ESTIMATOR DESIGN WITH PLS MODEL FOR CONSISTENT CONTROL OF REFINERY MAIN FRACTIONATORS

Dante Pastore * **Alessandro Brambilla** * **Gabriele Pannocchia** *,[1]

* *Department of Chemical Engineering – University of Pisa*
Via Diotisalvi 2, 56126 Pisa (Italy)

Abstract: In this paper the problem of designing product quality inferentials for refinery main fractionations is addressed by using the PLS regression. A simulated crude distillation unit is chosen as case study, and several linear steady-state estimators are designed and compared in terms of accuracy and consistency, i.e. the estimator ability of guaranteeing low closed-loop offset. The paper shows the importance of the auxiliary measurement choice in order to build an effective inferential control scheme. Moreover, it shows that the use of only temperature measurements is not sufficient to guarantee an acceptable estimator performance. Additional measurements as the operating pressure and internal flows have been used to improve the estimator accuracy and consistency.

Keywords: distillation control, estimator, consistency, inferential

1. INTRODUCTION

Product quality control is an important and difficult issue in many chemical applications and in particular in refinery main fractionators because of the effect of these units on the quality of final commercial products and on the efficiency of downstream operations. For such processes on-line quality analyzers are not widely used because they are expensive, because they require frequent maintenance work and because, for some properties, they are not available. Moreover, on-line analyzers suffer from large time delays which would make the product quality control a difficult task. Indeed, a common alternative is to use some auxiliary measurements (often tray temperatures) to infer the product properties, thus building an inferential control scheme. The issue of measurement selection is of crucial importance for the effectiveness of an estimator, and it has been the subject of extensive research in the chemical engineering community (Joseph and Brosilow, 1978; Morari and Stephanopoulos, 1980; Yu and Luyben, 1987; Mejdell and Skogestad, 1991).

In this work product quality estimators are investigated for a crude distillation unit (which is one of the most common main fractionators present in oil refineries) using the PLS regression technique. Comparisons among different linear steady-state estimators in terms of accuracy and achievable control performance are presented. To this aim the concept of estimator closed-loop consistency (Pannocchia and Brambilla, 2002) is used, and results show that estimators "apparently" well designed (i.e. in terms of accuracy) should not be used in an inferential control scheme because they would lead to improper control actions with significant closed-loop offset in the presence of disturbances and/or set-point changes.

2. BASIC BACKGROUND CONCEPTS

2.1 *PLS regression*

PLS is a multivariate regression technique that can easily handle large numbers of noisy and correlated data sets, and it is based on the extraction of a number of *latent variables* which are linear combinations of the original predictor variables. PLS has been extensively applied in chemometrics (see e.g. (Wold *et al.*, 2001*b*) and references therein) and, more recently, in the area of process engineering (Mejdell and Skogestad, 1991; Kresta *et al.*, 1994; Kano *et al.*, 2000);

[1] Corresponding author. Email: pannocchia@ing.unipi.it, Fax: +39 050 511266.

several extensions of PLS have been proposed during the last years (see e.g. (Wold *et al.*, 2001*a*) and references therein).

Linear PLS seeks a relation between the auxiliary variables $x \in \mathbb{R}^m$ and the response (dependent) variables $y \in \mathbb{R}^p$ of the type:

$$\hat{y} = Kx \ , \tag{1}$$

in which $\hat{y} \in \mathbb{R}^p$ is the $y-$estimate and $K \in \mathbb{R}^{p \times m}$ is the estimator gain. Notice that both the auxiliary variables and the response variables are centered around a reference (mean) value, so that no constant bias term appears in (1). Given a set of n training runs, i.e. n vectors x and the corresponding n vectors y, the matrices X and Y are built by stacking the corresponding x and y vectors as rows. Usually, each variable (i.e. each column of X and Y) is scaled to unit variance. Then, PLS generates a "few" (k in number) $X-$scores as a linear combination of the original variables:

$$T = XW \ , \tag{2}$$

in which $T \in \mathbb{R}^{n \times k}$ is the $X-$score matrix and $W \in \mathbb{R}^{m \times k}$ is an appropriate weight matrix. These weights are computed so that each of them maximizes the covariance between the response variables and the $X-$scores. The $X-$scores are, multiplied by an appropriate loading matrix $P \in \mathbb{R}^{m \times k}$, good "summaries" of X, that is:

