Contents lists available at ScienceDirect

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

The use of first principles model for evaluation of adaptive soft sensor for multicomponent distillation unit

Andrei Torgashov[a,*], *Sigurd Skogestad*[b]

[a] *Process Control Laboratory, Institute of Automation and Control Process FEB RAS, 5 Radio Str., Vladivostok, Russia*
[b] *Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, NO-7491, Norway*

ARTICLE INFO

Article history:
Received 30 November 2018
Received in revised form 19 August 2019
Accepted 26 August 2019
Available online 5 September 2019

Keywords:
Soft sensing
Process models
Distillation columns

ABSTRACT

Traditionally, soft sensors are developed based on measurement data only, but here we consider an adaptive soft sensor that uses data generated from a fitted, first principles model of the distillation columns. The contribution of the paper is a procedure for moving window soft sensor design that incorporates a priori knowledge, which is especially suitable when the training sample is small and contains measurement errors. In addition, we propose a continuous adaptation of all model parameters based on new data, instead of the usual procedure of only updating the bias. The accuracy of the predicted product quality is investigated by calculating the coefficient of determination and root mean squared error for the test sample. Several approaches were considered, and we found that a constrained optimization approach was superior. The constraints on the model parameters of soft sensors are derived from a fitted, rigorous distillation unit model. The improved estimator quality resulted in the successful industrial application of advanced process control systems.

1. Introduction

One efficient way for an industrial company to gain additional benefits or reduce production costs is to apply an advanced process control (APC) system. Soft sensors belong to the main functional part of the APC system's structure. The estimation of product quality via soft sensors (Funatsu, 2018; Kim et al., 2013) is an inexpensive and attractive technique in industrial automation. Improving soft sensor evaluation methods may be the subject of any innovative APC or real-time optimization platform (Amrit et al., 2015; Fayruzov et al., 2017).

In the present work, the sequence of industrial multicomponent distillation columns in the gas separation section of the fluidized catalytic cracking (FCC) unit is investigated. The widespread approach in industry for soft sensor evaluation is to use a data-driven approach (Fortuna et al., 2007; Kadlec et al., 2009, 2011). Most of them are based on the "black-box" concept with use of neural networks or statistical methods (Rogina et al., 2011; Kaneko and Funatsu, 2014; Kaneko and Funatsu, 2015; Shao and Tian, 2015; Xiong et al., 2017; Zhang et al., 2017). In practice, several obstacles are encountered, including a small training dataset, measurement errors, and an unsteady mode of distillation operation. Moreover, unmeasured feed-composition disturbances and variations in the efficiency of separation stages cause problems for obtaining a reliable training sample. Also, some of the key (informative) inputs have low variability ranges due to the operator seldom varying them (for instance, the set point for the top pressure of the distillation column). This leads to statistical insignificance (via t-test) of the corresponding coefficients in the regression model and may not be reconciled with the physico-chemical meaning of the parameters of soft sensors. A fitted rigorous multicomponent distillation model is helpful to overcome these difficulties. It allows the ranges of the soft sensor's parameter values to be estimated (Torgashov et al., 2016).

The difference from the previous research of the developed approach here is its using data from the first principal (rigorous) distillation model for the soft sensor design. One of its merits is that the rigorous model dataset may give a priori information on the soft sensor equation in terms of admissible parameter ranges, particularly for a

* *Corresponding author.*
E-mail addresses: torgashov@iacp.dvo.ru (A. Torgashov), skoge@ntnu.no (S. Skogestad).
https://doi.org/10.1016/j.cherd.2019.08.017

