Preprints, 10th IFAC International Symposium on
Advanced Control of Chemical Processes
Shenyang, Liaoning, China, July 25-27, 2018

Evaluation of Steady-State and Dynamic Soft Sensors for Industrial Crude Distillation Unit under Parametric Constraints

Andrei Torgashov[*], Anton Goncharov[*,], Evgeny Zhuravlev[***]**

*Institute of Automation and Control Process FEB RAS, 5 Radio St., and
Far-Eastern Federal University, Bld. 12(E), Campus FEFU, Russky Island
Vladivostok, Russia (Tel: +7-423-231-02-02; e-mail: torgashov@iacp.dvo.ru)

**JSC "Honeywell", 7 Kievskaya St., Moscow, Russia.

***JSC "Gazprom neftekhim Salavat", 30 Molodogvardeitsev st., Salavat, Russia.

Abstract: The parametric identification problem for industrial crude distillation unit (CDU) is considered. We take the a priori knowledge of the process into account by using a system of constraints for parameters of soft sensors models. The identification problem is transformed into a constrained optimization problem, which we solved using the active set method. The static and dynamic soft sensors are evaluated for industrial CDU located at JSC "Gazprom neftekhim Salavat" refinery. It was found that the model performed better when we used the proposed constrained optimization approach for identification instead of robust regression methods.

Keywords: Soft sensing, identification, constrained parameters, optimization, distillation columns.

1. INTRODUCTION

In the present, researchers are paying more attention to the identification problems of the model parameters of soft sensors for CDU (Bolf et al., 2010; Dam and Saraf, 2006; Macias-Hernandez et al., 2007; Napoli and Xibilia, 2011). The most widespread approach for soft sensor evaluation is based on the partial least-squares (PLS) method (Shang et al., 2015). The integration of the first principles (material and energy balances and phase equilibrium) of inferential modeling can be found in Chatterjee and Saraf (2003), Mahalec and Sanchez (2012), and Fujii and Yamamoto (2014). Johansen (1996) considers the more general framework of integrating available a priori information into the identification problem.

However, the methods developed from the previous research do not completely deal with practical obstacles such as small training datasets that cover all operating points, low variability ranges for key (informative) input variables, and an unknown feed composition (i.e. the feed distillation curves as TBP).

In order to overcome the abovementioned difficulties, Torgashov et al. (2016) has proposed the use of the system of parametric constraints. The system of parametric constraints is derived from the preliminary calibrated first principle (rigorous) distillation model (Torgashov and Zmeu, 2015). In the current work, we consider the extension of this technique in the case of CDU. We introduce and solve the statements of identification problems for static (steady-state) and dynamic soft sensors using the constraint optimization technique.

2. INDUSTRIAL CRUDE UNIT DESCRIPTION AND PROBLEM FORMULATION

The crude distillation unit considered in this paper is represented by two multicomponent distillation columns: K-1 and K-2 (Fig. 1). The plant is located at JSC "Gazprom neftekhim Salavat" refinery (Salavat, Russia). The feed flow (crude) comes from the oil desalting unit and enters under the 16th tray of K-1. The overhead products of K-1 are gas and naphtha. The naphtha is also withdrawn from the top of K-2. The bottom products of the 3 strippers of the column K-2 are gasoline, kerosene cut (KC) and diesel oil cut (DOC). The main product of the column K-2 is desired cut 1 (DC1), which is the mixture of kerosene cut (KC) and diesel oil cut (DOC). Desired cut 2 (DC2), which is the mixture of naphtha and gasoline. The atmospheric residue (AR) is withdrawn from the bottom part of column K-2. The main process variables of the industrial CDU are shown in Table 1 and may be considered informative inputs of the soft sensor model.

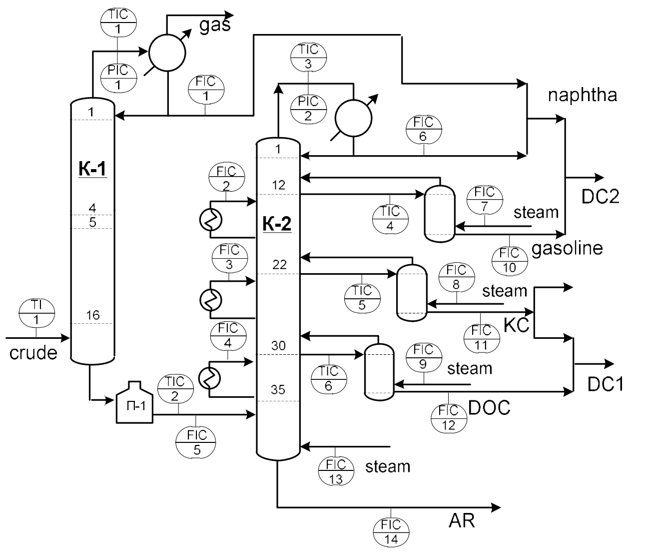

Fig. 1. The sequence of industrial multicomponent distillation columns (CDU).

