Chem. Eng. Comm., 1999, Vol. 178, pp. 221–248
Reprints available directly from the publisher
Photocopying permitted by license only

DESIGN OF A COMPOSITION ESTIMATOR FOR INFERENTIAL CONTROL OF DISTILLATION COLUMNS*

JOONHO SHIN[a,†], MOONYONG LEE[b] and SUNWON PARK[a,‡]

[a] *Dept. of Chem. Eng., KAIST, Taejon 305-701, Korea;*
[b] *School of Chem. Eng. and Tech., Yeungnam Univ., Kyongsan 712-749, Korea*

(Received 10 October 1996; In final form 25 June 1999)

In distillation column control, secondary measurements such as temperatures and flows are widely used in order to infer product composition. This paper addresses the design of the linear static estimators using the secondary measurements for estimating product compositions of distillation columns. Based on the unified framework for the estimator design, the relationships among various static estimators are discussed in terms of the estimator structure. It is shown that the projection estimator is equivalent to the regression estimators in the special cases. Since the projection estimator heavily depends on the measured inputs such as reflux flow and heat input to the reboiler due to its structural characteristic, the estimation performance is far more sensitive to measurement noise and nonlinearity of them, compared with the regression estimators based on the PCR or PLS method. It is also found that the use of the measured inputs leads to performance deterioration of both the projection and regression estimators because of their nonlinear effects on the product compositions especially in high-purity columns. Design guidelines for the PCR and PLS estimators are presented by analyzing the results of the simulation studies on a high-purity column example. The estimator based on the guidelines is robust to sensor noise and has a good predictive power.

Keywords: Composition estimator; PLS(Partial-Least-Squares); PCR(Principal-Component-Regression); distillation column

1. INTRODUCTION

Typical production objective in distillation is to deliver products meeting certain composition specifications. Therefore economics in distillation

* Condensed version of this paper was presented at CPC-V, Tahoe City, CA in 1996.

†Current address: SK Engineering and Construction Co. Ltd. R & D Technology Center.
‡Corresponding author. Tel.: 82-42-869-3920, Fax: 82-42-869-3910, e-mail: swpark@convex.kaist.ac.kr

process heavily depend on composition control. However, product quality measurement has been one of the major difficulties associated with the composition control of distillation columns. Although on-line analyzers such as Gas Chromatography(GC) have the advantage of directly measuring the product quality, the composition control by the analyzers has not been preferred yet because the on-line analyzers still suffer from large measurement delays, high investment/maintenance costs and low reliability. One of common alternatives to the analyzers is to use a secondary measurement which is able to infer product composition. In fact, the composition control using a single tray temperature is perhaps the most popular means of controlling product composition. For the binary column at constant pressure, the temperature at the column end is an exact indicator of the product composition. However, for the high purity distillation columns, the use of a single temperature at the column end section is generally not reliable because the temperature variation is too small to distinguish from measurement noise. In order to improve the sensitivity, one may try to use other tray temperature some distance away from the column end section. But in this case, the temperature does not correlate well with the product composition because the temperature is strongly influenced not only by the corresponding product composition but also by the feed disturbances and the product composition at the other column end. For these reasons, many workers (Weber and Brosilow, 1972; Joseph and Brosilow, 1978; Mejdell and Skogestad, 1991; Kresta *et al.*, 1994; Piovoso and Kosanovich, 1994) have studied the inferential models using multiple secondary measurements and presented some promising results.

In developing the composition estimator using the secondary measurements, we have to consider the following issues: (1) the selection of the estimator type; (2) the determination of the number of factors if the PCR or PLS regression is used; (3) the selection of the secondary measurements to be used; (4) the selection of the most effective variable transformation and scaling. The first step for the construction of an inferential model is to select the estimator type. In this work, two different approaches, *i.e.*, the projection type estimator (Weber and Brosilow, 1972; Joseph and Brosilow, 1978) and the regression type estimator (Kresta *et al.*, 1994; Mejdell and Skogestad, 1991; Piovoso and Kosanovich, 1994), will be concerned. In the regression type estimators using the PCR or PLS method, the determination of the number of factors (or components) is very important because it is closely related to the prediction ability and the sensitivity (or robustness) of the inferential model. The measurement selection including the measured inputs can largely affect the estimation performance especially depending on

the estimator type. Additional design problem is to find the effective transformation and scaling of the process variables in order to enhance the performance of the estimator.

The design problems mentioned above will be discussed and several design guidelines will be presented. The rest of this paper is structured as follows: In the next section, the estimator problem is presented, followed by a brief review of the projection estimator and the regression estimators. Next, the relationship between the two types of the estimators is discussed including the effect of the measured inputs. The estimator design issues are also discussed by analyzing the results of simulation studies on a high-purity binary column example, and finally conclusions are drawn.

2. STATIC ESTIMATOR DESIGN PROBLEM

Since the process outputs are the functions of the process inputs, the measured outputs θ and the unmeasured outputs $\mathbf{c}$ can be expressed in terms of the measured inputs $\mathbf{m}$ and the unmeasured inputs $\mathbf{u}$ as follows:

$$\theta = \mathbf{f}_\theta(\mathbf{u}, \mathbf{m}) \tag{1}$$

$$\mathbf{c} = \mathbf{f}_\mathbf{c}(\mathbf{u}, \mathbf{m}) \tag{2}$$

where $\mathbf{f}_\theta$ and $\mathbf{f}_\mathbf{c}$ are nonlinear vector functions.

Through this paper, we mean the term *inputs* as the variables that affect a process independently while the term *outputs* as the variables that are dependently determined by the inputs. The total number of inputs for a system is equal to the number of degrees of freedom of the system. Theoretically, we can freely select the input variables as far as the degrees of freedom are satisfied. However, in the case especially where only the actual operation data is available for the calibration set, the inputs and the outputs are fixed in accordance with the actual operation mode of the process.

Let us consider a simple binary column. The number of degrees of freedom for the system is seven and the typical inputs are feed flow rate F, feed temperature T_F, feed pressure P_F, feed composition z_F, and column pressure P. Remaining two input variables can be arbitrarily selected depending on the situation. For example, we can choose top product composition y_D and bottom product composition x_B, which we typically want to estimate, as the inputs in order to span all directions in the estimation variable space. On the other hand, if only the actual operation data is available for the calibration

set, we should select the inputs and the outputs based on the actual operation mode of the process. For example, in the column under manual mode, reflux ratio R and reboiler duty Q_B must be considered as the inputs while y_D and x_B as the outputs.

What we call the estimator design is to find an adequate function **g**, which is a kind of an inverse function of **f**, to estimate **c** and/or **u** by using the information of the secondary measurements θ and/or **m** as

$$\mathbf{u} = \mathbf{g}_u(\theta, \mathbf{m}) \tag{3}$$

$$\mathbf{c} = \mathbf{g}_c(\theta, \mathbf{m}) \tag{4}$$

Note that in general, to find a function **g** is much more difficult than to find **f**. One of the reasons is that the measured outputs θ are strongly correlated with each other.

If a process is assumed to behave linearly in certain operation range, Eqs. (1) and (2) can be described with the linear form as

$$\theta = \mathbf{F}_u\mathbf{u} + \mathbf{F}_m\mathbf{m} \tag{5}$$

$$\mathbf{c} = \mathbf{G}_u\mathbf{u} + \mathbf{G}_m\mathbf{m} \tag{6}$$

where $\mathbf{F}_u$, $\mathbf{F}_m$, $\mathbf{G}_u$, and $\mathbf{G}_m$ are process matrices with suitable dimension.

Furthermore, Eqs. (3) and (4) can also be described with the following linear form

$$\hat{\mathbf{y}} = \mathbf{K}\mathbf{x} \tag{7}$$

where the estimation variable vector $\hat{\mathbf{y}}$ consists of the unmeasured variables $\hat{\mathbf{u}}$ and/or $\hat{\mathbf{c}}$, and the secondary measurement vector **x** consists of the measured variables θ and/or **m**.

The constant matrix **K** will be referred to as the estimator. In this article, we will limit our scope to the linear static estimator case.

