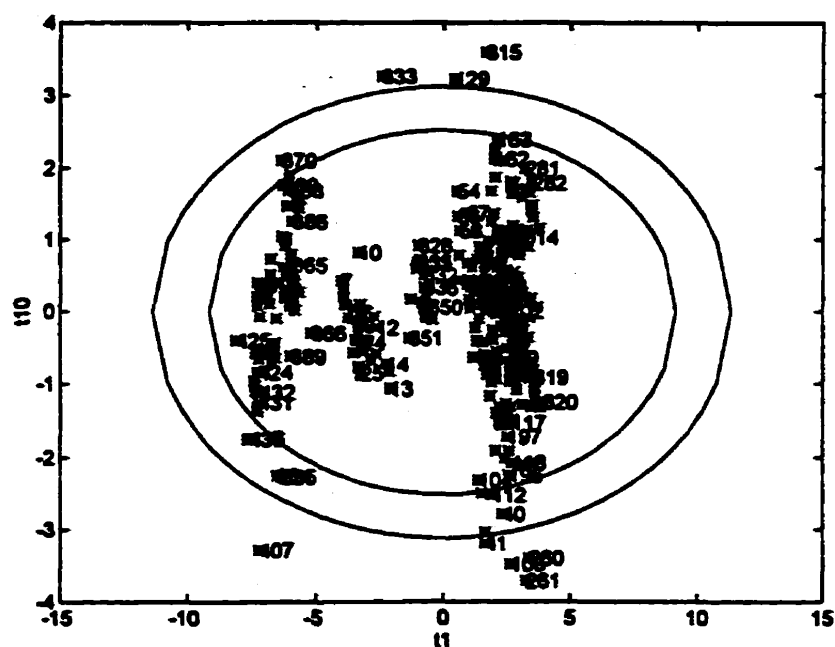


Figure 5.1.29: Graphique  $t_{10}/t_1$ , modèle PCA, grade 100.



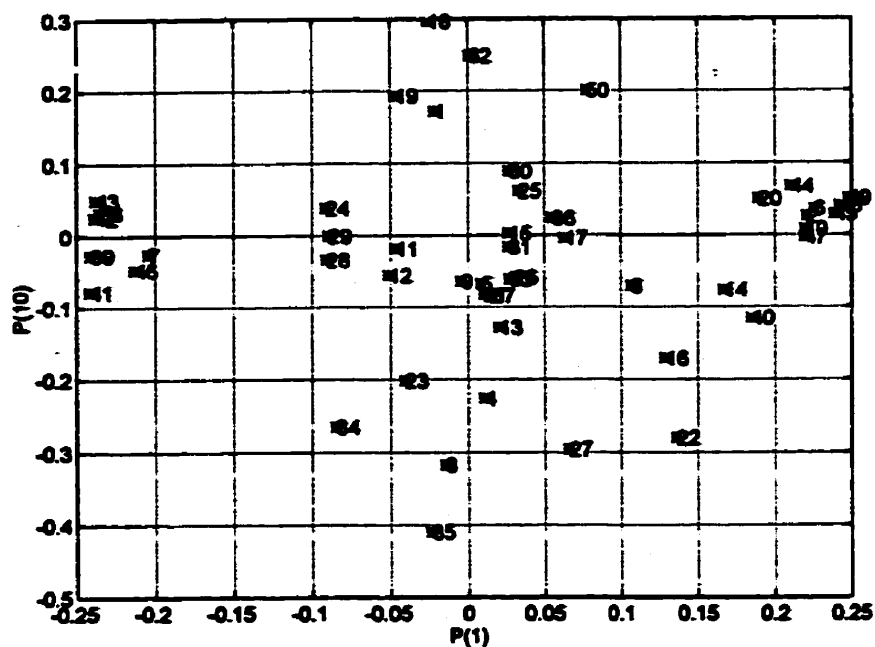


Figure 5.1.30: Graphique  $p_{10}/p_1$ , modèle PCA, grade 100.

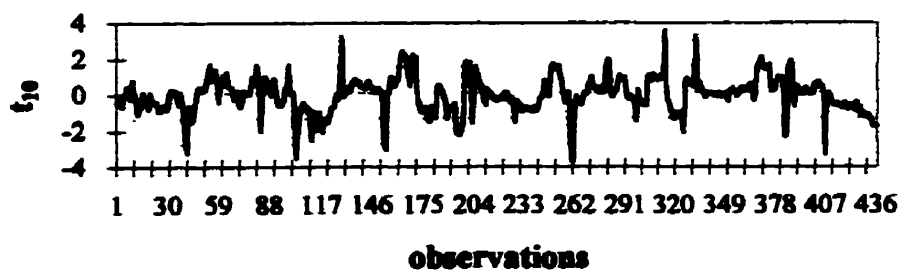


Figure 5.1.31:  $t_{10}$  en évolution dans le temps, modèle PCA, grade 100.

### Poids $p_{10}$

Les variables reliées au lavages sont:

- le niveau de la curve du filtre B;
- la vitesse de rotation du filtre B;
- la pression d'aspiration du filtre B;
- le débit de solvant froid au filtre A;

Cette dimension modélise le cinquième échangeur, et la prédilution une fois de plus:

- débit d'alimentation;
- débit de prédilution;
- débit de la cinquième dilution;
- température de la prédilution;
- température à la sortie de X-2437;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation.

#### **5.1.2.2 Discussion des résultats**

Les quatre premières variables latentes du modèle expliquent à elles seules 59,08 % de la variabilité de X, validée par la validation croisée. Elles contiennent donc les événements marquants de l'opération durant la période étudiée. Des regroupements de données sont observés sur les graphiques des variables latentes pour ces quatre premières dimensions. Ces regroupements signifient qu'il existe plusieurs zones d'opération du procédé pour le grade 100. Les principales causes du changement de zone d'opération relevées par le modèle PCA sont:

- le combinaison des filtres en fonction;
- le profil de température du circuit de cristallisation;
- le débit d'alimentation au procédé;
- les pressions dans le cycle de cristallisation, et;
- le niveau et la température de la botte du filtre A.

Le niveau et la température du filtre A sont relevés par le modèle PCA puisqu'en moyenne ils varient d'une période d'opération à l'autre (voir figures 5.35 et 5.36).

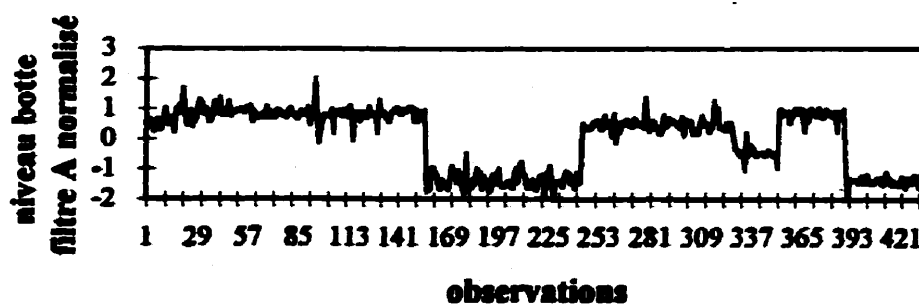


Figure 5.1.32: Niveau de la botte du filtre A normalisé en évolution dans le temps, grade 100.

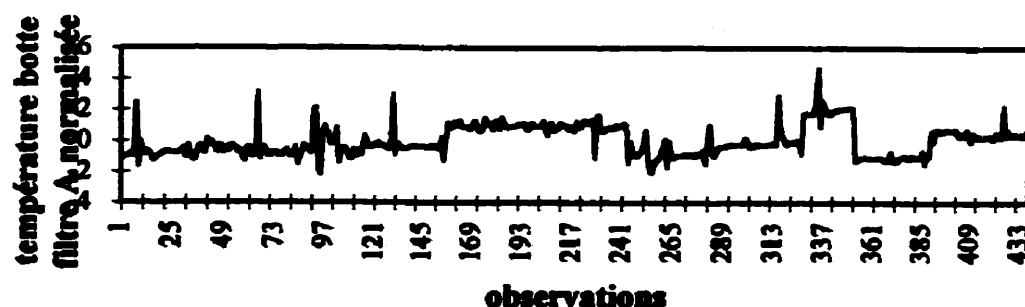


Figure 5.1.33: Température de la botte du filtre A en évolution dans le temps, grade 100.

Les paramètres sont normalisés sur ces figures de façon à conserver la confidentialité de ces données.

Le test échelon sur la quatrième dilution est modélisé par la quatrième dimension. Ce dernier a provoqué beaucoup de variabilité sur le procédé.

Les dimensions 5 à 10 accordent beaucoup d'importance à la variabilité du filtre A. Chacun leur tour les paramètres d'opération du filtre A ressortent comme variables importantes pour ces dimensions, et parfois sur plusieurs dimensions. Ces paramètres d'opération sont l'exemple parfait de l'effet de la dynamique sur les modèles PCA et

**PLS.** Les lavages chauds représentent la source de variation la plus considérable pour le filtre A. À cette occasion, le débit de solvant froid est coupé, la pression d'aspiration augmente à 760 mmHg, le niveau de la cuve du filtre diminue à zéro et la vitesse de rotation augmente. Or, le modèle n'arrive pas à modéliser cet événement en une seule dimension. Ceci s'explique par le fait que tous les paramètres d'opération du filtre n'ont pas la même dynamique. Ils ne reviennent pas à leur état initial exactement au même instant. C'est pourquoi plusieurs dimensions prennent en charge de modéliser cet événement.

La cinquième variable latente souligne l'importance du test échelon sur les vitesses de rotation effectué afin de déterminer la possibilité d'utiliser ces paramètres pour contrôler le niveau du réservoir d'alimentation aux filtres. Il s'agit du seul test sur les filtres considéré important par le modèle PCA. En effet, le traitement préliminaire des données atténue beaucoup la variabilité des paramètres d'opération des filtres B, C et D. Seul le filtre A est considéré avec importance au niveau de son opération de tous les jours. La variabilité d'opération des autres filtres est fortement atténuée par la variabilité causée par leur état: en service ou hors service. La variabilité du filtre B, qui n'est arrêté qu'une seule journée, est la moins affectée. Elle entre donc en considération au niveau des dernières variables latentes du modèle. Les filtres C et D ne présentent pas d'importance, sauf pour ce qui est de leur état. La figure 5.1.37 illustre bien ce point.

Les quatre dernières variables latentes expliquent la variabilité due à des événements distincts, comme les tests échelons. Le débit de prédilution et de la cinquième dilution semblent importants selon ces variables latentes. La plupart des tests échelons sont modélisés par plusieurs dimensions, au niveau de paramètres d'opération différents. Il s'agit une fois de plus de l'effet de la dynamique ou des comportements transitoires créés par les changements de type échelon sur les paramètres étudiés lors des tests.

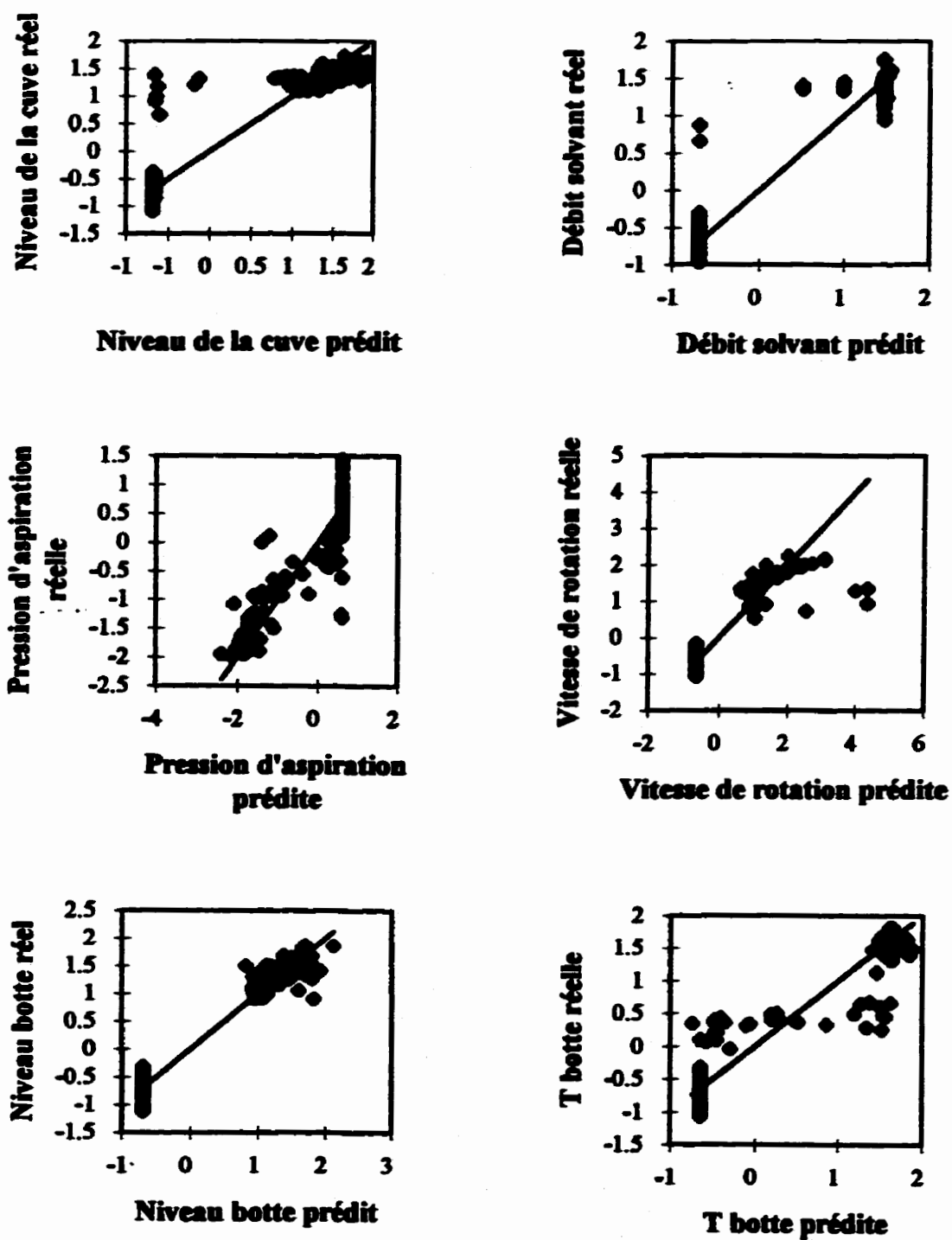


Figure 5.1.34: Paramètres d'opération du filtre C, prédiction en fonction de la valeur réelle.

### 5.1.3 Analyse par la méthode des moindres carrés partiels pour le grade 100

L'analyse des données par la méthode des moindres carrés partiels a pour objectif d'identifier quels sont les paramètres d'opération du procédé qui ont un impact important sur la qualité de la cire (contenu en huile). L'ensemble des données recueillies du 5 juillet au 10 août 1996 sont utilisées. La distribution des données est présentée au début du chapitre.

Le tableau suivant présente les statistiques relatives à l'analyse des moindres carrés partiels. La somme des carrés de l'erreur initiale pour l'espace X est de 21 800 et de 436 pour l'espace Y:

Tableau 5.1.3: Données statistiques de l'analyse PLS sur l'espace Y et l'espace X, grade 100

t	SS <sub>Yordonné</sub> %	SS <sub>Yordonné</sub> cumulée %	SS <sub>Yvalid</sub> croisé cumulée %	Différence SS <sub>Yordonné</sub> - SS <sub>Yvalid</sub> croisé	SS <sub>Xordonné</sub> %	SS <sub>Xordonné</sub> cumulée %	CSV/SD
1	52,38	52,38	52,17	0,21	25,30	25,30	0,692
2	21,56	73,94	73,70	0,24	11,88	37,17	0,743
3	6,86	80,80	80,52	0,28	10,24	47,41	0,865
4	2,62	83,42	83,10	0,32	8,90	56,31	0,938
5	1,74	85,16	84,59	0,57	5,77	62,08	0,964
6	1,50	86,66	86,16	0,50	4,11	66,19	0,966
7	0,62	87,28	86,78	0,50	4,14	70,34	0,996
8	0,57	87,85	87,18	0,67	2,5	72,84	1,004

La dimension du modèle selon le critère CSV/SD est de 7. Le modèle explique 86,78% de la variabilité de Y (validée par la validation croisée). Les trois premières variables

latentes sont les plus importantes. Elles expliquent 80,5 % de la variation de l'indice de réfraction de la cire. La variabilité de l'espace X est expliquée à 70,34 % à la septième variable latente. Ce qui représente un peu moins que ce que les sept premières dimensions du modèle PCA expliquent (74,77 %). En effet, le modèle PLS est construit de façon à expliquer le mieux possible la variabilité de Y au détriment de l'explication de l'espace X. Le graphique du carré de l'erreur de prédiction sur les paramètres d'opération X illustre bien cette situation (figure 5.1.38).

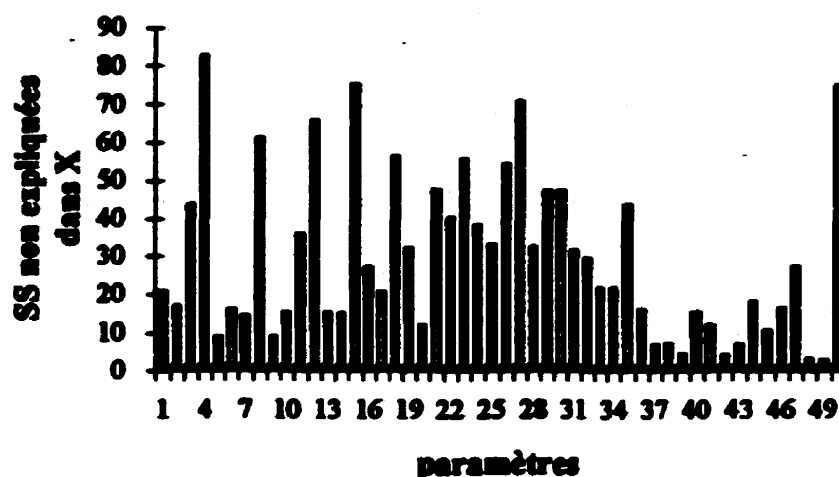


Figure 5.1.38: Pourcentage de la variabilité des paramètres de X non expliquée par le modèle PLS.

Le graphique du carré de l'erreur de prédiction sur l'indice de réfraction au niveau des observations pour le modèle PLS est présenté à la figure 5.1.39. Le modèle PLS obtenu semble bien prédire l'indice de réfraction. Seules quelques observations se retrouvent à l'extérieur de la limite de confiance du modèle. Il s'agit d'événements ponctuels et distincts. Le modèle peut donc être utilisé pour prédire l'indice de réfraction.

La prochaine étape de l'analyse consiste à analyser les résultats obtenus pour les graphiques  $u/t$ . La variable  $u$  représente l'indice de réfraction et  $t$  les événements qui se produisent dans le procédé.

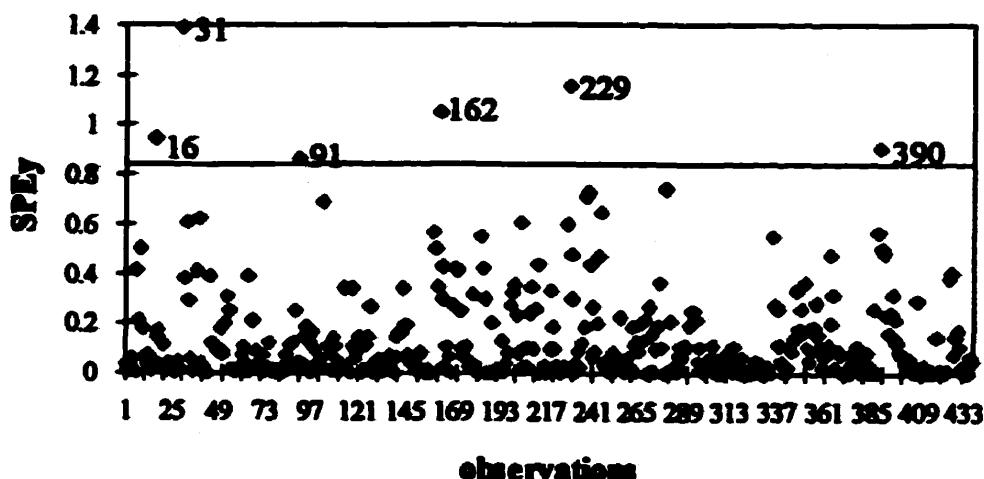


Figure 5.1.39: Somme des carrés des erreurs de prédiction sur l'indice de réfraction  $SPEy$ , modèle PLS, grade 100.

#### 5.1.3.1 Relation entre X et Y

Ces graphiques représentent la relation qui existe entre l'indice de réfraction et les paramètres d'opération du procédé pour la première dimension, et entre les matrices de résidus de Y et de X pour les dimensions subséquentes. La relation  $u/t$  que le modèle PLS calcule est linéaire. Il faut vérifier cette hypothèse. De plus, ces graphiques aident à déterminer si la relation explique des changements globaux ou des événements bien précis.

$$u_1/t_1$$

La relation entre  $u_1$  et  $t_1$  est bien linéaire. Les données se subdivisent en trois groupes qui sont le reflet d'une variation de l'opération: ces groupes représentent la variation de



l'indice de réfraction de la cire en fonction du profil de cristallisation et du débit de filtrat recyclé au réservoir d'alimentation des filtres.

$u_2/t_2$

La relation entre  $u_2$  et  $t_2$  est effectivement linéaire. Aucun regroupement ou événement distinct ne s'observent à première vue.

$u_3/t_3$

La relation entre  $u_3$  et  $t_3$  semble représenter une variabilité globale plutôt qu'un événement distinct. Aucune particularité n'est observée. La relation est linéaire.

$u_4/t_4$

Seule une petite partie des données contribue à cette relation: il s'agit des données du 14 septembre 1996. Pendant cette période d'opération le filtre B est hors service.

$u_5/t_5$

Rien de vraiment remarquable, à part peut-être l'observation 103 qui dévie la corrélation vers le bas. Cet objet représente un des lavages du filtre B.

$u_6/t_6$

Quelques points dirigent la corrélation: ces données représentent les lavages des filtres A et B.

$u_7/t_7$

Cette corrélation est influencée encore une fois par une petite quantité de données: il semble s'agir de la variabilité du débit de la cinquième dilution

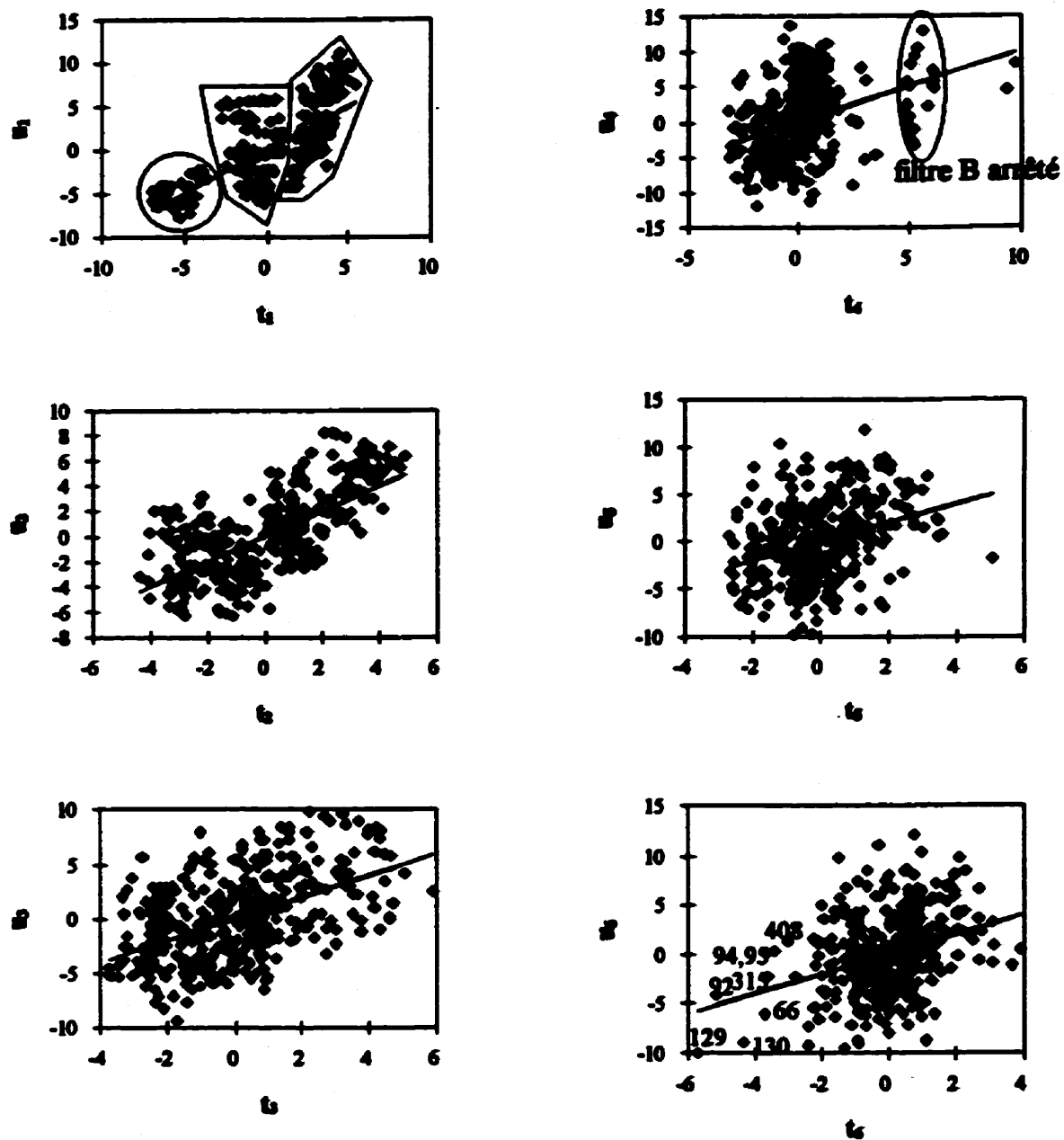


Figure 5.1.40: Graphiques  $u_i/t_i$ , modèle PLS, grade 100.

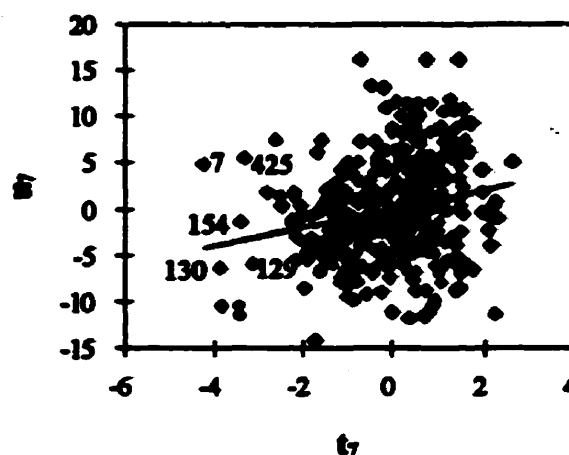


Figure 5.1.40: Graphiques  $u/t_i$ , modèle PLS, grade 100 (suite).

Toutes les relations  $u/t$  sont bien représentées par un modèle linéaire. L'analyse des variables latentes permet d'identifier plus précisément l'événement qui influence l'indice de réfraction dans chacune des dimensions. Les poids étudiés dans le cas du modèle PLS sont les poids  $w$ , définis dans le chapitre 3 sur la théorie des modèles multivariables. Les paramètres associés à un grand poids, sont corrélés avec  $u$ , donc avec l'indice de réfraction.

### 5.1.3.2 Analyse des variables latentes pour le modèle PLS de grade 100

#### Dimension 1

##### Variable latente $t_1$

Cette variable latente représente plusieurs changements d'opération très distincts.

Le débit de filtrat recyclé au réservoir d'alimentation des filtres ainsi que les températures du circuit de cristallisation sont les principales sources de variabilité retenues par cette

dimension. Les changements de période d'opération sont bien identifiées par cette variable latente. De plus, les tests échelons sur la prédilution et la quatrième dilution qui ont eu beaucoup d'impact sur les températures du circuit de cristallisation et sur le débit de filtrat recyclé sont bien modélisés par la première dimension.

La température du solvant froid est aussi modélisée par cette variable latente. Les températures du circuit de cristallisation sont corrélées avec celle-ci.

De plus, la variable  $t_1$  modélise conjointement avec  $t_2$  la combinaison des filtres en opération ABC ou ABD:

- le filtre D en fonction et le filtre C en arrêt: les données se situent dans le plan inférieur droit du graphique  $t_2/t_1$  (observations 37 à 351).
- le filtre C en fonction et le filtre D en arrêt: les données se trouvent dans le plan supérieur gauche de ce graphique (observations 1 à 36 et 352 à 437).

Les observations 327 à 351 se situent plus au centre puisque les filtres C et D sont en fonction au même moment.

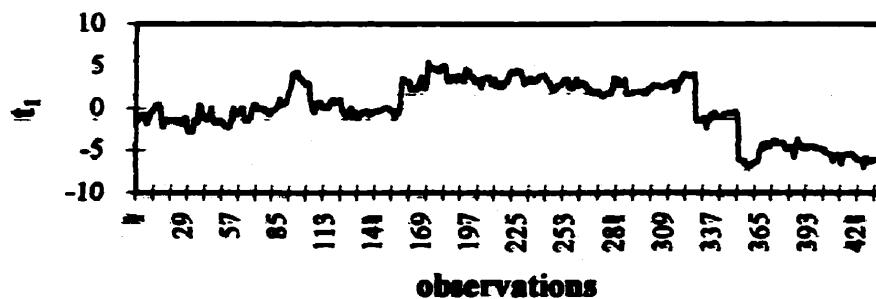


Figure 5.1.41:  $t_1$  en évolution dans le temps, modèle PLS, grade 100.

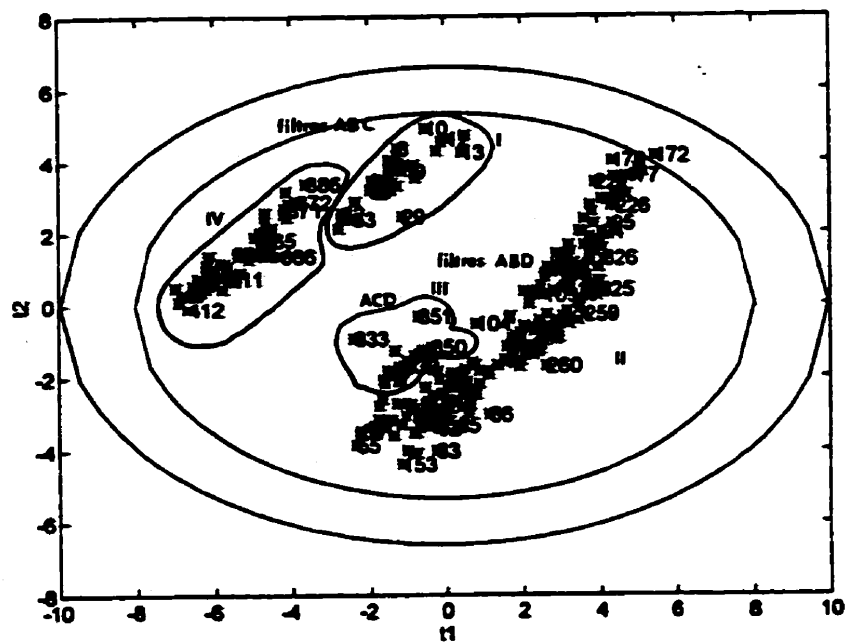


Figure 5.1.42: Graphique  $t_2/t_1$ , modèle PLS, grade 100

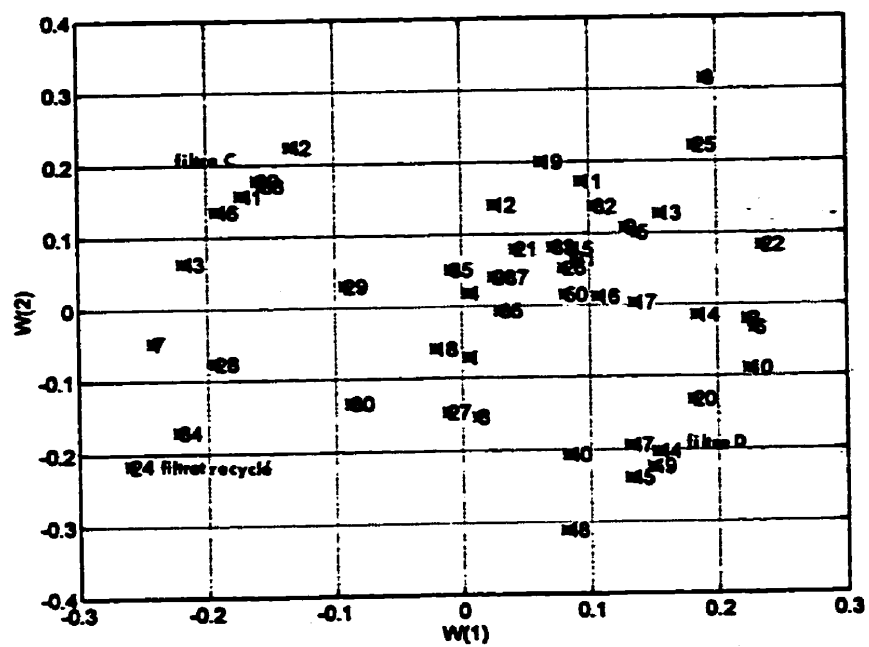


Figure 5.1.43: Graphique  $w_2/w_1$ , modèle PLS, grade 100

**Poids  $w_1$** 

Les paramètres d'opération les plus importants pour la variable latente  $t_1$  sont:

- le débit de filtrat recyclé
- la température de la vapeur 40 lbs/eau recyclée;
- la température à l'entrée de l'échangeur X-2401A;
- la température de la troisième passe de l'échangeur X-2401A;
- la température à la sortie de l'échangeur X-2401A;
- la température à la sortie de l'échangeur X-2401B;
- la température à la sortie de l'échangeur X-2437;
- la température du solvant froid;
- le débit de prédilution;

Les paramètres importants au niveau des filtres sont:

- la pression d'aspiration basse du filtre A;
- la pression d'aspiration basse du filtre B.