We consider the plant with several measured inputs $u_1,\ldots,u_i,\ldots,u_N$ and one output $y(\tau)$. We use the measured process variables (pressure, temperature, and flow) as inputs. A priori knowledge of the distillation process allows us to select the most informative variables (Table 1) from the thermodynamic essence.

Table 1. Main process variables

No	Notation	Process variable, u_i
1	TIC6	Temperature of the flow in the bottom stripper of column K-2, °C
2	PIC2	Top pressure of the column K-2, MPa
3	FIC13/ FIC5	The ratio of steam to feed flowrate for column K-2
4	TIC3	Top temperature of the column K-2, °C
5	FIC9/ FIC5	The ratio of steam to feed flowrate for bottom stripper of column K-2
6	FIC4/ FIC5	The ratio of bottom pumparound to feed flowrate of column K-2
7	PIC1	Top pressure of column K-1, MPa
8	TIC2	Feed temperature of column K-2, °C
9	FIC10/ FIC5	The ratio of the product of the top stripper to the feed flowrate of column K-2
10	FIC11/ FIC5	The ratio of the product of the middle stripper to the feed flowrate of column K-2
11	TI1	Feed temperature of the column K-1, °C
12	FIC12/ FIC5	The ratio of product of bottom stripper to feed flowrate of column K-2

We consider the identification problem of the soft sensor (SS) evaluation, which is best for predicting the quality of the products of crude distillation process. We obtain the model for the soft sensor in the form of a linear regression model based on the following equation:

$$y(\tau) = b_0 + b_1 u_1(\tau) + b_2 u_2(\tau) \ldots + b_N u_N(\tau), \tag{1}$$

where b_j is the j-th model coefficient, j=0,1,…,N, b_0 is the constant term, N is the number of input variables, τ is the irregular time points of output measurement $\tau_1, \tau_2, \tau_3, \ldots$, $\tau_i = \tau_{i-1} + \tau_0 + \varepsilon$, $i \geq 2$; $\tau_1 = \tau_0 + \varepsilon$, τ_0 is the constant term; and ε is the random variable restricted in the given range.

The dynamic SS accounts for the influence the process dynamics have on the quality of the products. The predictive model is represented as a sum of convolutions of plant inputs and a finite impulse response (FIR) h_i (discrete analogues of the first degree Volterra kernels):

$$y(\tau) = h_0 + \sum_{k=0}^{n_1-1} h_1(k+1)u_1(\tau-k) + \sum_{k=0}^{n_2-1} h_2(k+1)u_2(\tau-k) + \ldots \tag{2}$$
$$\ldots + \sum_{k=0}^{n_N-1} h_N(k+1)u_N(\tau-k),$$

where h_0 is a constant term.

We use the determination coefficient (a number that indicates the proportion of the variance in the dependent variable that is predictable from the independent variable)

$$R^2 = 1 - \sum_i (\bar{y}_i - y_i)^2 \Big/ \sum_i (\bar{y}_i - \bar{y}^a)^2 , \tag{3}$$

the root mean squared error (*RMSE*)

$$RMSE = \left(\sum_{i=1}^{M} (\bar{y}_i - y_i)^2 / M \right)^{1/2} , \tag{4}$$

the Akaike (1969) Information Criterion:

$$AIC = M \; In\left(\sum_i \left(y_i - y_i^m\right)^2\right) + 2(N+1) - M \; In(M), \tag{5}$$

the Schwarz Bayesian Criterion:

$$BSC = M \; In\left(\sum_i \left(y_i - y_i^m\right)^2\right) + (N+1)In(M) - M \; In(M) \tag{6}$$

as identification criteria on a given time interval, where $\bar{y}_i$ is the measured value of the output variable, y_i is the value obtained based on the SS, $\bar{y}^a$ is the mean value of the measured output variable, and M is the number of output measurements. The model is more consistent the closer to unity the value of the coefficient of determination R^2 is or the closer to zero the value of the *RMSE* is or then less the value of the *AIC* and *BSC* are.

The goal of the paper is to develop an approach for soft sensor model identification based on the industrial data while taking into account parametric constraints. These constraints can be derived from the rigorous modeling (Torgashov et al., 2016). The introduction of the system of constraints also allows us to overcome such difficulties as small training datasets and laboratory errors. The final boiling point (FBP) of DC1 and final boiling point (FBP) of DC2 are considered soft sensor outputs.