3. REVIEW OF TYPICAL ESTIMATORS

3.1. Projection Estimator (Brosilow Estimator)

The projection estimator (Weber and Brosilow, 1972; Joseph and Brosilow, 1978) will be reviewed in this section. In the projection estimator, the

unmeasured inputs are firstly estimated by seeking the function $\mathbf{g}_u$ from $\mathbf{f}_\theta$, and then the unmeasured outputs are indirectly estimated from $\mathbf{f}_c$. For the distillation column, they consider R and Q_B as the measured inputs and z_F as the unmeasured inputs while tray temperatures as the measured outputs and y_D and x_B as the unmeasured outputs. Assume that all other inputs are constant for the moment. When the number of unmeasured inputs is less than or equal to the number of measured outputs, the unmeasured disturbances are estimated from Eq. (5) as follows:

$$\hat{\mathbf{u}} = \mathbf{F}_u^{\dagger}(\theta - \mathbf{F}_m\mathbf{m}) \tag{8}$$

where $\mathbf{F}_u^{\dagger}$ is the left pseudoinverse of $\mathbf{F}_u$, *i.e.*, $\mathbf{F}_u^{\dagger} = (\mathbf{F}_u^{\mathrm{T}}\mathbf{F}_u)^{-1}\mathbf{F}_u^{\mathrm{T}}$.

The estimate of $\mathbf{u}$ is the best solution in the least-squares sense. By substituting the above equation into Eq. (6), the estimate of $\mathbf{c}$ are:

$$\hat{\mathbf{c}} = \mathbf{G}_u\mathbf{F}_u^{\dagger}(\theta - \mathbf{F}_m\mathbf{m}) + \mathbf{G}_m\mathbf{m} \tag{9}$$

Note that the effects of the measured inputs on the measured outputs are compensated and the compensated measured outputs are used in order to estimate the unmeasured inputs.

The above equations have two restrictions: (1) $\dim(\theta) \geq \dim(\mathbf{u})$ and (2) $\mathrm{rank}(\mathbf{F}_u) = \dim(\mathbf{u})$. Physical meaning of the second restriction is that every unmeasured disturbance should have different effect on the measured output. Otherwise, there exist a number of linear relationships between the unmeasured inputs corresponding to the difference between $\dim(\mathbf{u})$ and $\mathrm{rank}(\mathbf{F}_u)$. The first restriction had been handled by using the statistics of the disturbances and/or measurement noise (Joseph and Brosilow, 1978). When $\dim(\theta) < \dim(\mathbf{u})$, the estimator can be expressed by

$$\hat{\mathbf{c}} = (\Phi_{y\theta^T}\Phi_{\theta\theta}^{-1})(\theta - \mathbf{F}_m\mathbf{m}) + \mathbf{G}_m\mathbf{m} \tag{10}$$

where $\Phi_{\theta\theta} = E\{\theta\theta^T\} = \mathbf{F}_u\mathbf{F}_u^{\mathrm{T}}$, $\Phi_{c\theta} = E\{c\,\theta^T\} = \mathbf{G}_u\mathbf{F}_u^{\mathrm{T}}$ and $E\{\bullet\} \equiv$ the expected value of $\{\bullet\}$.

Note that the projection estimator can also be written in the same form as Eq. (9) by defining the matrix $\mathbf{F}_u^{\dagger}$ as the right pseudoinverse of $\mathbf{F}_u$, *i.e.*, $\mathbf{F}_u^{\mathrm{T}}(\mathbf{F}_u\mathbf{F}_u^{\mathrm{T}})^{-1}$. However, in this case, true $\mathbf{u}$ can not be exactly estimated by θ. The estimated unmeasured output vector from Eq. (10) is just the minimum norm solution among infinite numbers of input vectors to give the same response on the measured outputs. For this reason, it is very hard to expect good estimation performance with the estimator.

3.2. Regression Estimators

In the regression estimators, the unmeasured variables **u** and **c** are directly estimated by seeking the function **g** (*i.e.*, the matrix **K**) from the data matrices for the secondary measurements, $\mathbf{X}^{n\times q}$ and the estimation variables, $\mathbf{Y}^{n\times p}$. Note that in the projection estimator, every possible unmeasured disturbance should be included in **u** regardless of whether it actually needs to be estimated or not. On the other hand, in the regression estimators, only the variables which we want to estimate are selected for **Y** while **X** must contain information about all inputs affecting **x**.

3.2.1. Multiple Linear Regression (MLR)

The most popular method is called the MLR and it is well known as "least-squares method". The least-squares solution of **K** in Eq. (7) is given by

$$\mathbf{K}_{\mathbf{MLR}} = \mathbf{Y}^{\mathbf{T}}(\mathbf{X}^{\dagger})^{\mathbf{T}} \tag{11}$$

where $\mathbf{X}^{\dagger} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}$.

The most frequent problem in MLR is that the inverse of $(\mathbf{X}^{\mathbf{T}}\mathbf{X})$ may not exist or be almost singular due to collinearity of the measurements. It can cause a severe ill-condition problem when noise is not included in the calibration set, and give significant performance deterioration for the prediction in noisy environment. Selecting the independent measurements is a solution but we can use the collinear measurements effectively by using the PCR or PLS method. Nonlinearity and noise in the calibration set usually make the rank of **X** greater than input dimensions of the system and approach q. In this case, the error in the prediction may be very large because of the unnecessary incorporation of nonlinearity and random errors.

3.2.2. Principal Component Regression (PCR)

Before considering the PCR method, let's review singular value decomposition (SVD). The SVD of **X** may be written

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{\mathbf{T}} \tag{12}$$

or as a sum of r matrices of rank 1

$$\mathbf{X} = \mathbf{u}_1\sigma_1\mathbf{v}_1^{\mathbf{T}} + \mathbf{u}_2\sigma_2\mathbf{v}_2^{\mathbf{T}} + \cdots + \mathbf{u}_r\sigma_r\mathbf{v}_r^{\mathbf{T}} \tag{13}$$

where r is the rank of **X**.

Note that $\text{rank}(\mathbf{X}) = r \leq \min(n, q)$ and the matrix $(\mathbf{X}^T\mathbf{X})$ is singular if $r < \min(n, q)$. The pseudoinverse of $\mathbf{X}$ is obtained by

$$\mathbf{X}^\dagger = \mathbf{V}\Sigma^{-1}\mathbf{U}^T \tag{14}$$

Here $\Sigma^{-1} = \text{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_r^{-1})$. As mentioned by Mejdell and Skogestad (1991), the smallest singular value, σ_r becomes the largest in the pseudoinverse. The sensitivity of the pseudoinverse to small errors (*e.g.*, noise) in $\mathbf{X}$ may be large if the condition number

$$\gamma(\mathbf{X}) = \frac{\sigma_1}{\sigma_r} \tag{15}$$

is large. The SVD in Eq. (13) can be expressed as

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + \cdots + \mathbf{t}_r\mathbf{p}_r^T \tag{16}$$

where $\mathbf{t}_i = \mathbf{u}_i\sigma_i$ is the score vector (or latent variable) and $\mathbf{p}_i = \mathbf{v}_i$ is the loading vector for principal component i. In PCR, only the first k terms which may be distinguished from measurement noise are kept, and the matrices $\mathbf{P}^{q\times k}$ and $\mathbf{T}^{n\times k}$ include only these k most important directions. The overall estimator gain matrix, which we call principal component regression estimator, becomes

$$\mathbf{K}_{\mathbf{PCR}} = \mathbf{Y}^T\mathbf{T}(\mathbf{T}^T\mathbf{T})^{-1}\mathbf{P}^T \tag{17}$$

If all of the information in $\mathbf{X}$ not associated with random errors is necessary to predict the estimation variables, the PCR estimator should be quite satisfactory. However, in the PCR, the order of selection of singular vectors to be used does not have any relation to their significance for describing the estimation variables. They only describe in order of decreasing variance, the variance structure of $\mathbf{X}$. Therefore, for estimating a specific variable, the PCR uses all the singular vectors needed to describe the variation in the matrix of responses regardless of whether or not they contain any information relevant to the variable to be estimated.