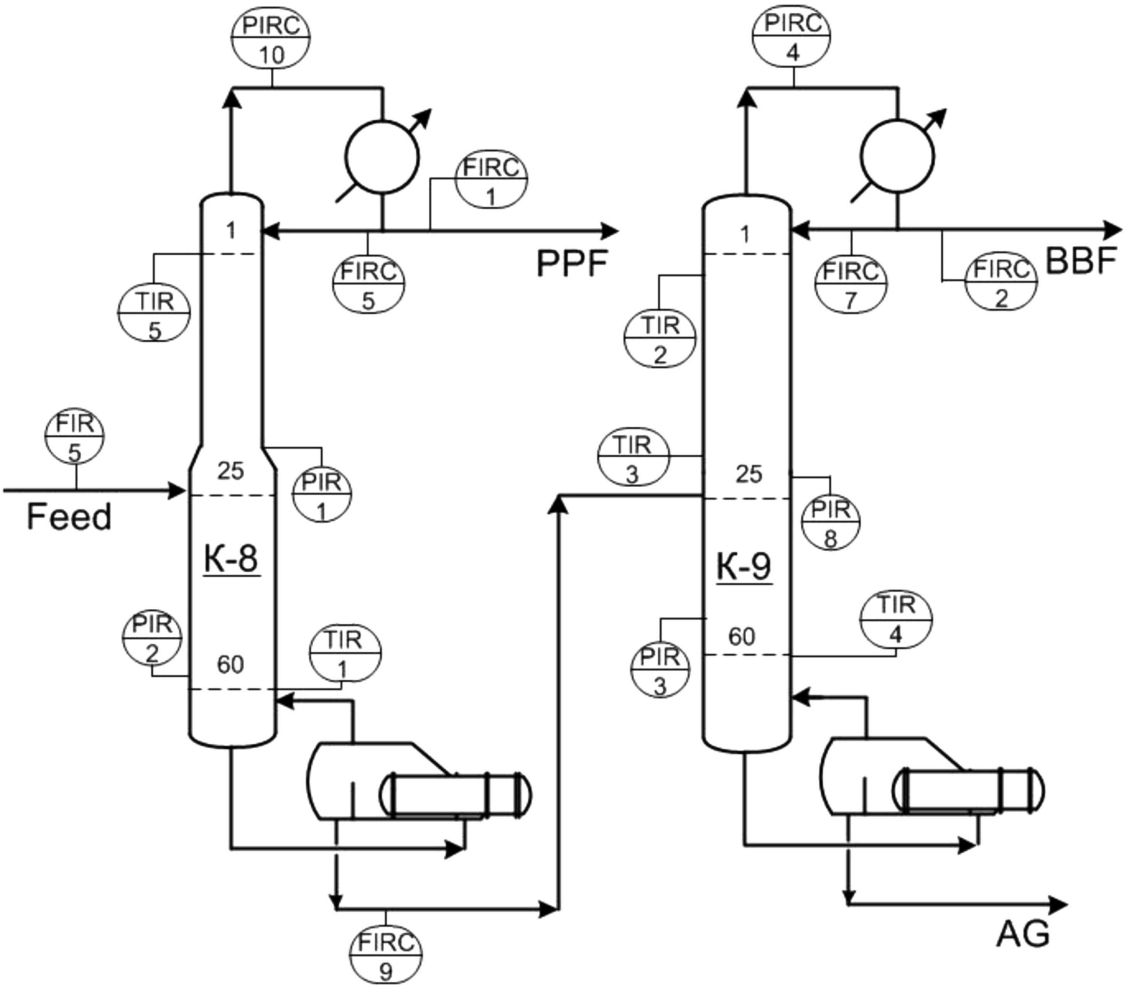

Fig. 1 – Industrial multicomponent distillation unit.

small industrial training sample. The accuracy of product quality predictions by soft sensors is investigated by calculating the coefficient of determination (R^2) and root mean squared error (*RMSE*) for the test sample.

It should be noted about an alternative approach for the estimation of compositions of products of chemical process systems with the use of observers or Kalman Filter based observers (Mohd Ali et al., 2015; Porru et al., 2013; Tronci et al., 2005). The observers are based on first principles models of chemical processes. To implement observer, it is necessary that the pair (A,C) be observable (for linearized plant) (Mohd Ali et al., 2015). However, as will be shown in this paper (Section 3), the multicomponent distillation process is unobservable. In addition, the feed composition in industrial conditions is not measured in real time, and this makes it even more difficult to use the rigorous model of the observer in practice, since the problem of evaluation an unknown input observer has no solution in the absence of observability of the plant. Also, the presence of nonstationarity and uncertainty of the efficiency values of the separation stages limits the use of observers in practice as well as in some cases the parameters of the phase equilibrium model of multicomponent mixture remain unknown.

The paper is organized as follows. The second section describes an industrial process and states the problem. The issues of observability analysis of multicomponent distillation are reported in Section 3 from a soft sensor evaluation point of view. The fitting procedure for a rigorous distillation model is presented in Section 4. A comparative analysis of the soft sensor parameters obtained from industrial and fitted rigorous model datasets is described in Section 5 as well as the constrained optimization approach for moving window (adaptive) soft sensor design is reported. A description of how the dataset was integrated from rigorous modelling into the soft sensor evaluation procedures based on the raw industrial data is given. In conclusion, the superiority of using both industrial and rigorous-model datasets for adaptive soft sensor design is summarized.

Table 1 – Main process variables.

Process variable	Notation	Soft sensor input variable
K-8 25th tray pres., kgf/cm^2	PIR_1	x_1
K-8 bot. pres., kgf/cm^2	PIR_2	x_2
K-8 bot. temp., °C	TIR_1	x_3
K-9 bot. pres., kgf/cm^2	PIR_3	x_4
K-9 top temp., °C	TIR_2	x_5
K-9 25th tray temp., °C	TIR_3	x_6
Sum of C_3 in BBF product, %	–	y

2. Industrial process description and statement of the problem

The gas-separation section of the FCC unit considered in this paper is represented by two multicomponent distillation columns, K-8 and K-9 (Fig. 1). The feed flow comes from the FCC absorption unit and enters on the 25th tray of K-8. The overhead product of K-8 is propane-propylene fraction (PPF). The butane-butylene fraction (BBF) is withdrawn from the top of K-9. Absorption gasoline (AG) is a residue of the gas-separation unit and is recycled in the absorption section of FCC unit. The main process variables of the industrial distillation unit are shown in Table 1 and may be considered as informative inputs of the soft sensor.

Table 2 – Feed composition and material balances of the multicomponent distillation unit.

No.	Component	Feed, kg/h	PPF, kg/h	BBF, kg/h	AG, kg/h
1	Ethane	23.47	23.47	0.0	0.0
2	Propylene	6837	6732	105	0.0
3	Propane	1629	1342	287	0.0
4	i-Butane	4610	233	4377	0.0
5	Butene-1	5051	50	5001	0.0
6	*n*-Butane	1043	0.0	1042	1.0
7	t-Butene-2	2611	0.0	2604	7.9
8	c-Butene-2	1876	0.0	1840	36
9	3-Methylbutene-1	95.9	0.0	0.0	95.9
10	i-Pentane	571	0.0	0.0	571
11	2-Methylbutene-1	182	0.0	0.0	182
12	*n*-Pentane	24.2	0.0	0.0	24.2
13	t-Pentene-2	71.8	0.0	0.0	71.8
14	c-Pentene-2	80.4	0.0	0.0	80.4
15	3.3-Dimeth.but.-1	45.1	0.0	0.0	45.1
16	2-Methylpentane	5.54	0.0	0.0	5.54
17	3-Methylpentane	1.97	0.0	0.0	1.97
18	*n*-Hexane	2.70	0.0	0.0	2.70
19	Benzene	0.89	0.0	0.0	0.89
20	2-Methylhexane	0.46	0.0	0.0	0.46

The material balances and feed composition for the nominal steady-state operating point are presented in Table 2. The mass balance in Table 2 for the nominal operating point is derived by averaging industrial data based on the available process statistics.

The main goal of this paper is to develop an approach to moving window soft-sensor design based on industrial data and rigorous modeling. It is necessary to overcome difficulties such as small training datasets and laboratory errors. The total concentrations of propylene and propane (C_3) in the BBF product is considered the soft sensor output.

3. Model analysis of multicomponent distillation for purpose of product quality estimation

The motivation for this analysis is to find out the limitations of the application of observers (an alternative approach to a soft sensor) for the process of multicomponent distillation. Therefore, the full observability will be analyzed. It will be shown that it is impossible to construct an observer or static estimator due to immeasurable feed composition disturbances and lack of full observability.