Les paramètres qui modélisent la combinaison de filtres sont:

- les niveaux des curves des filtres C et D;
- les débits de solvant froid aux filtres C et D;
- les pressions d'aspiration basse aux filtres C et D;
- les vitesses de rotation des filtres C et D;
- les niveaux et températures des bottes des filtres C et D;

## Dimension 2

### Variable latente $t_2$

Plusieurs variations sont expliquées par  $t_1$  et  $t_2$ :

- combinaison des filtres en opération;
- test échelon sur la prédilution (observations 157 à 198);
- variation de la température du solvant.

Quelques tests supplémentaires sont modélisés par  $t_2$ :

- le test échelon sur la troisième dilution (observations 199 à 243);
- le test échelon sur la quatrième dilution (observations 87 à 127);

Le graphique de  $t_2$  en fonction du temps montre les changements généraux d'opération liés aux variables importantes dans cette dimension.

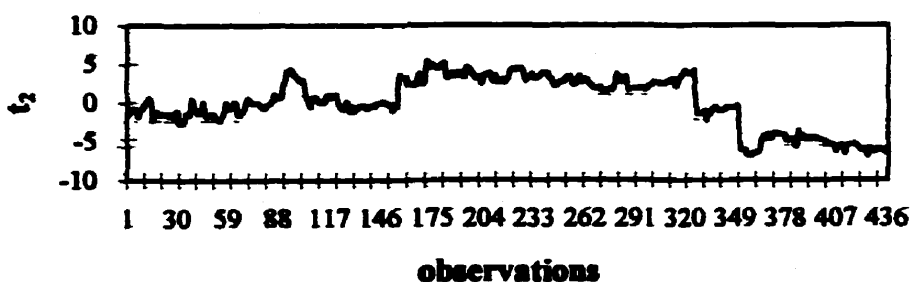


Figure 5.1.44:  $t_2$  en évolution dans le temps, modèle PLS, grade 100.

### Poids $w_2$

Conjointement avec la première dimension, les variables d'opération des filtres C et D sont importantes (voir le graphique 5.1.42).

Le débit de prédilution (en  $t_1$  aussi), le débit de la troisième dilution ont des poids significatifs dans la deuxième dimension. La différence de température entre l'alimentation et la cinquième dilution, la température du solvant froid (en  $t_1$  aussi) et le débit de filtrat recyclé (en  $t_1$  aussi) sont également associées à des poids élevés. Remarquez que les variables considérées importantes également dans la première dimension influencent le profil de cristallisation.

### **Dimension 3**

#### Variable latente $t_3$

Il n'y a pas de regroupement de données par rapport à  $t_3$ . Plusieurs événements sont modélisés par cette variable. D'abord des changements d'opération:

- les variations du débit d'alimentation et des pressions dans le circuit de cristallisation;
- la température à la sortie du cristalliseur X-2402A;
- la variation du niveau dans le réservoir V-2401.

Les tests échelons sur la cinquième dilution sont également modélisés: observations 65 à 86 et 287 à 326 (voir figures 5.1.45 et 5.1.46).



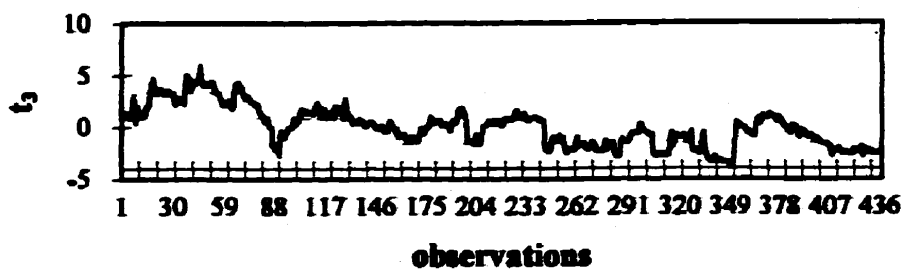


Figure 5.1.45:  $t_3$  en évolution dans le temps, modèle PLS, grade 100.

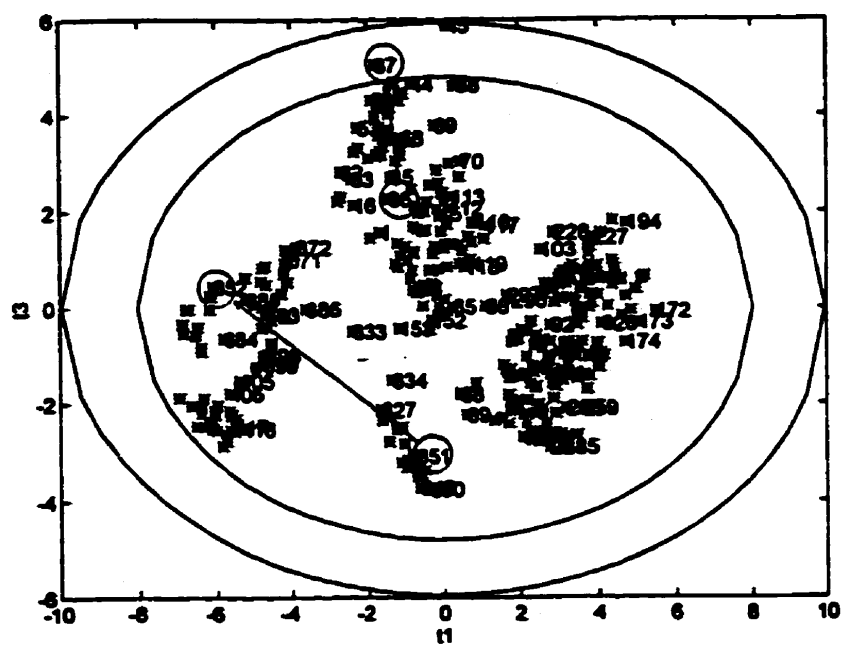


Figure 5.1.46: Graphique  $t_3/t_1$ , modèle PLS, grade 100.

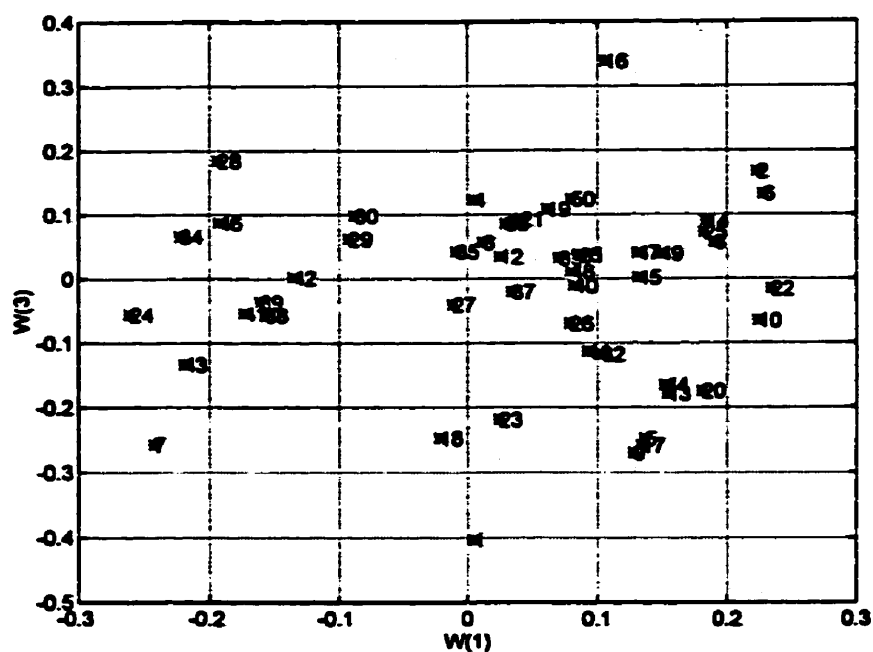


Figure 5.1.47: Graphique  $w_3/w_1$ , modèle PLS, grade 100 .

### Poids $w_3$

Les paramètres d'opération importants sont:

- le débit d'alimentation;
- le débit de la cinquième dilution;
- le niveau du réservoir d'alimentation aux filtres;
- la température à la sortie de l'échangeur X-2402A.

Les pressions à l'entrée de tous les cristalliseurs sont aussi modélisées par cette dimension puisqu'elles sont directement reliées avec les débits.

#### Dimension 4

##### Variable latente $t_4$

Cette variable latente modélise le comportement du filtre B. Sur le graphique de  $t_1/t_4$  (figure 5.1.48), les observations 327 à 351 sont à l'écart du groupe. En effet, pendant cette période d'opération le filtre B est en arrêt. Tous les changements de comportement du filtre B sont aussi reflétés par cette variable latente.

Les lavages sur les filtres A et B sont les pics observés sur le graphique de  $t_4$  en évolution dans le temps. Sur le graphique  $t_4/t_1$ , les lavages sont représentés par les déplacements isolés des observations vers le haut. La variation de  $t_4$  entre les observations 91 et 92 est due à la variation du débit de la troisième dilution.

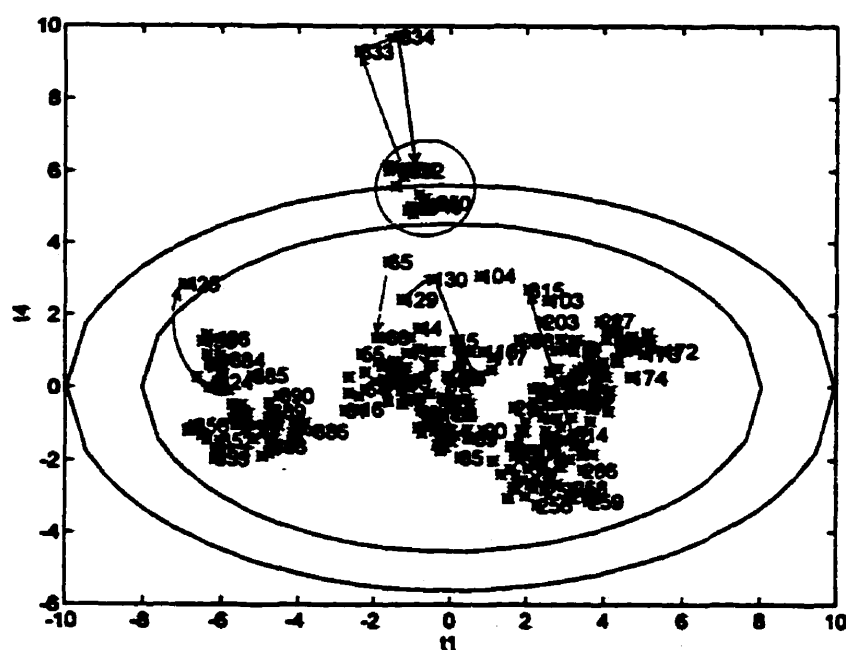


Figure 5.1.48: Graphique  $t_4/t_1$ , modèle PLS, grade 100.

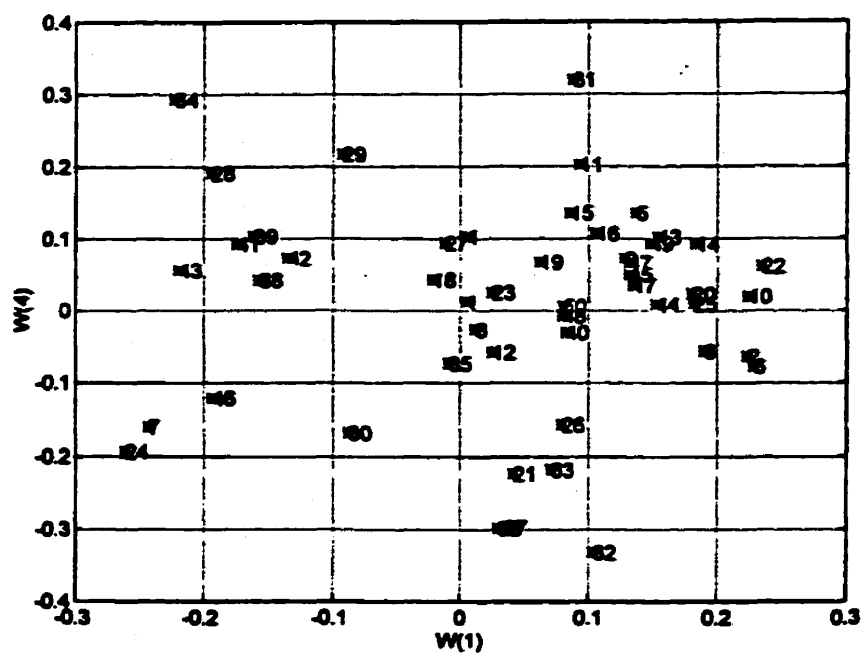


Figure 5.1.49: Graphique  $w_4/w_1$ , modèle PLS, grade 100.

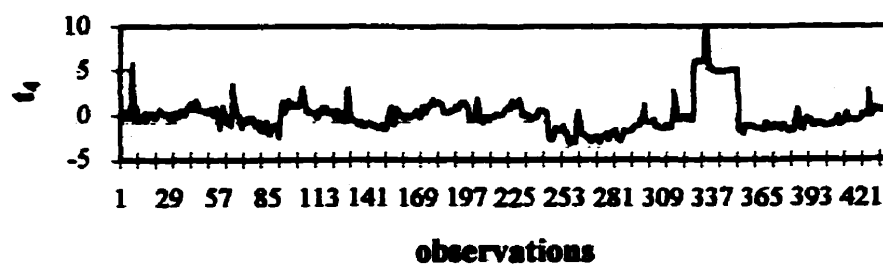


Figure 5.1.50:  $t_1$  en évolution dans le temps, modèle PLS, grade 100.

**Poids  $w_4$** 

Les paramètres d'opération du filtre B sont importantes dans cette dimension:

- niveau de la cuve du filtre B;
- débit de solvant froid au filtre B;
- pression d'aspiration basse au filtre B;
- vitesse de rotation du filtre B;
- niveau et température de la cuve du filtre B.

Certaines des variables du filtre A qui varient lors des lavages sont aussi importantes:

- vitesse de rotation du filtre A;
- pression d'aspiration basse du filtre A;
- niveau et température de la cuve du filtre A.

Le débit de filtrat recyclé qui varie aussi lors des lavages est important dans cette dimension.

Le débit de la troisième dilution ainsi que la température à la sortie de l'échangeur X-2402B ont des poids assez élevés aussi.

**Dimension 5****Variable latente  $t_3$** 

La variable latente  $t_3$  modélise les tests échelons sur les débits de solvant froid aux filtres A et B (observations 364 à 388). Les lavages des filtres A et B ressortent aussi dans cette dimension puisque le débit de solvant froid est coupé pendant cette opération. Le

test du contrôle de niveau du réservoir d'alimentation aux filtres à l'aide des vitesses de rotation des filtre est aussi un événement important.

Le test échelon sur la troisième dilution est aussi modélisé par cette dimension. D'autres variations diverses en relation avec les paramètres d'opération importants dans cette dimension sont identifiées par  $t_3$ .

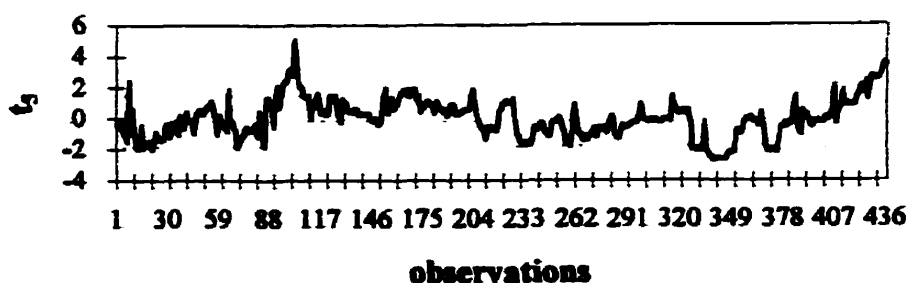


Figure 5.1.51:  $t_3$  en évolution dans le temps, modèle PLS, grade 580.

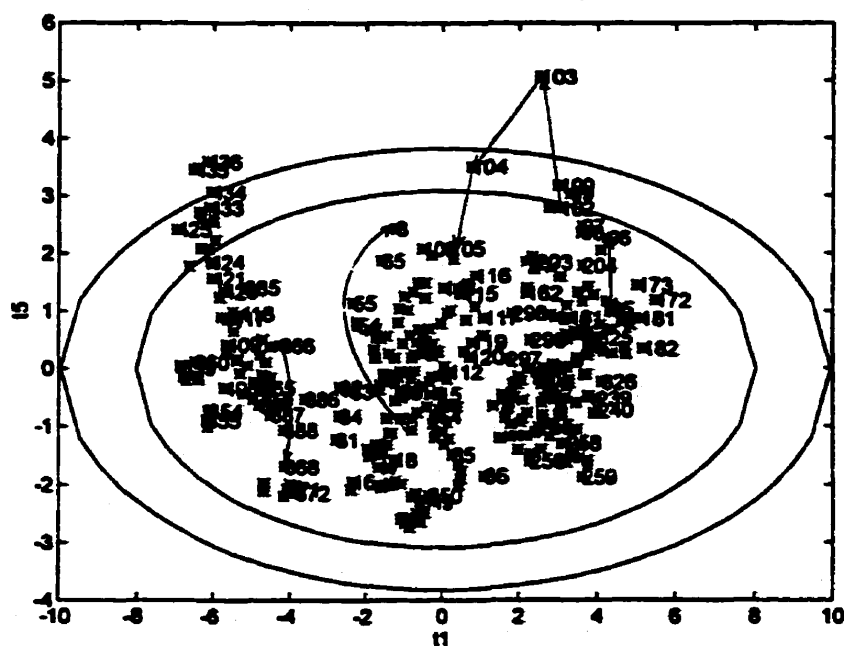


Figure 5.1.52: Graphique  $t_3/t_1$ , modèle PLS, grade 100.

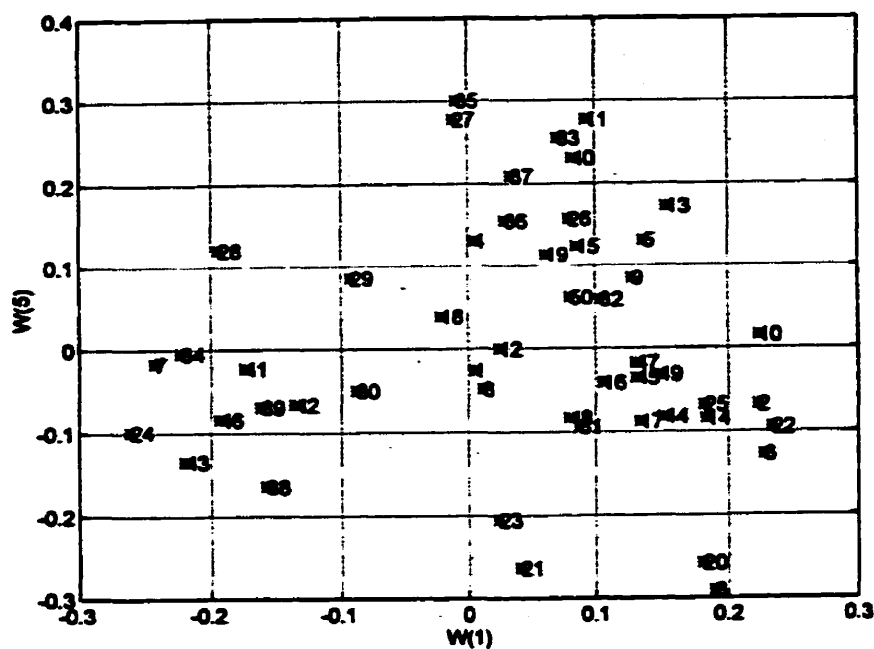


Figure 5.1.53: Graphique  $w_5/w_1$ , modèle PLS, grade 100.

#### Poids $w_1$

Les paramètres d'opération importants sont:

- le débit de solvant au filtre A;
- le débit de solvant au filtre B;
- la vitesse de rotation du filtre B;
- la température de la botte du filtre B;
- la pression d'aspiration basse du filtre C.
- le niveau du réservoir V-2401.

Les débits de la troisième dilution et de la prédilution sont aussi importants. La température à la sortie de l'échangeur X-2402B et la pression à l'entrée de l'échangeur X-2437 se démarquent également.

## **Dimension 6**

### **Variable latente $t_6$**

Plusieurs événements sont modélisés dans cette dimension:

- lavages du filtre A: observations 128 à 129, 332 à 333, et 424 à 425;
- lavages du filtre B: observations 383 à 384, et 407 à 408.

Encore quelques variations sur la température à la sortie de l'échangeur X-2401B, sur le différentiel de température entre la cinquième dilution et l'alimentation, sur la température du solvant froid et sur le débit de la troisième dilution sont modélisées dans cette dimension.

- observations 9 à 10;
- observations 31 à 32;
- observations 64 à 66;
- observations 91 à 92;
- observations 92 à 96;
- observations 199 à 243.

### **Poids $w_6$**

Les variables importantes dans cette dimension sont:

- le débit de la 3<sup>e</sup> dilution;
- la température à la sortie de l'échangeur X-2401B;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation;
- la température du solvant froid.

Aussi, les variables suivantes du filtre A sont modélisées: le niveau de la cuve, la vitesse de rotation et la température de la botte.



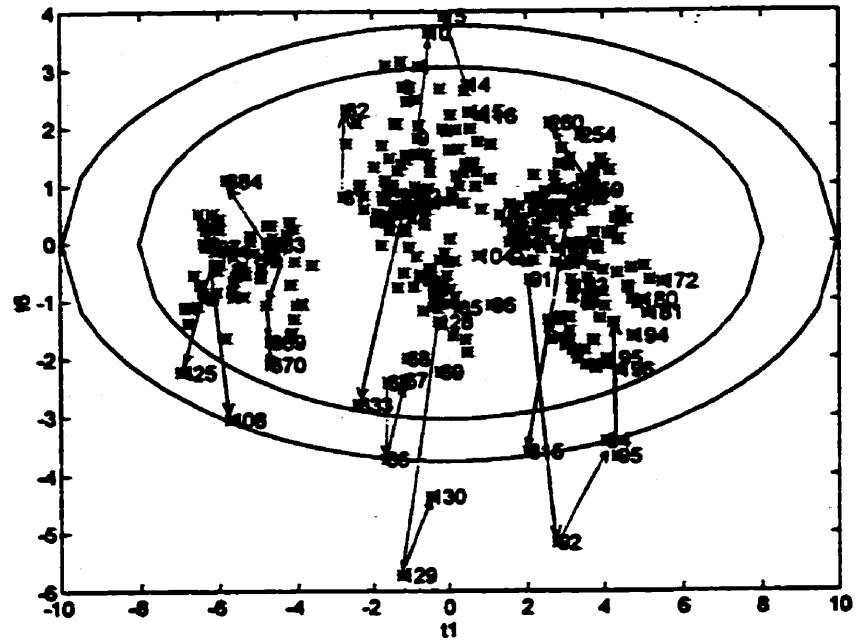


Figure 5.1.54: Graphique  $t_2/t_1$ , modèle PLS, grade 100.

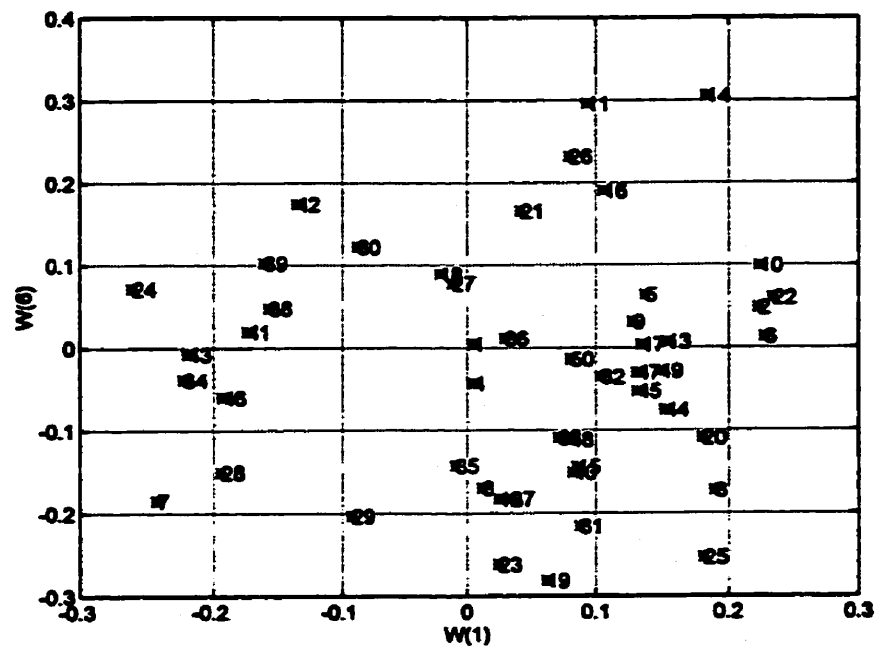


Figure 5.1.55: Graphique  $w_2/w_1$ , modèle PLS, grade 100.

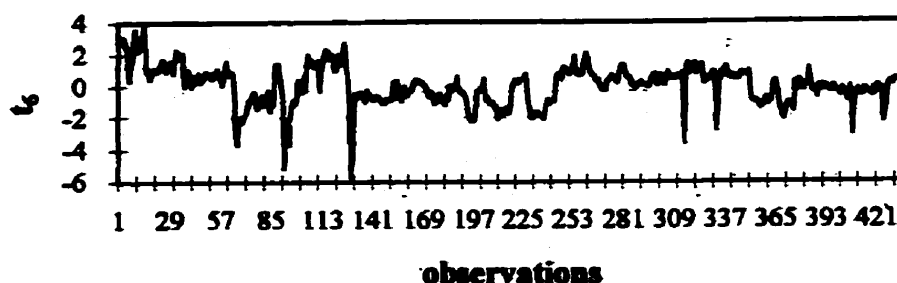


Figure 5.1.56:  $t_6$  en évolution dans le temps, modèle PLS, grade 100.

## Dimension 7

### Variable latente $t_7$

Cette variable latente modélise le test échelon sur la cinquième dilution (observations 65 à 86 et 287 à 326). De plus, les débordements de l'épouseur de la section de récupération du solvant dans la cire sont identifiés par cette dimension (observations 98, 247 à 255, 278 à 283).

Les lavages des filtres sont encore modélisés par cette variable latente ainsi que le test échelon sur les vitesses de rotation effectué en vue d'élaborer une nouvelle stratégie de contrôle de niveau sur le réservoir V-2401.

### Poids $w_7$

La variable la plus importante de cette dimension est le débit de la cinquième dilution. Les variables qui se distinguent avec moins d'importance sont : le niveau de l'épouseur de la section de récupération du solvant et les pressions à l'entrée des échangeurs X-2401B, X-2402A, X-2402B.

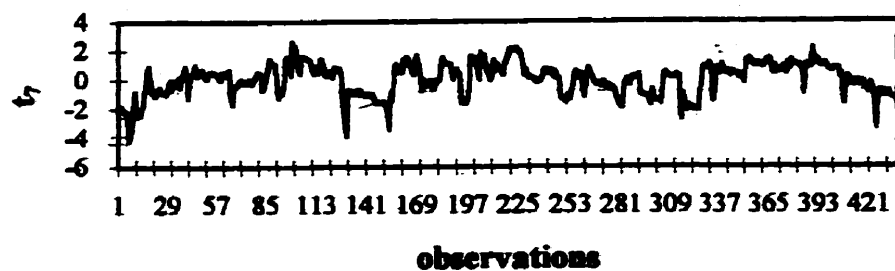


Figure 5.1.57:  $t_7$  en évolution dans le temps, modèle PLS, grade 100.

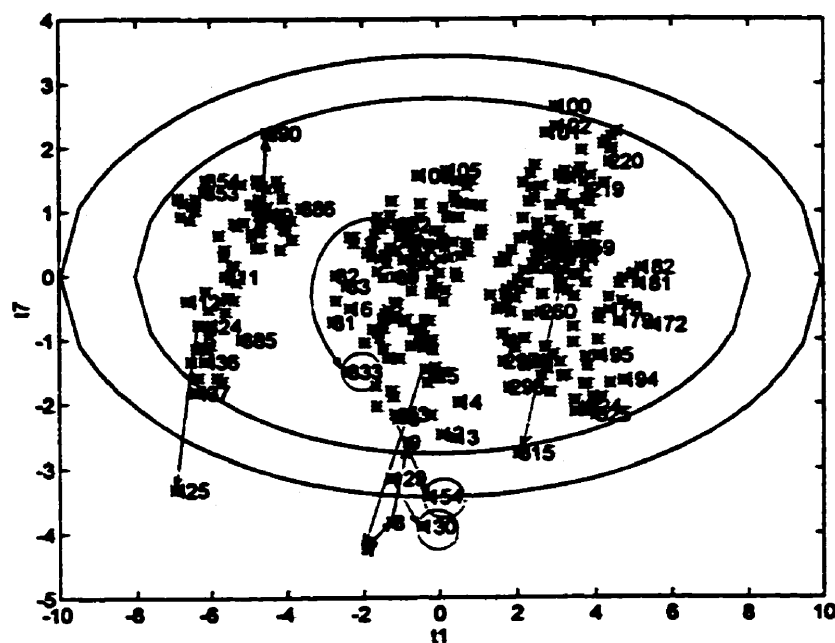


Figure 5.1.58: Graphique  $t_7/t_1$ , modèle PLS, grade 100.