3.2.3. Partial Least Squares Regression (PLS)

This method is a variation of the PCR which recently has become popular among analytical chemists (Geladi and Kowalski, 1986). The latent variables are determined in order to have the largest covariance with the

dependent variables. By this way, if a vector is not needed to describe a specific variable, the vector will not be used in the estimation of the variable. Therefore it is possible to describe a variable using fewer components than required by PCR. The PLS algorithm can be found in many articles (Geladi and Kowalski, 1986; Lorber *et al.*, 1987; Mejdell and Skogestad, 1991) and the estimator based on k factors is given by

$$\mathbf{K}_{\mathbf{PLS}} = \mathbf{Q}(\mathbf{P}^{\mathbf{T}}\mathbf{W})^{-1}\mathbf{W}^{\mathbf{T}} \tag{18}$$

where $\mathbf{Q}^{p\times k}$ is the loading matrix for the dependent variables (*e.g.*, the product compositions), $\mathbf{P}^{q\times k}$ is the loading matrix for the independent variables (*e.g.*, temperatures and flow rates), and $\mathbf{W}^{q\times k}$ is the matrix formed in the PLS algorithm.

4. RELATIONSHIP BETWEEN PROJECTION AND REGRESSION ESTIMATORS

4.1. Projection and MLR

Consider the case where (1) there is no measured inputs, *i.e.*, $\mathbf{m} = 0$ and $\mathbf{x} = \theta$, (2) $\dim(\theta) < \dim(\mathbf{u})$, (3) only the unmeasured outputs are estimated, *i.e.*, $\mathbf{y} = \mathbf{c}$, then the projection estimator can be written by

$$\hat{\mathbf{c}} = (\Phi_{\mathbf{c}\theta^{\mathrm{T}}}\Phi_{\theta\theta}^{-1})\theta \tag{19}$$

The covariance matrix $\Phi_{\theta\theta}$ and the cross correlation matrix $\Phi_{\mathbf{c}\theta}$ can be expressed in terms of Θ and $\mathbf{C}$ as

$$\Phi_{\theta\theta} = \mathbf{F}_{\mathbf{u}}\mathbf{F}_{\mathbf{u}}^{\mathbf{T}} = (1/n)\Theta^{\mathbf{T}}\Theta \tag{20}$$

$$\Phi_{\mathbf{c}\theta} = \mathbf{G}_{\mathbf{u}}\mathbf{F}_{\mathbf{u}}^{\mathbf{T}} = (1/n)\mathbf{C}^{\mathbf{T}}\Theta \tag{21}$$

Therefore, Eq. (19) can be expressed

$$\hat{\mathbf{y}} = \mathbf{Y}^{\mathbf{T}}\mathbf{X}(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{x} \tag{22}$$

Therefore, it is clear that in the case where $\dim(\theta) < \dim(\mathbf{u})$ and $\mathbf{m} = \mathbf{0}$, the MLR estimator is equivalent to the projection estimator, and thus it can not estimate the unmeasured variables exactly.

Since Eqs. (20) and (21) are still valid in the case where $\dim(\theta) \geq \dim(\mathbf{u})$, it is also clear that the inversion problem in the MLR estimator is inherent

in the case where $\dim(\theta) > \dim(\mathbf{u})$ because $\text{rank}(\mathbf{F_u}) = \text{rank}(\Theta) = \dim(\mathbf{u})$. One may try to use the PCR method in order to overcome the inversion problem instead of the MLR method. The number of factors used in the PCR method should be $k \leq \dim(\mathbf{u})$. If $k = \dim(\mathbf{u})$, the resulting estimator provides the exact estimate of the product composition and the estimator gain matrix using the PCR method is the same as the projection estimator using Eqs. (8) and (9). This result is obvious because the number of factors needed to describe a model is equal to the model dimensionality if the relationship between the variables is linear.

4.2. Projection and PCR or PLS

Firstly, let's consider the estimators which use only the measured outputs like tray temperatures as the secondary measurements. In this case, all inputs are considered as the unmeasured inputs regardless of whether they are actually measurable or not. Thus, Eqs. (5) and (6) can be expressed by augmenting the matrices $\mathbf{F_u}$, $\mathbf{F_m}$, $\mathbf{G_u}$, and $\mathbf{G_m}$:

$$\theta = \mathbf{F}\begin{bmatrix}\mathbf{u}\\ \mathbf{m}\end{bmatrix} \tag{23}$$

$$\mathbf{c} = \mathbf{G}\begin{bmatrix}\mathbf{u}\\ \mathbf{m}\end{bmatrix} \tag{24}$$

where $\mathbf{F} = [\mathbf{F_u} \cdots \mathbf{F_m}]$ and $\mathbf{G} = [\mathbf{G_u} \cdots \mathbf{G_m}]$.

If $\dim(\theta) \geq \dim(u) + \dim(m)$ and $\text{rank}(\mathbf{F}) = \dim(\mathbf{u}) + \dim(\mathbf{m})$, $\mathbf{u}$ and $\mathbf{m}$ can be estimated from Eq. (23) by

$$\begin{bmatrix}\hat{\mathbf{u}}\\ \hat{\mathbf{m}}\end{bmatrix} = \mathbf{F}^{\dagger}\theta \tag{25}$$

Therefore, the product compositions $\mathbf{c}$ are estimated by

$$\hat{\mathbf{c}} = \mathbf{K}_{\theta}\theta \tag{26}$$

where $\mathbf{K}_{\theta} = \mathbf{G}\mathbf{F}^{\dagger}$.

Note that the estimator using Eq. (26) is unique regardless of the estimation methods. Both the projection estimator and the PCR estimator of which number of factors are equal to system dimensionality give the same estimator gain matrix.

Now let's consider the estimators which use both the measured outputs and the measured inputs as the secondary measurements. The form of the estimator can be written by

$$\hat{\mathbf{c}} = \mathbf{K}_{\theta \mathbf{m}} \begin{bmatrix} \theta \\ \mathbf{m} \end{bmatrix} \tag{27}$$

The matrix $\mathbf{K}_{\theta \mathbf{m}}$ is composed as follows:

$$\mathbf{K}_{\theta \mathbf{m}} = [\mathbf{K}_1 \cdots \mathbf{K}_2] \tag{28}$$

The matrix $\mathbf{K}_1$ is for the measured outputs θ and the matrix $\mathbf{K}_2$ is for the measured inputs $\mathbf{m}$.

One can easily find that the matrix $\mathbf{K}_{\theta \mathbf{m}}$ is not unique (there are infinite number of sets for $\mathbf{K}_1$ and $\mathbf{K}_2$): since the matrix $\mathbf{K}_\theta$ in Eq. (26) is unique (Stewart, 1973), the following relation can be obtained from Eqs. (26) and (28).

$$\mathbf{K}_\theta \theta = \mathbf{K}_1 \theta + \mathbf{K}_2 \mathbf{m} \tag{29}$$

The above relation should be valid for arbitrary inputs. Thus, by substituting Eq. (5), we can get

$$(\mathbf{K}_\theta - \mathbf{K}_1)\mathbf{F}_\mathbf{u} = \mathbf{0} \tag{30}$$

$$(\mathbf{K}_\theta - \mathbf{K}_1)\mathbf{F}_\mathbf{m} - \mathbf{K}_2 = \mathbf{0} \tag{31}$$

Since dim (θ) is greater than dim ($\mathbf{u}$), it is clear that there are infinite number of sets for $\mathbf{K}_1$ which satisfy Eq. (30) for given $\mathbf{F}_\mathbf{u}$ and $\mathbf{K}_\theta$. The projection estimator in Eq. (9) is one of them.