$$X = TP^T + E \ , \tag{3}$$

in which $E \in \mathbb{R}^{n \times m}$ contains the $X-$residuals. Then, a linear regression model for Y is obtained as:

$$Y = TQ^T + F = XWQ^T + F \ , \tag{4}$$

in which $Q \in \mathbb{R}^{p \times k}$ is an appropriate matrix and $F \in \mathbb{R}^{n \times p}$ contains the $Y-$residuals. Finally, the PLS regression estimates can be written as:

$$\hat{Y} = XK^T \ , \tag{5}$$

in which the estimator gain is given by

$$K = QW^T \ . \tag{6}$$

Several PLS algorithms exist in literature (Wold *et al.*, 2001*b*) and the following by de Jong (1993) is used in this work.

Algorithm 1. (SIMPLS).
For each $h = 1, \ldots, k$ (where $A_1 = \Theta^T Y$, $M_1 = \Theta^T \Theta$, $C_1 = I$), repeat the following steps:

(1) compute q_h, the dominant eigenvector of $A_h^T A_h$,
(2) set $w_h = A_h q_h$, $c_h = w_h^T M_h w_h$, $w_h \leftarrow w_h / \sqrt{c_h}$, and store w_h into W as a column,
(3) set $p_h = M_h w_h$, and store p_h into P as a column,
(4) set $q_h = A_h^T w_h$, and store q_h into Q as a column,
(5) set $v_h = C_h p_h$ and $v_h \leftarrow v_h / \|v_h\|_2$,
(6) set $C_{h+1} = C_h - v_h v_h^T$ and $M_{h+1} = M_h - p_h p_h^T$,
(7) set $A_{h+1} = C_{h+1} A_h$.

Multivariate techniques as PLS can also be used to select a smaller number of auxiliary variables to be used in the estimator. In this work, several tray temperatures are used to estimate the product quality properties of interest (see Section 3), and the location choice of such temperatures is based on the method discussed in (Mejdell and Skogestad, 1991, Sec. 5.6). However, it is important to remark that several engineering considerations specific of refinery main fractionators will be made to exclude some trays.

An important issue in PLS is the determination of the number of latent variables to use. To this aim, the so-called "Explained Variance" (EV) is introduced. Let $y_{i,j}$ be the actual value of the $j-$th response variable in the $i-$th calibration run, and let $\hat{y}_{i,j}(k)$ be the corresponding estimate obtained by the PLS estimator with k latent variables. Then, the "Mean Square Error" (MSE) for the $j-$th response variable is given by:

$$\mathrm{MSE}_j(k) = \frac{1}{n} \sum_{i=1}^{n} (y_{i,j} - \hat{y}_{i,j}(k))^2 \ , \tag{7}$$

and the Explained Variance (EV) for the j-th product composition is

$$\mathrm{EV}_j(k) = 100 \left(1 - \frac{\mathrm{MSE}(k)}{\mathrm{MSE}(0)} \right) \ . \tag{8}$$

Typically, one increases the number of latent variables k until the increment in EV, i.e. $\mathrm{EV}(k+1) - \mathrm{EV}(k)$, is not significant (say less than 1-2%). In fact, if a latent variable which gives a small increment in the explained variance is used, the estimator accuracy does not improve significantly while the estimator becomes sensitive to data errors aligned with the direction of this latent variable, and the estimator prediction ability deteriorates. See e.g. (Wold *et al.*, 2001*b*, par. 3.8) for a detailed discussion on methods for choosing the number of latent variables.

2.2 *Estimator consistency*

Usually, the estimator is designed to fit the training data reasonably well and it is validated with additional (historical) data (as well as with real time plant data). After this stage, the estimator is inserted into the control loop, and no implications on the control performance, e.g. about the steady-state offset, are evaluated at the estimator design stage. The underlying assumption is that if an estimator fits the data sufficiently well the corresponding inferential control scheme will perform satisfactorily, as well. However, in general this is incorrect.