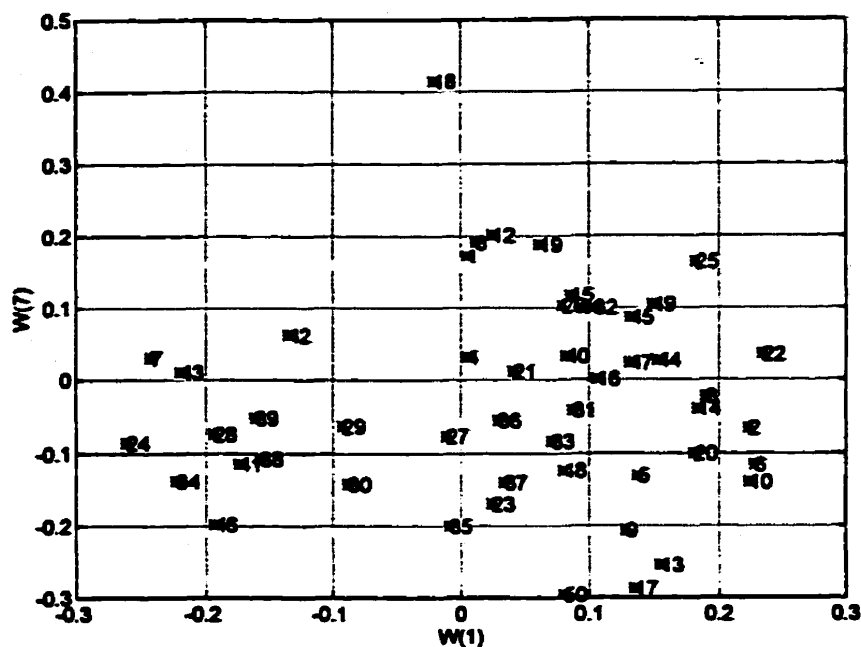


Figure 5.1.59: Graphique  $w_7/w_1$ , modèle PLS, grade 100.

### 5.1.3.3 Discussion des résultats

Les trois premières variables latentes du modèle PLS modélisent 80,52 % de la variabilité de l'indice de réfraction, validée par la validation croisée. Les causes principales de la variation de l'indice de réfraction sont donc identifiées par ces trois variables latentes. Il s'agit:

- du débit de filtrat recyclé au réservoir d'alimentation aux filtres;
- du profil de cristallisation;
- de la température du solvant froid;
- du choix des filtres, et;
- du débit d'alimentation au procédé.

Les tests échelons qui influencent le plus l'indice de réfraction (soulevés par ces trois dimensions) sont les tests échelon sur les ratios de la quatrième dilution, de la prédilution, de la troisième dilution, de la cinquième dilution et sur le débit de solvant de lavage froid alimenté aux filtres.

Les dernières variables latentes ( $t_4$  à  $t_7$ ) expliquent 6,26 % de la variabilité de l'indice de réfraction. Le comportement des filtres A et B est modélisé par ces variables latentes: lavage des filtres, variation de l'opération. De plus, certains événements spéciaux sont modélisés: le débordement de l'épouseur dans la section de récupération du solvant de la cire, le test échelon sur les vitesses de rotation des filtres (contrôle de niveau du réservoir V-2401) et la variabilité température à la sortie de l'échangeur X-2402B. Le reste de la variabilité expliquée par le modèle PLS semble représenter le comportement transitoire de certaines variables qui réapparaissent sur plusieurs dimensions: débits de la troisième et de la cinquième dilution, la température du solvant froid et les pressions à l'entrée des cristalliseurs, déjà modélisés par les premières variables latentes.

L'objectif du projet étant d'identifier les paramètres qui influent le plus la qualité de la cire, le calcul de l'influence des paramètres est effectué. Il consiste à additionner le poids de chacun des paramètres sur chacune des dimensions et permet ainsi d'identifier ceux qui sont globalement les plus importants. Le résultat de ce calcul est présenté dans la prochaine section.

#### 5.1.3.4 Variables les plus influentes sur la qualité de la cire de grade 100

L'indice de réfraction peut être calculé à l'aide du modèle linéaire qui découle du modèle PLS:

$$Y = X \cdot B + F$$

où

$$B = [b_1, b_2, b_3, \dots, b_{30}] = W \cdot (P^T W)^{-1} \cdot Q \quad (5.2)$$

ainsi

$$Y = b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_{30} \cdot x_{30} + F$$

L'importance d'un paramètre est déterminée à partir du calcul de l'influence des variables X sur la variable Y. Le calcul de l'influence s'effectue comme suit:

$$Influence = \sqrt{\frac{\sum w_{ak}^2 \cdot SSy_{expliquée(a)}}{SSy_{expliquée}} \cdot K} \quad (5.3)$$

- où  $w_{ak}$  est le poids du modèle PLS attribué au paramètre k pour la dimension a;  
 $SSy_{expliquée(a)}$  est la somme des carrés de l'erreur expliquée par la a<sup>ème</sup> dimension du modèle;  
 $SSy_{expliquée}$  est la somme des carrés de l'erreur expliquée par le modèle PLS complet;  
 K est le nombre de paramètres d'opération dans la matrice X.

Ainsi, la somme des carrés des influences est égale au nombre de paramètres d'opération dans la matrice X. Un paramètre X est important si l'influence qui lui est associée est supérieur à 0,8 selon Wold, Kettaneh Wold et MacGregor (1996).

Trente-trois variables sont importantes selon ce calcul (voir figure 5.1.60). Le tableau 5.1.4 classe ces variables en ordre décroissant de leur influence sur l'indice de réfraction. La flèche dans la colonne juxtaposée indique s'il faut augmenter ou diminuer ce paramètre pour diminuer l'indice de réfraction de la cire.

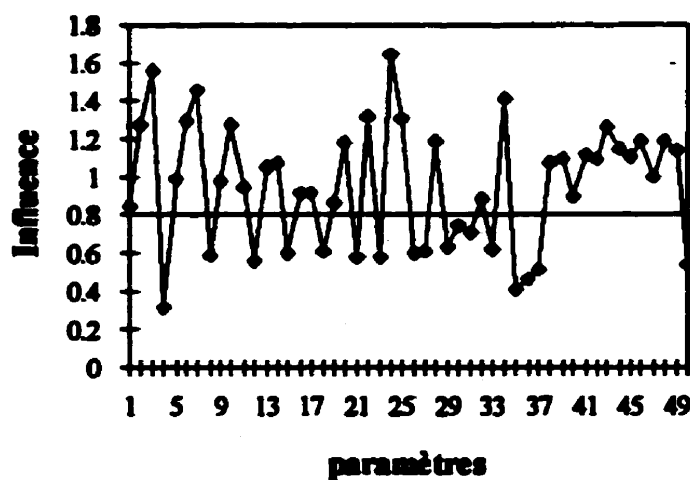


Figure 5.1.60: Influence des variables sur l'indice de réfraction.

**Tableau 5.1.4: Liste des paramètres d'opération qui influencent la qualité de la cire de grade 100**

<b>Paramètre d'opération</b>	<b>Action à prendre</b>
Débit de filtrat recyclé	↑
Débit de prédilution	↓
Température de la 3 <sup>e</sup> passe de X-2401A	↑
Pression d'aspiration basse du filtre B	↑
Température à la sortie de X-2437	↓
Température du solvant froid	↓
Température à l'entrée de l'échangeur X-2401A	↓
Température à la sortie de l'échangeur X-2401A	↓
Température de la vapeur 40lb/eau recyclée	↓
Température de la botte du filtre C	↑
Pression d'aspiration basse du filtre A	↑
Pression à l'entrée de X-2437	↑
Pression d'aspiration basse du filtre D	↑
Niveau de la botte du filtre D	↓
Niveau de la cuve du filtre D	↑
Température de la botte du filtre D	↓
Vitesse de rotation du filtre C	↑
Débit de solvant froid du filtre D	↑
Niveau de la botte du filtre C	↓
Débit de solvant froid au filtre C	↓
Niveau de la cuve du filtre C	↑
Température à la sortie de l'échangeur X-2401B	↓
Pression à l'entrée de l'échangeur X-2402A	↓



**Tableau 5.1.4: Liste des paramètres d'opération qui influencent la qualité de la cire de grade 100 (suite)**

<b>Paramètre d'opération</b>	<b>Action à prendre</b>
Vitesse de rotation du filtre D	↑
Pression à l'entrée de l'échangeur X-2401A	↓
Pression à l'entrée de l'échangeur X-2401B	↓
Débit de la troisième dilution	↓
Température à la sortie de l'échangeur X-2402A	↓
Pression à la sortie de l'échangeur X-2402B	↑
Pression d'aspiration basse du filtre C	↑
Niveau de la cuve du filtre B	↑
$\Delta T$ 5 <sup>e</sup> dilution/alimentation	↓
Débit d'alimentation	↑

Le modèle PLS pour le grade 100 accorde au débit de filtrat recyclé le plus grand rôle. Il s'agit du paramètre d'opération le mieux corrélé avec l'indice de réfraction. Le graphique représentant ce débit en parallèle avec l'indice de réfraction démontre très bien cette corrélation (figure 5.1.61). Lorsque le débit de filtrat recyclé est à son maximum, l'indice de réfraction est faible, et vice-versa.

Une relation causale entre l'indice de réfraction et ce débit est très plausible. Le contenu en huile dans la cire du grade 100 étant faible, un changement de concentration de la solution à filtrer peut avoir un impact important. En effet, le filtrat recyclé au réservoir d'alimentation des filtres est composé presque exclusivement de solvant. Ainsi, plus on ajoute de ce filtrat au débit provenant du cycle de cristallisation, plus on dilue la concentration en huile de la solution à filtrer. La concentration en huile du liquide dans les interstices du gâteau à son entrée dans la section de lavage du filtre est alors plus

faible. Le processus de lavage du gâteau parvient donc à mieux le laver et de meilleurs résultats sont obtenus. L'importance du ratio solide-liquide de la solution alimentée aux filtres est aussi un aspect très important selon McKetta (1989). Si la concentration de solides dans le liquide à filtrer est trop élevée, les cristaux s'agglomèrent et un gâteau épais se forme. Ce dernier rend la filtration plus difficile et un contenu en huile de la cire plus élevé en résulte.

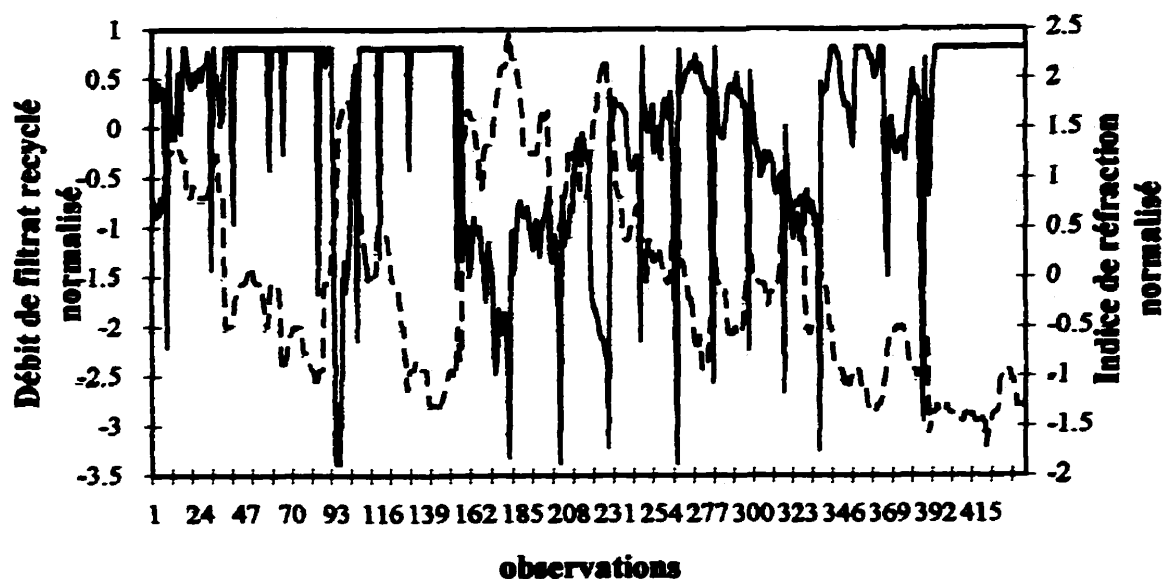


Figure 5.1.61: Débit de filtrat recyclé au réservoir d'alimentation des filtres et indice de réfraction de la cire de grade 100.

Cette relation causale est appuyée par plusieurs autres faits. D'abord, la figure 5.1.61 illustre bien que l'indice de réfraction moyen de la cire est plus élevé lors des périodes d'opération pendant lesquelles le débit de filtrat recyclé moyen est faible. Les tests échelons sur les débits de dilution démontrent de façon encore plus exhaustive cette relation de causalité. Les tests qui s'avèrent influents sur l'indice de réfraction sont ceux qui ont eu un impact sur le débit de filtrat recyclé. Comme le débit de filtrat recyclé sert à contrôler le niveau du réservoir d'alimentation aux filtres, les augmentations agressives d'un débit de dilution entraînent une diminution du débit de filtrat recyclé et provoquent

une augmentation de l'indice de réfraction. Par ailleurs, les tests échelons sur les dilutions jugés non significatifs selon le modèle PLS correspondent aux périodes d'opération pendant lesquelles le débit de filtrat recyclé demeure constant. Il appert donc qu'il est très important de stabiliser ce débit pour stabiliser le contenu en huile dans la cire de grade 100. La stratégie de contrôle du niveau dans le réservoir V-2401 devrait donc être modifiée. Le test sur le contrôle de ce niveau à l'aide des vitesses de rotation et des pressions d'aspiration effectué le 10 octobre 1996 a eu peu d'influence sur l'indice de réfraction en comparaison de l'impact du débit de filtrat recyclé. Il semble donc que ce soit une alternative intéressante pour contrôler le niveau de ce réservoir.

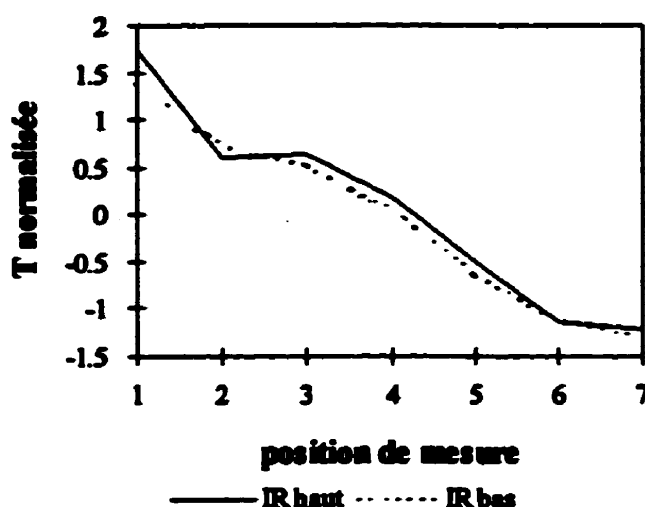


Figure 5.1.62: Profil de température du cycle de cristallisation lors de la production de cire à haut indice de réfraction et à bas indice de réfraction.

Le deuxième groupe de paramètres que le modèle PLS fait ressortir sont les températures du circuit de cristallisation. Les meilleurs résultats d'indice de réfraction sont obtenus lorsque les températures du cycle de cristallisation et du solvant froid sont plus froides. Le graphique à la figure 5.1.62 présente le profil de cristallisation pour des opérations à faibles indices de réfraction et le profil pour des opérations à indices de réfraction élevés.

Les températures du début du cycle sont les plus importantes. L'indice de réfraction est plus faible lorsque la température de l'alimentation est plus froide. Aussi, un refroidissement plus graduel est favorable. C'est pourquoi la température de la troisième passe de l'échangeur X-2401A a beaucoup d'impact selon le modèle PLS.

Les débits de prédilution et de la troisième dilution, ainsi que le débit d'alimentation sont aussi significatifs. Afin de réduire le contenu en huile dans la cire, selon le modèle PLS, il faut diminuer les débits de prédilution et de la troisième dilution et augmenter le débit d'alimentation. Les conclusions tirées sur les débits de dilution doivent être revérifiées une fois que le débit de filtrat recyclé sera stabilisé puisqu'ils sont corrélés avec ce débit. La conclusion sur le débit d'alimentation est quant à elle contraire aux attentes. En fait, la relation observée ici n'est pas causale. Elle est reliée aux tests échelons sur les dilutions. Ces tests échelons visaient à vérifier l'influence des ratios entre les débits de dilution et l'alimentation sur l'indice de réfraction. À cause des limites de capacité de filtration, le débit d'alimentation est réduit lors de ces tests pour donner plus de latitude sur la variation de ces ratios. C'est pourquoi cette relation entre l'alimentation et l'indice de réfraction est observée.

Les résultats du calcul de l'influence montre aussi que l'indice de réfraction de la cire de grade 100 est plus faible si les pressions dans le cycle de cristallisation sont plus faibles. De faibles pressions correspondent à de faibles débits d'alimentation et de dilution. Ce résultat confirme que la relation entre l'indice de réfraction et le débit d'alimentation n'est pas causale.

Les prochains paragraphes analysent l'impact des filtres sur l'indice de réfraction. En premier lieu, le modèle PLS spécifie que de meilleurs résultats sur le contenu en huile dans la cire sont obtenus lorsque les pressions d'aspiration aux filtres A et B sont élevées. En effet, lorsque ces pressions sont plus élevées la compaction du gâteau est moindre.

Ainsi, le gâteau offre moins de résistance au solvant de lavage qui passe plus efficacement à travers celui-ci. De plus, l'augmentation de la pression d'aspiration diminue les risques de craquage du gâteau, qui sont désastreux au point de vue du contenu en huile dans la cire. Les craques créent des courants préférentiels à travers lesquelles le solvant de lavage s'écoule. Les pressions d'aspiration peuvent également être influencées par la viscosité du fluide à filtrer. Moins le fluide est visqueux, plus la pression d'aspiration est difficile à créer. En effet, les pressions d'aspiration les plus élevées sont observées au moment où le débit de filtrat recyclé est maximum (figure 5.1.63), donc lorsque la solution à filtrer est la moins visqueuse. L'effet direct des pressions est remis en question à cause de sa corrélation avec le débit de filtrat recyclé. Par contre, il peut s'agir de l'effet combiné des deux paramètres, qui mènent tous deux à de meilleurs résultats lorsqu'ils sont élevés.

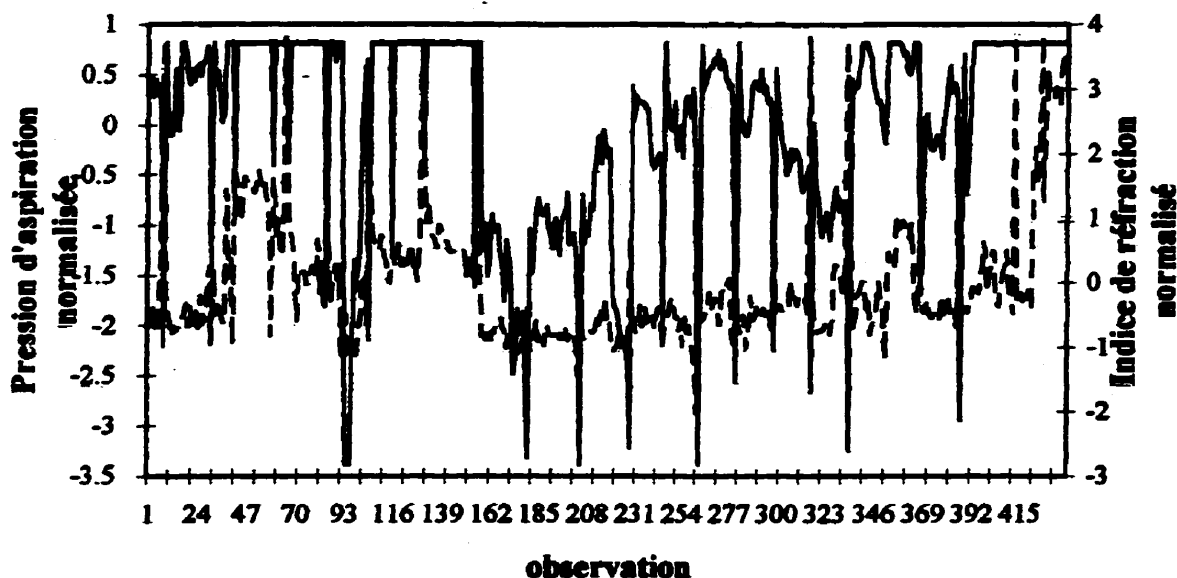


Figure 5.1.63: Débit de filtrat recyclé et pression d'aspiration du filtre A.

Le choix idéal des filtres à opérer est difficile à faire à partir des coefficients obtenus. Ils se contredisent. Ils ne permettent donc pas de conclure si un filtre est meilleur que

**l'autre. De bons résultats d'indice de réfraction sont obtenus pour toutes les combinaisons de filtres en opération. L'importance de ces paramètres peut probablement être reliée au fait que leur variabilité est considérable à cause des situations où ils sont hors service. En présence d'une structure de variation très forte dans la matrice X, le modèle PLS la considérera bien qu'elle n'explique pas nécessairement le comportement de l'indice de réfraction. En effet, la première dimension du modèle n'accorde pas toute l'importance à cet événement, mais aussi à plusieurs autres paramètres d'opération.**

**Finally, la température du solvant froid est significative selon le calcul de l'influence des paramètres. Cette température a un impact à plusieurs niveaux du procédé. Au niveau du circuit de cristallisation, plus cette température est froide, plus les températures à la sortie des trois derniers échangeurs sont froides. Ce paramètre est aussi directement corrélé avec le différentiel de température de la cinquième dilution. Or, toutes ces températures sont considérées importantes par le modèle et doivent toutes diminuer, comme la température du solvant froid. D'autre part, la température du solvant froid est très importante au niveau de la filtration. Lorsque le solvant de lavage froid est trop chaud, la cire fond et est récupérée dans le filtrat. Si cette température est trop froide, de l'huile supplémentaire est cristallisée dans les pores du gâteau. Donc, il faut s'assurer de bien contrôler cette température. Dans le cas présent, l'impact de la température du solvant froid semble plutôt avoir un effet sur les températures du circuit de cristallisation que sur la filtration.**

**À la lumière de ces résultats, on peut conclure que la production de la cire de grade 100 est très dépendante du circuit de cristallisation. Le profil de cristallisation est à travailler ainsi que les ratios des débits de dilution à y alimenter. Mais en premier lieu, il semble qu'une stabilisation du débit de filtrat recyclé diminuera la variabilité du contenu en huile dans la cire. L'effet des pressions d'aspiration et des débits de dilution pourra alors être revérifié.**

## 5.2 Analyse multivariable des données du grade 580

### 5.2.1 Distribution des données du grade 580

L'analyse multivariable de la production du grade 580 s'effectue à partir des données recueillies pendant 15 journées de tests, entre le 2 juillet et le 27 septembre 1996. Contrairement au grade 100, les courts-circuits ne sont pas nécessaires lors de la production du grade 580. L'ensemble des données recueillies a donc été utilisé. Le tableau suivant présente la distribution de ces données dans la matrice X. Les tests échelons effectués chaque journée sont indiqués.

Tableau 5.2.1: Distribution des données pour le grade 580

Dates	Test échelon	Plage d'observations
2 juillet 1996	Solvant froid au filtre A	1 à 35
3 juillet 1996	Solvant froid aux filtres A et C	36 à 83
9 juillet 1996	Vitesse de rotation du filtre A	84 à 130
13 juillet 1996	Vitesse de rotation des filtres	131 à 179
20 juillet 1996	Ratio 5 <sup>e</sup> dilution	180 à 230
25 juillet 1996	Ratio 4 <sup>e</sup> dilution	231 à 273
26 juillet 1996	Ratio 4 <sup>e</sup> dilution	274 à 316
31 juillet 1996	Ratio 2 <sup>e</sup> dilution	317 à 359
1 <sup>er</sup> août 1996	Ratio prédilution	360 à 402
8 août 1996	Ratio prédilution	403 à 446
9 août 1996	Niveaux des curves des filtres A et D	447 à 488
29 août 1996	Ratio 3 <sup>e</sup> dilution	489 à 532
3 septembre 1996	Ratio 5 <sup>e</sup> dilution	533 à 564
4 septembre 1996	Régime permanent	565 à 608
27 septembre 1996	Contrôle de niveau du réservoir V-2401	609 à 654

### 5.2.2 Analyse en composantes principales pour le grade 580

Les données statistiques de l'analyse des composants principaux pour le grade 580 sont présentées ci-dessous. La base de données X contient 50 paramètres d'opération et 654 observations. La somme des carrés des erreurs sur X par rapport à la moyenne initiale est de 32 650.

Tableau 5.2.2: Données statistiques de l'analyse PCA sur l'espace X du grade 580

t	SS <sub>X_ordinaires</sub> %	SS <sub>X_ordinaires</sub> cumulée %	SS <sub>X_val. croisés</sub> cumulée %	Différence SS <sub>X_ordinaires</sub> - SS <sub>X_val. croisés</sub>	CSV/SD
1	22,99	22,99	21,51	1,48	0,886
2	17,44	40,43	38,51	1,92	0,894
3	8,29	48,72	46,19	2,53	0,950
4	7,43	56,14	53,56	2,58	0,952
5	6,27	62,41	60,37	2,04	0,951
6	4,81	67,22	65,47	1,75	0,959
7	3,93	71,15	69,30	1,85	0,968
8	3,17	74,32	72,14	2,18	0,983
9	3,16	77,47	75,77	1,70	0,971
10	2,46	79,93	78,64	1,29	0,974
11	1,95	81,88	78,84	3,04	1,027

Selon le critère d'arrêt de la validation croisée, la dimension du modèle est de 10 variables latentes. Ce modèle semble avoir un bon pouvoir explicatif puisque la somme ordinaire du carré des erreurs expliquées par le modèle et la somme des carrés des erreurs expliquées par le modèle et validées par la validation croisée sont très semblables. Quelques paramètres d'opération restent mal expliqués, comme le démontre le graphique ci-dessous (figure 5.2.1). Le paramètre le moins bien expliqué est le débit de la quatrième dilution. La variabilité de la température de la prédilution, du débit de la deuxième



dilution, du débit de la cinquième dilution et du débit de solvant froid au filtre A est expliquée à moitié.

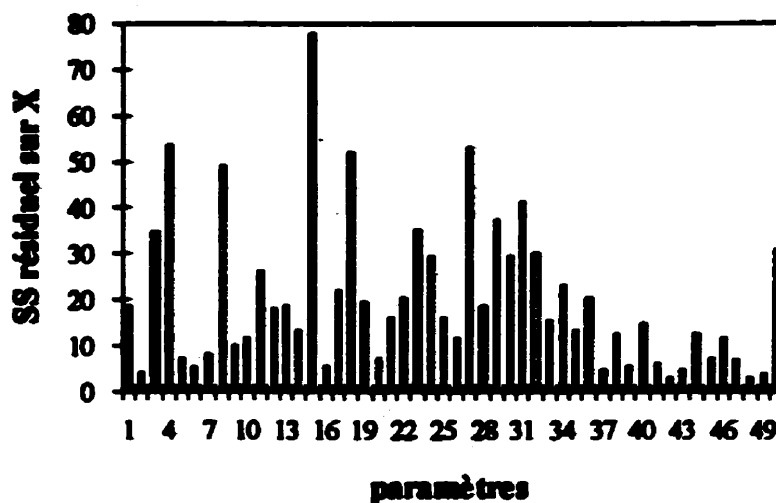


Figure 5.2.1: Fraction de la variabilité des paramètres non expliquée par le modèle PCA.

Le graphique du carré de l'erreur de prédiction présenté à la figure 5.2.2 identifie les observations pour lesquelles l'espace X n'est pas bien expliqué. Quelques observations sont à l'extérieur de la limite de confiance de 99%. Il s'agit d'événements précis pour les observations 86 et 304: respectivement le lavage du filtre A et la diminution du débit de la quatrième dilution.

Les observations 499 à 502 se retrouvent toutes à l'extérieur de la limite de confiance. Ceci signifie qu'un événement hors du commun qui persiste un certain temps s'est produit à cet instant. En effet, pendant cette période de temps l'alimentation de la cinquième dilution a été complètement coupée.

En général le modèle est acceptable. Il est donc conservé.

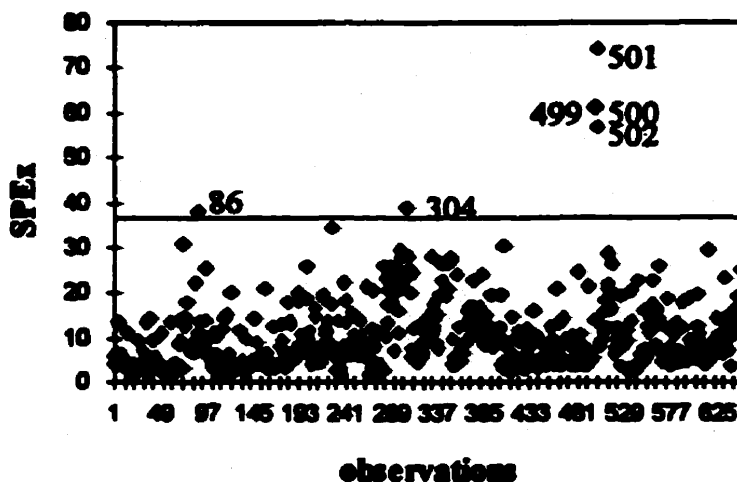


Figure 5.2.2: Somme des carrés de l'erreur de prédiction sur les observations pour le grade 580.

Le détail de l'analyse des variables latentes pour le grade 580 est présenté en annexe I. Les résultats obtenus sont très semblables à ceux du grade 100.

#### 5.2.2.1 Discussion des résultats

Les deux premières variables latentes du modèle PCA pour le grade 580 représentent à elles seules la moitié de la variabilité expliquée par le modèle (38,51 % sur 78,64%). En résumé, les événements qui créent le plus de variabilité dans le procédé lors de la production du grade 580 sont:

- la combinaison des filtres en opération;
- la mise en marche du filtre B;
- la variation des températures du circuit de cristallisation, et;
- le débit d'alimentation.

Les tests échelons sur la deuxième et la troisième dilution sont les plus significatifs selon le modèle PCA, puisqu'ils sont modélisés par les trois premières variables latentes. Par contre, tous les tests échelons sur les débits de dilution sont modélisés à l'aide des autres dimensions du modèle.

La modélisation des tests échelons sur les débits de dilution s'effectuent à l'aide de plusieurs variables latentes pour un même test. Ceci est dû en partie à la différence entre la dynamique des températures et celle des débits de dilutions.

Les tests échelons sur les filtres ne sont pas modélisés par le modèle PCA, à part le test sur les vitesse de rotation effectué en vue du développement de la nouvelle stratégie de contrôle du niveau du réservoir V-2401 (27 septembre 1996). La situation au niveau des filtres est la même que pour le grade 100. Le comportement de tous les jours des filtres B,C et D n'est pas capté par le modèle PCA à cause de la très grande variabilité que créent les périodes d'arrêt et de fonctionnement des filtres. Le comportement du filtre A est quant à lui modélisé puisqu'il est constamment en fonction. Plusieurs variables latentes se partagent la modélisation ( $t_6$  à  $t_8$ ).

Le niveau du réservoir V-2401 est modélisé par la variable latente  $t_9$ . Cette variable latente explique 3,16% de la variabilité totale de la matrice X, ce qui est peu.

### 5.2.3 Analyse par la méthode des moindres carrés partiels des données du grade 580

Les résultats de l'analyse effectuée à l'aide de la méthode des moindres carrés partiels pour la production du grade 580 sont présentés dans ce sous-chapitre. La somme des carrés des erreurs initiale dans l'espace X est de 32 650, et de 653 dans l'espace Y. L'espace Y représente l'indice de réfraction, comme pour le grade 100.