In the projection estimator, $\mathbf{K}_1$ and $\mathbf{K}_2$ are obtained separately and $\mathbf{K}_1$ is the matrix by which the unmeasured outputs are estimated from the measured outputs compensated by the measured inputs. On the other hand, in the regression estimators $\mathbf{K}_1$ and $\mathbf{K}_2$ are obtained simultaneously and they are the matrices by which the unmeasured outputs are estimated from both the measured inputs and outputs. In the projection estimator, the measured outputs such as temperatures are used only for estimating the unmeasured inputs rather than the unmeasured outputs. The effects of the measured inputs on the unmeasured outputs is directly described in terms of the measured inputs. It is easy to understand if considering the extreme case where there is no unmeasured disturbances, *i.e.*, $\mathbf{u} = 0$. The estimation of the projection estimator is done by $\hat{\mathbf{c}} = \mathbf{G}_\mathbf{m}\mathbf{m}$. On the other hand, if we construct

the PCR or PLS estimators using both θ and $\mathbf{m}$ as the secondary measurements, the estimator use both θ and $\mathbf{m}$ even when $\mathbf{u} = 0$. This structural difference between the regression estimators and the projection estimator results in quite different estimation characteristics.

Since the projection estimator highly depends on the information of a few measured inputs, it is very sensitive to process noise in measured inputs. Furthermore, if the measured inputs have nonlinearities and/or different dynamic characteristics from the unmeasured outputs to be estimated, which is often the case in high purity distillation columns, the estimation performance severely decreases. On the other hand, since the regression estimators such as PCR and PLS equally use both the measured outputs and inputs, they are relatively insensitive to the process noise and the nonlinearities of the measured inputs.

The projection estimator is valid under the condition that all of the related inputs and outputs are linear. Therefore, the projection estimator suffers from the modeling error if any of the variable has a nonlinear relationship. Even though a variable can have linear relationships with other variables by proper transformations, the resulting estimator may not have a linear form like Eqs. (5) and (6) unless the transformation between the measured output and the input is same as that between the unmeasured output and the input. On the other hand, the performance of the PCR or PLS estimator does not critically depend on the nonlinearity between the variables as far as the relationship between the variables which we want to estimate and the secondary measurements selected for estimation is linear.

The comparisons of two approaches will be presented by considering a linear case study in the later section.

5. EXAMPLE

5.1. Process Description

The steady state simulation of the binary column of normal-hexane and cyclo-hexane with 40 theoretical stages(including reboiler) was performed using the rigorous steady state simulator, Aspen-Plus^{TM}. The feed stream enters the column at stage 20 as saturated liquid. The nominal operation conditions and the variation of the inputs of the binary column are given in Table I. In the simulation, the reflux ratio and the heat duty of the reboiler are used as the manipulated variables. The composition of the feed is considered as a unknown disturbance.

TABLE I Simulation conditions for the binary distillation column

	Base case conditions	*Variation in steady state reference set*
F	1000.0 kmol/hr	Constant
T_F	346.57 K	Constant
z_F	0.5	$0.25 \sim 0.75$
D	500.0 kmol/hr	Varied
Q_B	0.130736×10^8 cal/sec	Varied
P	1 atm	Constant
y_D	0.99	$0.97 \sim 0.997$
x_B	0.01	$0.003 \sim 0.03$

5.2. Variable Transformation and Scaling Method

Generally, the product composition and tray temperatures are nonlinear functions of the operating variables (*e.g.*, reflux ratio R, heat duty of reboiler Q_B *etc.*). Since the use of transformed variable is a way to deal with nonlinearity, various transformations have been studied. It has been reported that logarithmic transformation of the product compositions can cope well with the nonlinearity between the product compositions and the operation variables (Joseph and Brosilow, 1978; Mejdell and Skogestad, 1991; Kresta *et al.*, 1994). For the binary mixture, the following logarithmic transformation on the distillate composition y_D is used

$$Y_D = \ln\left(\frac{y_D}{1 - y_D}\right) \tag{32}$$

Mejdell and Skogestad (1991) intensively tested several transformation methods and proposed that the logarithmic transformation of both the composition and the temperatures improved the estimation performance by linearizing the product composition and the tray temperature profile. In their method, the tray temperatures are transformed by

$$L_{T,i} = \ln\left(\frac{T_i - T_L^b}{T_H^b - T_i}\right) \tag{33}$$

where T_L^b and T_H^b are the boiling temperatures of pure light and heavy components, respectively.

In this work, several transformation methods are adopted and evaluated to select the most effective variable transformation method.

The PCR and PLS method is scaling dependent as SVD. The most common approach is to scale all variables to unit variance. Unit variance

scaling is done by

$$W_E = 1/S_D \tag{34}$$

where S_D is the standard deviation of the process variable. In this work, only the unit variance scaling method is considered and the effect of the scaling will be explained in terms of the collinearity: By taking unit variance scaling, collinear variables have the same values. It will be explained how the unit variance scaling enhances the performance of the estimator in the binary column example.

5.3. Evaluation Criteria

A cross validation procedure is adopted by splitting the data into two parts: the estimation data and the prediction data. The estimation data are used to build the regression model, and the prediction data are then used to study the predictive ability of the model. The Prediction Error Sum of Squares (PRESS) (Montgomery, 1992) is used to evaluate the absolute performance:

$$\text{PRESS} = \sum_{i=1}^{N_{\text{set}}} (\hat{y}_{D,i} - y_{D,i})^2 \tag{35}$$

The Explained Prediction Variance (EPV) (Mejdell and Skogestad, 1991) in percent is also used to evaluate the performance of the estimator:

$$\text{EPV}(k) = 100 \times \left(1 - \frac{\text{MSEP}(k)}{\text{MSEP}(0)}\right) \tag{36}$$

where the mean squared error of prediction $\text{MSEP}(k) = 1/N_{\text{set}} \sum_{i=1}^{N_{\text{set}}} (\hat{y}_i(k) - y_i)^2$ and N_{set} is the number of data sets (here, $N_{\text{set}} = 64$). The EPV measures in an approximate sense how much of the variability in new observations the model might be expected to explain.

6. RESULTS OF CASE STUDIES

If a system linear, the performance characteristics of each estimator depends only on the inherent property due to the estimator structure. However, because of nonlinearity, the estimation performance in the actual case is affected by the combined effect due to the two factors. This usually makes the analysis more difficult. In order to distinguish the effect of the estimator

structure from that of system nonlinearity, the analysis were performed both for the linear and nonlinear cases.

In order to capture the effect of the estimator structure on the performance, a linear model was obtained by the following procedure: (1) The data were obtained from the simulation of the binary distillation column. (2) The data were mean-centered and unit variance scaled. (3) A linearized model $\mathbf{F}_u$, $\mathbf{F}_m$, $\mathbf{G}_u$, and $\mathbf{G}_m$ in Eqs. (5) and (6) were obtained by the PCR method. (4) Based on the linear model, both unmeasured and measured inputs were randomly varied in order to obtain the calibration set for the regression estimators. The tray temperatures ($T_i, i = 1, \ldots, 40$) were used as the measured outputs. The reflux ratio R and the reboiler duty Q_B were used as the measured inputs.

6.1. Linear Case

6.1.1. Effect of Measured Inputs

When both the measured outputs and the measured inputs are used as the secondary measurements, the projection estimator is not equivalent to the regression estimators (PCR and PLS) as analyzed in the previous section. The model coefficients of the tray temperatures for the top product composition are shown in Figure 1 (see Eq. (28)). The model coefficients of

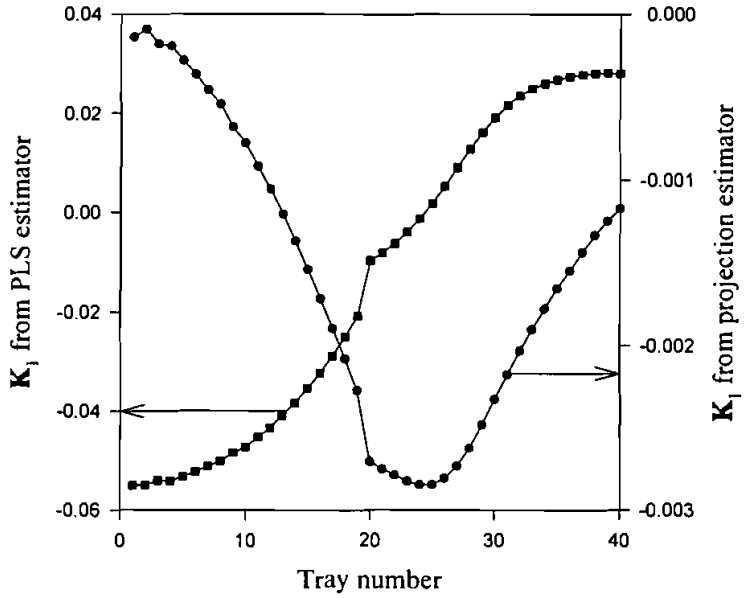

FIGURE 1 Comparison of the model coefficients of the projection and PLS estimator.