3. MAIN RESULTS

3.1 Steady-state model identification under constraints

Let $\mathbf{u} = [1, u_1(\tau), u_2(\tau), \ldots, u_N(\tau)]^T$ be a combined vector of the measured input variables and $\mathbf{b} = [b_0, b_1, \ldots, b_N]^T$ be a vector of coefficients of the same dimension, the components of which reflect the contributions of the corresponding input variables. Then the equation (1) takes the following form:

$$y = \mathbf{u}^T \cdot \mathbf{b} .$$

We form the vector $\mathbf{Y}$ of dimension q from the output value y dataset as

$$\mathbf{Y} = (y(\tau_1), y(\tau_2), \ldots, y(\tau_q))^T$$

and the matrix $\mathbf{U}$, containing the measured inputs u_j, corresponding to output value y from (1):

$$\mathbf{U} = \begin{bmatrix} 1 & u_1(\tau_1) & u_2(\tau_1) & \ldots u_N(\tau_1) \\ 1 & u_1(\tau_2) & u_2(\tau_2) & \ldots u_N(\tau_2) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & u_1(\tau_q) & u_2(\tau_q) & \ldots u_N(\tau_q) \end{bmatrix}$$

We consider the multicollinearity case, which occurs when there is an almost linear relationship between inputs. In this case, the matrix $\mathbf{C} = \mathbf{U}^T\mathbf{U}$ is close to singular, so it is the smallest eigenvalue $\lambda_{\min} = 0$, the condition number is infinitely increased, and it causes the instability of the solution. If $\lambda_{\min} = 0$, then it corresponds to the strict multicollinearity. In order to obtain a stable solution, it is necessary to reduce the condition number of the matrix **C** by, for example, adding thereto a diagonal matrix $\mathbf{B} = k\mathbf{I}$ (k> 0). Then we find the solution in a class of ridge parameter estimates:

$$\mathbf{b} = (\mathbf{U}^T\mathbf{U} + k\mathbf{I})^{-1}\mathbf{U}^T\overline{\mathbf{Y}} \ . \tag{5}$$

The quality, obtained using models (7), depends on the number of available output measurements. The length of the training sample is often insufficient to obtain reliable results. Also, the available data contains significant measurement errors in inputs and outputs, which are unmeasured influences. Taking into account constraints on the model coefficients b_j allows us to avoid these problems. When taking into account constraints on the model parameters, we solve the problem of least squares with simple constraints on the variables:

$$\min\left(\overline{\mathbf{Y}} - \mathbf{U}\mathbf{b}\right)^2 \\ \mathbf{b}^{\min} \le \mathbf{b} \le \mathbf{b}^{\max}. \tag{6}$$

The solution of the problem (8) is obtained by the active set numerical method (Gill et al., 1981). The given constraints are reduced to the form:

$$\mathbf{A}\mathbf{b} \ge \hat{\mathbf{b}},$$

where $\mathbf{A} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -1 & 0 & \ldots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{bmatrix}$, $\hat{\mathbf{b}} = \begin{bmatrix} \mathbf{b}^{\min} \\ -\mathbf{b}^{\max} \end{bmatrix}$.

Constraint $\mathbf{a}_i^T\mathbf{b} \ge \hat{b}_i$ is active in acceptable point b if $\mathbf{a}_i^T\mathbf{b} = \hat{b}_i$, and inactive if $\mathbf{a}_i^T\mathbf{b} > \hat{b}_i$, $\mathbf{a}_i^T$ is the i-th row of **A**. The sufficient minimum conditions for simple constraints are as follows:

$$\begin{aligned} &1.\ \mathbf{b}^{\min} \le \mathbf{b}^* \le \mathbf{b}^{\max} \quad \mathbf{b}_{FR}^{\min} < \mathbf{b}_{FR}^* < \mathbf{b}_{FR}^{\max} \\ &2.\ \mathbf{U}_{FR}^T(\overline{\mathbf{Y}} - \mathbf{U}\mathbf{b}^*) = 0 \\ &3.\ \boldsymbol{\lambda}^{\min} = \mathbf{U}_{\min}^T(\overline{\mathbf{Y}} - \mathbf{U}\mathbf{b}^*) \quad \lambda_i^{\min} > 0 \quad i = 1,\ldots,t^{\min} \\ &4.\ \boldsymbol{\lambda}^{\max} = \mathbf{U}_{\max}^T(\overline{\mathbf{Y}} - \mathbf{U}\mathbf{b}^*),\ \lambda_i^{\max} < 0,\ i = 1,\ldots,t^{\max} \\ &5.\ \mathbf{U}_{FR}^T\mathbf{U}_{FR} \text{ is the positive definite} \end{aligned} \tag{9}$$

where $\mathbf{b}^*$ is the minimum point of the solution of problem (8); subscript FR indicates that in the vector and matrices, the elements and columns with index numbers corresponding to the index numbers of **b** elements that have not met the boundary values of (8) are used; the subscript min and max indicate that in the matrix, only the columns with index numbers corresponding to the index numbers of **b** elements that take the appropriate minimum or maximum boundary value are used. $t^{\min}$ and $t^{\max}$ refer to the numbers of active upper and lower limits, respectively; $\boldsymbol{\lambda}^{\min}$ and $\boldsymbol{\lambda}^{\max}$ are vectors of Lagrange multipliers corresponding to the lower and upper active constraints.