Tableau 5.2.3: Données statistiques de l'analyse PLS sur l'espace Y et l'espace X, grade 580

t	SSy <sub>ordinaire</sub> %	SSy <sub>ordinaire</sub> cumulée %	SSy <sub>val. croisée</sub> cumulée %	Différence SSy <sub>ordinaire</sub> - SSy <sub>val. croisée</sub>	SSx <sub>ordinaire</sub> %	SSx <sub>ordinaire</sub> cumulée %	CSV SD
1	34,72	34,72	34,55	0,17	17,92	17,92	0,809
2	17,09	51,81	51,59	0,22	11,33	29,26	0,861
3	5,97	57,78	57,63	0,15	15,25	44,51	0,938
4	6,70	64,48	64,00	0,48	5,79	50,30	0,923
5	4,45	68,93	68,17	0,76	4,07	54,37	0,947
6	2,09	71,02	70,36	0,66	4,74	59,11	0,977
7	1,63	72,65	72,03	0,62	4,45	63,56	0,982
8	1,54	74,19	73,36	0,83	3,72	67,28	0,987
9	1,01	75,19	74,54	0,65	4,17	71,44	0,993
10	0,99	76,19	75,29	0,90	2,68	74,12	0,998
11	0,75	76,94	75,90	1,04	2,53	76,65	1,006

La dimension du modèle est de 10 selon le critère CSV/SD. Le modèle explique 75,9 % de la variabilité de Y. Le pouvoir explicatif du modèle PLS est bon puisque la différence entre la somme des carrés des erreurs expliquées ordinaire et celle validée par la validation croisée est très faible. La variabilité de l'espace X est expliquée à 74,12 % par

le modèle PLS comparativement à 79,93 % pour le modèle PCA pour un même nombre de variables latentes.

Sur le graphique des résidus sur les paramètres X, on observe que ces derniers sont effectivement moins bien expliqués par le modèle PLS que par le modèle PCA (figure 5.2.3). Le graphique de la somme des carrés des erreurs de prédiction sur Y est présenté à la figure 5.2.4. Quelques excursions à l'extérieur des limites de confiance du modèle y sont observées. Les excursions sont assez persistantes. Cela signifie que certaines variations de l'indice de réfraction ne sont pas expliquées par le modèle. Par contre, la majorité l'est très bien. Le modèle est donc conservé pour fin d'analyse.

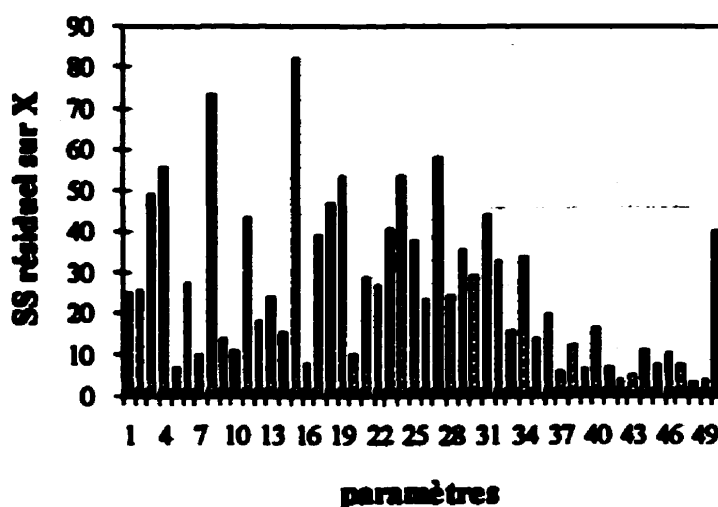


Figure 5.2.3: Fraction de la variabilité des paramètres non expliquée par le modèle PLS, grade 580.

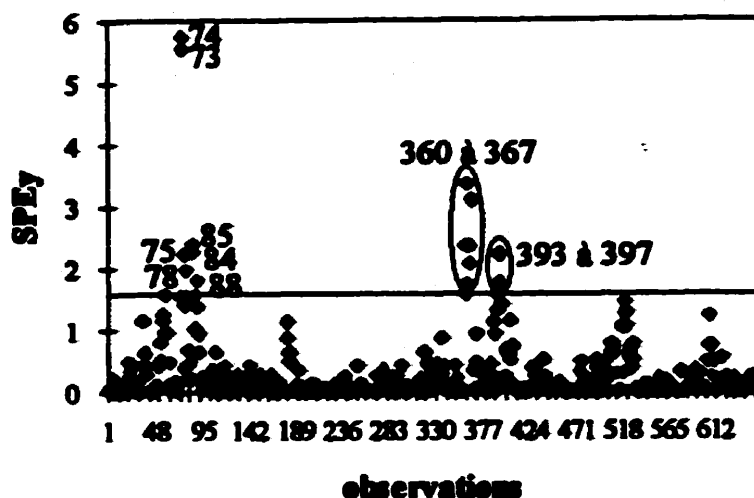


Figure 5.2.4: Somme des carrés des erreurs de prédiction sur les observations de Y, modèle PLS, grade 580.

L'analyse détaillée des variables latentes pour le modèle PLS du grade 580 est présentée à l'annexe II.

### 5.2.3.1 Discussion des résultats

Les deux premières variables latentes du modèle PLS expliquent 51,59 % de la variabilité de l'indice de réfraction à elles seules. Elles comportent donc les principaux événements qui influencent le contenu en huile de la cire:

- la combinaison des filtres en opération;
- le différentiel de température entre la troisième dilution et l'alimentation;
- les pressions dans le cycle de cristallisation

Les autres variables latentes expliquent les tests échelons effectués sur les débits de la prédilution, de la deuxième, de la troisième et la cinquième dilution. Le niveau du

réservoir V-2401 est également modélisé par le modèle PLS à l'aide des dimensions 9 et 10. Les tests échelons réalisés sur les filtres sont peu ou pas importants selon le modèle PLS.

### 5.2.3.2 Paramètres d'opération les plus influents sur l'indice de réfraction de la cire de grade 580

Trente-trois variables sont importantes selon le calcul de l'influence des paramètres pour le grade 580. Le tableau suivant classe ces variables en ordre décroissant de leur influence sur l'indice de réfraction. La flèche dans la colonne juxtaposée indique s'il faut augmenter ou diminuer ce paramètre pour diminuer l'indice de réfraction.

Tableau 5.2.4: Liste des paramètres d'opération qui influencent la qualité de la cire de grade 580

Paramètre	Action à prendre
Pression à l'entrée de X-2401A	↑
$\Delta T$ 3 <sup>e</sup> dilution/alimentation	↓
Niveau de la botte B	↑
Température sortie X-2402A	↓
Température botte B	↑
Débit de filtrat recyclé	↑
Solvant froid B	↓
Vitesse de rotation B	↓
Pression à l'entrée de X-2401B	↑
Niveau de la cuve du filtre B	↑
Niveau de la cuve du filtre D	↓
Pression à l'entrée de l'échangeur X-2402A	↓
Aspiration basse du filtre D	↓

**Tableau 5.2.4: Liste des paramètres d'opération qui influencent la qualité de la cire de grade 580 (suite)**

<b>Paramètre</b>	<b>Action à prendre</b>
Température à la sortie de l'échangeur X-2401A	↓
Pression à l'entrée de l'échangeur X-2402B	↑
Température 3 <sup>e</sup> passe X-2401A	↓
Débit de solvant froid au filtre D	↓
Vitesse de rotation du filtre D	↓
Température à la sortie de l'échangeur X-2402B	↑
Température à la sortie de l'échangeur X-2401B	↓
Niveau de la botte du filtre D	↓
Niveau de la botte du filtre A	↓
Température de la botte du filtre D	↑
Aspiration basse du filtre A	↑
Débit de la troisième dilution	↓
Débit de prédilution	↓
Niveau de l'épouseur de cire	↓
Température de solvant froid	↑
Débit de la cinquième dilution	↑
Vitesse de rotation du filtre A	↓
Niveau de la cuve du filtre A	↑
Pression à l'entrée de l'échangeur X-2437	↓
$\Delta T$ 5 <sup>e</sup> dilution / alimentation	↑

Le débit de filtrat recyclé au réservoir d'alimentation aux filtres est un paramètre considérablement important pour ce grade également. Il doit augmenter, comme dans le cas du grade 100, et pour les mêmes raisons.



Les paramètres d'opération des filtres sont très importants pour la production du grade 580. La moitié des paramètres classés dans ce tableau sont des paramètres de filtration. Le choix des filtres à utiliser est effectué à partir de ces résultats. Selon le calcul de l'influence des paramètres, il est préférable de ne pas utiliser le filtre D lors de la production de la cire de grade 580. Ceci peut s'expliquer en partie par la quantité de solvant de lavage froid maximale qui peut y être alimentée. Ce débit est inférieure à la quantité limite de solvant à utiliser selon les tests échelons effectués. Les paramètres d'opération du filtre B sont également importants selon l'analyse. Par contre, les coefficients obtenus pour ces paramètres ne permettent pas de conclure qu'il doive fonctionner ou pas.

Les paramètres d'opération du filtre A significatifs selon le modèle sont la pression d'aspiration, le niveau de la cuve, la vitesse de rotation et le niveau de la botte du filtre. La pression d'aspiration doit augmenter de façon à favoriser l'indice de réfraction de la cire. Les mêmes explications que celles données pour le grade 100 sont applicables ici. La vitesse de rotation du filtre doit quant à elle diminuer pour diminuer le contenu en huile dans la cire. Théoriquement, lorsque la vitesse est faible, le temps de lavage est plus long donc plus efficace. Par contre, en pratique, la vitesse de rotation est constante tout au long de la période étudiée, sauf lors du test sur le contrôle de niveau du réservoir V-2401 à l'aide des vitesses de rotation. Or, lors de cette journée de test, l'indice de réfraction était des plus bas observés sur toute la base de données. Il s'agit donc visiblement d'une coïncidence.

Le niveau de la cuve du filtre A doit augmenter pour réduire le contenu en huile dans la cire. Selon la théorie, et en pratique, lorsque la pression d'aspiration est élevée, l'augmentation du niveau dans les cuves des filtres permet d'obtenir la même capacité de filtration. Ceci explique la corrélation entre l'indice de réfraction et ce niveau. En effet,

des pressions d'aspiration élevées favorisent de meilleurs résultats au point de vue de la qualité de la cire, tel qu'expliqué auparavant.

Le niveau de la botte du filtre A doit quant à lui diminuer pour favoriser le contenu en huile dans la cire. Ce paramètre est relié à la température de la botte du filtre qui contrôle ce niveau. Il est aussi fonction de la production totale du mélange cire-solvant-huile. Ce niveau est plus faible lorsque la quantité de ce mélange est faible puisque la botte du filtre est plus libre et l'écoulement est facilité. Cette situation se produit entre autre lorsque la concentration des particules de cire est plus faible dans la solution à filtrer, ce qui correspond au moment où le débit de filtrat recyclé est élevé. Elle se produit aussi lorsqu'il reste moins du mélange solvant-huile dans les pores du gâteau, dû à une meilleure aspiration du liquide par le filtre.

L'opération du circuit de cristallisation est également significative. Les températures du circuit de cristallisation sont des plus importantes. Elles doivent être plus froides afin de diminuer l'indice de réfraction de la cire. Le différentiel de température entre la troisième dilution et l'alimentation doit aussi diminuer. Cela signifie que la température du solvant alimenté à la troisième dilution devrait être plus froide. Le différentiel de température entre la cinquième dilution et l'alimentation doit augmenter, donc le solvant qui y est injecté devrait être plus chaud.

La température du solvant froid est ainsi un paramètre important selon le paragraphe précédant. Cette température est directement reliée au différentiel de température de la cinquième dilution qui est significatif selon l'analyse de l'influence des paramètres. Cette température est en fait celle du solvant de dilution à l'entrée du dernier échangeur et celle du solvant de lavage froid aux filtres. Lorsque cette température est trop froide, elle peut provoquer la cristallisation supplémentaire d'huile dans les pores du gâteau lors de la filtration. Il est difficile d'identifier de quelle situation il s'agit ici. La situation la plus

plausible est que la température du solvant froid est importante à cause de son impact sur le différentiel de température de la cinquième dilution puisque la température de filtration est presque égale à la température du solvant froid, comme suggéré en théorie.

Les pressions dans le circuit de cristallisation sont aussi importantes selon le modèle. La pression à l'entrée de l'échangeur X-2401A est d'ailleurs le paramètre le plus influent sur l'indice de réfraction d'après l'analyse de l'influence des paramètres sur la qualité de la cire. Elle doit augmenter afin de favoriser de meilleurs résultats quant au contenu en huile dans la cire. Par contre, les autres pressions ne doivent pas toutes augmenter. Il est donc difficile de conclure quant à l'impact des pressions.

L'impact des débits de dilutions est significatif pour trois d'entre eux. Les débits de la prédilution et de la troisième dilution doivent diminuer et le débit de la cinquième dilution doit augmenter afin d'améliorer la qualité de la cire. La conclusion sur la cinquième dilution est difficile à confirmer puisqu'elle semble être issue de la période pendant laquelle ce débit a été complètement coupé. Cette coupure du débit a eu un grand impact sur l'indice de réfraction.

En conclusion, la qualité de la cire de grade 580 est autant affectée par l'opération des filtres que par les paramètres d'opération du circuit de cristallisation. La stabilisation du débit de filtrat recyclé est de mise ici aussi pour stabiliser le contenu en huile dans la cire.

Le prochain chapitre présente une comparaison entre les résultats obtenus pour le grade 100 et les résultats obtenus pour le grade 100.

### **5.3 Comparaison des résultats du grade 100 et du grade 580**

La production du grade 100 et du grade 580 est très différente à première vue. L'huile et la cire de grade 100 sont beaucoup plus légères que celles du grade 580. Au niveau de la production, le grade 100 est limité par le pouvoir réfrigérant du procédé. La production du grade 580 est quant à elle limitée par la filtration. Les résultats des analyses de l'influence des paramètres d'opération pour chacun des grades reflètent bien cette situation: L'analyse de l'influence des paramètres pour le grade 100 montre que ce sont les températures du circuit de cristallisation et du solvant froid qui sont les plus importantes, mis à part le débit de filtrat recyclé. L'analyse de l'influence des paramètres pour le grade 580 donne des résultats beaucoup plus partagée. Autant de paramètres d'opération du circuit de cristallisation que de paramètres d'opération des filtres se retrouvent dans les premières positions du classement. Par contre, aucun test échelon réalisé sur les filtres n'est significatif selon les deux modèles PLS. Les événements marquants ont trait à la combinaison des filtres en opération.

Il est intéressant de constater que plusieurs conclusions identiques sont tirées pour les deux grades. De meilleurs résultats d'indice de réfraction de la cire sont obtenus lorsque:

- les pressions d'aspiration basse des filtres sont élevées;
- le débit de filtrat recyclé est élevé;
- les températures du circuit de cristallisation sont plus basses aux premiers échangeurs, et;
- les débits de la prédilution et de la troisième dilution sont faibles.

L'analyse de la signification des variables latentes pour les modèles PCA et PLS du grade 580 s'est avérée plus difficile que pour le grade 100. Ceci s'explique par la qualité des données utilisées. Au niveau de l'indice de réfraction, sa mesure pour le grade 580 est moins précise que pour la cire de grade 100. De plus, la plage à l'intérieur de laquelle

l'indice de réfraction de la cire du grade 580 varie est beaucoup moins large que celle du grade 100. Ainsi, l'erreur de mesure est relativement beaucoup plus importante pour le grade 580. D'autre part, il semble que la production du grade 580 soit plus imperméable aux tests échelons effectués sur le procédé. C'est pourquoi l'indice de réfraction n'a pas beaucoup varié en comparaison avec celui du grade 100. À cause de cette situation, la détermination des délais de chacun des paramètres par rapport à l'indice de réfraction est moins précise pour les données du grade 580. La base de données est par conséquent moins efficace pour prédire le comportement du procédé et de l'indice de réfraction.

La qualité des modèles est donc très affectées par la condition des bases de données, ce qui se reflète bien dans les résultats obtenus au cours de ce travail: le modèle PLS du grade 580 n'explique que 75,29 % de la variation de l'indice de réfraction à l'aide de dix variables latentes, tandis que le modèle PLS pour le grade 100 explique 86,78 % de la variation de l'indice de réfraction à l'aide de sept variables latentes.

#### **5.4 Carte de contrôle de la qualité multivariable**

Les cartes de contrôle de la qualité de type Shewart sont communément utilisées en industrie. Les mesures de qualité des produits y sont tracées de façon individuelle. Lorsque le résultat obtenu pour une variable de qualité se trouve à l'intérieur de ses limites de variabilité statistique, le produit est considéré normal. Or, les mesures de qualité des produits ne sont souvent disponibles qu'à des intervalles de temps très long. Par contre, les paramètres d'opération du procédé sont accessibles très fréquemment, voir à quelques secondes ou minutes d'intervalle. Pourtant, ils ne sont pas utilisés à plein rendement pour le contrôle de la qualité des produits.

À l'aide des modèles multivariables, tels PCA ou PLS, des cartes de contrôle de la qualité en continu peuvent être construites. Les cartes de contrôle de la qualité multivariables

basées sur un modèle PLS permettent d'identifier les événements anormaux et prédire la qualité du produit. Par contre, si les mesures de qualité ne sont pas disponibles, un modèle PCA peut être construit. À l'aide d'un modèle PCA, les larges variations sur les paramètres d'opération seront identifiées, mais celles-ci ne refléteront pas nécessairement un changement de la qualité du produit.

La carte de contrôle multivariable est composée de plusieurs graphiques: certains graphiques de variables latentes (par exemple un graphique  $t_1/t_2$ ) qui semblent significatifs, et le graphique représentant la somme des carrés de l'erreur de prédiction sur les paramètres (SPE<sub>x</sub>). Les calculs de contribution à l'erreur de prédiction et de contribution à la variation d'une variable latente permettent d'identifier quels paramètres d'opération ont varié et éclairent ainsi la recherche des causes de la déviation.

Les événements anormaux se manifestent de plusieurs façons sur les graphiques mentionnés ci-haut. Lorsque les paramètres d'opération varient considérablement, mais que la structure relationnelle entre les paramètres d'opération reste la même que celle identifiée par le modèle servant à la carte, un déplacement des données sera observé sur le graphique des variables latentes ( $t_1/t_2$ ). La somme des carrés de l'erreur de prédiction restera quant à elle à l'intérieur des limites de confiance. Si l'événement anormal met en jeu des structures relationnelles différentes de celles contenues dans le modèle utilisé par la carte, les observations se déplaceront à l'extérieur des limites de confiance sur le graphique de la somme de carré de l'erreur de prédiction.

#### **5.4.1 Choix de la base de données de référence**

Le choix de la base de données qui servira de référence est cruciale dans l'élaboration de la carte de contrôle multivariable. Cette base de données doit refléter l'opération normale du procédé, pendant laquelle de bons résultats de qualité sont obtenus. Toute opération

jugée anormale doit être retirée de la base de données afin que celle-ci ressorte avec éclat sur les graphiques lors de son avènement.

**Tableau 5.4.1: Présentation des données utilisées**

<b>Utilisation</b>	<b>Dates</b>
<b>Construction du modèle PLS (indice de réfraction faible)</b>	<b>15, 16 juillet, 6 août, 14 septembre, 1<sup>er</sup>, 2 et 10 octobre 1996 sans les courts-circuits</b>
<b>Base de données avec courts-circuits</b>	<b>15, 16 juillet, 6 août, 1<sup>er</sup>, 2 et 10 octobre 1996</b>
<b>Bases de données avec un indice de réfraction élevé</b>	<b>5 juillet, 5, 12 et 13 août 1996, sans les courts-circuits</b>

À titre d'exemple, un modèle de référence PLS est construit à partir des données du grade 100 ayant un indice de réfraction faible. Deux séries de données contenant des événements anormaux sont testées: des données contenant des courts-circuits du réfrigérant aux premiers échangeurs et des données ayant un indice de réfraction élevé.

#### **5.4.2 Modèle PLS obtenu**

Le modèle PLS obtenu à partir de la base de données de référence contient six variables latentes et modélise un total de 81,4 % de la variabilité de l'indice de réfraction, validée par la validation croisée. On dénote encore une fois l'importance des températures du cycle de cristallisation, puis des débits d'alimentation et de dilution. L'analyse approfondie du modèle n'est pas présentée ici.

### 5.4.3 Données avec courts-circuits

Une base de données contenant des courts-circuits a d'abord été testée à l'aide de la carte de contrôle multivariable basée sur le modèle PLS obtenu. Lors des courts-circuits, l'indice de réfraction augmente de façon considérable. Les actions concrètes prises pour effectuer les courts-circuits sont de couper le réfrigérant aux deux premiers échangeurs et d'augmenter le débit de prédilution. Les courts-circuits ont pour effet de faire augmenter les températures du circuit de cristallisation. Ainsi, la glace qui encrasse les deux premiers échangeurs fond, et les pressions dans les échangeurs chutent. Une fois que les pressions ont chuté, les opérateurs réalimentent le réfrigérant aux deux échangeurs et replacent le débit de prédilution à sa position normale.

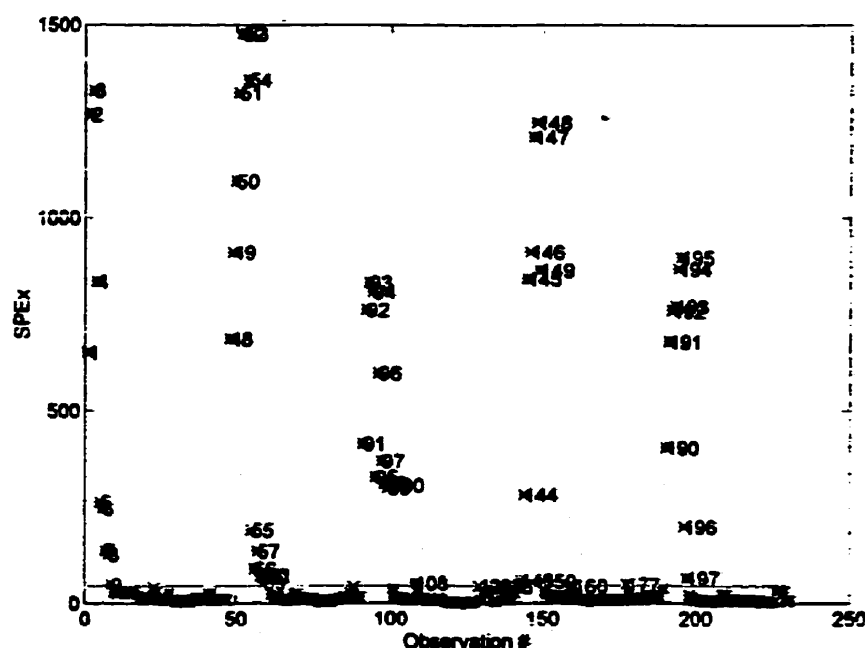


Figure 5.4.1: Somme des carrés des erreurs de prédiction pour les données avec courts-circuits.



Les observations pendant lesquelles se produisent des courts-circuits sortent largement des limites de confiance du modèle sur le graphique du SPEx présenté à la figure 5.4.1. Donc, la structure relationnelle entre les paramètres d'opération et l'indice de réfraction lors des courts-circuits est très différente de la structure en place lors de l'opération régulière. Les graphiques de contribution au SPEx (voir exemple figure 5.4.2), indiquent bien que le débit de prédilution et certaines températures du circuit de cristallisation ont subi des variations importantes. Sur les graphiques des variables latentes, les observations au cours desquelles ont lieu les courts-circuits sortent des limites de confiance du modèle pour revenir à l'intérieur par la suite. L'évolution du court-circuit peut ainsi être suivie tout au long de son avènement. Selon les résultats obtenus, une à deux heures sont nécessaires avant que la variabilité du procédé soit revenue à la normale lorsqu'un court-circuit est effectué.

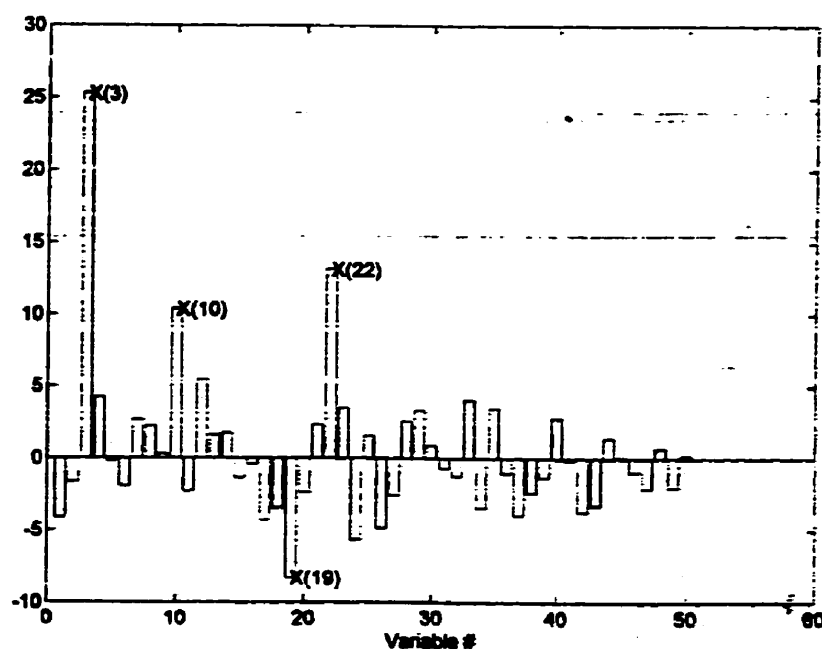


Figure 5.4.2: Contribution à l'erreur de prédiction.

#### 5.4.4 Données avec un indice de réfraction élevé

Les données d'opération du grade 100 pour lesquelles les indices de réfraction sont élevés sont également testées sur le modèle PLS. Les paramètres d'opération très différents au cours de ces périodes en comparaison avec le modèle de référence sont identifiés par la carte de contrôle de qualité.

Le carré de l'erreur de prédiction est inférieur à celui obtenu pour les courts-circuits, mais par contre, toutes les observations, ou presque, se retrouvent à l'extérieur de la limite de confiance du modèle sur le graphique du SPE<sub>x</sub> (figure 5.4.3). Les observations qui sont à l'intérieur des limites de confiance sont celles dont l'indice de réfraction est faible, compris dans la plage de variation des indices de réfraction du modèle de référence. Ainsi, une structure relationnelle différente de celle du modèle de référence existe entre les paramètres d'opération et l'indice de réfraction lors de ces périodes de production.

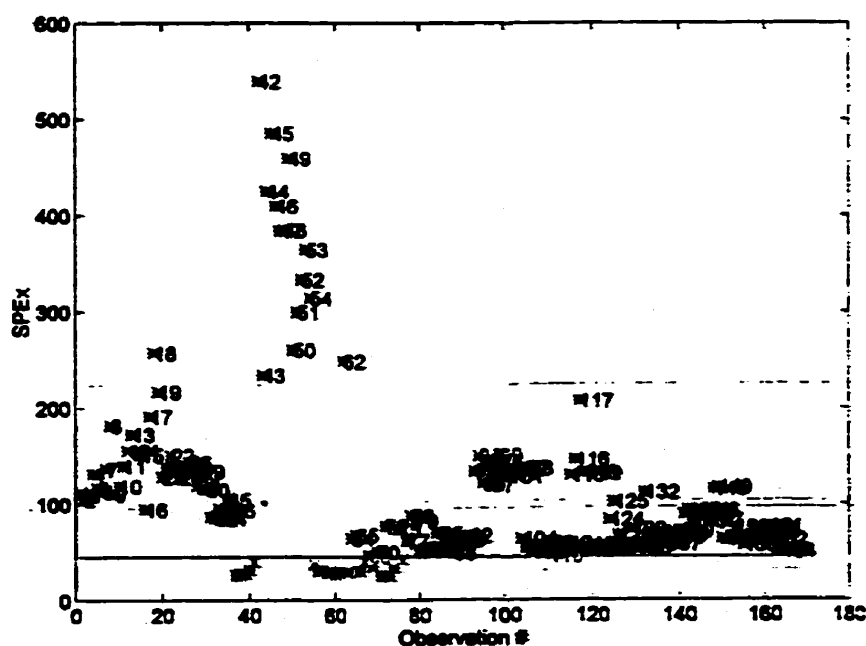


Figure 5.4.3: Somme des carrés des erreurs de prédiction pour les données à haut contenu en huile.

À l'aide du calcul des contributions au SPEX les paramètres qui diffèrent le plus peuvent être identifiés. Le paramètre le plus différent est le débit de filtrat recyclé. Les autres variables identifiées par le calcul de la contribution sont reliées aux tests échelons qui ne sont pas modélisés par les variables latentes du modèle de référence.

Cette démonstration de l'utilisation des modèles PLS comme carte de contrôle de la qualité est très intéressante. Elle permet d'évaluer la puissance de tels outils au point de vu du contrôle de la qualité. Par contre, il n'est pratique que lorsque le procédé est stabilisé et optimisé de façon à ce qu'il se trouve constamment dans la même zone d'opération. Dans le cas présent, ces essais permettent d'appuyer la conclusion quant à l'importance de stabiliser le débit de filtrat recyclé au réservoir d'alimentation des filtres, et ils donnent une évaluation du temps requis au procédé pour revenir à son opération normale suite à un court-circuit (entre une et deux heures).

## **6. Conclusion et Recommandations**

### **6.1 Conclusion**

L'objectif du présent projet était d'identifier les paramètres d'opération les plus influents sur la qualité de la cire à l'aide des méthodes de modélisation multivariable PCA et PLS dans le but ultime d'en faire le contrôle.

Afin d'atteindre cet objectif, plusieurs tests échelons ont été réalisés. Au cours de ces tests, une mesure de la qualité de la cire était nécessaire. La mesure du contenu en huile de la cire étant laborieuse, longue et coûteuse, une mesure alternative a été identifiée: l'indice de réfraction. Les régressions calculées entre les deux paramètres démontrent qu'il existe une relation linéaire entre eux. Les coefficients de détermination pour le grade 100 et le grade 580 sont respectivement de 97 % et 83 %. Ce qui valide l'utilisation de la mesure de l'indice de réfraction pour approximer le contenu en huile de la cire.

Les tests échelons ont été effectués sur les paramètres susceptibles d'avoir un impact sur la qualité de la cire. Ces tests ont permis de créer de la variabilité au niveau des paramètres qui sont très stables lors de l'opération en régime permanent. L'objectif étant d'identifier les paramètres ayant un impact sur la qualité de la cire, les paramètres soupçonnés d'en avoir un doivent varier pour en connaître l'étendu.

Les données d'opération recueillies lors des tests échelons ont dû faire l'objet d'un processus de prétraitement afin s'assurer de l'applicabilité des méthodes et de la validité des résultats. Les modèles PCA et PLS étant efficaces pour des données d'opération en régime permanent, le traitement de données consistait à vérifier la présence de dynamique trop importante dans la base de données à l'aide des autocorrélations et à évaluer les délais entre les paramètres et l'indice de réfraction à l'aide des corrélations croisées.

Aucune dynamique importante n'a été identifiée. Les délais évalués varient entre 30 et 80 minutes. Les données ont donc été alignées dans le temps. Finalement, les données ont été centrées sur leur moyenne respective et ramenées à une variance de un pour rendre chacun des paramètres aussi importants les uns que les autres face aux méthodes de modélisation PCA et PLS.

La modélisation PCA et PLS des bases de données des grades 100 et 580 a ensuite été entreprise. À l'aide de la modélisation PCA, les sources les plus importantes de variabilité au sein de la base de données ont pu être identifiées. L'importance des tests échelons effectués sur le procédé par rapport au bruit a pu être vérifiée. Les modèles PLS ont permis d'identifier les événements et les paramètres d'opération qui ont un effet important sur la qualité de la cire.

Les modèles PCA pour les grades 100 et 580 modélisent environ 80% de la variabilité des bases de données à l'aide de 10 variables latentes. Ils ont identifié la combinaison des filtres en opération et les températures du circuit de cristallisation comme les sources de variabilité les plus importantes dans les bases de données. La variation du débit d'alimentation est également significative, mais moins importante que les premiers paramètres. Les tests échelons sur les dilutions sont significatifs par rapport au bruit selon les modèles PCA puisqu'ils sont modélisés par ces derniers.