Traditional observability analysis considers whether all the process states are identifiable. However, in our case, only the top and bottom compositions (quality indicators) are interesting states for analysis. We consider a simplified model of the multicomponent distillation process for one column but cover important principles of the industrial unit.

$j = 1$:

$$\frac{d\tilde{x}_{1i}}{dt} = (R + D)(\tilde{y}_{2i} - \tilde{x}_{1i}); \quad (1)$$

$j = 2. \ldots f\text{-}1$:

$$\frac{d\tilde{x}_{ji}}{dt} = R(\tilde{x}_{j-1,i} - \tilde{x}_{ji}) + (R + D)(\tilde{y}_{j+1,i} - \tilde{y}_{ji}); \quad (2)$$

$j = f$:

$$\frac{d\tilde{x}_{ji}}{dt} = R(\tilde{x}_{j-1,i} - \tilde{x}_{ji}) + F(\tilde{x}_{F,i} - \tilde{x}_{ji}) + (R + D)(\tilde{y}_{j+1,i} - \tilde{y}_{ji}); \quad (3)$$

$j = f\text{+}1. \ldots K\text{-}1$:

$$\frac{d\tilde{x}_{ji}}{dt} = (R + F)(\tilde{x}_{j-1,i} - \tilde{x}_{ji}) + (R + D)(\tilde{y}_{j+1,i} - \tilde{y}_{ji}); \quad (4)$$

$j = K$:

$$\frac{d\tilde{x}_{Ki}}{dt} = (R + F)\tilde{x}_{K-1,i} - (F - D)\tilde{x}_{Ki} - (R + D)\tilde{y}_{Ki}, \quad (5)$$

where $i = 1 \ldots c$; c – number of separated compounds; K – total number of separation stages (trays); $\tilde{x}_{ji}$, $\tilde{y}_{ji}$ – concentrations of ith compound on the jth tray in the liquid and vapor phases, respectively;

$\tilde{y}_{ji} = E_j Y_{ji} + \sum_{p=j+1}^{K} E_p Y_{pi} \prod_{q=j}^{p-1}(1 - E_q)$ for $j = 2 \ldots K\text{-}1$; $\tilde{y}_{Ki} = E_K Y_{Ki}$;

$E_j = \frac{\tilde{y}_{ji} - \tilde{y}_{j+1,i}}{Y_{ji} - \tilde{y}_{j+1,i}}$ – the Murphree vapor phase tray efficiency, $\forall i \in [1; c]$; $Y_{ji} = \frac{\alpha_i \tilde{x}_{ji}}{\alpha_c + \sum_{i=1}^{c-1} \tilde{x}_{ji}(\alpha_i - \alpha_c)}$ – vapor composition under equilibrium condition; $\tilde{x}_{F,i}$ – concentration of ith compound in the feed (the feed is entered to the column in the liquid phase); R – reflux flowrate; F – feed flowrate; D – distillate flowrate; α – relative volatility.

The system of differential Eqs. (1)–(5) may be linearized and represented in the state-space form. The state vector is formulated as $\mathbf{x} = (\tilde{x}_1 \ldots \tilde{x}_1 \ldots \tilde{x}_{1,c-1}; \ldots \tilde{x}_{K,1} \ldots \tilde{x}_{K,i} \ldots \tilde{x}_{k,c-1})^T$.

The temperature on each stage j may be expressed via Antoine equation:

$$T_j = \frac{B_b + C_b lg\left(\frac{P_j}{\gamma_j}\right) - C_b A_b}{A_b - lg\left(\frac{P_j}{\gamma_j}\right)}$$

where $\gamma_j = \alpha_c + \sum_{i=1}^{c-1} \tilde{x}_{ji}(\alpha_i - \alpha_c)$; P_j – stage pressure; A_b, B_b and C_b – constants of Antoine equation for base compound.

The system (1)–(5) can be linearized in the neighborhood of the operating point and presented in the following continuous state-space form:

$$\begin{aligned}\dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u} + D\mathbf{d}\\ \mathbf{y} &= C\mathbf{x}, \quad \mathbf{z} = Z\mathbf{y}\\ \mathbf{h} &= H\mathbf{x},\end{aligned} \tag{6}$$

where $\mathbf{z}$ is available temperature measurements; $\mathbf{h}$ contains elements of state vector of top and bottom composition; Z and H are transforming matrices. The matrix $C = \left(\frac{\partial T_j}{\partial \tilde{x}_{k,i}}\right)_{k,i}^{K,K(c-1)}$ is obtained as

$$k = j: \ \frac{\partial T_j}{\partial \tilde{x}_{k,i}} = -\frac{(\alpha_i - \alpha_c)\left(T_j + C_b\right)}{\gamma_j\left[A_b - lg\left(\frac{P_j}{\gamma_j}\right)\right] ln10}$$

$$k \neq j: \ \frac{\partial T_j}{\partial \tilde{x}_{k,i}} = 0$$

The observability analysis of system (6) is proper to do for its analog in the discrete time because of digital implementation of control algorithms in industry:

$$\begin{aligned}\mathbf{x}_{k+1} &= A_\delta \mathbf{x}_k + B_\delta \mathbf{u}_k + D_\delta \mathbf{d}_k\\ \mathbf{y}_k &= C_\delta \mathbf{x}_k, \quad \mathbf{z}_k = Z\mathbf{y}_k\\ \mathbf{h}_k &= H\mathbf{x}_k.\end{aligned} \tag{7}$$

The following proposition defines the conditions of observability of multicomponent distillation process (with states related to product purity).

Proposition. The system (7) with state vector $\mathbf{h}_k$ is observable if:

1) the $rank\left(ZC_\delta|ZC_\delta A_\delta|\ldots|ZC_\delta A_\delta^{K(c-1)-1}\right)^T = K(c-1)$

2) the disturbance vector $\mathbf{d}_k$ is measurable.