Below is the algorithm for searching the minimum point $\mathbf{b}^*$ for each iteration k.

1. Find the starting point according to (7) in order to initialize the method of the active set.
2. Verify the performance of the stop conditions. (Reaching the performance errors of conditions (9), which are constraints on the number of iterations).
3. Select a logic branch. Does it make sense to remove any constraint from the list of the set of active constraints? The condition of performance 3 in (9) is checked. If the condition is not satisfied for some of the vector elements, the constraint is excluded from the list of active constraints.
4. Calculate the search direction $\mathbf{p}_k$. Equation (5) solves the problem $\min\left(\overline{\mathbf{Y}} - \mathbf{U}\mathbf{b}_k - \mathbf{U}_{FR}\mathbf{p}_{FR}\right)^2$. Calculate the non-zero $(N+1-t_k)$, the dimensional vector $\mathbf{p}_{FR}$, and the direction of search $\mathbf{p}_k = \left(\mathbf{A}^T\right)_{FR}\mathbf{p}_{FR}$, where t_k is the number of active constraints on k iterations.
5. Calculate the step length α_k. We calculate the diagonal matrix $\boldsymbol{\Psi}$ from $\begin{bmatrix} \mathbf{b}_{FR} \\ -\mathbf{b}_{FR} \end{bmatrix} + \boldsymbol{\Psi}\begin{bmatrix} \mathbf{p}_{FR} \\ -\mathbf{p}_{FR} \end{bmatrix} = \hat{\mathbf{b}}_{FR}$, $\hat{\mathbf{b}}_{FR}$ consists of the elements $\hat{\mathbf{b}}$, which aren't active constraints. The elements $\hat{\mathbf{b}}$, which are opposite boundary values in (8) for constraints in the active set, are excluded from $\hat{\mathbf{b}}_{FR}$. The $\overline{\alpha}_k = \min\{\boldsymbol{\Psi}_{ii}\}$ is an available maximum positive step from $\mathbf{b}_k$ along $\mathbf{p}_k$. We remember the index j of minimum positive diagonal element $\boldsymbol{\Psi}$. If $\overline{\alpha}_k > 1$, then $\alpha_k = 1$; otherwise, $\alpha_k = \overline{\alpha}_k$.
6. Add a constraint to the list of active constraints. If $\alpha_k = \overline{\alpha}_k$, then j constraint $\hat{\mathbf{b}}_{FR}$ becomes active and necessary to add to the list.
7. Approximate a recalculation. After $\mathbf{b}_{k+1} = \mathbf{b}_k + \alpha_k\mathbf{p}_k$ is calculated, return to step 2 of the algorithm.

3.2 Dynamic model identification under constraints

Let $\mathbf{u} = [1, u_1(\tau), \ldots, u_1(\tau - n_1 + 1), \ldots, u_N(\tau), \ldots, u_N(\tau - n_N + 1)]^T$ be the combined vector of the measured input variables of dynamic SS (DSS) with dimensionality $q = 1 + \sum_{k=1}^{N} n_k$, where n_k is a number of values of the k-th input variable and $\mathbf{h} = (h_0,\ h_1(1), \ldots, h_1(n_1), \ldots, h_N(1), \ldots, h_N(n_N))^T$ is the vector FIR of the same dimension, the components of which reflect the contributions of the respective input variables of DSS. Then the equation (2) takes the following form: $y = \mathbf{u}^T \cdot \mathbf{h}$.

We write the vector **Y** of dimension q from the output value y as

$\mathbf{Y} = (y(\tau_1), y(\tau_2), \ldots, y(\tau_q))^T$

and matrix **U**, containing the measured inputs u_j, corresponding to output value y from (2) as

$$\mathbf{U}=\begin{bmatrix} 1 & u_1(\tau_1) & \dots & u_1(\tau_1-n_1+1) & \dots & u_\mathrm{N}(\tau_1) & \dots \\ 1 & u_1(\tau_2) & \dots & u_1(\tau_2-n_1+1) & \dots & u_\mathrm{N}(\tau_2) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & u_1(\tau_q) & \dots & u_1(\tau_q-n_1+1) & \dots & u_\mathrm{N}(\tau_q) & \dots \end{bmatrix}.$$

$$\begin{bmatrix} \dots & u_\mathrm{N}(\tau_1) & \dots & u_\mathrm{N}(\tau_1-n_N+1) \\ \dots & u_\mathrm{N}(\tau_2) & \dots & u_\mathrm{N}(\tau_2-n_N+1) \\ \vdots & \vdots & \vdots & \vdots \\ \dots & u_\mathrm{N}(\tau_q) & \dots & u_\mathrm{N}(\tau_q-n_N+1) \end{bmatrix}$$