In order to address this "myth" Pannocchia and Brambilla (2002) recently introduced the concept of estimator consistency, which is the ability of an estimator to guarantee low steady-state offset in the unmeasured controlled variables in the presence of disturbances (and/or setpoint changes). It is shown that the estimator consistency is not necessarily related to its accuracy, and it is shown how the steady-state offset is related to the estimator consistency. The problem of closed-loop consistency arises even in single-input single-output (SISO) systems and it is more dramatic for multi-input multi-output (MIMO) systems, as the

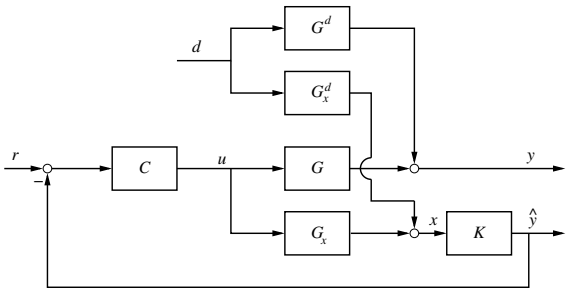

Fig. 1. Inferential control scheme

process studied in this work. This concept is quantified through the introduction of an appropriate consistency measure, which is briefly recalled.

Consider the generic inferential control scheme reported in Fig. 1, in which $u \in \mathbb{R}^p$ is the manipulated variable, $y \in \mathbb{R}^p$ is controlled variable (unmeasurable), $x \in \mathbb{R}^m$ is the auxiliary variable (measurable), $\hat{y} \in \mathbb{R}^p$ is the estimate of the controlled variable, and $d \in \mathbb{R}^q$ is the disturbance input. Notice that the feedback controller C operates on the estimate of the controlled variable, i.e. on $\hat{y}$. For a generic disturbance $d \neq 0$ (assuming $r = 0$), the consistency matrix $\xi = \{\xi_{i,j}\}$ is defined as (Pannocchia and Brambilla, 2002):

$$\xi_{i,j} = \frac{\left(\frac{\partial u_i}{\partial d_j}\right)_{\hat{y}=0}}{\left(\frac{\partial u_i}{\partial d_j}\right)_{y=0}} = \frac{\left\{\begin{array}{c}\text{variation of } u_i \text{ when} \\ \text{rejecting } d_j \text{ on } \hat{y}\end{array}\right\}}{\left\{\begin{array}{c}\text{variation of } u_i \text{ when} \\ \text{rejecting } d_j \text{ on } y\end{array}\right\}} \quad (9)$$

Similarly, for a generic setpoint change $r \neq 0$ (assuming $d = 0$) the consistency matrix $\varphi = \{\varphi_{i,j}\}$ is defined as:

$$\varphi_{i,j} = \frac{\left(\frac{\partial u_i}{\partial r_j}\right)_{\substack{\hat{y}_j = r_j \\ \hat{y}_{l\neq j}=0}}}{\left(\frac{\partial u_i}{\partial r_j}\right)_{\substack{y_j = r_j \\ y_{l\neq j}=0}}} = \frac{\left\{\begin{array}{c}\text{variation of } u_i \text{ to bring} \\ \hat{y}_j \text{ to } r_j \text{ with } \hat{y}_{l\neq j} = 0\end{array}\right\}}{\left\{\begin{array}{c}\text{variation of } u_i \text{ to bring} \\ y_j \text{ to } r_j \text{ with } y_{l\neq j} = 0\end{array}\right\}} \quad (10)$$

In a "small" neighborhood around the nominal steady-state, the process behavior can be linearized leading to the following expressions for ξ and φ:

$$\xi = \left[(KG_x)^{-1}(KG_x^d)\right] ./ \left[G^{-1}G^d\right] \quad (11)$$

$$\varphi = (KG_x)^{-1} ./ G^{-1} \,, \quad (12)$$

in which $./$ means element-by-element division, and G_x, G_x^d, G, G^d represent the gain matrices of the blocks shown in Fig. 1. It is clear from the definition that a desirable property of an estimator K is that $\xi_{i,j} \approx 1$ and $\varphi_{i,j} \approx 1$. Moreover, the steady-state offset $e_{\mathrm{CL}} = r - y$ can be expressed as (Pannocchia and Brambilla, 2002):

$$e_{\mathrm{CL}} = \epsilon_d d + \epsilon_r r \,, \quad (13)$$

in which $\epsilon_r = I - G\left(KG_x\right)^{-1}$ and $\epsilon_d = G\left(KG_x\right)^{-1} KG_x^d - G^d$. It is important to notice that these consistency matrices do not depend on the type of controller used (decentralized or multivariable) as long as the system is square and integral action is used.