The disturbance vector $\mathbf{d}$ is unknown and associated with the unmeasured feed composition fluctuations. Based on the vectors $\mathbf{u}$ and $\mathbf{d}$, the steady-state solution of $\mathbf{x}$ can be obtained as

$$\mathbf{h} = H(-A)^{-1}(B\mathbf{u} + D\mathbf{d}). \tag{8}$$

From Eq. (8), it follows that the solution of $\mathbf{h}$ is not unique, i.e., under fixed <H,A,B,**u**,D>, we have infinite numbers of $\mathbf{h}$ for the case of infinite numbers of values of unknown vector $\mathbf{d}$. This means that it is impossible to observe $\mathbf{h}$ via input $\mathbf{u}$.

From the other side, consider the inference of $\mathbf{h}$ via matrix C and vector of measurements y (temperatures). The equation for output y may be written via block-diagonal form of matrix C as

$$\begin{pmatrix} T_1 \\ -- \\ T_j \\ -- \\ T_K \end{pmatrix} = \begin{pmatrix} \mathbf{c}_1 & | & \mathbf{0} & | & \mathbf{0} \\ -- & \ddots & -- & -- & -- \\ \mathbf{0} & | & \mathbf{c}_j & | & \mathbf{0} \\ -- & | & -- & \ddots & -- \\ \mathbf{0} & | & \mathbf{0} & | & \mathbf{c}_K \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{x}}_1 \\ -- \\ \tilde{\mathbf{x}}_j \\ -- \\ \tilde{\mathbf{x}}_K \end{pmatrix},$$

where $\mathbf{c}_j = \left(\frac{\partial T_j}{\partial \tilde{x}_{j,1}} \cdots \frac{\partial T_j}{\partial \tilde{x}_{j,c-1}}\right)$ – single string matrix; $\tilde{\mathbf{x}}_j = \left(\tilde{x}_{j,1} \ldots \tilde{x}_{j,c-1}\right)^T$ – vector of liquid composition on stage j. For the case of top and bottom composition,

$$\begin{pmatrix} T_1 \\ -- \\ T_K \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{c}_1 & | & \mathbf{0} \\ -- & -- & -- \\ \mathbf{0} & | & \mathbf{c}_K \end{pmatrix}}_{G} \underbrace{\begin{pmatrix} \tilde{\mathbf{x}}_1 \\ -- \\ \tilde{\mathbf{x}}_K \end{pmatrix}}_{\mathbf{h}}. \tag{9}$$

From Eq. (9), it is obvious that $\mathbf{h}$ cannot be reconstructed (observed) because matrix G is non-square and has dimension $2 \times K(c\text{-}1)$. We can find the solution only via pseudo-inversion of G (to compute G^+) in order to get $\mathbf{h}$. It is well known that the pseudoinverse provides a least squares solution for Eq. (9) (Albert, 1972; Lawson and Hanson, 1974). This motivates the building of regression models of soft sensor and shows the limitation of direct use of rigorous model or its linearized analog for product quality estimation.

4. Fitting of rigorous distillation model on industrial data

In order to build adequate regression models of soft sensors the large samples including data of steady-states for a wide range are required. In practice, it is a problem to have a big training sample. A rigorous tray-by-tray (with physico-chemical essence) distillation model was used because it is necessary to extend the training sample with input variables that have low variability ranges and investigate the soft sensor model without measurement errors.

The principles of mass, energy balances, and phase-equilibrium equations of a rigorous model of multicomponent distillation are well known. The fitting procedure involves selecting the value of tray efficiency that minimizes the mismatch between industrial data and the rigorous model. In order to reduce the number of optimized variables involved in the fitting of rigorous model, the sectional Murphree efficiencies are used: $E_{j_1,j_2}^{<\text{column name}>}$, where the indexes j_1 and j_2 correspond to the trays numbers in the section in such a way that the following inequality should hold: $j_2 > j_1$, $j_1 = 2\ldots K\text{-}1$, $j_2 = 3\ldots K$. For example, for the section above the feed tray of the column model (1)–(5): $j_1 = 2$, $j_2 = f$ and $E_2 = \ldots = E_f = E_{2,f}^{<\text{column name}>}$; for the stripping section: $E_{f+1} = \ldots = E_K = E_{f+1,K}^{<\text{column name}>}$.

The average (nominal) steady-state operating point is involved when estimating the Murphree efficiency with industrial data. The data-reconciliation problem of reflux ratio (RR) for each column is solved in conjunction with the residual functions minimization. For K-8 and K-9 columns the residual functions J_{C4}^{PPF} and J_{C4}^{AG} are utilized in the following optimization problems:

$$J_{C4}^{PPF}(E_{2,25}^{K\text{-}8}, E_{26,60}^{K\text{-}8}, RR^{K\text{-}8}) = \left(\frac{\tilde{h}_{C4,PPF} - \tilde{h}_{C4,PPF}^{meas}}{\sigma_{\tilde{h}_{C4,PPF}}}\right)^2 + \left(\frac{RR^{K\text{-}8} - RR^{K\text{-}8,meas}}{\sigma_{RR^{K\text{-}8}}}\right)^2 \rightarrow \min$$

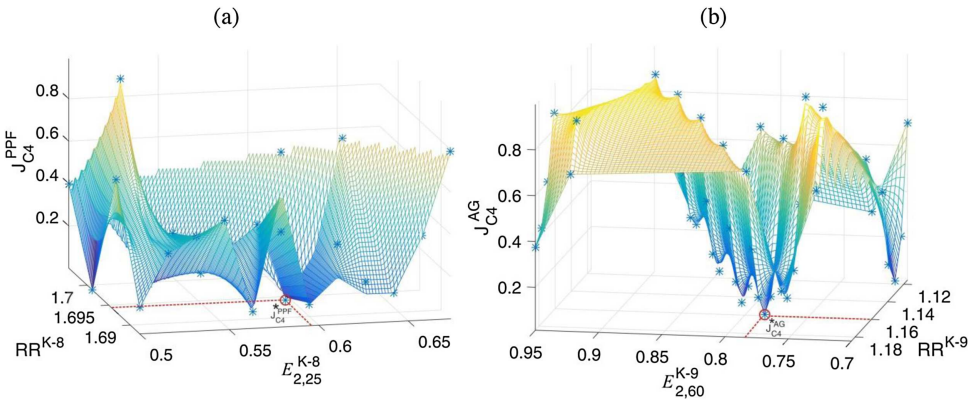

Fig. 2 – The results of fitting the sectional efficiencies on industrial data. (a) for column K-8 under $E_{26,60}^{K-8} = 0.54$. (b) for column K-9.