We write the matrix equation as $\mathbf{Y}=\mathbf{U}\,\mathbf{h}$. We introduce the error function:

$\mathbf{E}=\overline{\mathbf{Y}}-\mathbf{Y}=\overline{\mathbf{Y}}-\mathbf{U}\,\mathbf{h}$,

where $\overline{\mathbf{Y}}$ is the actual measurement of output, and minimize the following objective function:

$$\mathbf{\Psi}=\mathbf{E}^2=(\overline{\mathbf{Y}}-\mathbf{U}\mathbf{h})^2. \tag{7}$$

The constraints on transient response components are written as

$$\mathbf{s}^{\min}\le\mathbf{s}\le\mathbf{s}^{\max}, \tag{8}$$

where $\mathbf{s}=\left(s_1(1),...,s_1(n_1),...,s_N(1),...,s_N(n_N)\right)^T$, $\mathbf{s}^{\min}=\left[\mathbf{s}_1^{\min},...,\mathbf{s}_N^{\min}\right]^T$, $\mathbf{s}^{\max}=\left[\mathbf{s}_1^{\max},...,\mathbf{s}_N^{\max}\right]^T$. We obtain the constraint (11) based on the value of b_j $(j>0)$ of the steady-state soft sensor model $y(\tau)=b_0+b_1u_1(\tau)+...+b_ju_j(\tau)+...+b_Nu_N(\tau)$. That model is derived prior based on the steady-state industrial data of CDU. We select the parametric constraints for each FIR coefficient using the following values of a_1, a_2 and a_3.

a_1 is a fraction of the FIR length n_j from which the convergence of elements of $\mathbf{s}_j^{\min}$ and $\mathbf{s}_j^{\max}$ to $b_j^{\min}$, $b_j^{\max}$ is started.

a_2 is a fraction of $b_j^{\min}>0$ ($b_j^{\max}<0$) from which the smooth increasing (or decreasing) of elements of $\mathbf{s}_j^{\min}$ (or $\mathbf{s}_j^{\max}$) begins.

a_3 is a fraction of $b_j^{\max}>0$ ($b_j^{\min}<0$) from which the smooth decreasing (or increasing) of elements of $\mathbf{s}_j^{\min}$ (or $\mathbf{s}_j^{\max}$) begins.

For $b_j>0$, the following equations are valid for constraints

$$\mathbf{s}_j^{\min}=\begin{bmatrix} s_j^{\min}(1)=0 \\ \vdots \\ s_j^{\min}(\mathrm{fl}(a_1n_j))=0 \\ s_j^{\min}(\mathrm{fl}(a_1n_j)+1)=b_j^{\min}\left(a_2+(1-a_2)\left(1/(n_j-\mathrm{fl}(a_1n_j))\right)\right) \\ s_j^{\min}(\mathrm{fl}(a_1n_j)+2)=b_j^{\min}\left(a_2+(1-a_2)\left(2/(n_j-\mathrm{fl}(a_1n_j))\right)\right) \\ \vdots \\ s_j^{\min}(n_j)=b_j^{\min} \end{bmatrix}$$

$$\mathbf{s}_j^{\max}=\begin{bmatrix} s_j^{\max}(1)=a_3b_j^{\max} \\ \vdots \\ s_j^{\max}(\mathrm{fl}(\mathrm{a}_1n_j))=a_3b_j^{\max} \\ s_j^{\max}(\mathrm{fl}(\mathrm{a}_1n_j)+1)=b_j^{\max}\left(1+(a_3-1)\left(1-1/(n_j-\mathrm{fl}(\mathrm{a}_1n_j))\right)\right) \\ s_j^{\max}(\mathrm{fl}(\mathrm{a}_1n_j)+2)=b_j^{\max}\left(1+(a_3-1)\left(1-2/(n_j-\mathrm{fl}(\mathrm{a}_1n_j))\right)\right) \\ \vdots \\ s_j^{\max}(n_j)=b_j^{\max} \end{bmatrix}.$$