Les modèles PLS modélisent 87% de la variabilité de l'indice de réfraction de la cire de grade 100, et 75% dans le cas du grade 580. Ils ont identifié le débit de filtrat recyclé au réservoir d'alimentation des filtres comme le paramètre qui a le plus d'effet sur la qualité de la cire. Il est préférable de fixer ce débit à son maximum selon les modèles. En deuxième lieu, les modèles identifient les températures du circuit de cristallisation comme paramètres importants. Elles ont un plus grand impact sur la qualité de la cire de grade 100, mais sont tout de même significatives pour le grade 580. Un plus faible contenu en

huile de la cire est observé lorsque ces températures sont à leur point d'opération le plus faible. Quant aux paramètres d'opération des filtres, ils ont un plus grand effet sur la qualité de la cire de grade 580 que sur la cire de grade 100. Le principal résultat stipule que l'utilisation du filtre D doit être évitée lors de la production du grade 580. Finalement, les tests échelons qui ont eu un impact sur l'indice de réfraction de la cire sont ceux qui ont fait varier le débit de filtrat recyclé, ce qui appuie l'importance de ce débit sur la qualité de la cire. Selon les modèles PLS des deux grades, les débits de la prédilution et de la troisième dilution doivent être maintenus faibles. L'effet sur la qualité de la cire de la plupart des paramètres d'opération considérés importants par les modèles PLS a pu être expliqué par la théorie connue sur les unités de déparaffinage au solvant.

Ces conclusions cachées dans les bases de données auraient été difficiles à identifier sans les modèles PCA et PLS. Bien que certaines relations étaient connues, l'importance de leur impact sur la qualité de la cire restait nébuleux. L'analyse statistique multivariable a donc permis d'identifier les actions à entreprendre pour améliorer la qualité de la cire.

L'utilisation des chartes de contrôle statistique multivariable s'est avérée très instructive quant au pouvoir qu'elles ont pour identifier les événements qui sortent de l'ordinaire, tels les courts-circuits du réfrigérant aux deux premiers échangeurs. Les modèles PLS et PCA développés au cours de ce projet ne peuvent être utilisés pour effectuer le suivi en ligne le l'UDS. Les modèles ne sont pas représentatifs de l'opération normale du procédé. Lorsque ce dernier sera stabilisé et optimisé, de telles chartes pourront être construites à partir de nouveaux modèles PCA ou PLS. Beaucoup de travail reste à faire d'ici là.

## **6.2 Recommandations**

Suite aux conclusions tirées à partir des modèles PCA et PLS, plusieurs recommandations ont pu être faites. En premier lieu, il est nécessaire de stabiliser le débit de filtrat recyclé pour stabiliser le contenu en huile de la cire. Il semble que le point d'opération optimal pour ce débit soit son maximum pour obtenir la meilleure qualité de cire. Par contre, selon la théorie, il devrait être fixé en ratio par rapport à l'alimentation au procédé.

Le débit de filtrat recyclé sert à contrôler le niveau du réservoir d'alimentation aux filtres. Une nouvelle stratégie de contrôle pour ce niveau a dû être définie. L'utilisation des vitesses de rotation et des pressions d'aspiration des filtres peuvent servir comme variables manipulées vu leur impact nul sur la qualité de la cire lors des tests échelons sur les vitesses de rotation et les pressions d'aspiration effectués à cet effet (10 octobre pour le grade 100 et 27 septembre pour le grade 580).

Quelques autres points sont à travailler:

- Stabiliser le contrôle des températures du circuit de cristallisation, ce qui permettra de stabiliser d'avantage la qualité de la cire;
- Vérifier la relation entre les débits de dilution et la qualité de la cire;
- Travailler la procédure des courts-circuits du réfrigérant aux deux premiers échangeurs afin de minimiser leur impact sur la qualité finale du réservoir de cire.

Les méthodes PCA et PLS sont des instruments qui peuvent servir lors de l'optimisation du procédé. Des chartes de contrôle de la qualité multivariable peuvent également être construites à partir de bases de données recueillies lors de périodes d'opération satisfaisantes. Les méthodes de modélisation multivariables peuvent être très efficaces lorsqu'elles sont utilisées adéquatement en connaissant leur limites et en utilisant son bon jugement d'ingénieur chimique.

## RÉFÉRENCES

- BOX, G.E., JENKINS, G.M. (1976). *Time Series Analysis, Forecasting and Control*. Édition révisée. Holden-Day Inc. San Francisco.
- DAYAL, B.S., MACGREGOR, J.F., TAYLOR, P.A., KILDAW, R., MARCIKIC, S., (1994). Application of feedforward neural networks and partial least squares regression for modelling Kappa number in a continuous Kamyr digester. *Pulp and Paper Canada, Journal of Record, Technical Section CPPA*. 95:1 (1994). 7-13.
- GELADI, P., KOWALSKI, B.R. (1986). Partial least-squares regression: a tutorial. *Analitica Chimica Acta*. Elsevier Science Publisher B.V., Amsterdam, 185(1986). 1-17.
- JACKSON, J.E. (1991). *A User's Guide to Principal Components*. John Wiley and Sons Inc. New York.
- JOLLIFFE, I.T. (1986). *Principal Component Analysis*. Springer-Verlag. New York.
- KOSANOVICH, K.A., DAHL, K.S., PIOVOSO, M.J. (1996). Improved process understanding using multiway principal component analysis. *I. & E.C. Research*. Volume 35. Numéro 1. 138-146.
- KOURTI, T., MACGREGOR, J.F. (1995). Process analysis, monitoring and diagnosis, using multivariate projection methods, *Chemometrics and Intelligent Laboratory Systems*. Elsevier Science B.V. 28(1995). 3-21.
- KRESTA, J.V., MACGREGOR, J.F., MARLIN, T.E. (1991). Multivariate statistical monitoring of process operating performance. *The Canadian Journal of Chemical Engineering*. Volume 69. Février 1991.



MARTENS, H., NAES, T. (1991). *Multivariate Calibration*. John Wiley and Sons Inc. New York.

MACGREGOR, J.F., JAECKLE, C., KIPARISSIDES, C., KOUTOUDI, M. (1994). Process monitoring and diagnosis by multiblock PLS method. *American Institute of Chemical Engineering Journal*. Volume 40. Numéro 5. 926-838.

MCCABE, W.L., SMITH, J.C., HARRIOTT, P. (1985). *Unit Operations of Chemical Engineering*, Fourth edition. McGraw-Hill Publishing Company. Chemical Engineering Series. États-Unis.

MCKETTA, J.J. (1989). Dewaxing, solvent. *Encyclopedia of Chemical Processing and Design*. Volume 15. Marcel Dekker Inc. New York.

NOMIKOS, P., MACGREGOR, J.F. (1994). Monitoring batch processes using multiway principal component analysis. *AIChE Journal*. Volume 40. Numéro 8. 1361-1375.

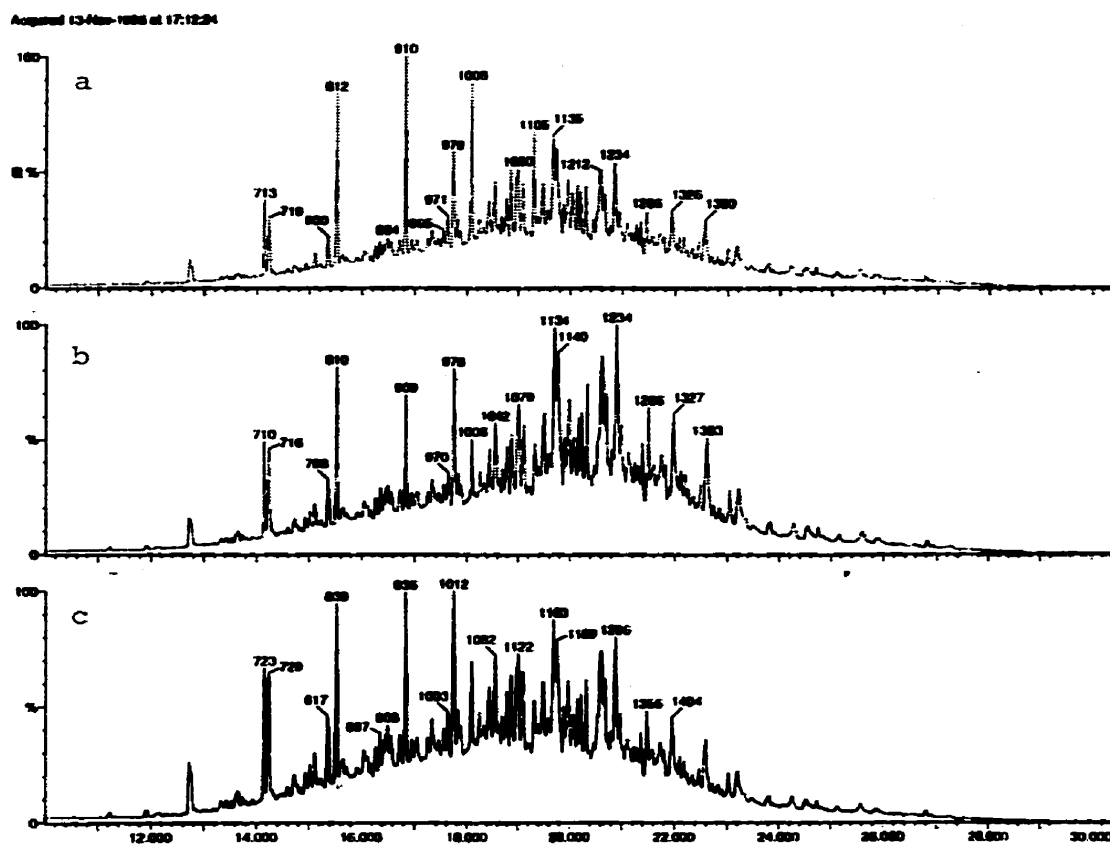
PIOVOSO, M.J., KOSANOVICH, K.A. (1994). Applications of multivariate statistical methods to process monitoring and controller design. *Int. J. Control*. Volume 59. Numéro 3. 743-765.

PIOVOSO, M.J., KOSANOVICH, K.A., YUK, J.P. (1992). Process data chemometrics. *IEEE Transactions on Instrumentation and Measurements*. Volume 41. Numéro 2. April 1992.

SLAMA, F.C. (1991). *Multivariate Statistical Analysis of Data from an Industrial Fluidized Catalytic Cracking Process Using PCA and PLS*. Master of Engineering Thesis, McMaster University, Ontario, Canada.

- TYUMKIN, S.V., SHABALINA, T.N., ZANOZINA, I.I., DISKINA, D.E. (1989). Selection of the conditions and prediction of product quality of preliminary dewaxing. *Chemistry and Technology of Fuels and Oils*. Volume 25. Numéro 3-4. Novembre 1989. 148-151.
- WOLD, S. (1978). Cross-validatory estimation of the number of components in factor and principal components models. *Tecnometrics*. Volume 20. Numéro 4. 397-405.
- WOLD, S. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*. Elsevier Science Publisher, B.V., Amsterdam. 2(1987). 37-52.
- WOLD, S.(1989). Multivariate data analysis: Converting chemical data tables to plots. Computer Applications in Chemical Research and Education. *Intelligent Instruments & Computers*. September/October 1989. 197-216.
- WOLD, S., KETTANEH WOLD, N., MACGREGOR J.F. (1996). *Multivariate Statistical Methods for the Analysis, Monitoring and Optimization of Processes*. Document de référence pour le cours Multivariate process modelling d'octobre 1996. ©MDS et Umetri Inc.
- WUTTHIER, P. (1965). *Le pétrole, raffinage et génie chimique*. Collection Science et technique du pétrole. Numéro 5. Éditions Technip. Paris.

**Annexe I : Résultats des analyses chromatographiques effectuées sur la  
cire et l'huile**



**Figure I.I: Résultats des analyses chromatographiques sur les différents échantillons d'huiles recueillis lorsque l'indice de réfraction de la cire était:**

**a) élevé, b) moyen, c) faible**

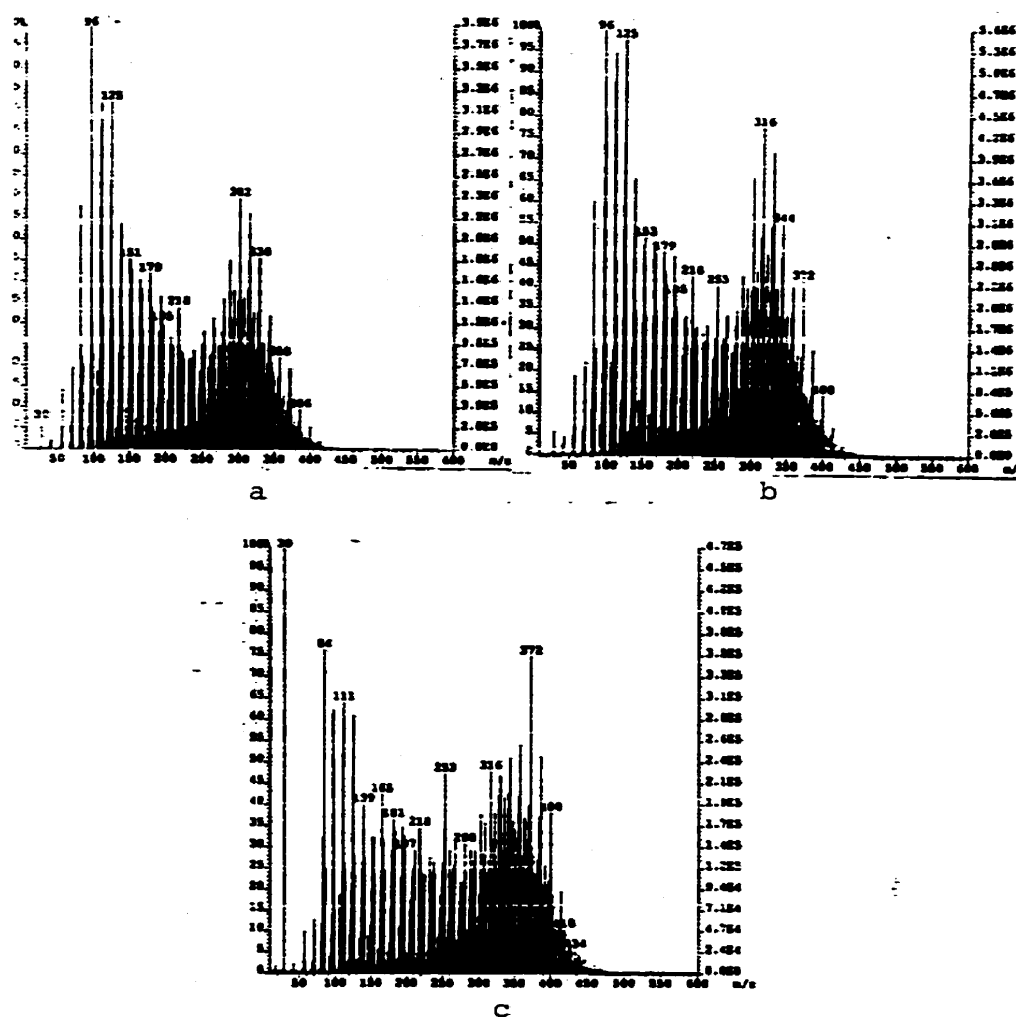
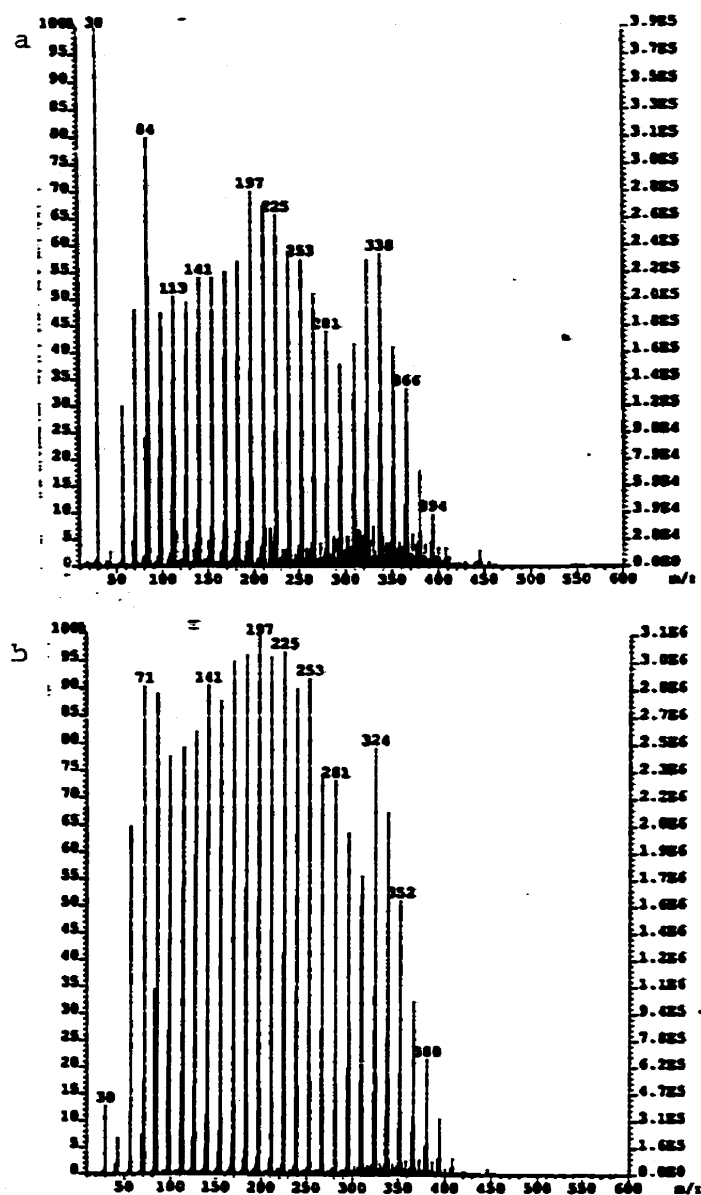


Figure I.II: Résultats des analyses de spectrométrie de masse (bombardement au Krypton) sur les échantillons d'huiles recueillis lorsque l'indice de réfraction de la cire était: a) élevé, b) moyen. c) faible



**Figure I.III: Résultats des analyses de spectrométrie de masse (bombardement au Krypton) sur les échantillons de cire recueillis lorsque l'indice de réfraction de la cire était:**  
**a) élevé b) faible**

**Annexe II : Analyse détaillée des variables latentes du modèle PCA  
pour le grade 580**

L'analyse détaillée des variables latentes pour le modèle PCA du grade 580 est présentée dans cette annexe.

### Dimension 1

#### Variable latente $t_1$

Sur le graphique de l'évolution de  $t_1$  dans le temps, deux niveaux d'opération se distinguent. En effet, cette dimension identifie quel filtre est en opération entre les filtres C et D: Si  $t_1$  est supérieure à zéro, c'est que le filtre D est arrêté et le filtre C fonctionne (observations 74 à 129 et 179 à 402). Si  $t_1$  est inférieure à zéro, le filtre D fonctionne et le filtre C est arrêté (observations 131 à 179 et 403 à 608). Si les deux fonctionnent, les données se retrouvent autour de  $t_1 = 0$  (observations 1 à 73 et 609 à 654).

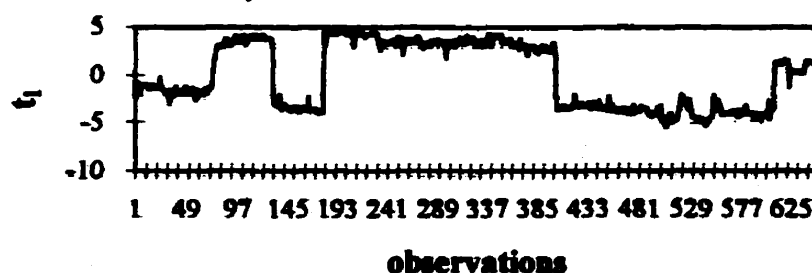


Figure II.I:  $t_1$  en évolution dans le temps, modèle PCA, grade 580.



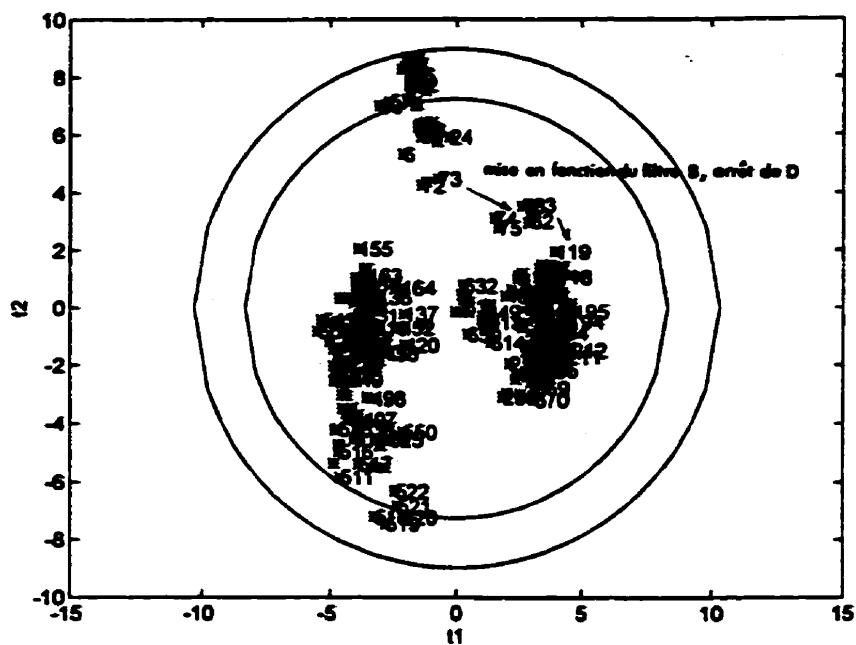


Figure II.II: Graphique  $t_2/t_1$ , modèle PCA, grade 580.

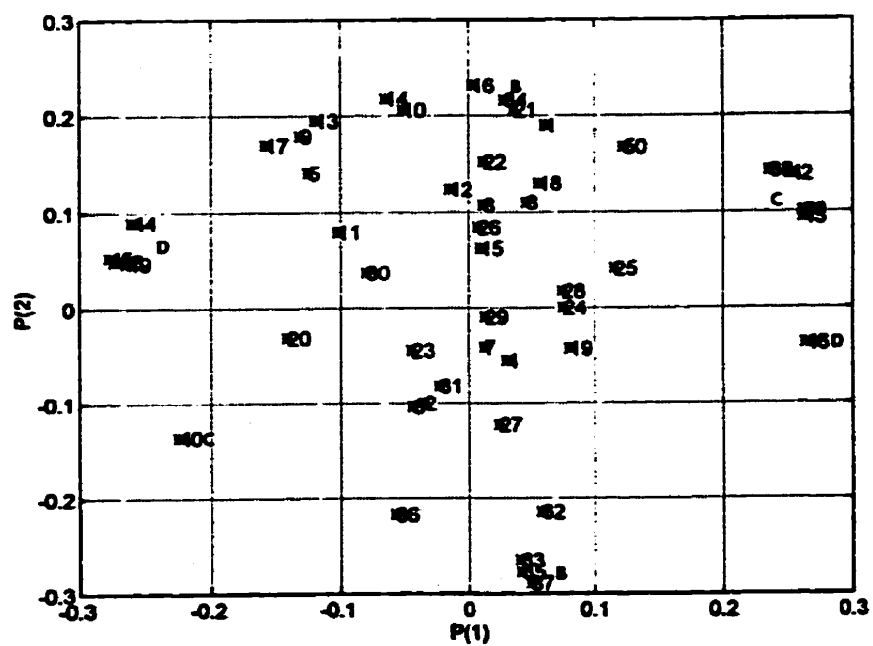


Figure II.III: Graphique  $p_2/p_1$ , modèle PCA, grade 580

**Poids  $p_1$** 

Les poids attribués aux paramètres d'opération des filtres C et D sont nettement les plus élevés dans cette dimension:

- niveaux des courbes des filtres C et D;
- débits de solvant froid alimentés aux filtres C et D;
- pressions d'aspiration basse des filtres C et D;
- vitesses de rotation des filtres C et D;
- niveaux et températures des bannes des filtres C et D.

**Dimension 2****Variable latente  $t_2$** 

Cette dimension modélise le fonctionnement du filtre B. Ce dernier est mis en service à l'observation 70. Le filtre D est complètement hors service à l'observation 73. La période de transition est clairement identifiable sur le graphique de  $t_2/t_1$  (figure 5.2.3). Les observations comprises dans la seule période durant laquelle le filtre B ne fonctionne pas se retrouvent avec des valeurs de  $t_2$  élevées (observations 1 à 70).

D'autre part, les températures du cycle de réfrigération ainsi que le débit d'alimentation au procédé sont modélisés par cette variable latente. La variabilité de ces paramètres d'opération est causée par le changement de journée d'opération ( observations 83 à 84 ) et par certains tests échelons: tests sur la prédilution, la deuxième et la troisième dilution: observations 360 à 402, 317 à 359 et 489 à 532 respectivement. Les observations 508 à 523 ont une valeur de  $t_2$  faible parce que le débit d'alimentation est faible lors de cette période d'opération.

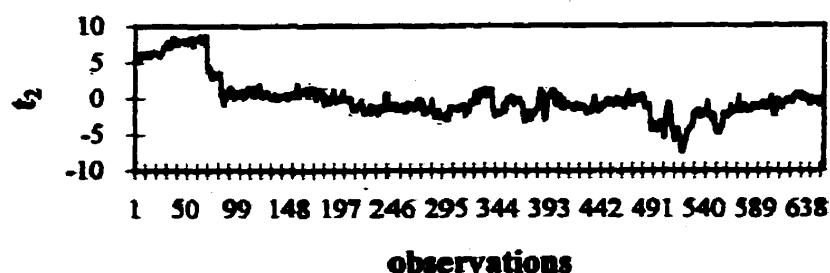


Figure II.IV:  $t_2$  en évolution dans le temps, modèle PCA, grade 580.

#### Poids $p_2$

Les poids attribués aux paramètres d'opération du filtre B sont les plus importants:

- niveau de la cuve du filtre B;
- débit de solvant froid alimenté au filtre B;
- pression d'aspiration basse du filtre B;
- vitesse de rotation du filtre B;
- niveau et température de la botte du filtre B.

Les paramètres d'opération du cycle de cristallisation ayant des poids considérables sont:

- le débit d'alimentation;
- la température à la sortie de l'échangeur X-2401A;
- la température à la sortie de l'échangeur X-2401B;
- la température à la sortie de l'échangeur X-2402A;
- la température à la sortie de l'échangeur X-2402B.

### Dimension 3

#### Variable latente $t_3$

La variable latente  $t_3$  représente plusieurs événements divers. D'abord elle modélise la variabilité de certaines températures du cycle de cristallisation au niveau des différentes périodes d'opération:

- observations 532 à 533;
- observations 130 à 131;
- observations 608 à 609.

De plus, le test échelon sur la troisième dilution est de nouveau modélisé par cette dimension: observations 489 à 532.

Finalement le comportement des filtres A et B est modélisée par cette variable latente: observations 145 à 146 (lavage du filtre A), 70 à 75 (arrêt d'opération du filtre B), 534 à 535 (lavage filtre A), etc.

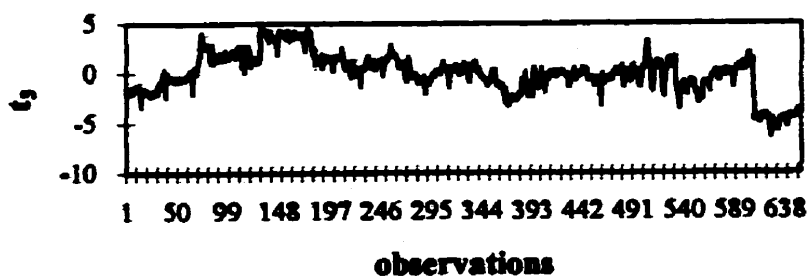


Figure II.V:  $t_3$  en évolution dans le temps, modèle PCA, grade 580.

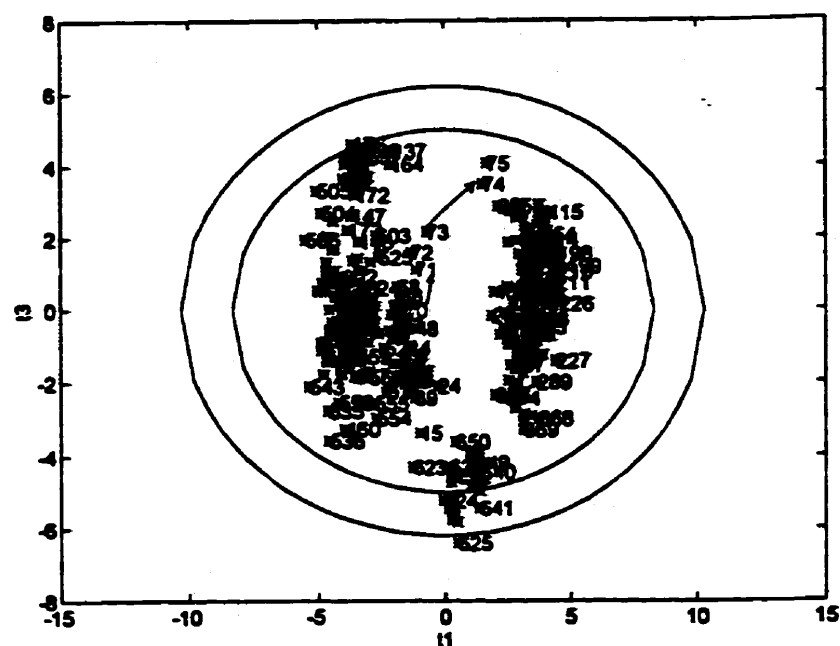


Figure II.VI: Graphique  $t_3/t_1$ , modèle PCA, grade 580.

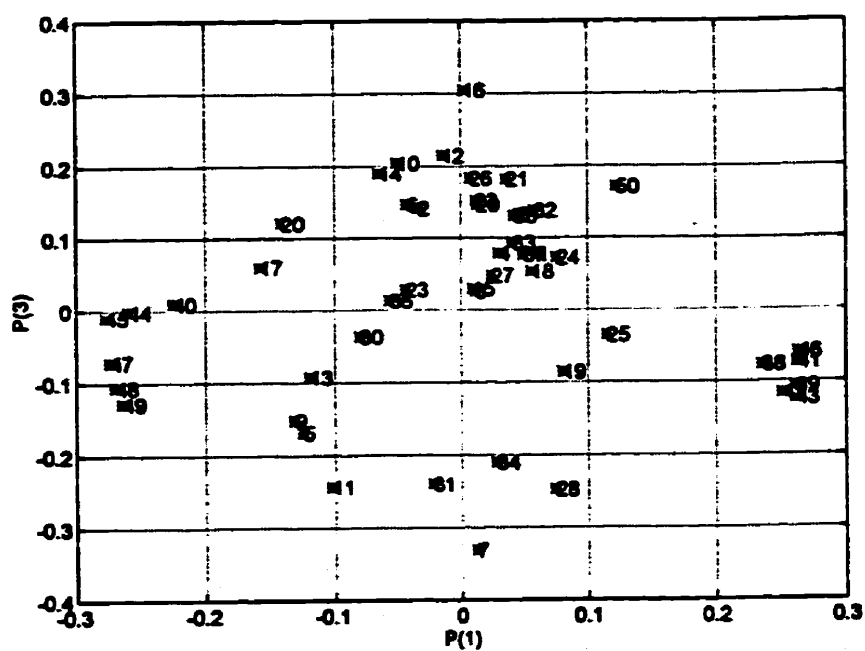


Figure II.VII: Graphique  $p_3/p_1$ , modèle PCA, grade 580.

**Poids  $p_2$** 

Les températures du cycle de réfrigération sont encore importantes dans cette dimension:

- température de la troisième passe de X-2401A;
- température à la sortie de X-2402A;
- température à la sortie de X-2401A;
- température à la sortie de X-2401B;
- température à la sortie de X-2402B;
- $\Delta T$  3<sup>e</sup> dilution/alimentation;

Les paramètres d'opération des filtres importants dans cette dimension sont la pression d'aspiration des filtres A et B, le niveau de la cuve du filtre A ainsi que la température de la botte du filtre A.

**Dimension 4****Variable latente  $t_4$** 

Les tests échelons de la deuxième, quatrième et cinquième dilution sont modélisés par cette variable latente:

- observations 295 à 297: quatrième dilution;
- observations 337 à 338: deuxième dilution;
- observations 349 à 355: deuxième dilution;
- observations 498 à 499: cinquième dilution;
- observations 502 à 503: cinquième dilution.

Les changements de période d'opération suivants sont identifiés par  $t_4$  à cause de l'importance accordée aux paramètres associés aux tests échelons: observations 83 à 84, 130 à 131, 315 à 317, 488 à 489 et 532 à 533.