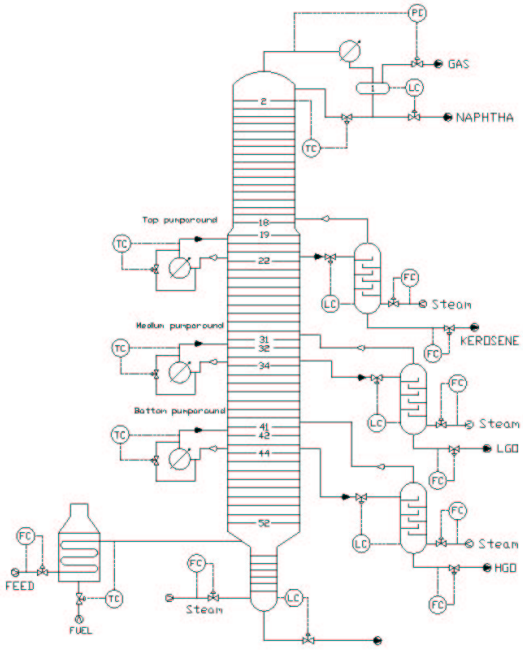

Fig. 2. Crude distillation unit layout with basic control loops

3. PROCESS DESCRIPTION

The crude distillation unit (CDU), which represents the first stage of separation of the crude oil, has the purpose of fractionating the crude oil into a number of products (4-5) with defined boiling temperature ranges. The most common CDU process scheme shows two separation stages: the first (Pre-flash unit) is a partial vaporization of lighter components present in the crude oil, and the second one (Main Fractionator unit) is the separation into products of the components vaporized after the furnace. This unit shows the presence of pumparounds for internal vapor flow reduction and heat recovery, and side strippers that increase the complexity of the unit. The process examined is depicted in Fig. 2, in which only the basic control loops are shown, and it has been simulated by means of the rigorous steady-state simulator Aspen Plus 10.2®. The crude oil, after pre-heating, undertakes a flash (pressure reduction to 4 atm) to remove the lightest components, thus reducing the load of the furnace. Then it is heated in the furnace up to 355 ℃ and, after a further pressure reduction, it enters the column almost at the bottom (this section of the column is known as "flash zone"). The column top product partly condensed constitutes the Naphtha, while several side-stream products as kerosene, light gas oil (LGO) and heavy gas oil (HGO) are drawn off from the column. Each side stream enters a stripper in which usually steam is used to remove the lighter components. From the bottom of the column a stream with heavy components defined as "atmospheric residue" is obtained, which usually undertakes a vacuum distillation. Several liquid streams are taken from the column, subcooled (usually with the fresh feed prior to entering the heat exchanger) and returned to the column a few trays up to condensate part of the vapor going up.

These external cooling systems called “pumparounds” are useful because, reducing the internal vapor flow, they maintain a uniform load of the column by recovering heat.

The product quality is usually defined in terms of boiling range, for example in terms of ASTM D86 95%, which roughly represents the temperature at which 95% of a product is evaporated. Some products as Kerosene and LGO have to meet quality also in terms of “Freezing point” and “Cloud point” (cold properties) which are related to the boiling range, but also to the type of crude processed. In the present work only ASTM D86 95% has been considered as product quality (for simplicity denoted with T95). The setpoint value for T95 of each product is as follows: 173℃ for naphtha, 236℃ for kerosene, 345℃ for LGO and 394℃ for HGO. The top temperature setpoint and side product flow-rate setpoints are the manipulated variables used to achieve these product quality targets. A complete description of all column parameters is available in (Pastore, 2002) and not presented here for the sake of space.

4. ESTIMATOR DESIGN

In this section several estimators for T95 of the four products are built by using the PLS regression with different inputs as auxiliary (measured) variables. As it will be shown, the input selection is the key step of an effective estimator design particularly in terms of closed-loop consistency. It is important to notice that an estimator is designed for each product T95 independently, and this implies that each product property is estimated by means of different inputs. In fact, the product properties are not strictly correlated and a unique “centralized” estimator may lead, in general, to a more interacting multivariable control system and in some cases to infeasibilities.