Table 3 – Correlation coefficients of feed components in the groups C3 and C4.

	Propylene	Propane	i-Butane	*n*-Butane	*t*-Butene-2
Propane	0,967				
i-Butane	0,096	0,106			
n-Butane	0,067	0,155	0,803		
t-Butene-2	−0,006	0,036	0,703	0,817	
c-Butene-2	0,040	0,087	0,662	0,782	0,833

$$\text{s.t.}\quad \begin{aligned} \underline{E_{2,25}^{K-8}} \le E_{2,25}^{K-8} \le \overline{E_{2,25}^{K-8}},\\ \underline{E_{26,60}^{K-8}} \le E_{26,60}^{K-8} \le \overline{E_{26,60}^{K-8}};\\ \underline{RR^{K-8}} \le RR^{K-8} \le \overline{RR^{K-8}} \end{aligned}$$

$$J_{C4}^{AG}(E_{2,60}^{K-9}, RR^{K-9}) = \left(\frac{\tilde{h}_{C4,AG} - \tilde{h}_{C4,AG}^{meas}}{\sigma_{\tilde{h}_{C4,AG}}}\right)^2 + \left(\frac{RR^{K-9} - RR^{K-9,meas}}{\sigma_{RR^{K-9}}}\right)^2 \to \min$$

$$\text{s.t.}\quad \begin{aligned} \underline{E_{2,60}^{K-9}} \le E_{2,60}^{K-9} \le \overline{E_{2,60}^{K-9}},\\ \underline{RR^{K-9}} \le RR^{K-9} \le \overline{RR^{K-9}}. \end{aligned}$$

where $\tilde{h}_{C4,PPF}$ and $\tilde{h}_{C4,AG}$ are the sum of C4 content in PPF and AG, respectively, calculated using the rigorous model, $\tilde{h}_{C4,PPF}^{meas}$ and $\tilde{h}_{C4,AG}^{meas}$ are the sum of C4 content in PPF and AG, respectively, (industrial data), $\sigma_{(.)}$ is the standard deviation of the optimized variable.

The calculation results of fitting procedure and Murphree efficiencies' values are shown in Fig. 2. The points J_{C4}^{*PPF} and J_{C4}^{*AG} correspond to the best fitted values of sectional efficiencies with optimal reconciled RR values.

The behavior of functions J_{C4}^{PPF} and J_{C4}^{AG} were studied visually in the neighborhood of optimal solutions due to the presence of local minima. For the K-8 column, two sectional efficiencies were used since the diameter and type of tray differ for the absorption and stripping sections. To show the behavior of function J_{C4}^{PPF} in three-dimensional space on Fig. 2a in the neighborhood of the optimum the value of Murphree efficiency for stripping section was fixed at the optimum value $E_{26,60}^{K-8} = 0.54$.

It will be noted that the fitting of efficiency values for each individual separation stages did not provide significant improvements of fitting results.

Based on the fitted rigorous model of industrial distillation unit, it will be possible to generate a training dataset. But the main problem for multicomponent mixture was to generate the feed composition vectors for simulation, which are close to industrial situations. For that reason, we study the redistribution of individual components among the main separated groups C3 and C4 in the feed composition (Table 3). It was helpful to reduce the numerical experiments with a fitted rigorous model. For highly correlated concentrations of components in the feed (e.g., propylene and propane), it is possible to express variation of both compounds via one.

5. Evaluation of moving window soft sensor based on the constrained optimization

The soft sensor model equation is considered in the following linear form:

$$\hat{y} = b_0 + \sum_{k=1}^{m} b_k x_k,$$

where $\hat{y}$ is the soft sensor's output (prediction), x_k is the measured input of the soft sensor (Table 1); b_k is the model parameter, and m is the number of inputs.

The comparative analysis of the soft sensor model parameters for predicting C_3 concentration in the BBF product may be done from Table 4.

The parameters of the model 1 are obtained using the bootstrapping regression approach (Chernick, 2008) because it was found that the ordinary regression gave the positive signs of the parameters b_1 and b_4 (they should be negative due to the first principles of distillation). t-statistics (t_{ind}) are calculated for each model 1 parameter and are given in Table 4. The total number of the model parameters (including bias b_0) is 7. The training sample size is 70. Therefore the number of degrees of freedom is 63. From a table of values of Student's t-distribution

Table 4 – Comparison of model parameters.

Model parameter	Model 1 (based on industrial data): b_{ind}	Model 1: |tind| t-statistic	Model 2 (based on rigorous simulation data): b_{rig}
b_1	−3.21	2.69	−0.06
b_2	5.66	6.44	2.63
b_3	−0.75	12.60	−0.97
b_4	−0.30	2.23	−3.82
b_5	−0.95	5.88	−1.02
b_6	1.03	15.77	1.53
b_0 (bias)	26.30	3.33	47.6

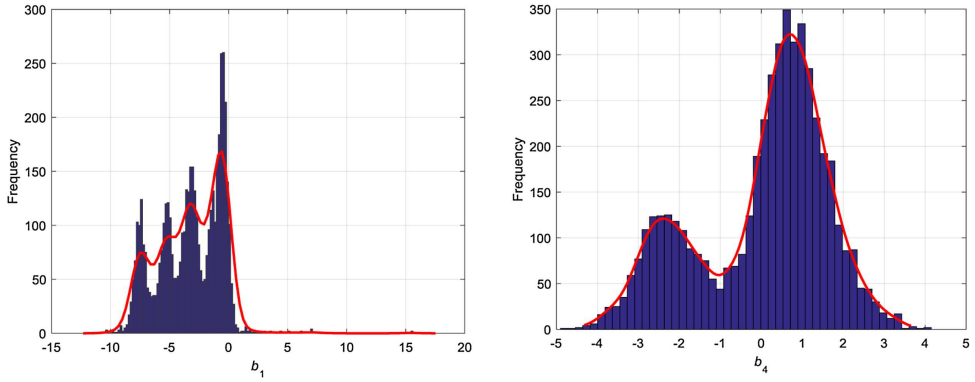

Fig. 3 – The histograms and kernel density estimates of the parameters b_1 (multimodal pattern) and b_4 (bimodal pattern).