L'arrêt du débit de la cinquième dilution entre les observations 498 et 503 est aussi modélisée.

Finalement, la mise en marche du filtre B est encore modélisé par cette dimension.

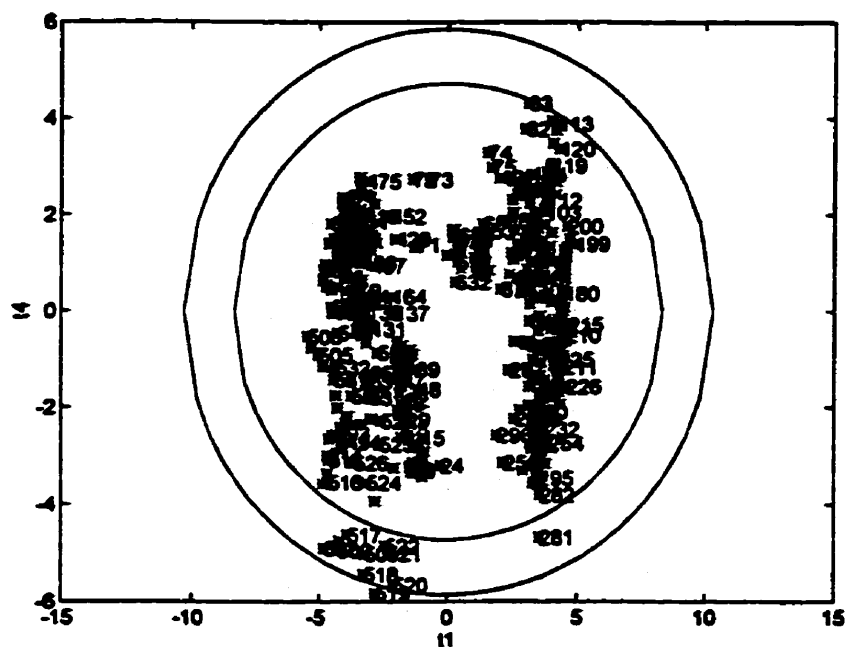


Figure II.VIII: Graphique  $t_4/t_1$ , modèle PCA, grade 580.

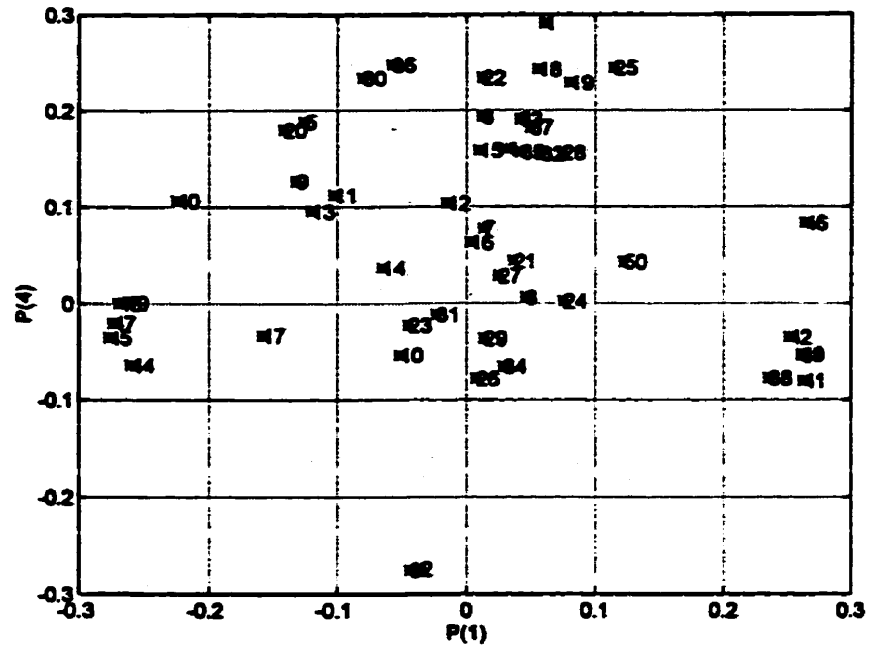


Figure II.IX: Graphique  $p/p_1$ , modèle PCA, grade 580.

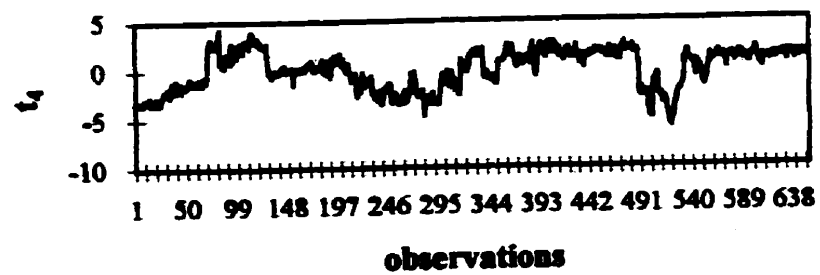


Figure II.X:  $t_4$  en évolution dans le temps, modèle PCA, grade 580.



**Poids  $P_4$** 

Les paramètres d'opération associés aux tests échelons sont :

- le débit de la deuxième dilution;
- le débit de la cinquième dilution;
- la température à la sortie de l'échangeur X-2437;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation;
- la température du solvant froid;
- la température de la vapeur 40 lbs/ eau recyclée;
- la température à l'entrée de l'échangeur X-2401A.

Le niveau des boîtes des filtres A et B sont aussi importants.

**Dimension 5****Variable lente  $t_5$** 

Les changements de période d'opération au niveau des températures du cycle de cristallisation sont modélisés par cette variable lente.

- observations 83 à 84;
- observations 130 à 131;
- observations 179 à 180;
- observations 230 à 231;
- observations 316 à 317;
- observations 608 à 609;

Certaines variations de ces paramètres à l'intérieur d'une journée d'opération sont aussi modélisées par cette dimension:

- observations 69 à 72;
- observations 120 à 122;

Le test sur la prédilution du 1<sup>er</sup> août est aussi modélisé:

- observations 367 à 369;
- observations 382 à 384;

De plus, le test sur les vitesses de rotation du 27 septembre est identifié  $t_3$ : objets 608 à 654

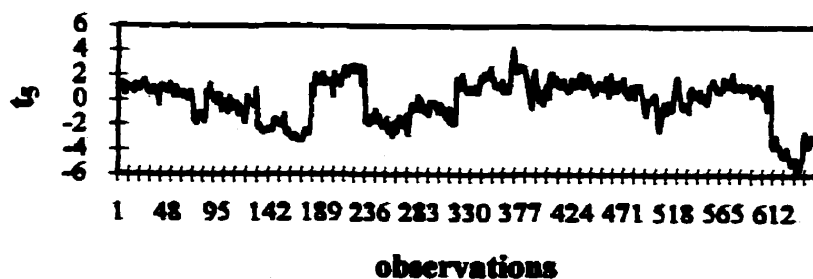


Figure II.XI:  $t_3$  en évolution dans le temps, modèle PCA, grade 580.

#### Poids $p_1$

Les variables importantes du cycles de réfrigération sont:

- la température de la troisième passe de l'échangeur X-2401A;
- la température de la vapeur 40 lbs / eau recyclée;
- la température à l'entrée de X-2401A;

- la température à la sortie de l'échangeur X-2401A;
- la température à la sortie de X-2401B;
- la température à la sortie de l'échangeur X-2437;
- la pression à l'entrée de l'échangeur X-2437;
- le  $\Delta T$  3<sup>e</sup> dilution/alimentation;
- la température prédilution.

Le débit de prédilution a aussi un poids important.

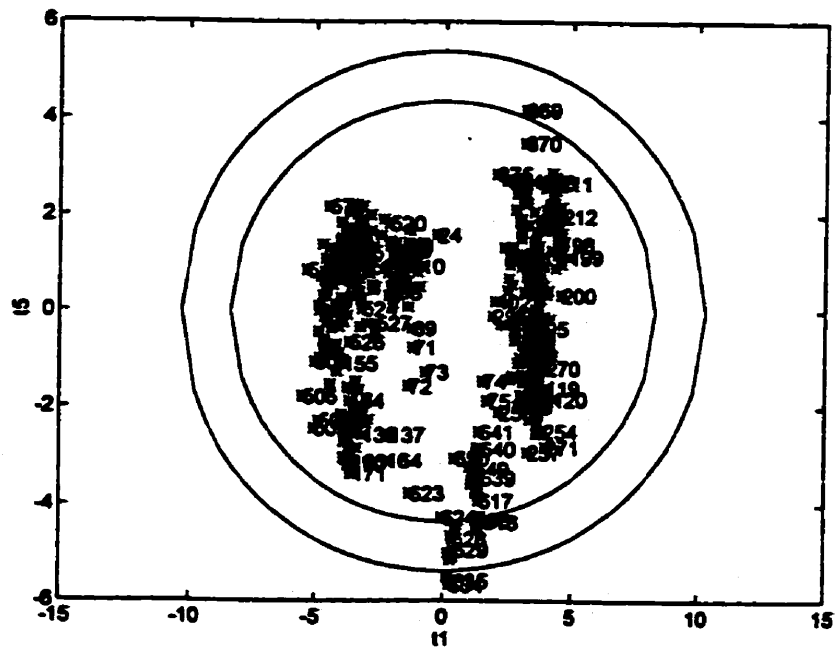


Figure II.XII: Graphique  $t_2/t_1$ , modèle PCA, grade 580.

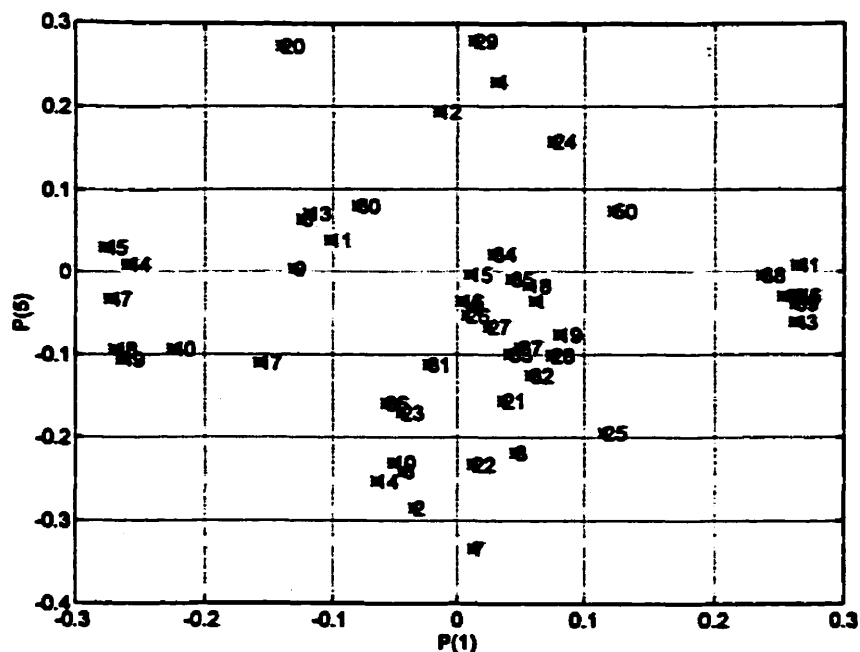


Figure II.XIII: Graphique  $p_5/p_1$ , modèle PCA, grade 580.

## Dimension 6

### Variable latente $t_6$

Cette variable modélise le comportement du filtre A et du débit de filtrat recyclé:

- lavages du filtre A: observations 312 à 313, 329 à 331, 390 à 393, 269 à 270, 587 à 588, 557 à 558, etc.;
- test sur les vitesses de rotation du 27 septembre: observations 609 à 654;

L'effet des tests échelons sur la troisième dilution et la cinquième dilution au niveau des pressions et des températures est aussi important dans cette dimension:

- observations 489 à 532: troisième dilution;
- observations 533 à 564: cinquième dilution.

Les changements d'opération au niveau des températures du cycle de cristallisation et du débit de prédilution sont aussi modélisés: observations 229 à 231, et 608 à 609.

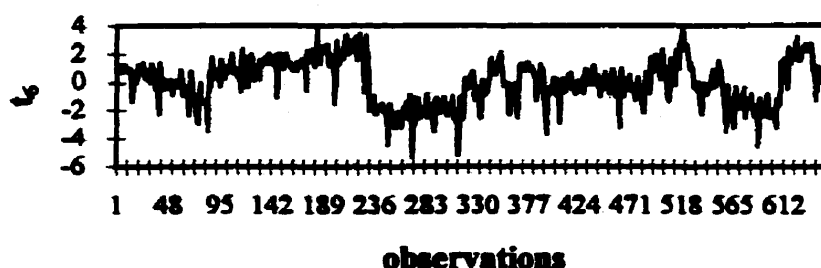


Figure II.XIV:  $t_c$  en évolution dans le temps, modèle PCA, grade 580.

#### Poids $p_i$

Les poids les plus élevés sont attribués aux variables suivantes:

- température à la sortie de X-2401A;
- $\Delta T$  3<sup>e</sup> dilution/alimentation;
- débit de filtrat recyclé;
- température et niveau de la botte du filtre A;
- vitesse de rotation du filtre A;
- pressions à l'entrée des échangeurs.

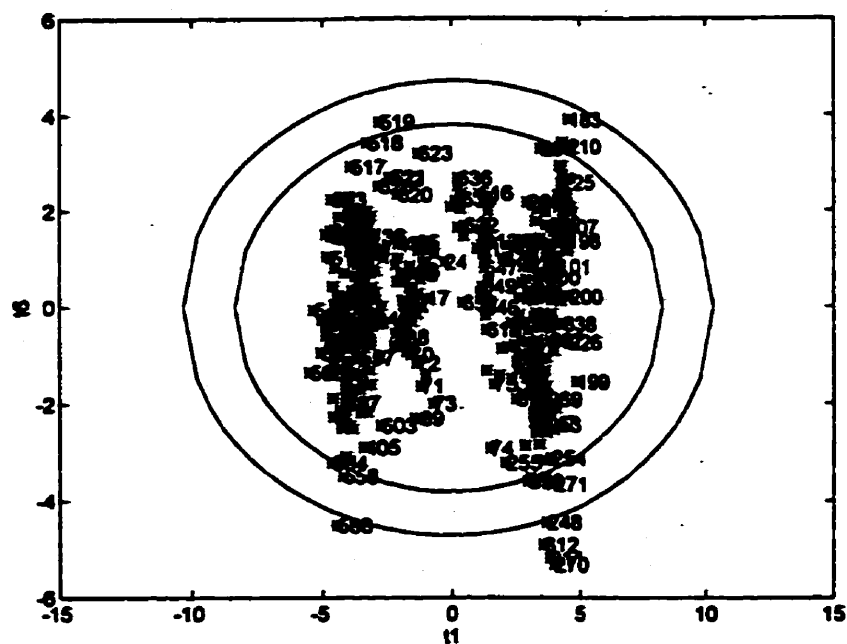


Figure II.XV: Graphique  $t_2/t_1$ , modèle PCA, grade 580.

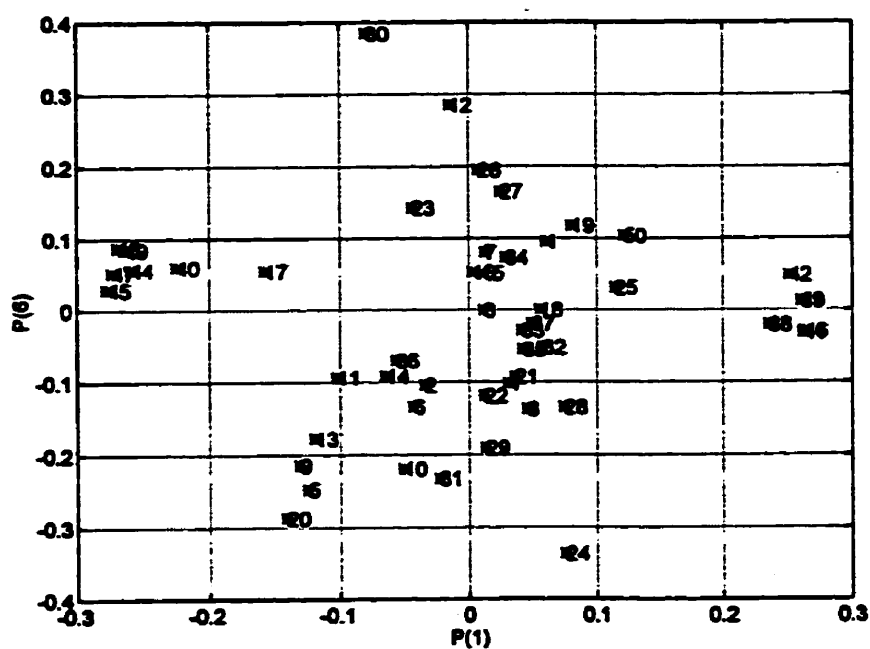


Figure II.XVI: Graphique  $p_2/p_1$ , modèle PCA, grade 580.

## Dimension 7

### Variable latente $t_7$

Les lavages des filtres A sont encore modélisés par cette dimension. Lorsqu'un lavage est enclenché,  $t_7$  augmente:

- observations 14 à 15;
- observations 111 à 113;
- observations 145 à 146;
- observations 225 à 226;
- observations 521 à 522.

Les tests échelons suivants sont aussi modélisés par cette dimension:

- prédilution: observations 360 à 402;
- deuxième dilution: observations 317 à 359;
- quatrième dilution: observations 274 à 316;
- cinquième dilution: observations 533 à 564.

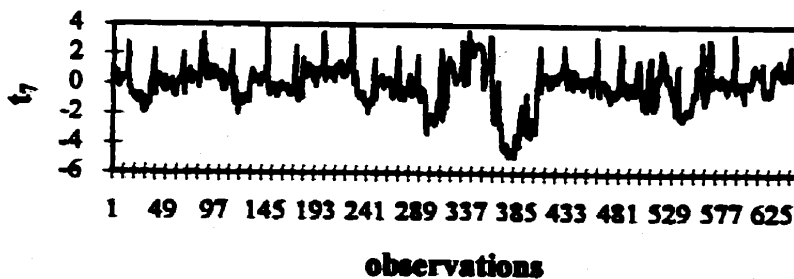


Figure II.XVII:  $t_7$  en évolution dans le temps, modèle PCA, grade 580.

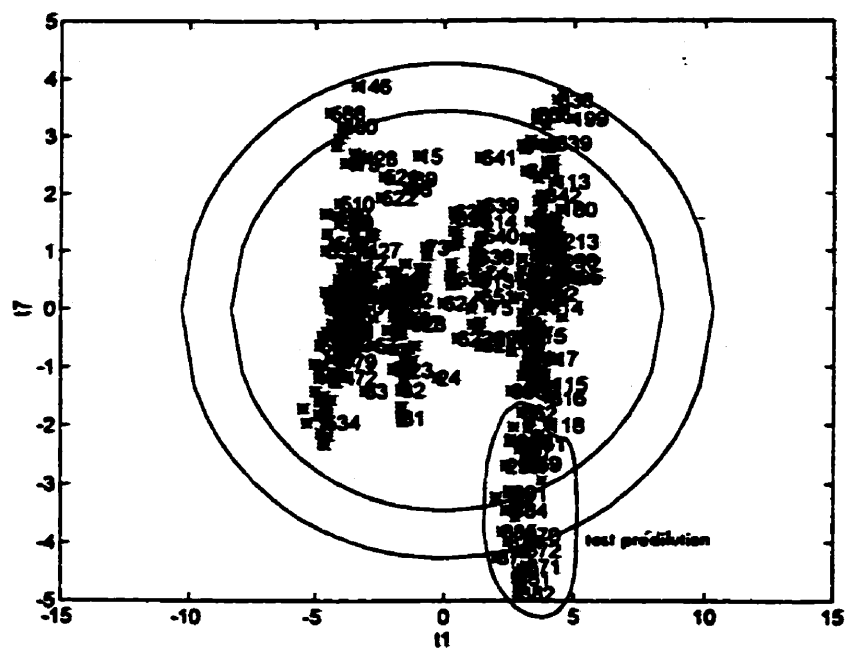


Figure II.XVIII: Graphique  $t_2/t_1$ , modèle PCA, grade 580.

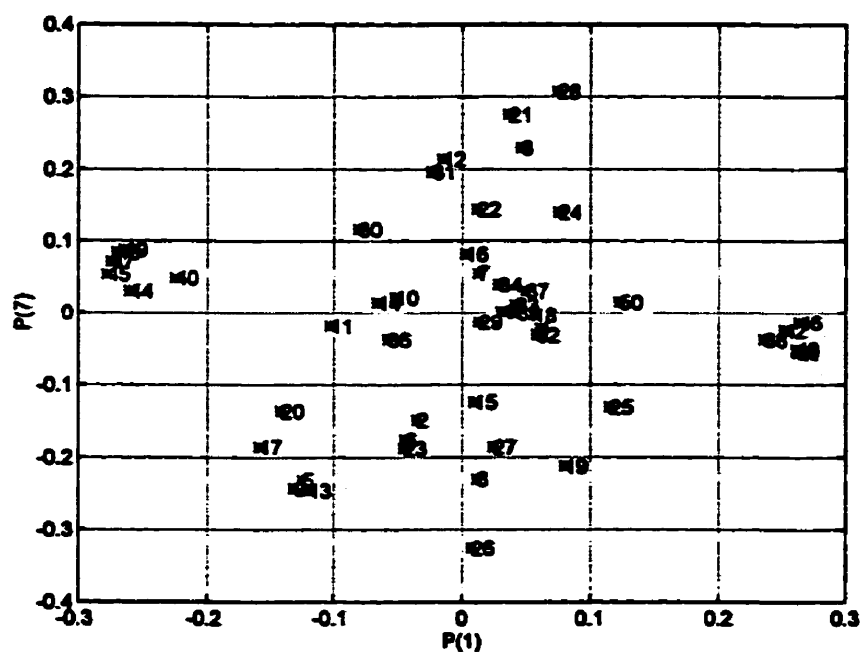


Figure II.XIX: Graphique  $p_2/p_1$ , modèle PCA, grade 580.



**Poids  $p_7$** 

Les paramètres d'opération ayant plus de poids dans cette dimension sont certains paramètres d'opération du filtre A:

- niveau de la cuve du filtre A;
- pression d'aspiration basse du filtre A;
- température de la botte du filtre A.

Pour le circuit de cristallisation, les paramètres importants sont:

- les pressions du cycle;
- le débit de la deuxième dilution;
- le débit de prédilution;
- la température à l'entrée de l'échangeur X-2402B;
- le  $\Delta T$  3<sup>e</sup> dilution/alimentation;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation;

Ces paramètres varient lors des tests échelons et entre chaque journée d'opération.

**Dimension 8****Variable latente  $t_g$** 

Cette dimension modélise l'opération du filtre A, notamment les lavages qu'on peut observer sur le graphique de  $t_g$  en fonction du temps.

Les tests sur la deuxième dilution sont aussi importants dans cette dimension: observations 317 à 359.

### Poids $p_1$

Les variables du filtre A qui ont du poids sont:

- le niveau de la curve du filtre A;
- le débit de solvant froid alimenté au filtre A;
- la pression d'aspiration du filtre A;
- la vitesse de rotation du filtre A;
- le débit de filtrat recyclé.

Les variables du cycle de réfrigération importantes sont:

- le débit de la deuxième dilution;
- la température à l'entrée de l'échangeur X-2401A;
- la température de la vapeur 40lbs / eau recyclée.

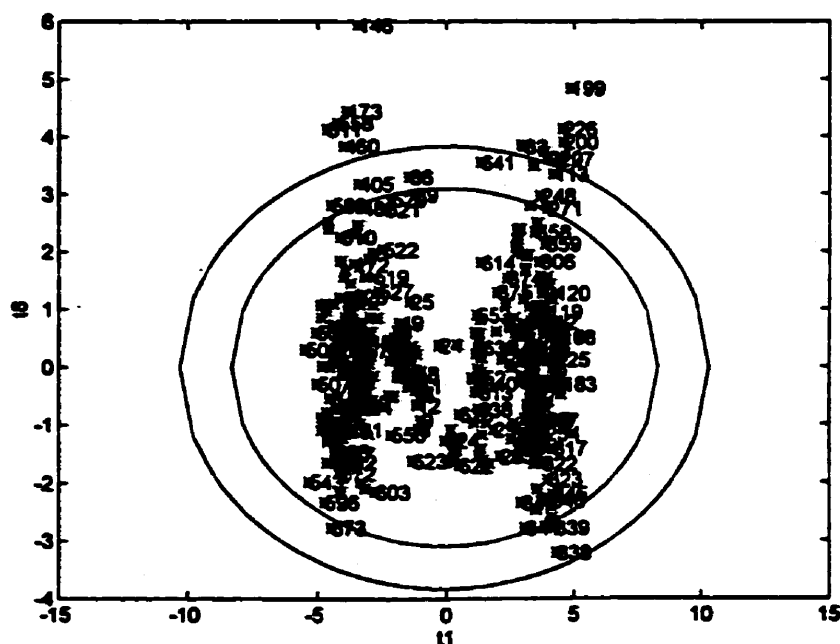


Figure II.XX: Graphique  $t_0/t_1$ , modèle PCA, grade 580.

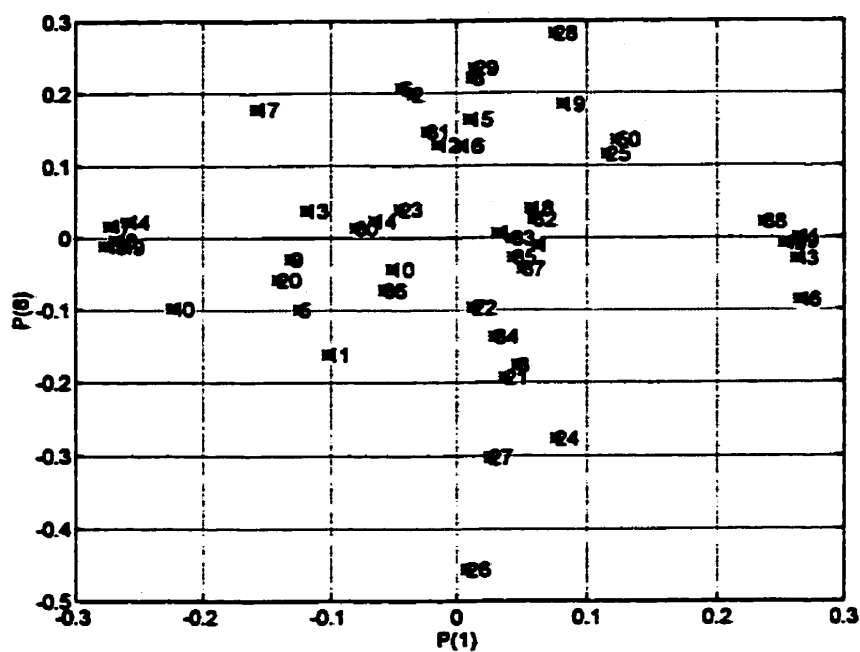


Figure II.XXI: Graphique  $p_2/p_1$ , modèle PCA, grade 580.

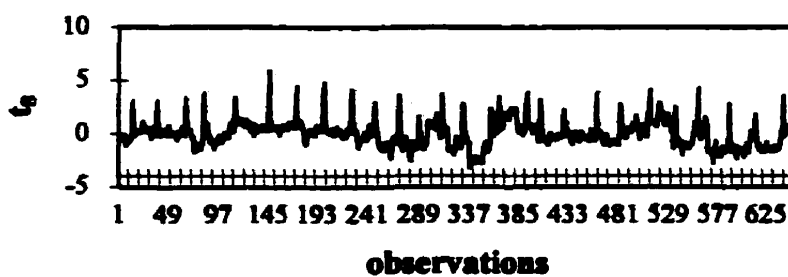


Figure II.XXII:  $t_s$  en évolution dans le temps, modèle PCA, grade 580.

## Dimension 9

### Variable latente $t_9$

Cette dimension modélise la variabilité du niveau du réservoir d'alimentation aux filtres:

- observations 30 à 40;
- observations 545 à 555;
- observations 532 à 533;
- observations 385 à 388.

Elle modélise aussi les variations de la température du solvant froid et du différentiel de température entre la cinquième dilution et l'alimentation.

- observations 95 à 110;
- observations 129 à 131: changement de période du 9 au 13 juillet;
- observations 316 à 319: changement de période du 26 au 31 août;
- observations 333 à 340: test sur la deuxième dilution;
- observations 385 à 388: test sur la prédilution.

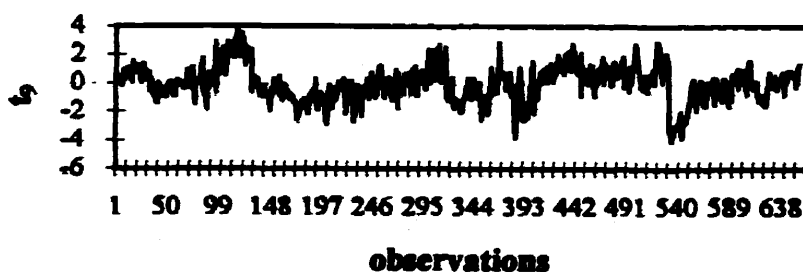


Figure II.XXIII:  $t_9$  en évolution dans le temps, modèle PCA, grade 580.

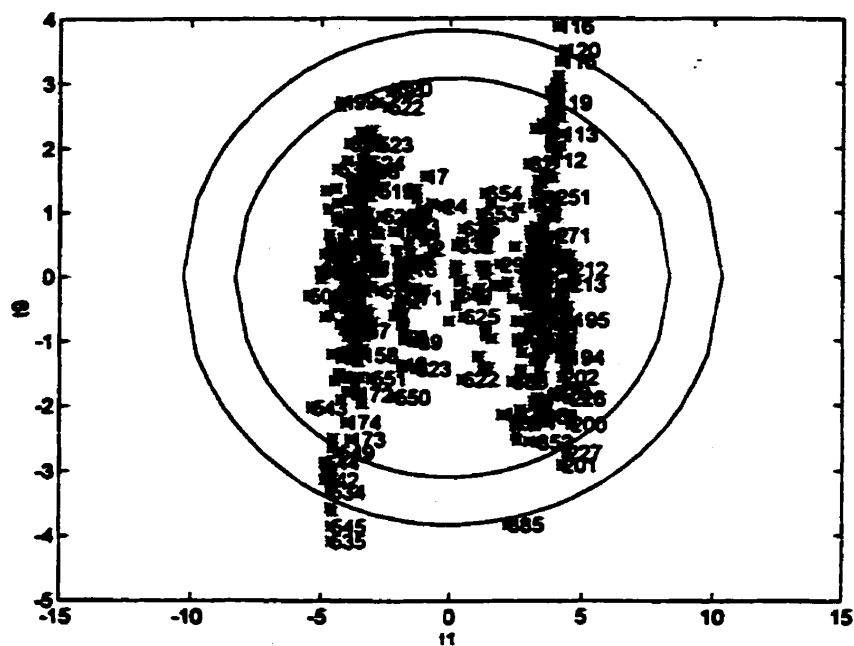


Figure II.XXIV: Graphique  $t_2/t_1$ , modèle PCA, grade 580.