The training set consists of 152 simulation runs in which the manipulated variables and several disturbances are varied (positive and negative steps) one by one. Moreover, 40 runs correspond to varying one manipulated variable with the other product quality loops closed. In fact, this training set has led to the choice of inputs that guarantee better estimator performance both in terms of accuracy and consistency (Pastore, 2002). It is important to remark that the last set of runs could be difficult to be done on the actual plant, because on-line analyzers for such properties are usually not available. However, rigorous simulators can be used to build a training set that includes closed-loop runs (as well as open-loop ones), and this allows one to choose the most appropriate input locations. Then, the actual estimator coefficients are calculated through a PLS regression on real plant data.

4.1 *Input selection*

For each product quality estimator several inputs have been considered, which consist of a number of tray temperatures, as well as other additional inputs as the operating pressure. The first estimator considered, referred to as E0, uses for each product quality the temperature of the tray where the product is drawn off from the column and fed to the corresponding stripper. Notice that this estimator uses inputs that are always available on a CDU.

Then, for each product quality, an “optimal” location of the temperature measurements is found by using the method proposed by Mejdell and Skogestad (1991, sec. 5.6), which is based on the PLS regression on “all” tray temperatures. It is worth noticing that the column sections included between the draw tray and pumparound return are not considered in this search because the function of those trays is mainly to exchange heat between the two phases by condensing vapor. In fact, the pumparounds present in those section (see Fig. 2) cause intense upsets (due to the condensation of the upcoming vapor) that render the trays in those sections far away from the equilibrium, thus making the temperature measurement not reliable to estimate the product properties. Thus, for each product quality a PLS regression is carried out using all the “potential” temperatures with a number of latent variables chosen by means of the explained variance as discussed in Sec. 2.1. Then, the trays are ranked in descending order of the absolute value of the corresponding PLS estimator coefficient in (6) (for the mean centered and covariance scaled data), and this corresponds to ranking the tray temperatures in order of correlation importance. In fact, since the tray temperatures are centered and scaled with respect to their covariance, a larger estimator coefficient indicates a temperature more correlated to the product quality property.

A large number of estimators have been examined (Pastore, 2002), and the following ones are shown and compared with E0.

- E1, estimator that uses the first three most correlated temperatures.
- E2, estimator that uses the first two most correlated temperatures and the top pressure.
- E3, estimator that uses the first three most correlated temperatures, the top pressure and the mean molar liquid-to-vapor ratio of the section below the product side-stream extraction.

Regarding the pressure the actual input used is the logarithm of the top pressure, and this nonlinear transformation is chosen assuming a relationship between boiling temperature and pressure as for the Antoine’s law. The mean molar liquid-to-vapor ratio, which is computable with a higher or lower accuracy depending on the process scheme of the column (e.g. the presence of unmeasured external flows entering the column) and on the available measurements (e.g. missing temperature of subcooled reflux), is considered as input to improve the estimator consistency, as shown in the next paragraph.

Table 1. Naphtha T95: estimator comparison

ID	Inputs	EV
E0	T_2	82.3
E1	T_2, T_3, T_{23}	92.3
E2	T_2, T_3, p	96.6
E3	T_2, T_3, T_{23} $p, L/V$	97.2

Table 2. LGO T95: estimator comparison

ID	Inputs	EV
E0	T_{34}	49.9
E1	T_{35}, T_{36}, T_{S10}	97.4
E2	T_{35}, T_{36}, p	98.2
E3	T_{35}, T_{36}, T_{S10} $p, L/V$	99.1

4.2 *Accuracy and consistency evaluation*

The four candidate estimators (i.e. E0, E1, E2 and E3) are first evaluated in terms of accuracy in fitting the training data. The inputs used by each estimator of T95 for naphtha and LGO and the corresponding explained variance are reported in Tables 1 and 2, respectively (the corresponding estimators for kerosene and HGO T95 are not shown for space limitations). Moreover, comparisons of E0 and E2 in the training data fitting of T95 for naphtha and LGO are reported in Fig. 3 and 4, respectively. From these results it is