R and Q_B for the top product composition are also obtained as:

$$\mathbf{K}_2 = [-0.1604 \quad 0.9455] \quad \text{for the projection estimator}$$
$$\mathbf{K}_2 = [0.0599 \quad 0.0680] \quad \text{for the PLS estimator}$$

In the projection estimator, the effects of the measured inputs on the product composition are described by only the measured inputs. Therefore, the magnitude of the elements in matrix $\mathbf{K}_2$ from the projection estimator is much larger than those from the PLS estimator while the magnitude of the elements in matrix $\mathbf{K}_1$ from the projection estimator is much smaller than those from the PLS estimator. Furthermore, in the projection estimator the magnitude of the elements in $\mathbf{K}_2$ is much larger than those in $\mathbf{K}_1$ while the elements in $\mathbf{K}_1$ and $\mathbf{K}_2$ from the PLS have the same order of magnitude.

Figure 2 shows the performances of the projection and PLS estimators when the standard deviation of the noise in the measured inputs is 10% of the signal. The dependency of the estimation on the measured inputs is so high that the projection estimator is much more sensitive to the noise in the measured inputs than the regression estimator.

Additionally, this dependency may cause the degradation of the estimation performance in the real situation. The measured inputs R and Q_B

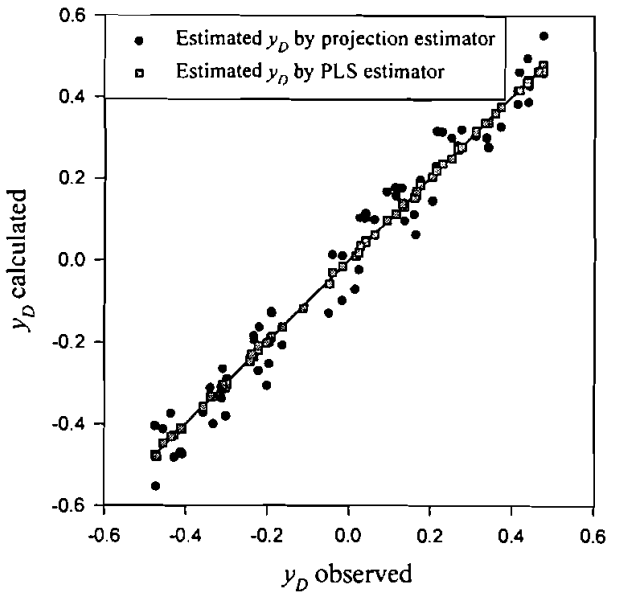

FIGURE 2 Comparison of the predictions of the top product composition using the projection and PLS estimator when the noise is present in the measured inputs.

have different dynamics from those of the product composition while the measured outputs the tray temperatures have very similar dynamics (Mejdell and Skogestad, 1993). For these reasons, the use of the measured inputs is not desirable in a static estimator even if the system is perfectly linear. This is more critical in the projection estimator because of heavy dependency on the measured inputs.

6.1.2. Equivalency of the Estimators

The projection estimator using the measured outputs only is equivalent to the PCR estimator with the number of factors $k = \dim(\mathbf{u}) + \dim(\mathbf{m})$ as mentioned above. The model coefficients for the linear case are shown in Figure 3. Note that the magnitude of the elements from the estimator matrix is similar with those from $\mathbf{K}_2$ using both the measured inputs and the measured outputs.

When several unmeasured disturbances have the same or similar effects on the measured outputs, the two estimators do not have equivalency any more. In that case, the projection estimator has a sensitivity or an inversion problem due to collinear inputs. A related topic is the selection of sensor locations. One may try to select the locations of the sensors in order to reduce

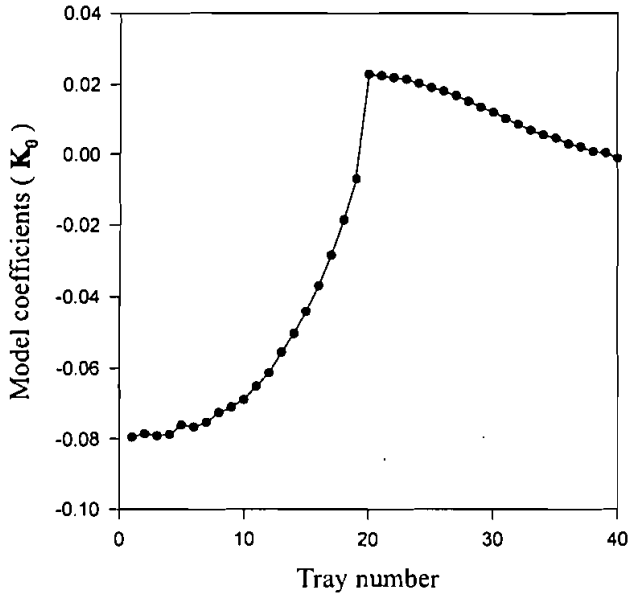

FIGURE 3 Model coefficients of estimator using the secondary measurements only.

the sensitivity and to distinguish the effects of inputs on the secondary measurements. The optimum selection of the sensor locations has been studied by many authors (Weber and Brosilow, 1972; Joseph and Brosilow, 1978; Morari and Stephanopoulos, 1980; Moore, 1986). Even though a proper selection of the sensor locations are taken, the same problem may still remain. To avoid the problem, one may use the pseudoinverse of $\mathbf{F_u}$ by deleting directions with singular values equal or close to zero (Mejdell and Skogestad, 1993). In the PCR or PLS method, this problem can be easily handled by selecting the number of factors as $k < \dim(\mathbf{u}) + \dim(\mathbf{m})$.

6.2. Nonlinear Case

In actual cases, in addition to the effect of the estimator structure, the nonlinearity effect should also be taken into account on the development of the estimator. Design guidelines will be presented by analyzing the binary column case study in this section. The results are based on all tray temperature measurements.

6.2.1. Determination of the Number of Factors in the PLS Estimator

There is an optimal model dimension (here the number of factors in the PLS estimator) which minimizes the prediction error (Kresta *et al.*, 1994). In a linear case, the optimal number of factors is clearly less than or equal to the system dimensionality (or the degrees of freedom) as shown in the previous section. In the actual case, extra factors may be needed to describe the nonlinearity. However, since the extra components essentially lead to performance deterioration in the presence of noise measurements, we should determine the number of factors by considering system nonlinearity and expected noise level simultaneously.

Since four input variables (*i.e.*, F, T_F, P_F, and P) are assumed to be constant, the number of degrees of freedom is three for our example column. Therefore, the column is fully defined by specifying three independent variables such as z_F, y_D, and x_B, or by specifying z_F, R, and Q_B. The temperature profile is also uniquely determined by specifying three independent variables, and thus dominant directions in the profile are clearly three as far as the independent variables have different effects on the temperatures. To get insights about the directionality of the temperature profile, let's consider the following inverse estimation problem:

$$\hat{\theta} = \mathbf{F_u} [z_F \;\; y_D \;\; x_B]^T \tag{37}$$

The above equation is from Eq. (5) by considering z_F, y_D, and x_B as unmeasured inputs. From the above equation, one can see that the temperature space is completely spanned by the three input variables. The projection estimator for the estimation of the compositions is:

$$[\hat{z}_F \;\; \hat{y}_D \;\; \hat{x}_B]^T = \mathbf{K}_{\mathbf{proj}}\,\theta \tag{38}$$

where $\mathbf{K}_{\mathbf{proj}} = \mathbf{F}_{\mathbf{u}}^{\dagger}$.