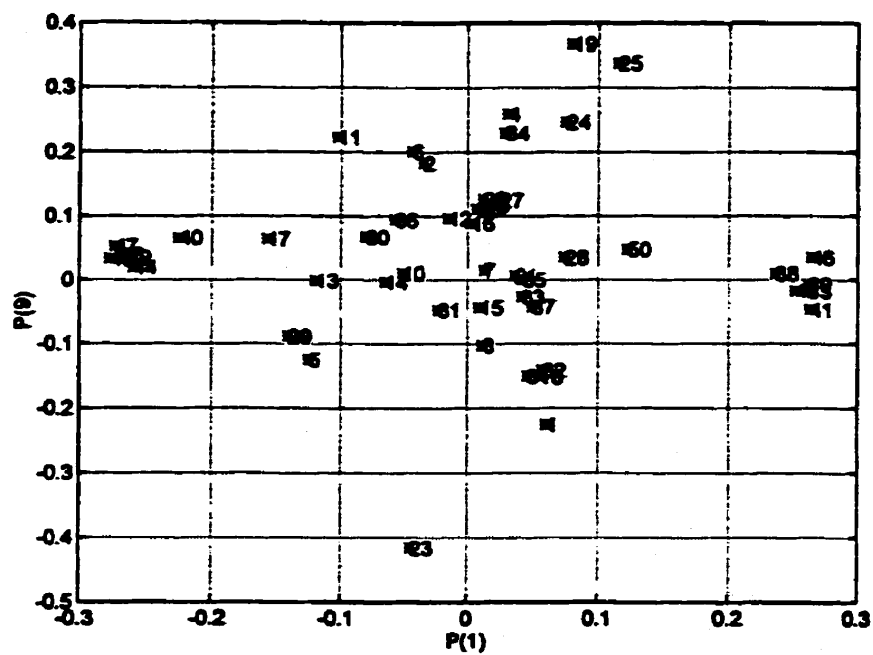


Figure II.XXV: Graphique  $p_2/p_1$ , modèle PCA, grade 580.

**Poids  $p_2$** 

Les paramètres d'opération qui se distinguent dans la neuvième dimension sont:

- le niveau du réservoir d'alimentation aux filtres;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation;
- la température du solvant froid;

**Dimension 10****Variable latente  $t_{10}$** 

Cette variable modélise plusieurs événements déjà modélisés auparavant:

- test sur la prédilution (observations 360 à 402);
- test sur la deuxième dilution (observations 317 à 359);
- test sur la troisième dilution (observations 489 à 532);
- test sur la cinquième dilution (observations 533 à 564);
- test sur les vitesses de rotation des filtres (observations 609 à 654).

**Poids  $p_{10}$** 

Les variables importantes sont:

- le débit de la troisième dilution;
- le niveau V-2401;
- la vitesse de rotation filtre A;
- le  $\Delta T$  5<sup>e</sup> dilution/alimentation;
- la température solvant froid;
- le débit de prédilution;
- le débit de deuxième dilution;
- la température de la vapeur 40 lbs / eau recyclée.

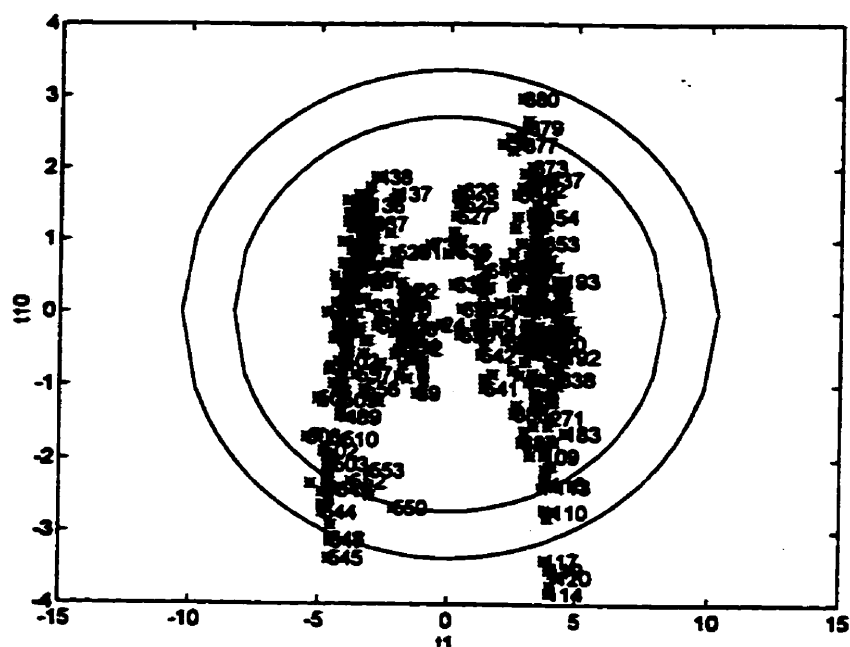


Figure II.XXVI: Graphique  $t_{10}/t_1$ , modèle PCA, grade 580.

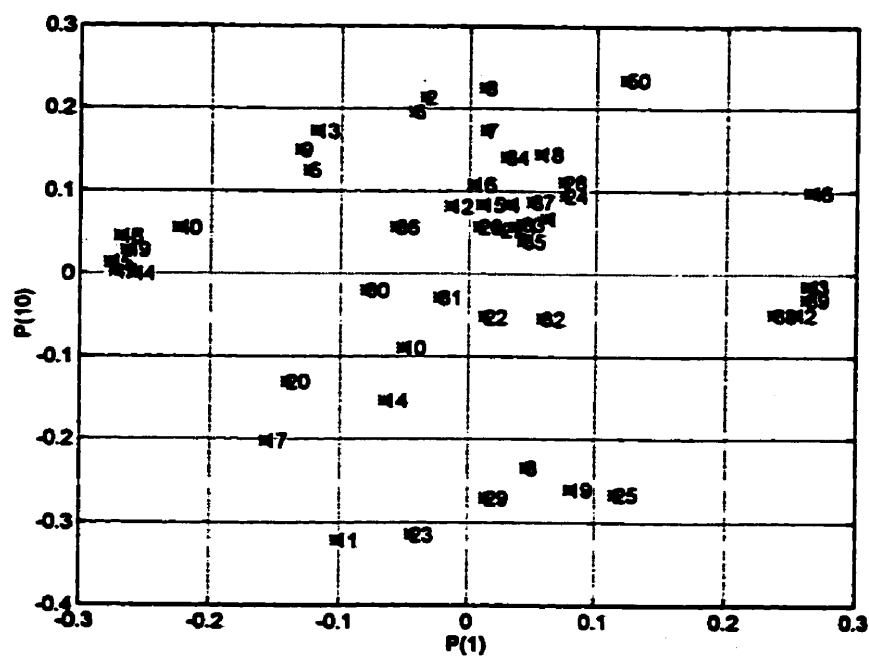


Figure II.XXVII: Graphique  $p_{10}/p_1$ , modèle PCA, grade 580.

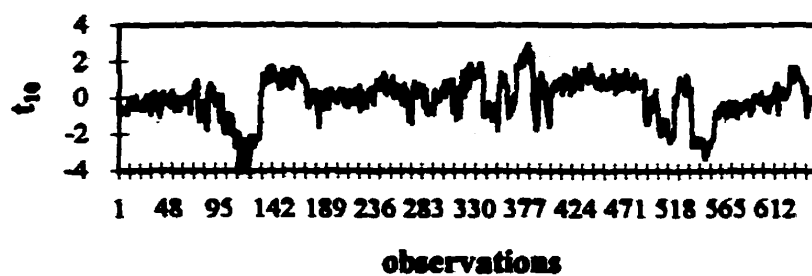


Figure II.XXVIII:  $t_{10}$  en évolution dans le temps, modèle PCA, grade 580.



**Annexe III : Analyse détaillée des variables latentes du modèle PLS  
pour le grade 580.**

## **Relation entre X et Y**

La première étape de l'analyse consiste à identifier les relations qui existent entre X et Y à travers les variables u et t.

### **$u_1/t_1$**

La relation  $u_1/t_1$  explique la différence qui existe entre le groupe de données 1 à 70 et le reste des données. Il s'agit de la seule période d'opération pendant laquelle le filtre B n'est pas en fonction. Cet événement explique 34,5% de la variabilité de l'indice de réfraction.

### **$u_2/t_2$**

La relation  $u_2/t_2$  semble assez linéaire. Les groupes d'observations 180 à 230, 499 à 501 et 518 à 521 se distinguent des autres données. Il s'agit de certains événements précis: test échelon sur la cinquième dilution, arrêt de la cinquième dilution et test échelon sur la troisième dilution.

### **$u_3/t_3$**

Un petit groupe de données orientent cette relation: observations 489 à 503 et 508 à 530. Ces observations correspondent respectivement à l'arrêt de la cinquième dilution et au test échelon sur la troisième dilution.

### **$u_4/t_4$**

Cette relation semble bien linéaire. Aucune particularité ne peut être identifiée à partir de ce graphique.

### **$u_5/t_5$**

Quelques données assez dispersées orientent cette dimension. Il s'agit des lavages du filtre A ainsi que de la variation des températures du cycle de cristallisation.

$n_6/t_6$

Les observations 84 à 179 se distinguent des autres observations. Celles-ci correspondent à une période durant laquelle les températures du cycle de réfrigération sont plus élevées.

$n_7/t_7$

La relation est bien linéaire. Les données 74 à 81 qui s'écartent de la droite correspondent à la mise en service du filtre B et à l'arrêt de filtre D.

$n_8/t_8$

Rien de particulier n'est observé sur ce graphique.

$n_9/t_9$

Un groupe de données orientent la relation: Il s'agit des observations 609 à 639 et 499 à 502. Ces observations correspondent respectivement au test de contrôle du niveau du réservoir V-2401 à l'aide des vitesses de rotation et des pressions d'aspiration des filtres et, une fois de plus, à l'arrêt de la cinquième dilution.

$n_{10}/t_{10}$

La dixième dimension explique très peu la variabilité de l'indice de réfraction (0,75%). Il s'agit en fait de la variabilité du niveau du réservoir d'alimentation aux filtres. Les lavages des filtres sont aussi modélisés par cette dimension.

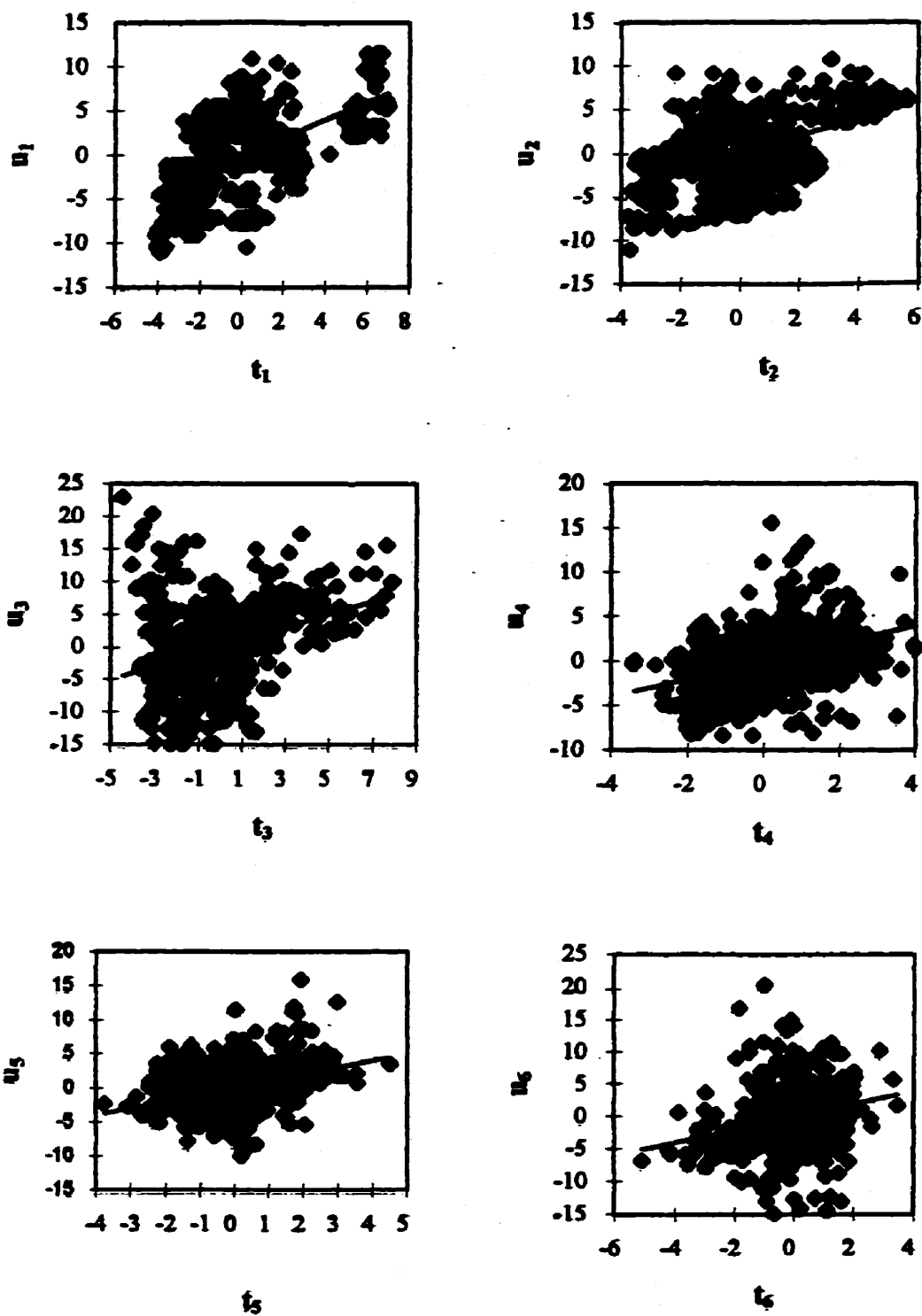


Figure III.I: Graphiques  $u/t$ , modèle PLS, grade 580.

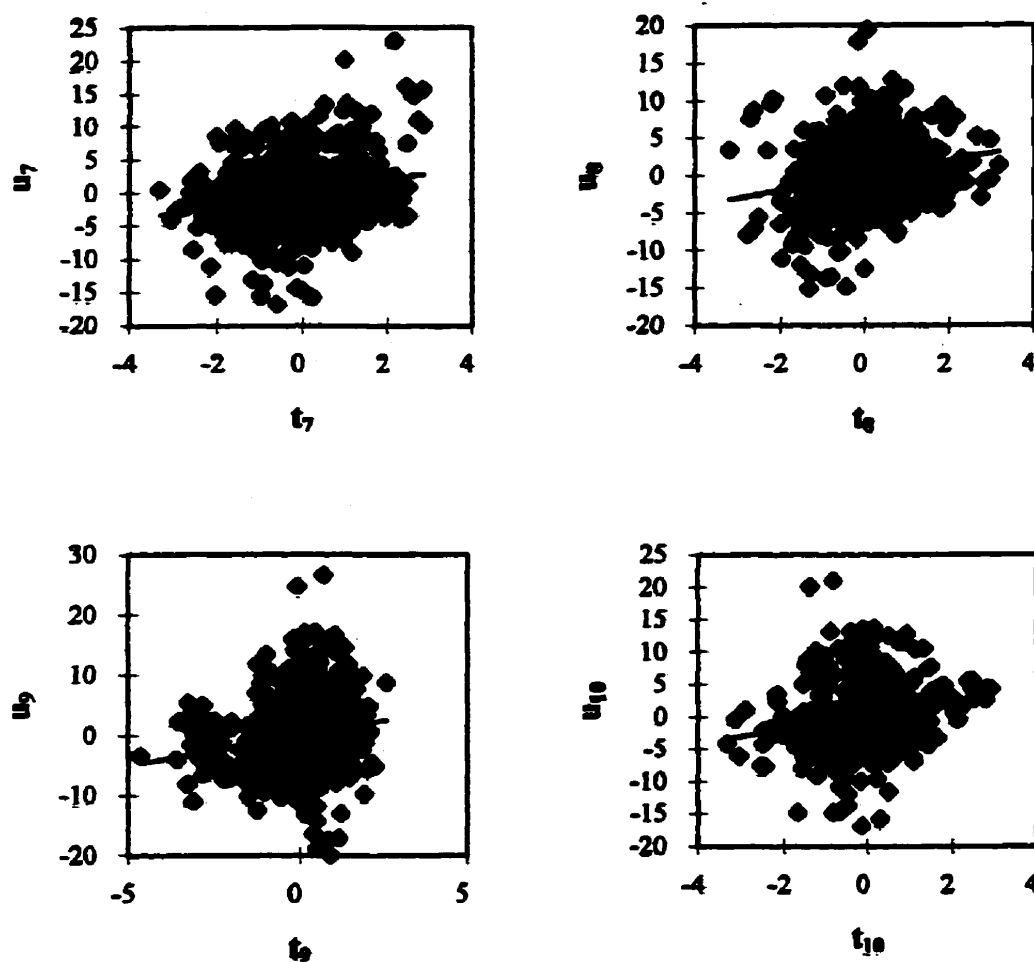


Figure II.XXIX: Graphiques  $u_i/t_i$ , modèle PLS, grade 580 (suite).

### Analyse des variables latentes

Les variables latentes  $t$  et les poids  $w$  permettent d'identifier à quel événement est associé chaque dimension. Elles permettent d'identifier les événements qui expliquent la variabilité de l'indice de réfraction.

### Dimension 1

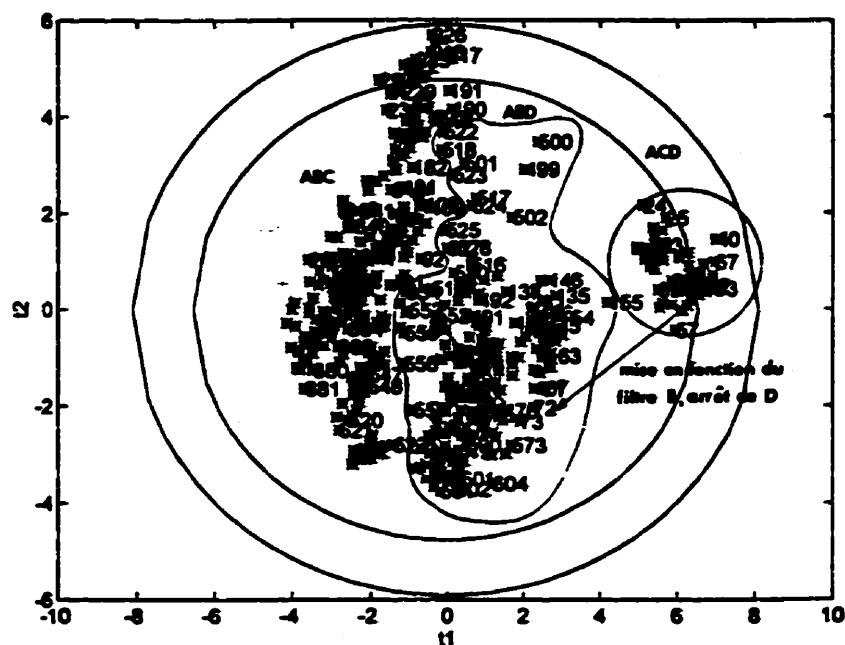


Figure III.II: Graphique  $t_2/t_1$ , modèle PLS, grade 580.

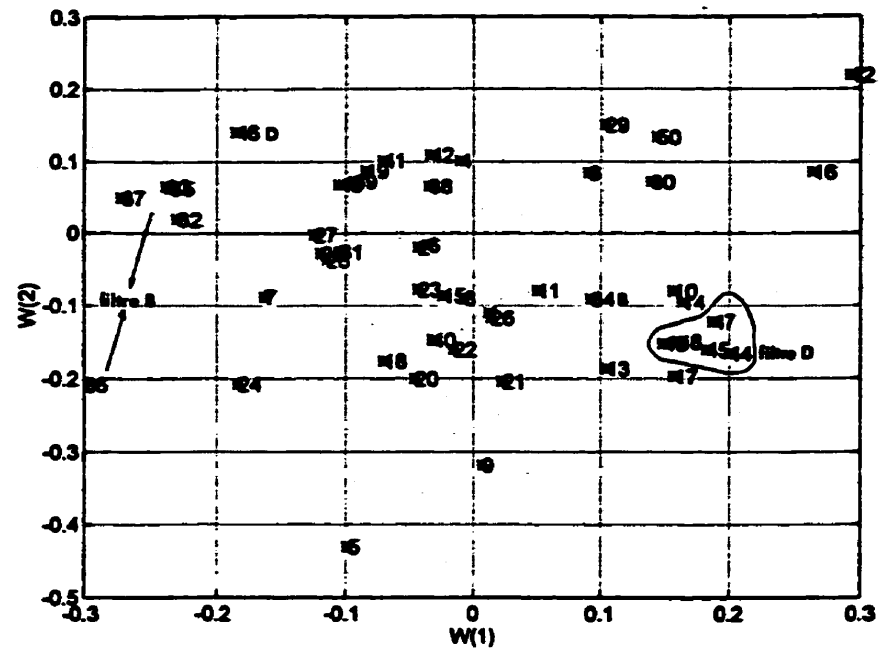


Figure III.III: Graphique  $w_2/w_1$ , modèle PLS, grade 580.

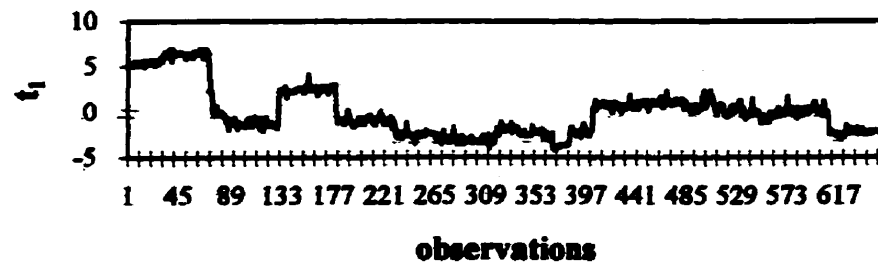


Figure III.IV:  $t_1$  en évolution dans le temps, modèle PLS, grade 580.

**Variable latente  $t_1$** 

La variable latente  $t_1$  regroupe les observations selon les filtres qui sont en opération:

- $t_1$  élevée: observations 1 à 70, filtres A,C,D en service;
- $t_1$  moyenne: observations 131 à 179 et 403 à 608, filtres A,B,D en service;
- $t_1$  faible: observations 74 à 130, 180 à 402 et 609 à 654, filtres A,B,C en fonction.

La variabilité du différentiel de température entre la troisième dilution et l'alimentation est aussi expliquée dans cette dimension, mais de façon moins importante que les filtres.

**Poids  $w_1$** 

Les variables les plus importantes dans cette dimension sont les paramètres d'opération des filtres B et D:

- niveaux des cuves des filtres B et D;
- débits de solvant froid alimentés aux filtres B et D;
- pressions d'aspiration basse des filtres B et D;
- vitesses de rotation des filtres B et D;
- niveaux et températures des bannes des filtres B et D.

La différence de température entre l'alimentation et la troisième dilution et la température à la sortie de l'échangeur X-2402A sont aussi importantes.



## Dimension 2

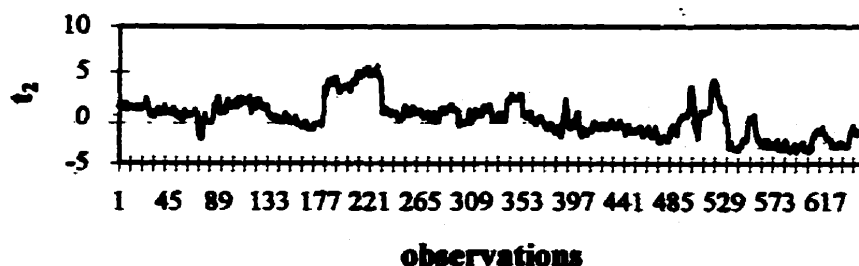


Figure III.V:  $t_2$  en évolution dans le temps, modèle PLS, grade 580.

### Variable latente $t_2$

La variable latente  $t_2$  modélise la variation des pressions du cycle de cristallisation. Elle identifie soit un changement de période d'opération (observations 179 à 180, 230 à 231, 488 à 489 et 608 à 609) ou des tests échelons (tests échelons sur la prédilution, deuxième dilution, troisième dilution et cinquième dilution). Ces variations de la variable latente  $t_2$  sont en grande partie causées par la variation des pressions à l'entrée des deux premiers échangeurs. La valeur de la variable latente  $t_2$  augmente lorsque la pression diminue.

Lorsque les débits de dilution augmentent, les pressions augmentent et  $t_2$  diminue. Par contre, le changement de pression entre les différentes journées d'opération ne correspondent pas nécessairement à une augmentation des débits, mais plutôt à l'encrassement des échangeurs. Lorsque les échangeurs sont plus encrassés, les pressions sont plus élevées, et  $t_2$  est plus faible.

### Poids $w_2$

Les variables importantes dans cette dimension sont les pressions du début du cycle de réfrigération. Ce sont les échangeurs les plus affectés par l'encrassement dû à l'accumulation de glace puisqu'on y alimente du solvant humide.

- Pression à l'entrée de l'échangeur X-2401A
- Pression à l'entrée de l'échangeur X-2401B

### Dimension 3

#### Variable latente $t_3$

La variable latente  $t_3$  modélise le fonctionnement du filtre C. Lorsque le filtre C fonctionne, la valeur de  $t_3$  est plus faible (observations 1 à 130, 179 à 402 et 609 à 654).

Les tests échelons sur la troisième dilution et l'arrêt de la cinquième dilution sont clairement identifiables sur cette dimension, comme sur la deuxième dimension.

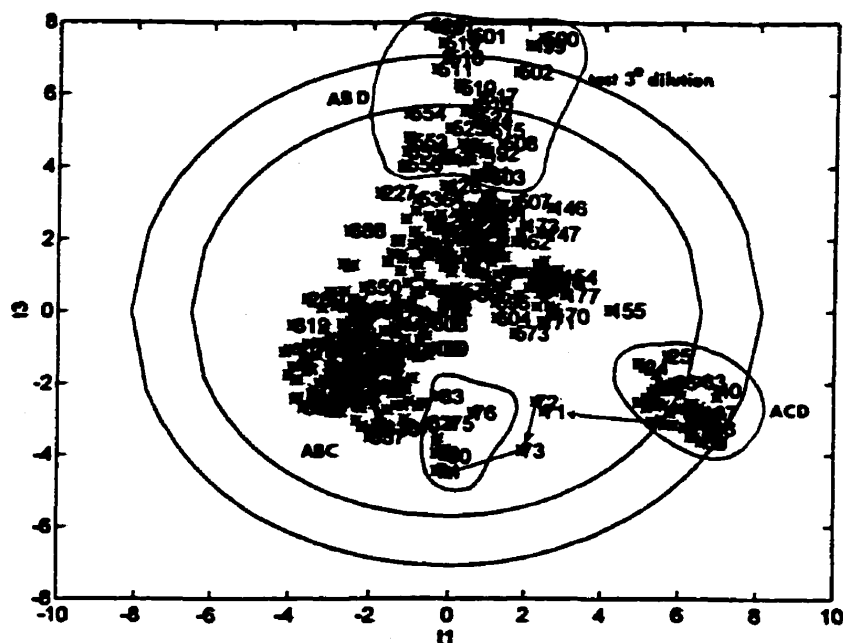


Figure III.VI: Graphique  $t_3/t_1$ , modèle PLS, grade 580.

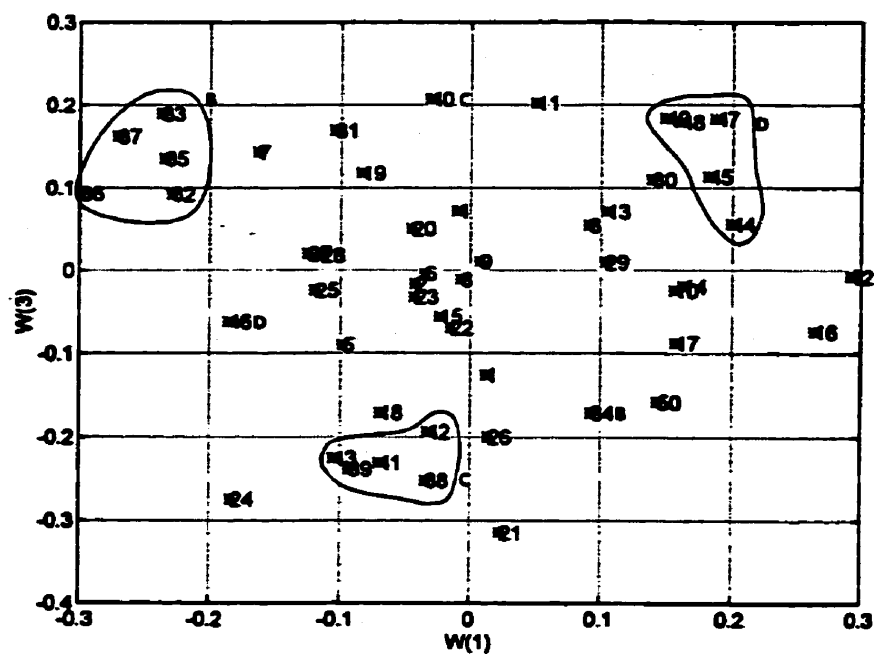


Figure III.VII: Graphique  $w_3/w_1$ , modèle PLS, grade 580.

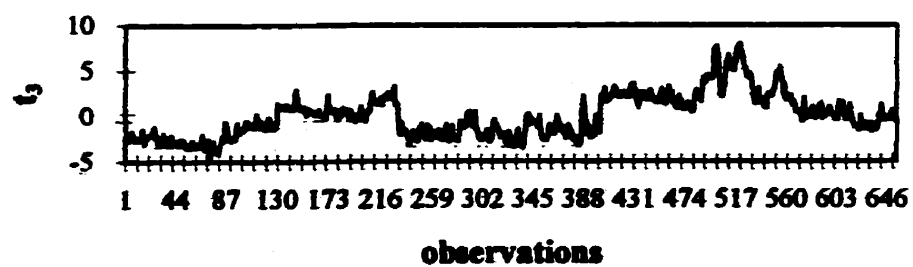


Figure III.VIII:  $t_3$  en évolution dans le temps, modèle PLS, grade 580.

**Poids  $w_3$** 

Les variables importantes pour cette dimension sont les paramètres d'opération du filtre C:

- Niveaux des curves des filtres C;
- Débits de solvant froid alimentés aux filtres C;
- Pressions d'aspiration basse des filtres C;
- Vitesses de rotation des filtres C;
- Niveaux et températures des bottes des filtres C.

Aussi le débit de la troisième dilution, la température à la sortie de l'échangeur X-2402B, le débit de filtrat recyclé et le niveau de la cuve du filtre A sont attribués de poids importants.

**Dimension 4****Variable latente  $t_4$** 

Cette variable latente modélise les tests échelons sur la prédilution et la troisième dilution. Les changements de période d'opération sont identifiés par cette variable latente à cause de l'importance des variables associées aux tests échelons déjà cités.

**Poids  $w_4$** 

Les paramètres les plus importantes sont:

- le débit de la troisième dilution;
- le débit de prédilution;
- la pression à l'entrée de l'échangeur X-2402A;
- la pression à l'entrée de l'échangeur X-2401B;
- la température de la troisième passe de l'échangeur X-2401A.
- la température à l'entrée de l'échangeur X-2401A;
- la température à l'entrée de l'échangeur X-2401B;

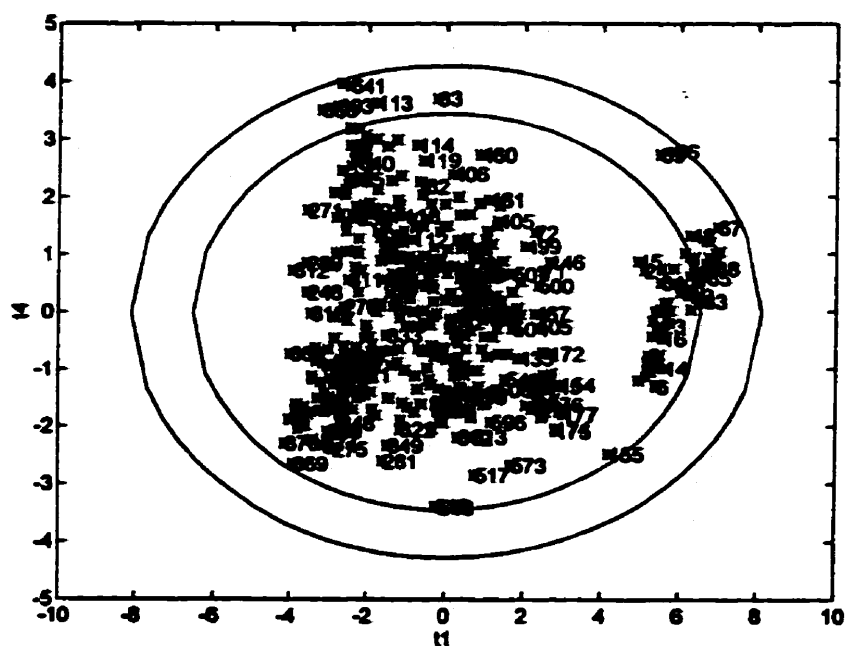


Figure III.IX: Graphique  $t_2/t_1$ , modèle PLS, grade 580.