For $b_j<0$:

$$\mathbf{s}_j^{\min}=\begin{bmatrix} s_j^{\min}(1)=a_3b_j^{\min} \\ \vdots \\ s_j^{\min}(\mathrm{fl}(\mathrm{a}_1n_j))=a_3b_j^{\min} \\ s_j^{\min}(\mathrm{fl}(\mathrm{a}_1n_j)+1)=b_j^{\min}\left(1+(a_3-1)\left(1-1/(n_j-\mathrm{fl}(\mathrm{a}_1n_j))\right)\right) \\ s_j^{\min}(\mathrm{fl}(\mathrm{a}_1n_j)+2)=b_j^{\min}\left(1+(a_3-1)\left(1-2/(n_j-\mathrm{fl}(\mathrm{a}_1n_j))\right)\right) \\ \vdots \\ s_j^{\min}(n_j)=b_j^{\min} \end{bmatrix}$$

$$\mathbf{s}_j^{\max}=\begin{bmatrix} s_j^{\max}(1)=0 \\ \vdots \\ s_j^{\max}(\mathrm{fl}(a_1n_j))=0 \\ s_j^{\max}(\mathrm{fl}(a_1n_j)+1)=b_j^{\max}\left(a_2+(1-a_2)\left(1/(n_j-\mathrm{fl}(a_1n_j))\right)\right) \\ s_j^{\max}(\mathrm{fl}(a_1n_j)+2)=b_j^{\max}\left(a_2+(1-a_2)\left(2/(n_j-\mathrm{fl}(a_1n_j))\right)\right) \\ \vdots \\ s_j^{\max}(n_j)=b_j^{\max} \end{bmatrix},$$

where fl means rounding to the nearest integer number in the direction of -∞.

An example for determining the constraint system in terms of $\mathbf{s}_j^{\min}$ and $\mathbf{s}_j^{\max}$ of $\mathbf{s}_j$ under $b_j>0$ is shown on the Fig. 2. We select the parameters a_1=0.5, a_2=0.8, and a_3=1.2.

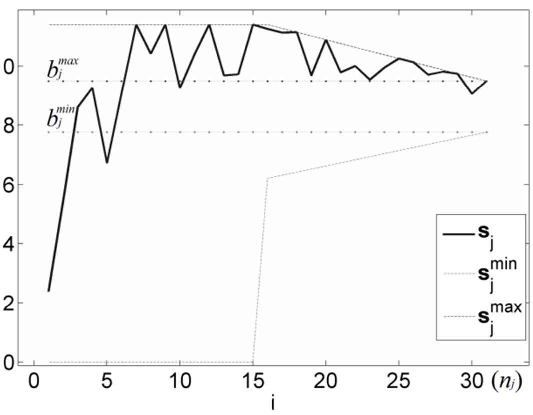

Fig. 2. The assignment of constraints for each step response function of the dynamic soft sensor model.

The coefficients of the finite step response (FSR) s_j are related to the components of the FIR h_j by the relations

$$s_j(k) = \sum_{i=1}^{k} h_j(i), \quad j = 1, 2, \ldots, N, \quad k = 1, \ldots, n_j. \tag{9}$$

The constraints (11) can be written as

$$\mathbf{A}\tilde{\mathbf{h}} \geq \hat{\mathbf{s}}, \tag{10}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -1 & 0 & \cdots & 0 \\ -1 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{bmatrix}, \quad \tilde{\mathbf{h}} = \begin{bmatrix} h_1(1) \\ \vdots \\ h_1(n_1) \\ \vdots \\ h_N(1) \\ \vdots \\ h_N(n_N) \end{bmatrix}, \quad \hat{\mathbf{s}} = \begin{bmatrix} \mathbf{s}^{\min} \\ -\mathbf{s}^{\max} \end{bmatrix}.$$

The sufficient minimum conditions are as follows for the current optimization problem (parametric identification of the dynamic model):

$$\begin{aligned} &1.\ \mathbf{A}\tilde{\mathbf{h}} \geq \hat{\mathbf{s}},\ \mathbf{A}_{ACT}\tilde{\mathbf{h}} = \hat{\mathbf{s}}_{ACT} \\ &2.\ \mathbf{Z}^{\mathbf{T}} * \mathbf{U}^T(\bar{\mathbf{Y}} - \mathbf{U}\mathbf{h}^*) = 0 \\ &3.\ \boldsymbol{\lambda} = \left(\mathbf{A}_{ACT}\mathbf{A}_{ACT}^T\right)^{-1}\mathbf{A}_{ACT}\mathbf{U}^T(\bar{\mathbf{Y}} - \mathbf{U}\mathbf{h}^*),\ \boldsymbol{\lambda}_i > 0,\ i = 1, \ldots, t \\ &4.\ \mathbf{Z}^T\mathbf{U}^T\mathbf{U}\ \mathbf{Z} \text{ is a positive definite,} \end{aligned} \tag{14}$$

where $\mathbf{h}^*$ is the minimum point of the solution of problem (10) with constraints (13); subscript ACT indicates that in the vector matrix, we only use the rows with index numbers corresponding to the element index numbers of active constraint in (13); t is the number of active constraints; $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers corresponding to the active constraints; $\mathbf{Z}$ is the matrix of the columns that are the basis of the feasible direction of the search for equality constraints (13). The matrix $\mathbf{Z}$ is formed by the variable-reduction technique (Gill et al., 1981).