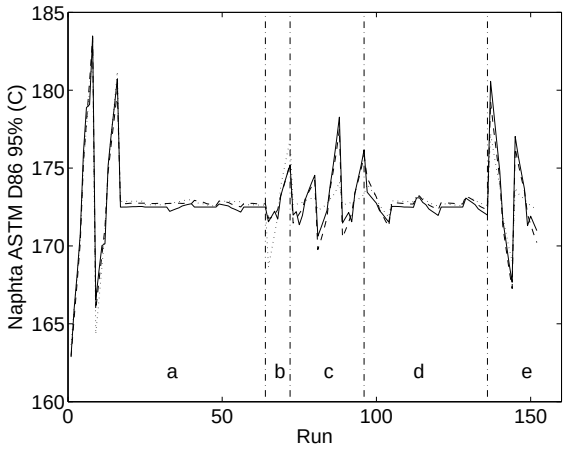

Fig. 3. Naphtha T95: training data fitting (– observ., ··· E0 estim., - - E2 estim.). Data groups: a, external flow changes; b, pressure changes; c, internal flow changes; d, stripper vapor changes; e, feed changes.

clear that an appropriate input selection can improve the estimator accuracy significantly, in particular for changes in column pressure and in internal flows (due to change of furnace temperature and pumparound duties).

Next, the four estimators are compared in terms of consistency for the rejection of several disturbances. A number of disturbances have been considered (Pastore, 2002) and results are presented for variations of the feed flow rate (d_F), of the bottom (d_{BPA})

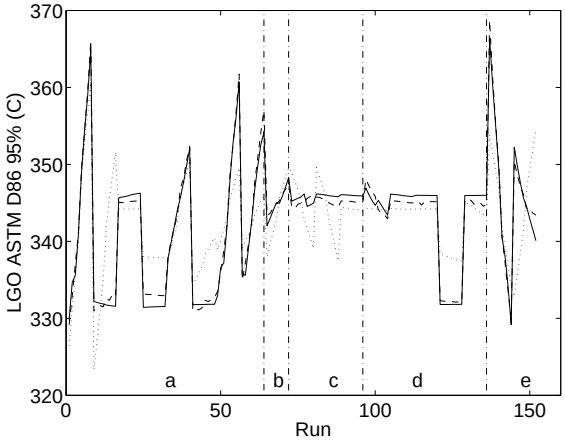

Fig. 4. LGO T95: training data fitting (see Fig. 3 for legend)

and top (d_{TPA}) pumparound duty, of the column pressure (d_P) and of the furnace outlet temperature (d_T). The disturbance consistency parameters for each estimator are shown in Tables 3-6, in which u_1 is the top temperature, u_2 is the kerosene flow rate, u_3 is the LGO flow rate, and u_4 is the HGO flow rate. These

Table 3. Disturbance consistency matrix ξ for E0

	d_F	d_{BPA}	d_{TPA}	d_P	d_T
u_1	0.52	-0.14	-1.10	2.19	-0.24
u_2	1.77	-0.24	-1.27	0.44	0.61
u_3	-4.16	7.15	0.56	3.09	32.18
u_4	19.25	186.16	1.14	3.33	21.44

Table 4. Disturbance consistency matrix ξ for E1

	d_F	d_{BPA}	d_{TPA}	d_P	d_T
u_1	3.35	9.76	-0.77	4.97	14.83
u_2	11.99	-102.20	2.42	4.95	247.65
u_3	55.36	-58.10	14.00	-17.54	-260.83
u_4	-254.66	-2622.21	56.87	-37.88	-373.32

Table 5. Disturbance consistency matrix ξ for E2

	d_F	d_{BPA}	d_{TPA}	d_P	d_T
u_1	0.95	16.97	4.16	5.49	-1.76
u_2	0.98	15.68	0.14	-0.7	-1.54
u_3	0.89	2.56	-1.77	0.92	0.24
u_4	1.05	-2.49	-4.36	0.88	0.59

Table 6. Disturbance consistency matrix ξ for E3

	d_F	d_{BPA}	d_{TPA}	d_P	d_T
u_1	0.98	0.57	1.19	1.03	0.86
u_2	0.85	0.18	1.40	1.01	-1.21
u_3	1.00	0.60	-0.31	0.89	1.4
u_4	1.48	1.80	-0.73	1.4	2.05

results show that E0 is not consistent for almost all disturbances and also that E1 is not very consistent,

even though the accuracy results previously shown suggested that E1 was a well designed estimator. Notice that in Tables 3 and 4 there are several negative parameters $\xi_{i,j}$ that indicate a manipulated variable change of opposite sign than the one actually required to remove offset. The most consistent estimator is E3 for which the parameters $\xi_{i,j}$ are quite close to 1 for almost all the disturbances. The estimator consistency parameters are also evaluated for setpoint changes, and results are shown in Tables 7-10. Notice that the upper part of the consistency matrix is not shown because the 4×4 product quality control system is nearly lower triangular. These results show that E0 and E1