Matrix $\mathbf{F}_{\mathbf{u}}$ is obtained in the least squares manner with unit variance scaling. The EPV for temperature is 93.87. The relatively good EPV value shows that the temperatures can be properly estimated from z_F, y_D, and x_B in the linear form. Figure 4(a) compares matrix $\mathbf{K}_{\mathbf{proj}}$ and matrix $\mathbf{K}_{\mathbf{PLS}}$ with three factors. As expected, two matrices are almost identical. Since the dominant directions of the temperature profile exactly reflect the dominant combined effects of the input variables z_F, y_D, and x_B on the temperatures, we can expect that if the system follows a linear relationship well, the SVD (Singular Value Decomposition) of $\mathbf{F}_{\mathbf{u}}$ should give us the same eigenvectors as the loading vectors of Θ which is the matrix of the temperatures for calibration. The SVD of matrix $\mathbf{F}_{\mathbf{u}}$ is performed and the largest three eigenvectors ($\mathbf{p}_{\mathbf{F1}}$, $\mathbf{p}_{\mathbf{F2}}$, and $\mathbf{p}_{\mathbf{F3}}$) are plotted with loading vectors ($\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$) from the PLS in Figure 4(b). The figure also demonstrates that input variables z_F, y_D, and x_B have linear relationships with tray temperatures.

The inverse estimation can be used for selecting the locations of temperature measurements and the measured inputs. One can select the temperatures which show good EPV values with sufficiently high gain values in $\mathbf{F}_{\mathbf{u}}$ when they are inversely estimated with the inputs. One can also determine the use of measured inputs by using R and Q_B instead of y_D and x_B to estimate tray temperatures. In our example, the EPV using z_F, R, and Q_B is only 50.4. The poor performance of temperature estimation shows that the directions of the temperature profile cannot be explained in terms of z_F, R, and Q_B in the linear form.

From the eigenvalues of the data matrices, we can extract some useful information for the optimal number of factors. The eigenvalues of $\Theta^T\mathbf{Y}\mathbf{Y}^T\Theta$ for the PLS are plotted in Figure 5(a) where $\mathbf{Y}$ is the matrix of product composition y_D for calibration runs. The magnitude of eigenvalues drastically decrease as the number of factors increases. One can easily find that the condition number of $\Theta^T\mathbf{Y}\mathbf{Y}^T\Theta$ is directly related to the sensitivity of the PLS estimator with corresponding factors. One should select the first three factors because the condition number increases by the order of 10^3 if the number of factors increases from 3 to 4. Figure 5(b) compares EPVs for

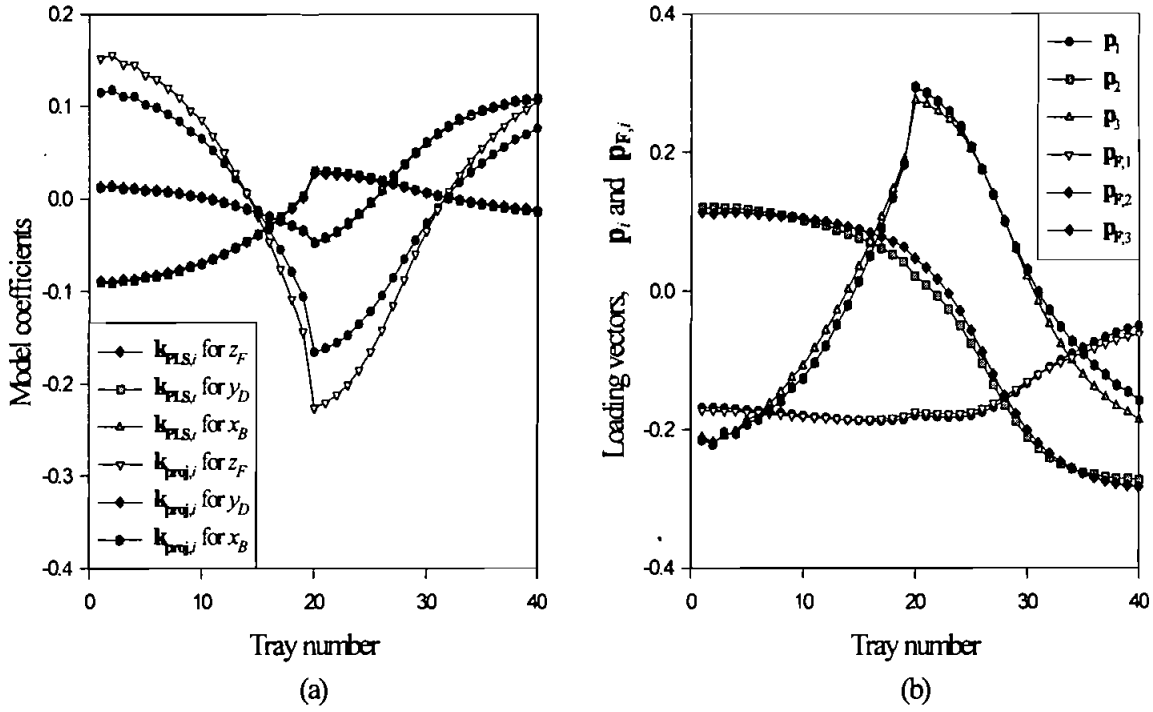

FIGURE 4 Comparison of (a) $\mathbf{K}_{PLS}$ and $\mathbf{K}_{proj}$ and (b) loading vectors of $\mathbf{F}_u$ and $\mathbf{K}_{PLS}$.

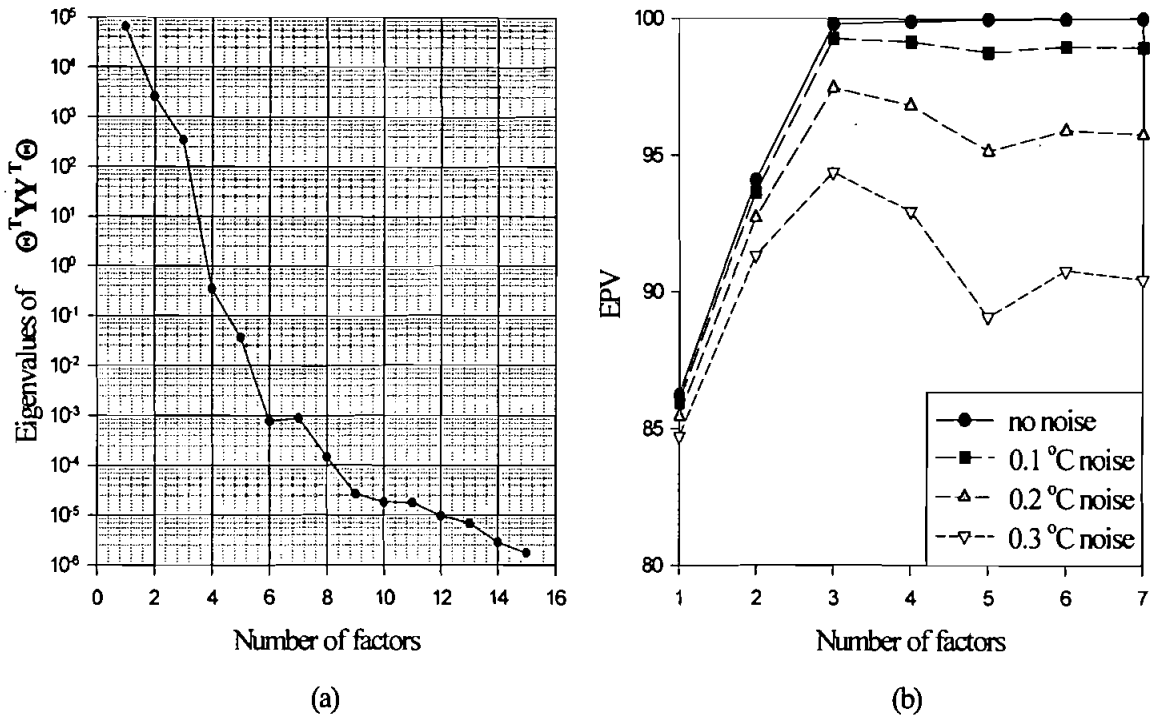

FIGURE 5 (a) Plot of eigenvalues of $\Theta^T YY^T\Theta$ and (b) Effect of measurement noise on EPV.

different levels of measurement noise on the temperatures in the prediction set. As shown in Figure 5(b), the EPV increases monotonically as the number of factors increases. However, when the inferential model with more than 4 factors is applied to noise corrupted data, the EPV decreases. It means that the positive effect of describing nonlinearity by increasing the number of factors is overshadowed by the negative effect due to noise corruption.