The search algorithm of minimum point $\mathbf{h}^*$ for iteration k is as follows:

1. In order to start the method of the active set, it is necessary to determine the starting point (using a solution of the problem (10) without any constraints, with subsequent correction of coefficients h_i that do not fall under the constraints (13)).
2. Verify the performance of the stop conditions (reaching the performance errors of conditions (14) and constraints on the number of iterations).
3. Select a logic branch. Does it make sense to remove any constraint from the set of active constraints list? The condition of performance of a condition 3 in (14) is checked. If the condition is not satisfied for some of the vector elements, constraint is excluded from the list of active constraints, and it is necessary to recalculate $\mathbf{Z}_k$.
4. Calculate the search direction $\mathbf{p}_k$. Use an equation like (5) to solve the problem $\min\left(\bar{\mathbf{Y}} - \mathbf{U}\mathbf{h}_k - \mathbf{U}\mathbf{Z}_k\mathbf{p}_Z\right)^2$. Calculate the non-zero $\left(1 + \sum_{k=1}^{N} n_k - t_k\right)$ dimensional vector $\mathbf{p}_Z$ and the direction of search $\mathbf{p}_k = \mathbf{Z}_k\mathbf{p}_Z$, where t_k is the number of active constraints on the k iteration.
5. Calculate the step length α_k. From $\mathbf{A}\left(\tilde{\mathbf{h}} + \boldsymbol{\Psi}\tilde{\mathbf{p}}_k\right) = \hat{\mathbf{s}}$, the diagonal matrix $\boldsymbol{\Psi}$ is calculated. The calculated $\bar{\alpha}_k = \min\{\boldsymbol{\Psi}_{ii}\}$ is a minimum non-negative available step from $\mathbf{h}_k$ along $\mathbf{p}_k$, where i is the index number of the element, which is not an active constraint in (13) or an element of opposite boundary values (11) for constraints in the active set. Also, $\tilde{\mathbf{p}}_k$ consists of elements of $\mathbf{p}_k$ without the first element. We memorize the index j of the minimum positive diagonal element of $\boldsymbol{\Psi}_{ii}$. If $\bar{\alpha}_k > 1$, then $\alpha_k = 1$; otherwise, $\alpha_k = \bar{\alpha}_k$.
6. Add the constraint to the list of active constraints. If $\alpha_k = \bar{\alpha}_k$, then j constraint $\hat{\mathbf{s}}_{FR}$ becomes active and $\mathbf{Z}_k$ is recalculated.
7. Calculate the $\mathbf{h}_{k+1} = \mathbf{h}_k + \alpha_k\mathbf{p}_k$ and return to step 2.

4. INDUSTRIAL CASE STUDY

The CDU (Fig.1) is considered a case study for evaluating static (steady-state) and dynamic soft sensors based on the proposed identification algorithm under parametric constraints. The final boiling point temperature of the DC1 when we obtained a model on the training sample is considered. The number of output observations in the training sample is 70. The length of the test dataset is equal to 30 observations. Fig. 3 and Table 2 show the results of the performance of the static models on the test sample.

Table 2. Results of the performance of the static models (test dataset)

	R^2	*RMSE*	*AIC*	*BSC*
Without use constraints	0,64	4,47	232,7	255,4
With use constraints	0,84	2,99	175,5	198,2

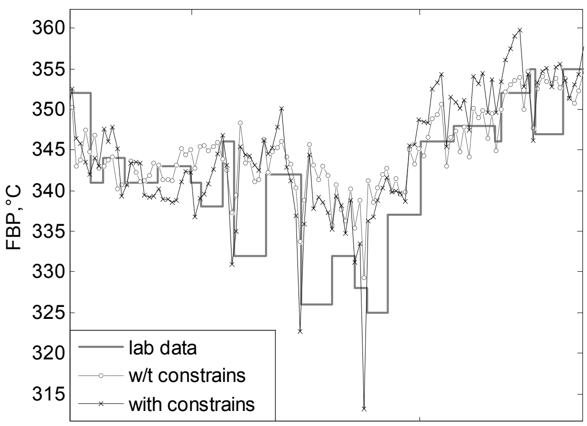

Fig. 3. Comparative study of static soft sensor model's performance.

The improvement of the prediction quality by the criterion *RMSE* of the identified static model obtained with the constraints on the parameters of SS is 100×(4,47 – 2,99) / 4,47 ≈ 33% compared to the case without constraints. In order to investigate the influence of constraints on the performance of the dynamic soft sensor model we obtained, we compared solutions of the optimization problem (10) without constraints and optimization problem (10) with constraints (13). We use the same values of the ridge coefficients. Fig. 4 and Table 3 show the results of the performance of the dynamic models on the test sample for the final boiling point temperature of DC2 when a model is obtained based on the training sample. The number of measurements in the training sample is 280. The size of the test dataset is 120 for the case with the dynamic model of the soft sensor.