Table 7. Setpoint change consistency matrix φ for E0

	r_1	r_2	r_3	r_4
u_1	0.69	-	-	-
u_2	0.55	0.60	-	-
u_3	-1.42	0.11	0.02	-
u_4	-0.67	-0.7	0.01	0.03

Table 8. Setpoint change consistency matrix φ for E1

	r_1	r_2	r_3	r_4
u_1	1.19	-	-	-
u_2	1.14	4.12	-	-
u_3	15.83	8.07	-11.95	-
u_4	1.30	1.55	1.21	2.97

Table 9. Setpoint change consistency matrix φ for E2

	r_1	r_2	r_3	r_4
u_1	1.16	-	-	-
u_2	1.67	2.10	-	-
u_3	1.86	1.14	2.15	-
u_4	5.677	6.64	5.19	12.22

Table 10. Setpoint change consistency matrix φ for E3

	r_1	r_2	r_3	r_4
u_1	1.00	-	-	-
u_2	1.05	1.27	-	-
u_3	1.30	0.94	1.44	-
u_4	1.12	1.34	1.07	1.63

are not very consistent even for setpoint changes and that E3 is very consistent for setpoint changes in any of the controlled variables.

5. CONCLUSIONS

The design of product quality estimators for refinery main fractionators has been considered, and the PLS regression technique has been used to choose the most appropriate auxiliary measurements. A crude distillation unit was chosen as case study, and several linear static estimators have been designed and compared both in terms of accuracy and consistency for disturbance rejection and/or setpoint changes. An estimator is consistent (Pannocchia and Brambilla, 2002) if it guarantees a low closed-loop offset when inserted in an inferential control scheme, and this property is not necessarily related to the estimator accuracy in fitting the data. It has been shown that the simple use of the temperatures always available on a CDU (i.e. the temperatures of the trays where the side products are drawn) does not lead to a well designed estimator particularly in terms of consistency. The PLS regression has been used to select more appropriate temperature measurement locations but additional measurements, as the operating pressure and the liquid-to-vapor molar ratio, were used to improve the estimator consistency significantly.

6. REFERENCES

de Jong, S. (1993). An alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* **18**, 251–263.

Joseph, B. and C. B. Brosilow (1978). Inferential control of processes. *AIChE J.* **24**, 485–509.

Kano, M., K. Miyazaki, S. Hasebe and I. Hoshimoto (2000). Inferential control system of distillation compositions using dynamic partial least squares regression. *J. Proc. Cont.* **10**, 157–166.

Kresta, J. V., T. E. Marlin and J. F. MacGregor (1994). Development of inferential process models using PLS. *Comput. Chem. Eng.* **18**, 597–611.

Mejdell, T. and S. Skogestad (1991). Estimation of distillation composition from multiple temperature measurements using partial-least squares regression. *Ind. Eng. Chem. Res.* **30**, 2543–2555.

Morari, M. and G. Stephanopoulos (1980). Optimal selection of secondary measurements with the framework of state estimation in the presence of persistent unknown disturbances. *AIChE J.* **26**, 247–258.

Pannocchia, G. and A. Brambilla (2002). Consistency of property estimators in multicomponent distillation control. Accepted for publication in *Ind. Eng. Chem. Res.*

Pastore, D. (2002). Estimator design for the product quality of complex columns (in italian). MS Thesis in Chemical Engineering, University of Pisa.

Wold, S., J. Trygg, A. Berglund and H. Antti (2001*a*). Some recent developments in PLS modeling. *Chemometrics and Int. Lab. Systems* **58**, 131–150.

Wold, S., M. Sjöström and L. Eriksson (2001*b*). PLS-regression: a basic tool of chemometrics. *Chemometrics and Int. Lab. Systems* **58**, 109–130.

Yu, C.-C. and W. L. Luyben (1987). Control of multicomponent distillation columns using rigorous composition estimators. In: *Distillation and Adsorption*. IChemE Symposium Series. The Institution of Chemical Engineers. Brighton, UK. pp. A29–A69.