6.2.2. Effects of Variable Scaling and Transformation

The main objective of scaling is to make the variables with different units to be on the same basis. However, even with the same unit, the scaling has effects on the performance of the estimator by giving different weight factors. Unit variance scaling makes the variables with small change more important by large weighting. Therefore, if the secondary measurements with small changes have closer linear relationships with the variables for estimation than those with large changes, unit-variance scaling gives better performance than no scaling. In high-purity distillation columns, the relationship between the product composition and the tray temperature is more linear toward the end of the column. Furthermore the change of the temperatures is small toward the end of the column, thus the estimator with unit scaling can provide good performance in the noise-free condition. But since the noise is also large in relative terms close to the ends, there should be a break even point between the performance improvement due to weighting effect on linear measurements and the performance deterioration due to noise amplification effect.

Another tip by unit variance scaling is that it makes all collinear measurements have the same value. By collinear variables, we mean the variables which are not only linear with each other but also affected by every disturbance in the same manner. Collinear variables actually provide the same information. For example, let's consider a simple case. The tray temperatures T_1 and T_2 are collinear, $T_1 = \alpha T_2 + \beta$. We can easily find that the mean-centered and unit variance scaled temperatures $\tilde{T}_1$ and $\tilde{T}_2$ have the same value, that is, $\tilde{T}_1 = \text{sign}(\alpha)\tilde{T}_2$.

In the linear case, only the variables which are not collinear have different kind of information. Additional measurements are redundant and play a noise averaging role. Therefore, scaling makes it easy to select the minimum independent measurements if needed. It can also be used for locating the single temperature measurement because we can see what secondary measurements are collinear to T_1 or T_N which is highly related with y_D or

x_B (note that $y_D = f(T_1)$ for the binary column). The plot of matrix $\mathbf{F_u}$ for input variables z_F, y_D and x_B is shown in Figure 6(a). One can easily see what tray temperatures are collinear to what inputs. However, it should be noted that the scaling has a side-effect of disregarding physical information such as sensitivity of the variable.

The performances of the inferential models based on the scaled variables are almost the same except for the case of the logarithmic transformation on the composition only. For $k = 3$, EPV = 98.048 and PRESS = 0.575×10^{-4} without variable transformation and scaling. When the transformation $\ln(y_D/(1-y_D))$ on y_D is applied, the relationship between $\ln(y_D/(1-y_D))$ and T_i is linear around the middle of the column. But the performance of the transformation is not good because the linear relationship between $\ln(y_D/(1-y_D))$ and T_i around the feed tray becomes distorted when the feed disturbances such as z_F is introduced. The logarithmic transformation on the composition is not desirable when the change of z_F is frequent. The performance of the logarithmic transformation is not good (EPV = 92.698, PRESS = 0.531×10^{-3} with $k = 3$.) because the change in z_F is very large for our column. The performance of the estimator is improved by using the logarithmic transformation both on the composition and the

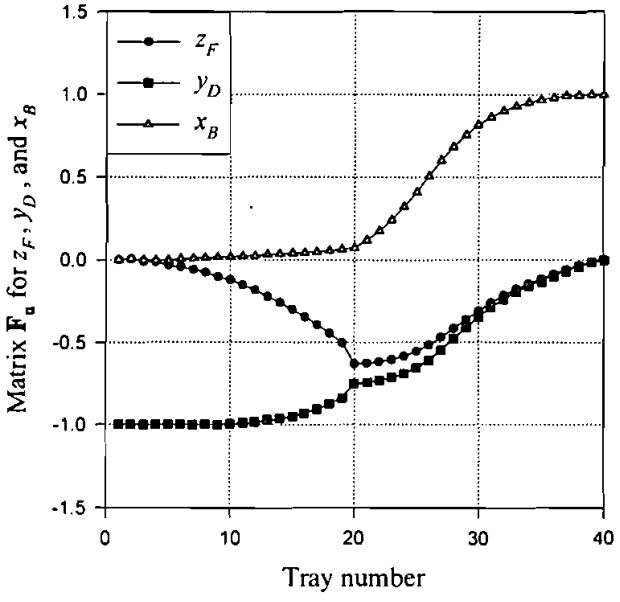

FIGURE 6 Plot of $\mathbf{F_u}$ for input variables z_F, y_D, and x_B.

temperature by $\ln(y_D/(1-y_D))$ and $\ln((T_i - T_L^b)/(T_H^b - T_i))$ (Mejdell and Skogestad, 1991). In addition to its linearizing effect, the above transformation gives more weighting to the temperatures toward the end of the column where the relationship is more linear. As a consequence, the EPV with no noise is close to 100% after only 3 factors due to both the linearizing and the weighting effect of the transformation (EPV = 99.943, PRESS = 0.958×10^{-6} with $k = 3$.). But the transformation becomes more sensitive to noise because of the weighting effect (*e.g.* $(T_i - T_L^b)/(T_H^b - T_i)$ term becomes zero or infinite at the end of the column). Estimation without variable transformation is best from the view points of prediction accuracy and robustness for our binary column.

6.2.3. Use of Other Secondary Measurements

The relationship between the auxiliary measured inputs (*e.g.*, R and Q_B) and the tray temperatures cannot be described in linear forms, as mentioned before. Therefore the estimation using the auxiliary measured inputs is not desirable for both types of estimators. Performance deterioration due to the use of the measured inputs is far more severe in the projection estimator than in the PCR or PLS estimator. The projection estimator using the measured inputs such as R and Q_B explicitly has been reported to be sensitive to noise and modeling error. This is due to several reasons. One of the reasons could be ill-conditioned property in the gain matrix of the measured inputs to the unmeasured outputs (Mejdell and Skogestad, 1993). This is often the case in the typical high-purity columns. The more important reason is high dependency on the measured inputs which are nonlinear to the product compositions and the temperatures. Furthermore, peculiar relation between Q_B and z_F makes the performance of the projection estimator poor.

Figure 7 shows the responses of D, L_1, R, Q_B, V_N, and H_F when only z_F is varied while the product compositions and the feed temperature are fixed at their nominal values are shown in Table I. From Figure 7, one can see why the projection estimator using the measured inputs shows poor performance. As can be seen, z_F can not be uniquely estimated with Q_B even in the nonlinear sense. The strange response is due to the effect of the total enthalpy of the feed. The total enthalpy of the feed is varied as shown in Figure 7 even though the feed temperature is fixed. It is also clear from the figure that secondary measurements L_1, R, and V_N are not desirable. Distillate flow D shows good linear relationship with z_F, it also is inadequate as the secondary measurement because y_D is very sensitive to D.

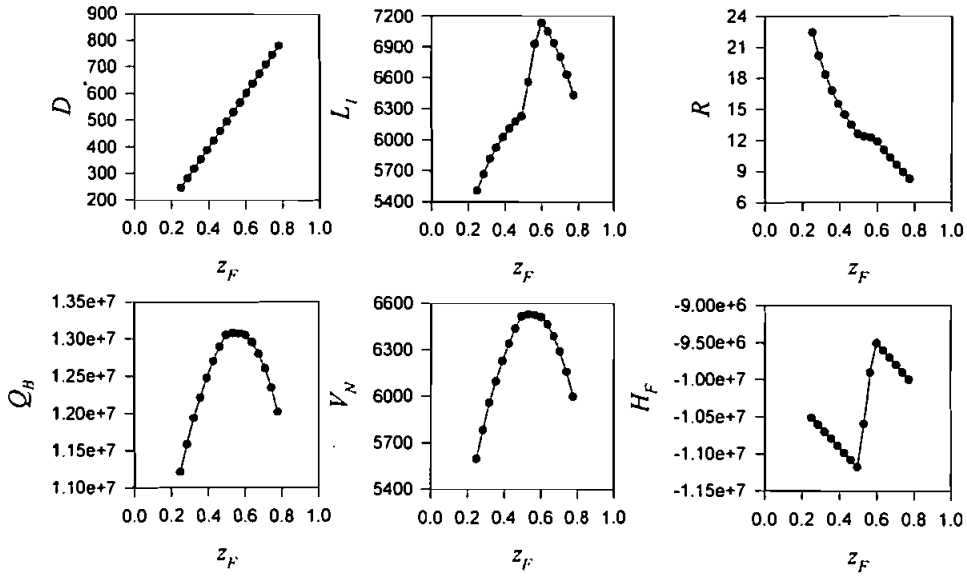

FIGURE 7 Effects of feed composition on various variables.