Table 5 – Moving window soft-sensors performance.

	Robust reg. with all prev. data	Robust reg. moving window	Bootstrap moving window	Constrained moving window
R^2	0.65	0.68	0.70	0.81
RMSE	0.62	0.60	0.53	0.45

Table 6 – Variance (std.) of the soft sensors' parameters.

Model parameter	Robust reg. with all prev. data	Robust reg. moving window	Bootstrap moving window	Constrained moving window
b_1	2.56	4.61	4.67	0.02
b_2	1.74	3.12	3.26	0.48
b_3	0.05	0.16	0.15	0.09
b_4	1.11	5.31	5.22	0.63
b_5	0.11	0.29	0.29	0.30
b_6	0.06	0.24	0.22	0.15
b_0 (bias)	40.63	54.22	50.63	19.45

the critical value of t-statistic $t_c = 1.998$ may be found in case of 95% confidence level and degrees of freedom 63. The condition of statistical significance of the model 1 parameters (differ from zero) is fulfilled, since for each parameter the $t_{ind} > t_c$ inequality is valid.

The industrial dataset was used to evaluate model 1. The bootstrap analysis provides the interesting facts of the multimodal (b_1) and bimodal (b_4) patterns of kernel densities estimates of these parameters (Fig. 3). The reason for multimodal patterns is a small training sample. For example, the training dataset does not contain operating point data for the entire range of plant load changes. Therefore, multimodal patterns may indicate a poor performance of soft sensor on the test sample.

The training sample for model 2 in Table 4 was generated based on simulations with a fitted, rigorous model of the industrial distillation unit (Section 4). The model 2 was derived by use of the robust regression *M*-estimator (Maronna et al., 2006) without constraints on the parameter values. The main advantage of model 2 is the absence of influence of laboratory errors and incorrect input values (due to sampling time uncertainty) on the soft sensor parameters. Interestingly, the 1st and 4th parameters (b_1 and b_4) are different in several times and maybe statistically significant for both models. However, the values of b_1 and b_4 are only consistent with the physico-chemical essence of distillation for model 2.

The model parameters of the soft sensor are subject to change. This is due to the nonstationarity of the plant: reboilers can become fouled; variations of the feed flow rate and composition can lead to a change in the hydrodynamic mode of operation of the trays and the mass transfer efficiency. To adjust the soft sensor model, the moving window method is well proven. It is that the model is updated based on the selected dataset from the process history (window). As new laboratory data occurrence, the window "slides" over the history data in such a way that new data is included in the training sample and old data is excluded.

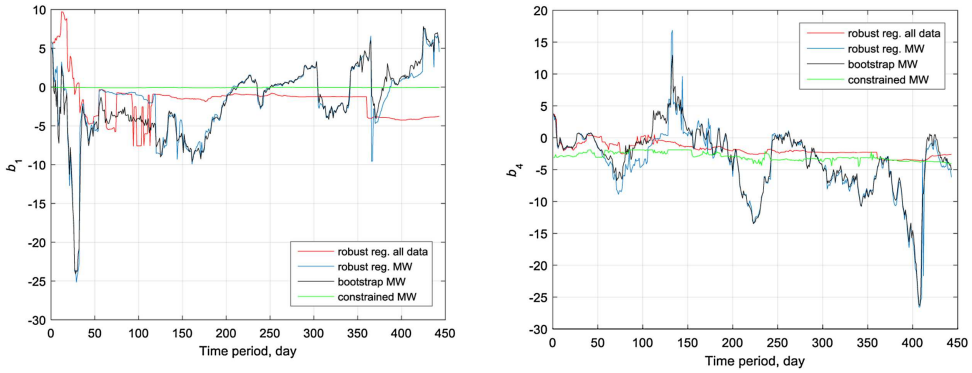

Fig. 4 – Adaptation of parameters b_1 and b_4 (test dataset).

The parameters of the soft sensor's model are recalculated after occurrence of new laboratory data. The new lab measurement values are added to the training sample and the last ones are removed. Thus, we get a moving training dataset or moving window. The least-squares fitting procedure (within the moving training dataset) under restriction of the model parameters is transformed into an optimization problem:

$$\min_{\mathbf{b}} \frac{1}{2} \left| |\mathbf{X}\mathbf{b} - \mathbf{y}| \right|_2^2 \quad \text{such that } \mathbf{l}_b \le \mathbf{b} \le \mathbf{u}_b, \tag{10}$$

where $\mathbf{b} = (b_1 \ldots b_m)^T$ is the vector of the model's parameters, $\mathbf{X}$ is the input data matrix, $\mathbf{y}$ is the vector of output measurements, and $\mathbf{l}_b$ and $\mathbf{u}_b$ are the lower and upper bounds of the soft sensor parameter values, respectively. Vectors $\mathbf{l}_b$ and $\mathbf{u}_b$ play an important role in integrating a fitted first-principle (rigorous) model into the soft sensor design procedure. The values of these vectors may only be obtained numerically because of the high dimensionality of the multicomponent distillation model.

The comparative summary of the application of soft sensor for the industrial distillation process can be seen in Tables 5 and 6. Table 5 compares the *RMSE* and R^2 values of the *M*-estimators for a conventional unconstrained (robust reg.) case and constrained optimization approach. The vectors of the lower and upper bounds of model 2's ($\mathbf{b}_{rig}$) parameters from Table 4 are derived as follows: $\mathbf{l}_b = (1 - v) \cdot \mathbf{b}_{rig}$ and $\mathbf{u}_b = (1 + v) \cdot \mathbf{b}_{rig}$, where the value of v (variation range parameter) is assigned for vector $\mathbf{b}_{rig}$ from the solution the of optimization problem (*RMSE* is criterion), as given below.