Table 3. Results of the performance of the dynamic models (test dataset)

	R^2	*RMSE*	*AIC*	*BSC*
Without use constraints	0,40	3,79	6418,4	6482
With use constraints	0,74	2,52	4453,2	4516,8

The improvement of the prediction quality by the criterion *RMSE* of the dynamic model obtained using the constraints on the parameters of SS is 100×(3,79 – 2,52)/ 3,79 ≈ 33,5 % compared to the identification without constraints.

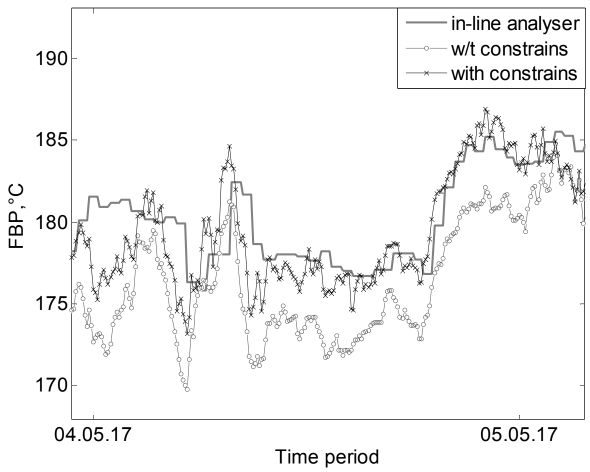

Fig. 4. Comparative study of dynamic soft sensors.

6. CONCLUSIONS

The introduction of a system of constraints into the identification algorithm improves the quality of derived models, especially on the test dataset. The reason is related to the integration of available a priori knowledge via the constraint identification problem statement and solution.The use of the active set method, taking into account constraints on the model coefficients, can improve the quality of the evaluated SS models, in particular in the case of small training datasets containing lab errors. The test of the proposed approach to solving the problem of obtaining a soft sensor model for industrial crude oil distillation unit showed that the reduction of the root mean square error on the test sample can be more than 33%.

7. ACKNOWLEDGEMENTS

This work was partially supported by the Russian Foundation for Basic Research (Project No. 17-07-00235 A) and by the Ministry of Education and Science of the Russian Federation, through Government Contract No. 02.G25.31.0173.

REFERENCES

Akaike, H. (1969). Fitting Autoregressive Models for Prediction. *Annals of the Institute of Statistical Mathematics*, 21, pp. 243-247.

Bolf, N., Galinec, G. and Baksa, T. (2010). Development of soft sensor for diesel fuel quality estimation. *Chemical Engineering & Technology*, 33, pp.405-413.

Chatterjee, T., and Saraf, D. N. (2004). On-line estimation of product properties for crude distillation units. *Journal of Process Control*, 14, pp. 61–77.

Dam M. and Saraf D.N. (2006) Design of neural networks using genetic algorithm for online property estimation of crude fractionator products. *Computers & Chemical Engineering*, 30 (4) pp. 722–729.

Fujii K. and Yamamoto T. (2014). Development of a nonlinear soft sensor using a GMDH network for a refinery crude distillation tower, *Electrical Engineering in Japan*, Vol. 188, No. 3. pp. 31-38.

Gill, P.E., W. Murray and Wright M.H. (1981). *Practical Optimization*. New York, USA: Academic Press.

Johansen, T. (1996). Identification of non-linear systems using empirical data and prior knowledge-an optimization approach. *Automatica*, Vol. 32, No. 3, pp. 337–356.

Macias-Hernandez J. J., P. Angelov and Zhou, X. (2007). Soft sensor for predicting crude oil distillation side streams using Takagi-Sugeno evolving fuzzy models. *Proc. 2007 IEEE Intern. Conf. on Systems, Man, and Cybernetics*, 7-10 Oct., 2007, Montreal, Canada, ISBN 1-4244-0991-8/07, pp.3305-3310

Mahalec, V. and Sanchez, Y. (2012). Inferential monitoring and optimization of crude separation units via hybrid models. *Computers & Chemical Engineering*, 45, pp. 15-26.

Napoli G. and Xibilia M.G. (2011). Soft sensor design for a topping process in the case of small data sets. *Computers and Chemical Engineering*, Vol. 35. No. 11, pp. 2447–2456.

Shang C., Yang F., Gao X. and Huang D. (2015) A comparative study on improved DPLS soft sensor models applied to a crude distillation unit. *IFAC-PapersOnLine*, Vol. 48. Issue 8, pp. 234–239.

Torgashov A. and Zmeu K. (2015). Nonparametric soft sensor evaluation for industrial distillation plant. *Computer Aided Chemical Engineering*. Vol. 37, pp. 1487-1492.

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