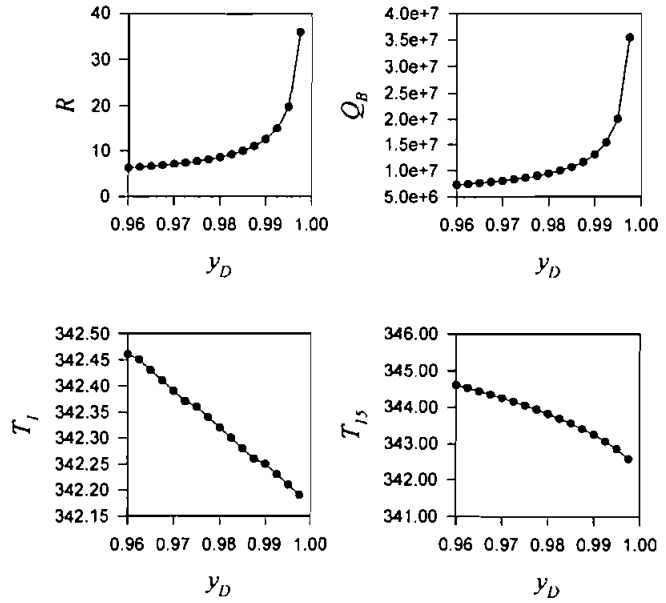

FIGURE 8 Effects of distillate composition on various variables.

For the regression estimators, the nonlinearities between the unmeasured input z_F and the secondary measurements do not cause any critical problem as far as the relationships between the product composition and the secondary measurements are linear. Figure 8 shows the plots of temperatures (here only T_1 and T_{15} are shown), R, and Q_B when only y_D is varied. The temperatures are close to linear with respect to y_D while R and Q_B are not. Thus we can expect that the estimation performance is not improved by adding the measured inputs as the secondary measurements. Actually the estimation performance of the PLS estimator decreases when we use R and Q_B in addition to all tray temperatures. The results using all tray temperatures, R, and Q_B with three factors were EPV = 99.119 and PRESS = 0.263×10^{-4} with unit variance scaling while EPV = 99.805 and PRESS = 0.588×10^{-5} were obtained when only the temperatures were used. Similar results were also reported by previous researchers (Mejdell and Skogestad, 1993; Piovoso and Kosanovich, 1994).

7. CONCLUSIONS

Two approaches (*i.e.*, projection estimator and regression estimator) for the design of the estimator have been discussed. The inversion problem in the

MLR estimator has been shown from the equivalency with the projection estimator in the special case. The projection estimator using measured outputs only is equivalent to the PCR estimator. When both the measured outputs and inputs are used as the secondary measurements, the projection estimator is not equivalent to the PCR estimator any more. In this case, the structural dependency on the measured inputs makes the two estimators have very different characteristics. It makes the projection estimator more sensitive to measurement noise, thus the projection estimator does not take the full benefit of using multiple measurements with a high degree of redundancy by the averaging effect of the process measurements noise. Furthermore since the relationship between the inputs and outputs are generally nonlinear, the high dependency on the inputs of the projection estimator leads to performance deterioration in the actual case. The control structure has no effects on the inferential model in the linear case but gives different results in the actual case due to nonlinearity.

Based on the analysis, the guidelines on the design of composition estimator *via* PLS have been presented: The recommended number of factors is equal to the number of degrees of freedom, *i.e.*, the number of independent variables (*e.g.*, z_F, y_D, x_B, and column pressure P) which uniquely affect the temperature profile. The additional factors do not consistently cope with nonlinearity while they lead to severe performance deterioration in the presence of noise. The relationships between the auxiliary measured inputs (R and Q_B) and the product composition cannot be described in the linear form. Thus, the estimation using the measured inputs is not desirable. Both variable transformation and scaling have effects on the performance of the estimator by giving a different weighting. The performance of the logarithmic transformation only on the composition is not good if the change of z_F is frequent. The transformation $\ln(y_D/(1 - y_D))$ and $\ln((T_i - T_L^b)/(T_H^b - T_i))$ is most effective but also somewhat sensitive to measurement noise. Unit variance scaling gives good insights into the collinearities among the measurements by making the collinear measurements with T_1 have the same information as T_1. It also enhances the estimation performance but makes the estimators sensitive to noise.

Acknowledgments

Financial support from the Korea Science and Engineering Foundation through the Automation Research Center at Postech is gratefully acknowledged. This research was also supported by Korea Research Foundation.

NOMENCLATURE

$\mathbf{c}$	vector of unmeasured outputs
$\mathbf{C}$	data matrix of unmeasured outputs
k	number of factors used in PCR or PLS estimator
$\mathbf{K}$	inferential model matrix
L_T	logarithmic temperature
$\mathbf{m}$	vector of measured inputs
R	reflux ratio
Q_B	heat duty of reboiler
$\mathbf{t}$	vector of latent variables
$\mathbf{T}$	matrix of latent variables
T_i	ith tray temperature
$\mathbf{u}$	vector of unmeasured inputs or disturbances
x_B	bottom product composition
y_D	distillate composition
Y_D	logarithmic distillate composition
$\mathbf{x}$	vector of secondary measurements
$\mathbf{X}$	data matrix of secondary measurements
$\mathbf{y}$	vector of outputs to be estimated
$\mathbf{Y}$	data matrix of outputs to be estimated
z_F	feed composition

Greek Symbols

γ	condition number
σ_i	singular value
θ	vector of measured outputs
Θ	data matrix of measured outputs
Φ	covariance matrix

Subscripts

H	heavy component
L	light component

Superscripts

b	boiling point of pure component
$\hat{}$	estimated variable
$\dagger$	pseudoinverse

References

Geladi, P. and Kowalski, B. R. (1986) "Partial least-squares regression: A tutorial", *Analytica Chimica Acta*, **185**(3), 1–17.

Weber, R. and Brosilow, C. B. (1972) "The Use of Secondary Measurements to Improve Control", *AIChE*, **18**(3), 614–623.

Joseph, B. and Brosilow, C. B. (1978) "Inferential control of processes: Part I. steady state analysis and design", *AIChE*, **24**(3), 485–492.

Kresta, J. V., Marlin, T. E. and MacGregor, J. F. (1994) "Development of inferential process models using PLS", *Computers Chem. Eng.*, **18**(7), 597–611.

Lorber, A., Wangen, L. E. and Kowalski, B. R. (1987) "A Theoretical Foundation For the PLS Algorithm", *J. Chemometrics*, **1**, 19–31.

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

Mejdell, T. and Skogestad, S., "Output estimation using multiple secondary measurements: High-purity distillation", *AIChE J.*, **39**, 1641–1653, October 1993.

Montgomery, D. C. and Peck, E. A. (1992) "Introduction to Linear Regression Analysis", John Wiley & Sons, Inc., New York.

Moore, C., Hackney, J. and Canter, D. (1986) "Selecting Sensor Location and Type for Multivariable Processes", The Shell Process Control Workshop, Prett, D. M. and Morari, M. Eds., Butterworths, Toronto.

Morari, M. and Stephanopoulos, G., "Optimal Selection of Secondary Measurements within the Framework of State Estimation in the Presence of Persistant Unknown Disturbances", *AIChE J.*, **26**(2), 247–260, March 1980.

Piovoso, M. J. and Kosanovich, K. A. (1994) "Application of multivariate statistical methods to process monitoring and controller design", *Int. J. Control*, **59**(3), 743–765.

Stewart, G. W. (1973) "Introduction to Matrix Computation", Academic Press, Inc., Orlando.