For the testing dataset of R^2 and *RMSE*, soft sensor performance improved compared to bootstrap moving window (Table 6) by $100 \cdot (0.81 - 0.70)/0.81 \approx 14\%$ and $100 \cdot (0.53 - 0.45)/0.53 \approx 15\%$, respectively.

The variability of the moving window soft-sensor parameters was lower in the case of constrained optimization (Fig. 4) for the test dataset. Therefore, the use of vectors $\mathbf{l}_b$ and $\mathbf{u}_b$ made the performance of the adaptive soft sensor more stable and reliable.

The deep comparative analysis of the moving window soft sensors performance was done (Table 6) in order to validate the superiority of proposed soft sensor (constrained moving window).

The width of moving window (w) is an important parameter of adaptive soft sensors. The calculation of w is proposed by the following formula:

$$w = q \cdot m, \tag{11}$$

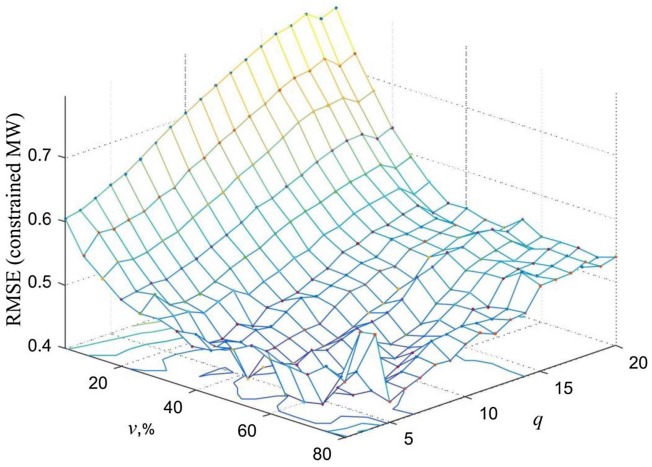

Fig. 5 – Selection of optimal values of v and q for constrained moving window soft sensor.

where q is the window width parameter.

The optimal parameters of the moving window soft sensor (based on the constrained optimization) are shown in Fig. 5. The function RMSE(v,q) may have a well-defined global minimum (Fig. 5). This allows to select the optimal parameters v^* and q^* of adaptive soft sensor. The optimal values (in the sense of the *RMSE* criterion) of the parameters are $v^* = 50\%$ and $q^* = 10$ in Eq. (11) for the considered adaptive soft sensor (v is variation range parameter considered above).

The evaluation of a constrained moving window soft sensor can be described by the following stepwise procedure:

- Step 1. Divide the industrial data into a training dataset and a test dataset.
- Step 2. Evaluate model 1 based on the training dataset and obtain the vector of parameters $\mathbf{b}_{ind}$ (Table 4).
- Step 3. Fit (jointly with data reconciliation) the rigorous distillation unit model with the training dataset, as described in Section 4.
- Step 4. Evaluate model 2 based on the dataset from the simulator with the fitted, rigorous distillation unit model and obtain the vector of parameters $\mathbf{b}_{rig}$ (Table 4).
- Step 5. Calculate bounds $\mathbf{l}_b$ and $\mathbf{u}_b$ based on the vector $\mathbf{b}_{rig}$ and study parameter v belonging to the interval 0.3–0.5. The use of $\mathbf{b}_{ind}$ is not valid because the signs and values of b_1 and b_4 (in model 1) are inconsistent with the physico-chemical essence of distillation.
- Step 6. Analyze the performance of the adaptive soft sensor with the test dataset. Calculate the prediction error (PE) for each new point of the test dataset to solve the optimization

problem (10) based on the previous historical data, **X**,**y**. The mean value of the set of PEs is the *RMSE* for the test dataset.

For each point of the test dataset, vector **b** is updated using Eq. (10). The application of a rigorous distillation unit model is realized via bounds $\mathbf{l}_b$ and $\mathbf{u}_b$.

6. Conclusions

This article improves soft sensor evaluation based on the moving window adaptation technique. A way to account for all available a priori information about a multicomponent distillation process, along with the physico-chemical meaning in the frame of a rigorous model, was proposed. This method can reduce the *RMSE* of the test sample in conditions when the training sample is small and contains measurement errors.

The fitted, rigorous multicomponent distillation model is useful for checking the correctness of the soft sensor's parameter values and for obtaining $\mathbf{b}_{\text{rig}}$. The introduction of vectors $\mathbf{l}_b$ and $\mathbf{u}_b$ may be considered an indirect use of data of rigorous modeling for evaluating soft sensors in multicomponent distillation columns. The selection of soft sensor inputs also becomes reliable in that case and is reconciled with the thermodynamic essence of the distillation. Finally, the use of constraints $\mathbf{l}_b$ and $\mathbf{u}_b$ reduces the variability of the moving window soft-sensor parameters and leads to more stable and accurate estimations of product quality for industrial distillation processes.

Acknowledgment

This work was partially supported by the Russian Foundation for Basic Research (Project No. 17-07-00235 A).