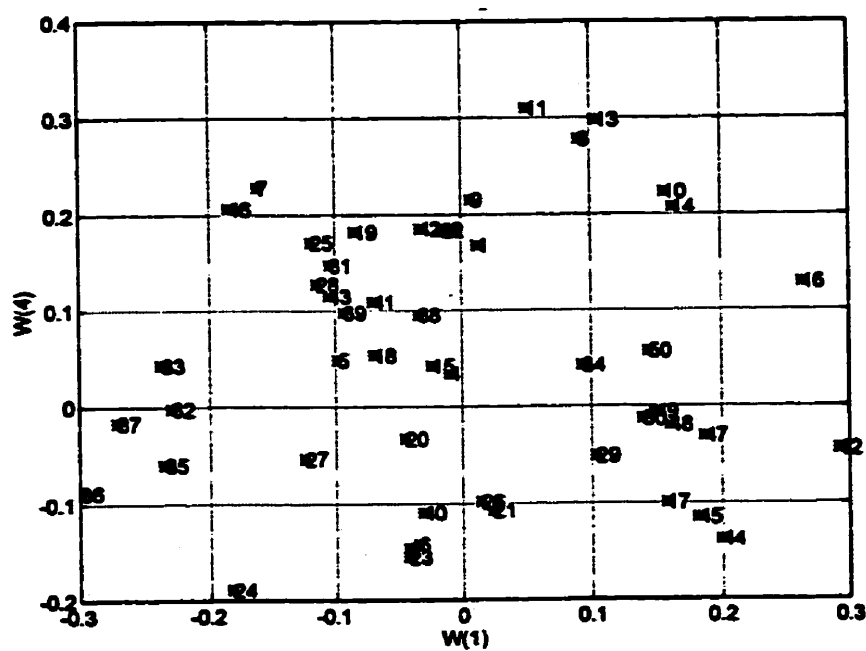


Figure III.X: Graphique  $w_2/w_1$ , modèle PLS, grade 580.

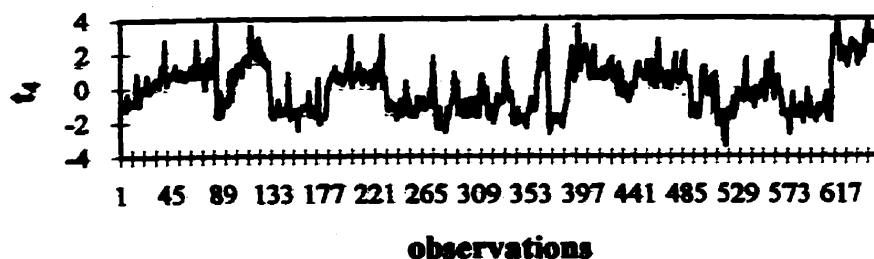


Figure III.XI:  $t_4$  en évolution dans le temps, modèle PLS, grade 580.

## Dimension 5

### Variable latente $t_5$

Cette variable latente modélise la variabilité des paramètres d'opération du filtre A. (observations 74 à 75, 500 à 501, 402 à 403, 425 à 428, etc.). De plus, les variations des températures du cycle de cristallisation sont modélisés par cette variable latente: observations 315 à 317, 230 à 231, 359 à 360 (changements de période d'opération).

### Poids $w_5$

Les variables les plus importantes sont:

- la pression d'aspiration basse du filtre A;
- le niveau de la botte du filtre A;
- le niveau de la cuve du filtre A.

Les températures à l'entrée et à la sortie de l'échangeur X-2401A, la différence de température entre la cinquième dilution et l'alimentation, la température du solvant froid

ainsi que la pression à l'entrée du troisième échangeur sont aussi des paramètres importants dans cette dimension.

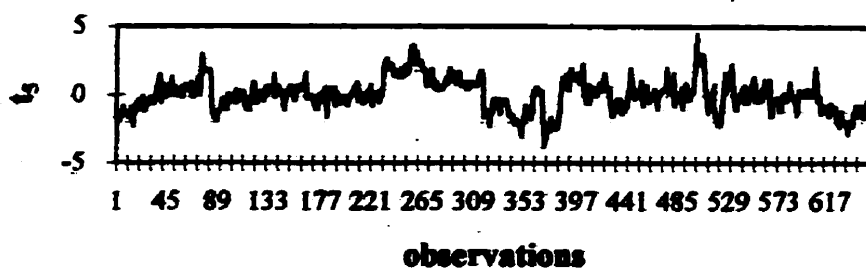


Figure III.XII:  $t_3$  en évolution dans le temps, modèle PLS, grade 580.

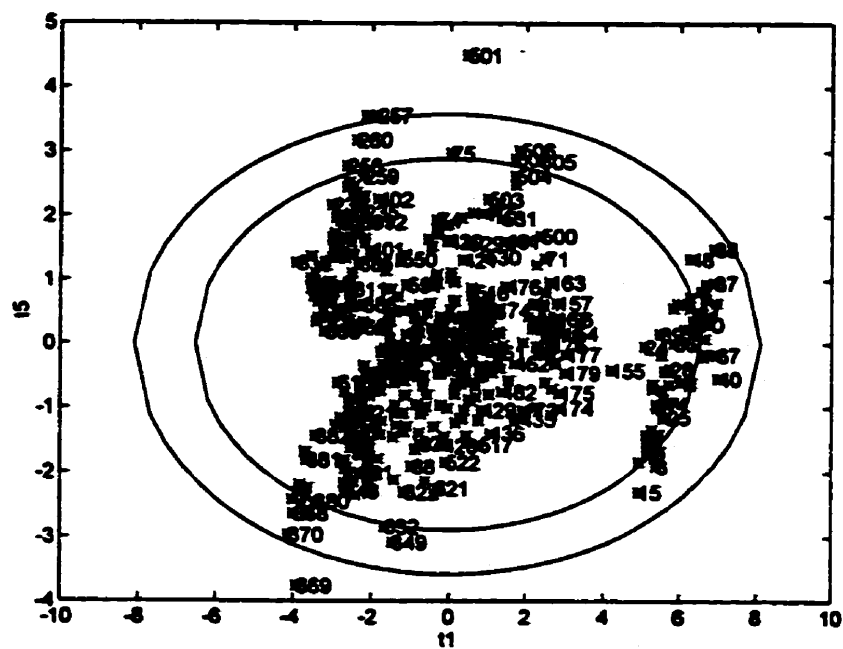


Figure III.XIII: Graphique  $t_2/t_1$ , modèle PLS, grade 580.

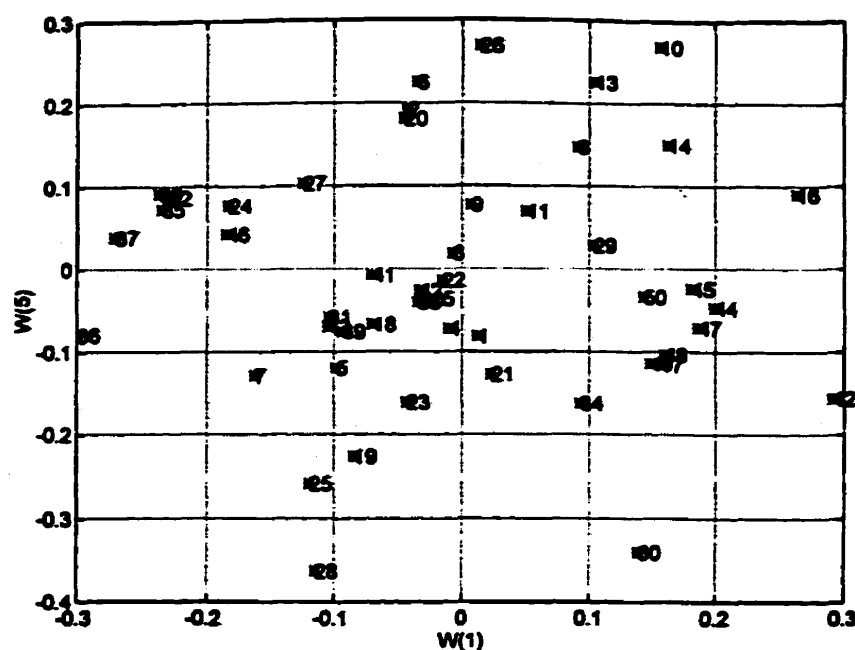


Figure III.XIV: Graphique  $w_5/w_1$ , modèle PLS, grade 580.

## Dimension 6

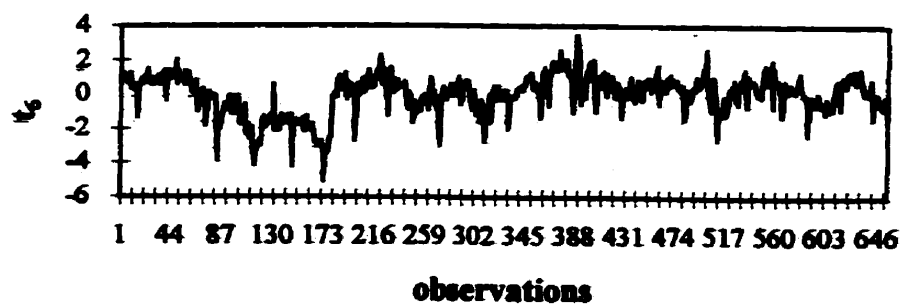
### Variable latente $t_6$

Cette dimension modélise plusieurs événements. D'abord, la variation du niveau du réservoir V-2401: observations 82 à 83, 145 à 146, 311 à 312, 388 à 389, etc.

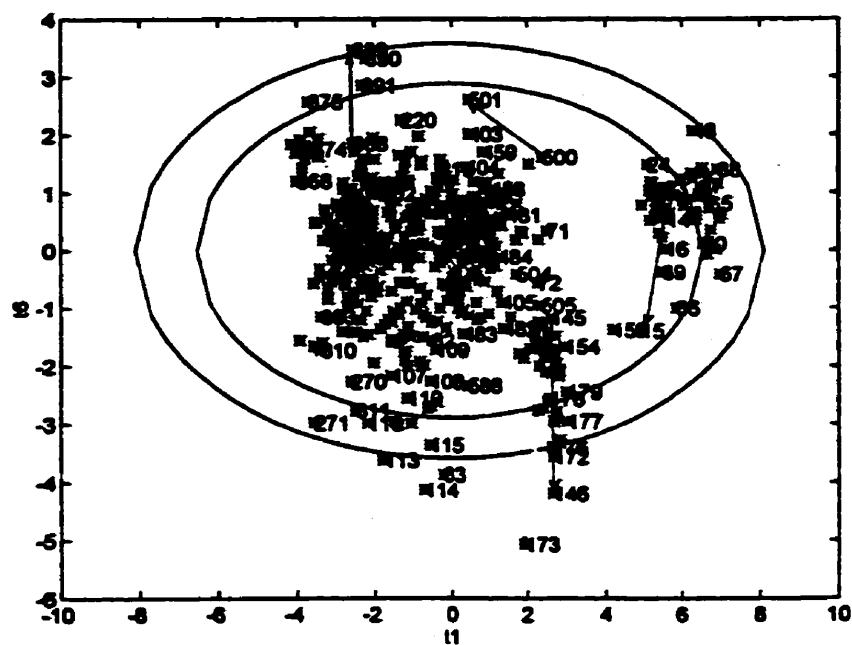
Certains changements de période d'opération sont aussi modélisés: observations 14 à 15, 179 à 180, 109 à 114, etc.

Les tests échelons sur la prédilution et sur la troisième dilution sont aussi modélisés par  $t_6$ .





**Figure III.XV:  $t_g$  en évolution dans le temps, modèle PLS, grade 580.**



**Figure III.XVI: Graphique,  $t_0/t_1$ , modèle PLS, grade 580.**

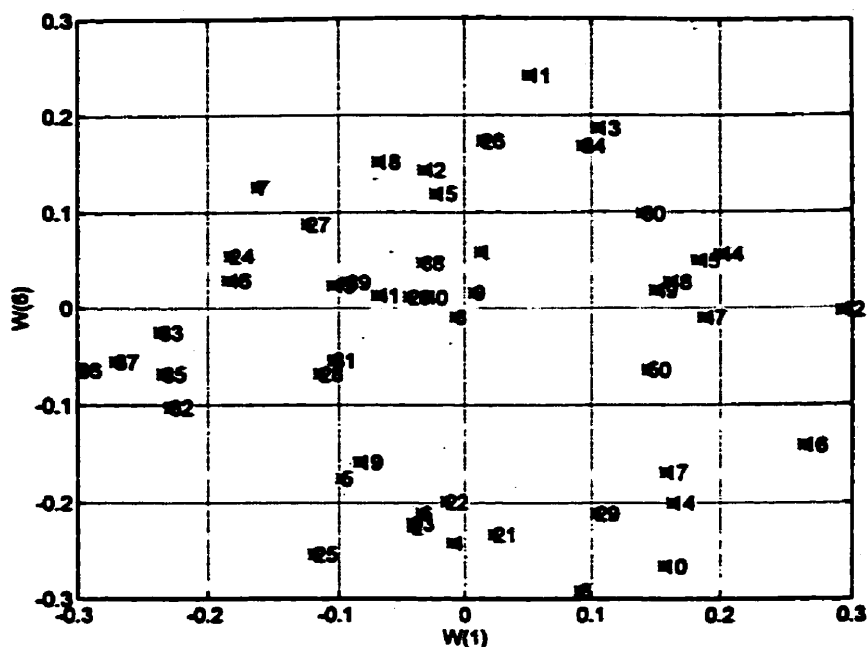


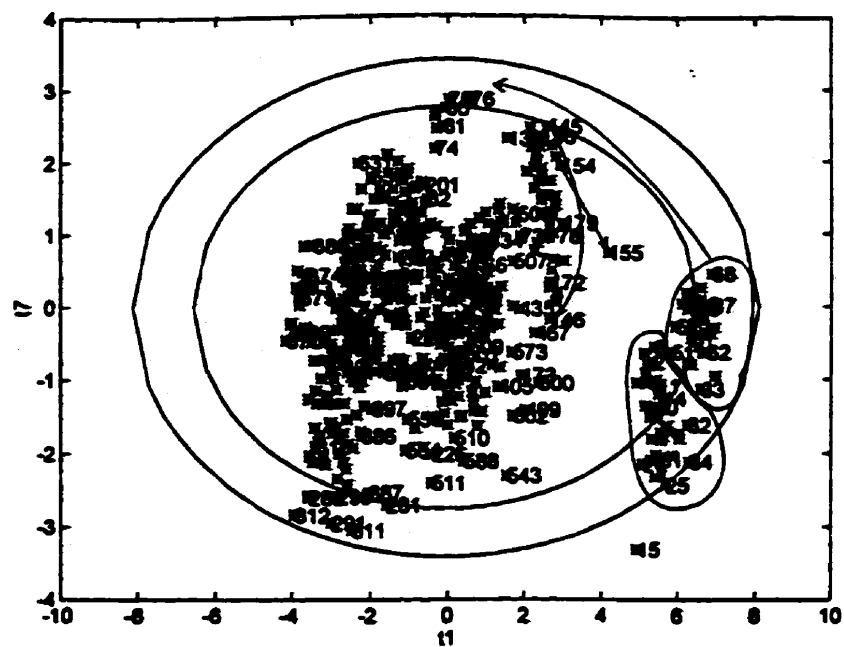
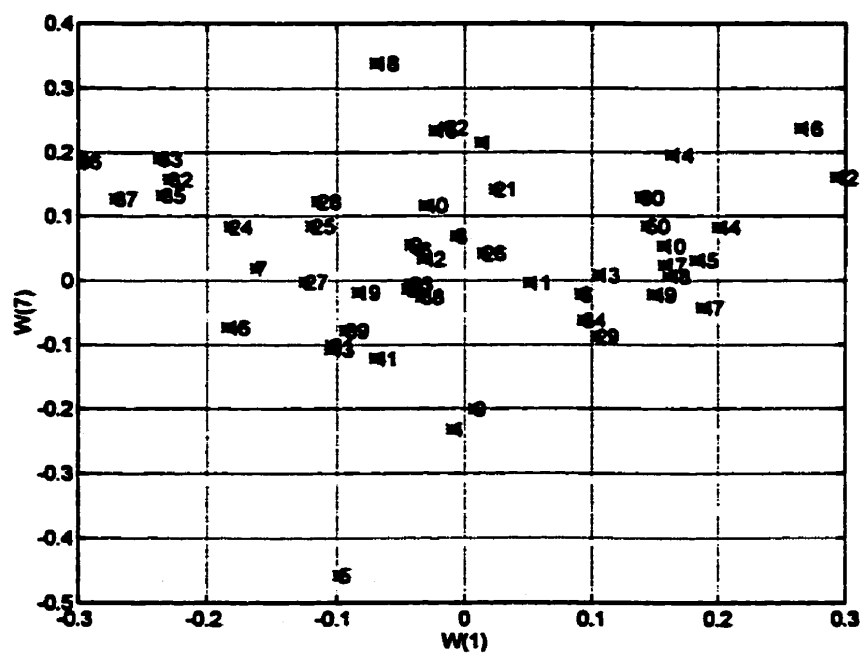
Figure III.XVII: Graphique  $w_6/w_1$ , modèle PLS, grade 580.

#### Poids $w_6$

Plusieurs variables sont importantes dans cette dimension:

- le débit de prédilution;
- le débit de la troisième dilution;
- les températures à la sortie des échangeurs X-2401A, X-2401B, X-2402B et X2437;
- les températures de la vapeur et de l'alimentation à l'échangeur X-2401A;
- la température du solvant froid;
- la température de la prédilution et;
- le niveau du réservoir d'alimentation au filtre.

## Dimension 7

Figure III.XVIII: Graphique  $t_7/t_1$ , modèle PLS, grade 580.Figure III.XIX: Graphique  $w_7/w_1$ , modèle PLS, grade 580.

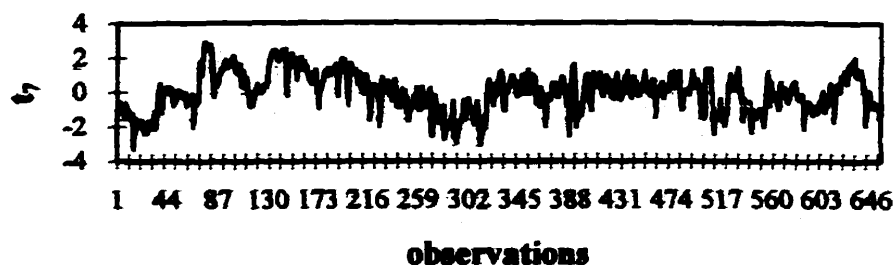


Figure III.XX:  $t_7$  en évolution dans le temps, modèle PLS, grade 580.

#### Variable latente $t_7$

La variable latente  $t_7$  explique la variabilité de la pression à l'entrée de l'échangeur X-2401A et du débit de la cinquième dilution. La chute du débit de la cinquième dilution qui a eu lieu le 29 août 1996 est bien identifiée.

#### Poids $w_7$

Les variables les plus importantes sont: la pression à l'entrée de l'échangeur et le débit de la cinquième dilution.

#### **Dimension 8**

#### Variable latente $t_8$

La variable  $t_8$  explique à nouveau la variabilité du filtre A et de la pression à l'entrée de l'échangeur X-2401A.

#### Poids $w_8$

Les variables importantes sont la pression d'entrée à l'échangeur X-2401A, l'aspiration basse du filtre A et finalement le niveau de la cuve du filtre A.

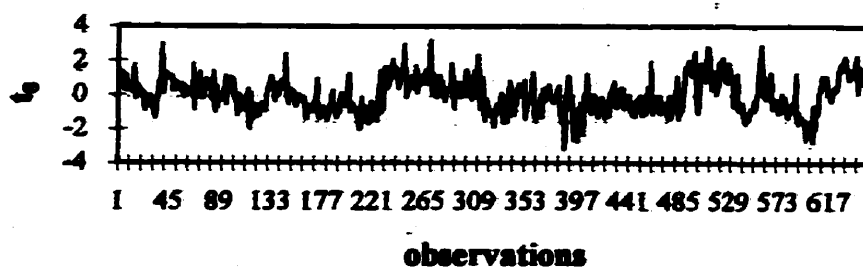


Figure III.XXI:  $t_2$  en évolution dans le temps, modèle PLS, grade 580.

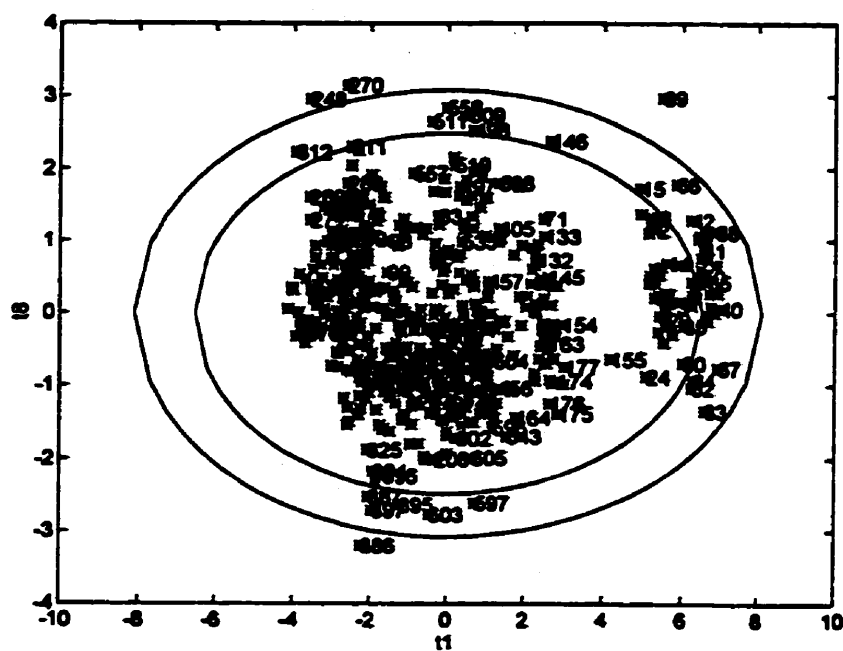


Figure III.XXII: Graphique  $t_2/t_1$ , modèle PLS, grade 580.

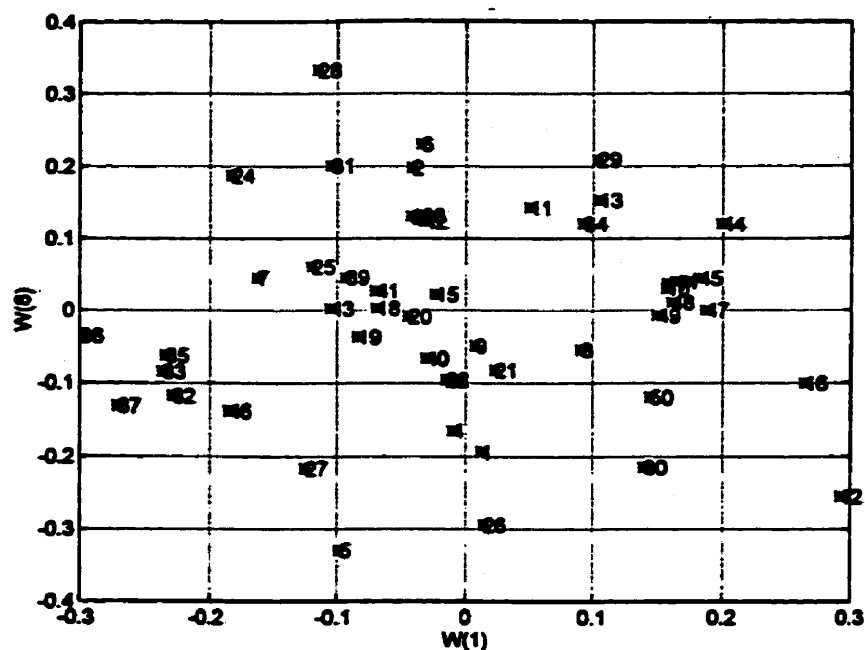


Figure III.XXIII: Graphique  $w_2/w_1$ , modèle PLS, grade 580.

### Dimension 9

#### Variable latente $t_2$

Des événements distincts sont expliqués par ces trois dimensions:

- la variabilité de la pression à l'entrée de l'échangeur X-2437 (observations 33 à 36, 130 à 131, 498 à 501 (chute du débit de cinquième dilution), 532 à 535, etc.);
- la température de la prédilution (observations 130 à 131, 498 à 501, 516 à 517, 532 à 535, 608 à 609);
- la vitesse de rotation du filtre A (170 à 200, 608 à 609 et 636 à 640), et;
- le niveau dans le réservoir V-2401 (observations 516 à 517, 532 à 535, 501 à 503).



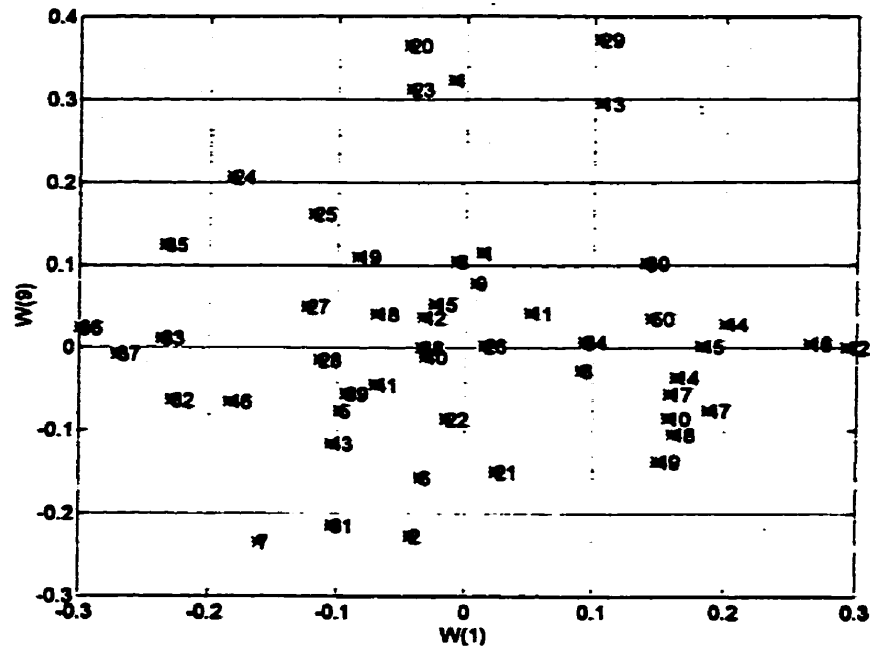


Figure III.XXVI: Graphique  $w_9/w_1$ , modèle PLS, grade 580.

### Dimension 10

#### Variable latente $t_{10}$

Cette dimension explique clairement la variation du niveau dans le réservoir d'alimentation aux filtres:

- observation 315 à 318;
- observations 495 à 501;
- observations 501 à 503;
- observations 519 à 520;
- observations 532 à 538;
- observations 545 à 558 et;
- observations 608 à 609.



**Poids  $w_{10}$** 

Le paramètre d'opération important pour cette dimension est le niveau dans le réservoir V-2401.

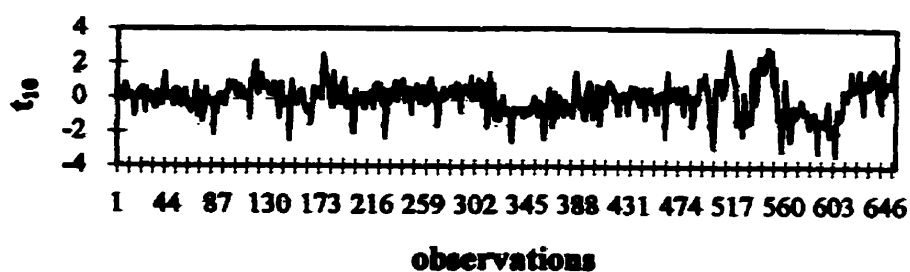


Figure III.XXVII:  $t_{10}$  en évolution dans le temps, modèle PLS, grade 580.

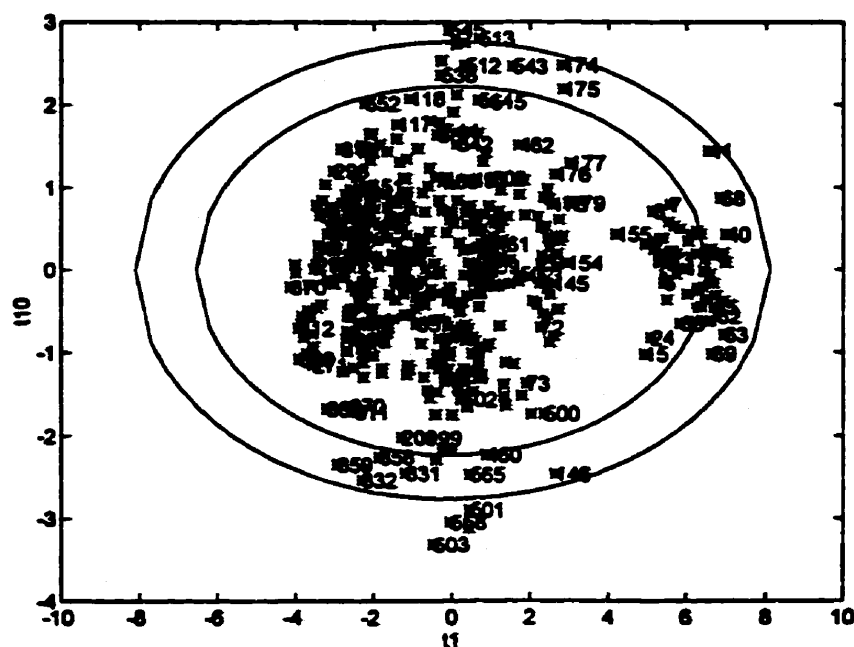


Figure III.XXVIII: Graphique  $t_{10}/t_1$ , modèle PLS, grade 580.

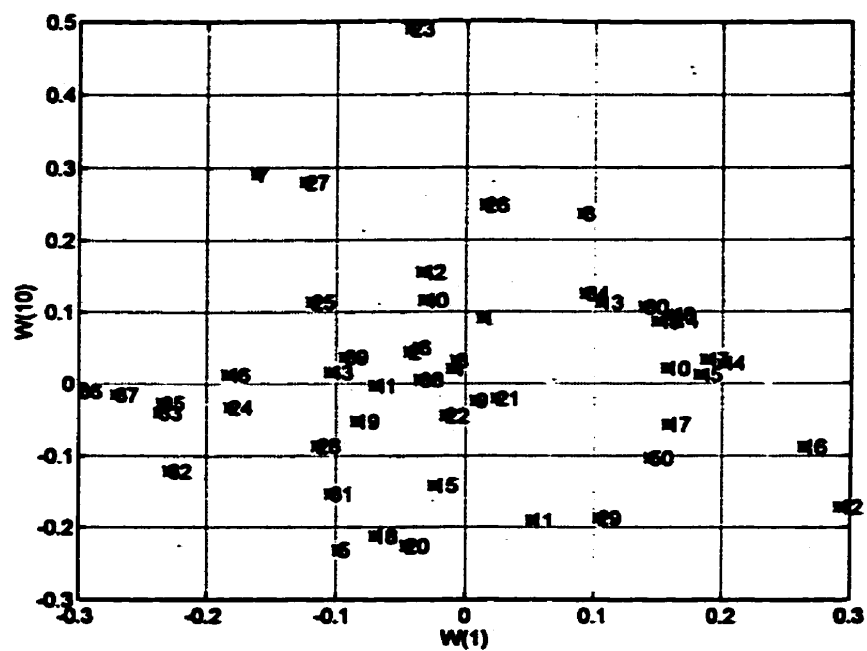


Figure III.XXIX: Graphique  $w_{10}/w_1$ , modèle PLS, grade 580.