Appendix A. Proof the proposition

Consider the m observations of the each element of the output vector $\mathbf{z}_k$ of system (7):

$$\mathbf{z}_{k+m-1} = Z\mathbf{y}_{k+m-1} = ZC_\delta A_\delta^{m-1} + \sum_{i=1}^{m-1} ZC_\delta A_\delta^{i-1}\tilde{U}_\delta \bar{\mathbf{u}}_{k+m-1-i} \qquad \text{(A1)}$$

where $\tilde{U}_\delta = (B_\delta|D_\delta)^T$; $\bar{\mathbf{u}}_{(.)} = \left(\mathbf{u}_{(.)}|\mathbf{d}_{(.)}\right)^T$. Introduce the following vectors:

$\mathbf{z_m} = (\mathbf{z}_k|\mathbf{z}_{k+1}|\ldots|\mathbf{z}_{k+m-1})^T$; $\bar{\mathbf{u}}_\mathbf{m} = (\bar{\mathbf{u}}_{k+m-1}|\ldots|\bar{\mathbf{u}}_{k+1}|\bar{\mathbf{u}}_k)^T$. Then the Eq. (A1) can be written as

$$\mathbf{z_m} = Q\mathbf{x}_k + S\bar{\mathbf{u}}_k \qquad \text{(A2)}$$

where $Q = \left(ZC_\delta|ZC_\delta A_\delta|\ldots|ZC_\delta A_\delta^{m-1}\right)^T$;

$$S = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & ZC_\delta\tilde{U}_\delta \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & ZC_\delta\tilde{U}_\delta & ZC_\delta A_\delta\tilde{U}_\delta \\ \mathbf{0} & \mathbf{0} & ZC_\delta\tilde{U}_\delta & ZC_\delta A_\delta\tilde{U}_\delta & ZC_\delta A_\delta^2\tilde{U}_\delta \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & ZC_\delta\tilde{U}_\delta & ZC_\delta A_\delta\tilde{U}_\delta & \ldots & ZC_\delta A_\delta^{m-2}\tilde{U}_\delta \end{pmatrix}$$

From Eq. (A2) it will be possible to obtain $\mathbf{x}_k$ and after that to find interesting state vector $\mathbf{h}_k$:

$$\mathbf{h}_k = HQ^{-1}(\mathbf{z_m} - S\bar{\mathbf{u}}_m) \qquad \text{(A3)}$$

In case of the state vector $\mathbf{x}_k$ in Eq. (A2) the m is to equal to $K(c\text{-}1)$ and the solution of Eq. (A3) exists if the disturbance vector is known and $rank(Q) = K(c\text{-}1)$.

References

Albert, A., 1972. Regression and the Moore-Penrose Pseudoinverse. Elsevier, Burlington, MA.

Amrit, R., Canney, W., Carrette, P., Linn, R., Martinez, A., Singh, A., Skrovanek, T., Valiquette, J., Williamson, J., Zhou, J., Cott, B.J., 2015. Platform for advanced control and estimation (PACE): Shell's and Yokogawa's next generation advanced process control technology. IFAC-PapersOnLine 48 (8), 1–5.

Chernick, M.R., 2008. Bootstrap Methods: A Guide for Practitioners and Researchers. John Wiley & Sons, Hoboken, New Jersey.

Fayruzov, D.Kh., Bel'kov, Yu.N., Kneller, D.V., Torgashov, A.Yu., 2017. Advanced process control system for a crude distillation unit. A case study. Autom. Remote Control 78 (2), 357–367.

Fortuna, L., Graziani, S., Rizzo, A., Xibilia, M.G., 2007. Soft Sensors for Monitoring and Control of Industrial Processes. Springer-Verlag, London.

Funatsu, K., 2018. In: Engel, Th., Gasteiger, J. (Eds.), Process Control and Soft Sensors, Applied Chemoinformatics: Achievements and Future Opportunities. Wiley-VCH, Weinheim.

Kadlec, P., Gabrys, B., Strandt, S., 2009. Data-driven soft sensors in the process industry. Comput. Chem. Eng. 33 (4), 795–814.

Kadlec, P., Grbic, R., Gabrys, B., 2011. Review of adaptation mechanisms for data-driven soft sensors. Comput. Chem. Eng. 35 (1), 1–24.

Kaneko, H., Funatsu, K., 2014. Adaptive soft sensor based on online support vector regression and Bayesian ensemble learning for various states in chemical plants. Chemom. Intell. Lab. Syst. 137 (10), 57–66.

Kaneko, H., Funatsu, K., 2015. Moving window and just-in-time soft sensor model based on time differences considering a small number of measurements. Ind. Eng. Chem. Res. 54 (2), 700–704.

Kim, S., Kano, M., Hasebe, Sh., Takinami, Ak., Seki, T., 2013. Long-term industrial applications of inferential control based on just-in-time soft-sensors: economical impact and challenges. Ind. Eng. Chem. Res. 52 (35), 12346–12356.

Lawson, C., Hanson, R., 1974. Solving Least Squares Problems. Prentice-Hall, Englewood Cliffs, NJ.

Maronna, R.A., Martin, R.D., Yohai, V.J., 2006. Robust Statistics: Theory and Methods. John Wiley & Sons, New York.

Mohd Ali, J., Ha, Hoang N., Hussain, M.A., Dochain, D., 2015. Review and classification of recent observers applied in chemical process systems. Comput. Chem. Eng. 76 (5), 27–41.

Porru, M., Alvarez, J., Baratti, R., 2013. Composition estimator design for industrial multicomponent distillation column. Chem. Eng. Trans. 32, 1975–1980, http://dx.doi.org/10.3303/CET1332330.

Rogina, A., Sisko, I., Mohler, I., Ujevic, Z., Bolf, N., 2011. Soft sensor for continuous product quality estimation (in crude distillation unit). Chem. Eng. Res. Des. 89 (10), 2070–2077.

Shao, W., Tian, X., 2015. Adaptive soft sensor for quality prediction of chemical processes based on selective ensemble of local partial least squares models. Chem. Eng. Res. Des. 95 (3), 113–132.

Torgashov, A., Skogestad, S., Kozlov, A., 2016. Comparative study of multicomponent distillation static estimators based on industrial and rigorous model datasets. IFAC-PapersOnLine 49 (7), 1187–1192.

Tronci, St., Bezzo, F., Barolo, M., Baratti, R., 2005. Geometric observer for a distillation column: development and experimental testing. Ind. Eng. Chem. Res. 44 (26), 9884–9893.

Xiong, W., Li, Y., Zhao, Y., Huang, B., 2017. Adaptive soft sensor based on time difference Gaussian process regression with local time-delay reconstruction. Chem. Eng. Res. Des. 117 (1), 670–680.

Zhang, X., Kano, M., Li, Y., 2017. Locally weighted kernel partial least squares regression based on sparse nonlinear features for virtual sensing of nonlinear time-varying processes. Comput. Chem. Eng. 104 (9), 164–171.