ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 7, pp. 628–643.

CONTROL IN TECHNICAL SYSTEMS

Developing a Multi-Output Soft Sensor Based on an Error Predictor for Quality Control of the Target Product of a Complex Distillation Column

D. V. Shtakin[*,a], **A. A. Plotnikov**[*,b], **O. Yu. Snegirev**[*,c], **and A. Yu. Torgashov**[*,d]

Institute of Automation and Control Processes, Far Eastern Branch, Russian Academy of Sciences, Vladivostok, Russia
e-mail: [a] *shtakin@iacp.dvo.ru,* [b] *plotnikov_aa@iacp.dvo.ru,*
[c] *snegirevoleg@iacp.dvo.ru,* [d] *torgashov@iacp.dvo.ru*

Received January 31, 2025
Revised April 30, 2025
Accepted May 5, 2025

Abstract—This paper addresses the problem of improving the accuracy of a multi-output soft sensor (SS). It is demonstrated that implementing a predictor for the vector time series allows for consideration of the dynamic interdependence of process components, thereby enhancing the SS accuracy. The development of a multi-output error predictor is performed using vector autoregressive models and a set of autoregressive distributed lag models, with their optimal structures and parameters determined by numerical methods. The proposed approach to developing a multi-output SS is compared with traditional methods based on the sequential development of single-output SS within the quality control system of the target product (light diesel fraction) of a complex industrial distillation column. The effectiveness of the proposed approach is also demonstrated for the class of adaptive SS.

Keywords: multi-output soft sensor, multi-output error predictor, vector autoregressive model, autoregressive distributed lag model, adaptation, complex distillation column

DOI: 10.31857/S0005117925070038

1. INTRODUCTION

Soft sensors (SS) are key elements of modern process control systems, responsible for providing feedback on the quality of the final product [1]. Fractional distillation (or fractionation) is one of the most common and energy-intensive continuous technological processes in oil refining and the petrochemical industry. In industrial fractionating columns, crude oil feedstock is separated into various fractions of petroleum products. Typically, the separation occurs in complex columns featuring several intermediate circulating refluxes and the withdrawal of multiple products. The target, or most valuable, product is identified, and stabilizing its quality provides significant economic benefits for the production facility. Important quality indicators for the target product, such as the light diesel fraction of a hydrocracking process unit, are the key points of the fractional composition (initial boiling points, 10, 50, 90, 95%). These are determined in a plant laboratory at a frequency of once or twice per day, which is insufficient for real-time optimal control tasks. Therefore, the implementation of a SS primarily serves to estimate the quality indicators of the products at each control step. A SS typically includes statistical models that correlate hard-to-measure quality indicators (output variables) with easily accessible measurements (input variables) of process parameters (temperature, pressure, flow rates, etc.). The efficiency of process control largely depends on the accuracy of the models within the SS. In this regard, improving the accuracy of the SS is crucial and directly related to enhancing the efficiency of production process

control [2, 3]. It is worth noting that the task of identifying dependencies between the quality indicators of fractionating column products and a set of input variables has been extensively studied for the case of a single-output variable [4–8]. Methods for developing SS based on neural networks have also become widespread. Study [9] is devoted to the development of a dynamic soft sensor based on a convolutional neural network that accounts for changes in the characteristics of the analyzed process over time. The use of deep learning methods for developing SS is also known [10]. The work [11] focuses on a neural network training method combining backpropagation with partial least squares. The key advantage of neural network methods is their ability to approximate nonlinear dependencies with high accuracy on the training sample; however, significant errors can be observed on the test sample.

It should be noted that the task of developing a multi-output SS (MOSS) is highly relevant. In comparison with traditional single-output models, multi-output regression analysis better explains the dependency between the evaluated variables and the input data, due to the correlation among the output variables when considering complex relationships between the input and output data [12]. The problem of developing a MOSS can be addressed by developing new methods and models, as in [13], which describes a method for developing a multi-output SS based on multi-output tree chaining. The essence of this method lies in constructing an ensemble of trees that evaluate several output variables simultaneously. Existing methods can also be improved for developing a MOSS. For instance, [14] presents a modification of the Least-Squares Support Vector Machine method (LS-SVM) for developing models with multiple outputs and accounting for the nonlinear relationship between different output data, along with a proposed training method for such a model.

A widely adopted class of methods for improving SS accuracy for non-stationary plants involves the development of adaptive SS. The paper [15] is dedicated to the development of ensemble adaptive SS based on several Gaussian process models. In [16], the combination of several adaptation mechanisms to achieve the best SS accuracy is discussed. Methods for developing adaptive SS based on the Just-In-Time (JIT) approach are also common [17]. The most significant advantage of methods for developing adaptive soft sensors is the ability to account for process evolution over time, not only at the level of updating the model parameters but also at the level of updating the model structure, as shown in works [18, 19] devoted to SS based on the JIT approach combined with online selection of SS input variables. It should be noted that the effectiveness of the JIT approach decreases with less data in the training sample, which is a significant limitation in real-world conditions and prevents its use for developing MOSS in the presence of significant data gaps or small training samples.

There is also a class of methods for improving the accuracy of SS through bias updates (BU) of the regression model. In [20], a strategy for updating the bias term of the regression model is described, based on a Bayesian approach that considers the mean value and standard deviation of the SS error. A weighted strategy for BU of the regression model is known from [21]. This strategy is based on considering the previous values of the regression model's bias term.

The common drawback of the above methods and approaches is their failure to account for the mutual influence between the evaluated quality indicators, which in some cases can be of significant importance. There are a number of works treating the task of BU as the task of predicting the SS error, which allows accounting for the dynamic influence of the error on SS accuracy. For example, in [22], the bias term update strategy for the MOSS regression model is considered as the task of predicting the soft sensor error using a moving-average autoregressive model. The work [23] is devoted to the development of a predictive filter for adaptive SS. It should be noted that the works [19–22] do not account for the mutual influence of correlated SS errors. In [24], the problem of predicting the vector of cross-correlated SS errors is solved using analytical methods. However,

the choice of the predictor model order from auto- and cross-spectral forms does not allow for structural adaptation of the predictor in online mode at each discrete time step, and also increases the complexity of developing SS for the case of n-outputs.

This paper proposes a method for developing a MOSS with a predictor, distinguished by taking into account the interdependence of errors using vector autoregressive (VAR) models [25] or autoregressive distributed lag (ARDL) models [26] These models are developed using numerical methods, which allows for both structural and parametric adaptation of the SS error predictor at each discrete moment in time in a closed loop of the quality control system for the target product of a continuous technological process.

2. PLANT DESCRIPTION AND PROBLEM STATEMENT

The object of the study is a complex distillation column of a hydrocracking process unit for separating the feed into the following products: gasoline fraction (GF), kerosene fraction (KF), light diesel fraction (LDF), heavy diesel fraction (HDF), and hydrocracking residue (HR). LDF is subsequently used as a component of arctic diesel fuel (ADF). Figure 1 shows the flow chart of the distillation process and the control system with a controller based on the predictive model (PM) for the LDF fractional composition. The following conventions are used: K1 – complex distillation

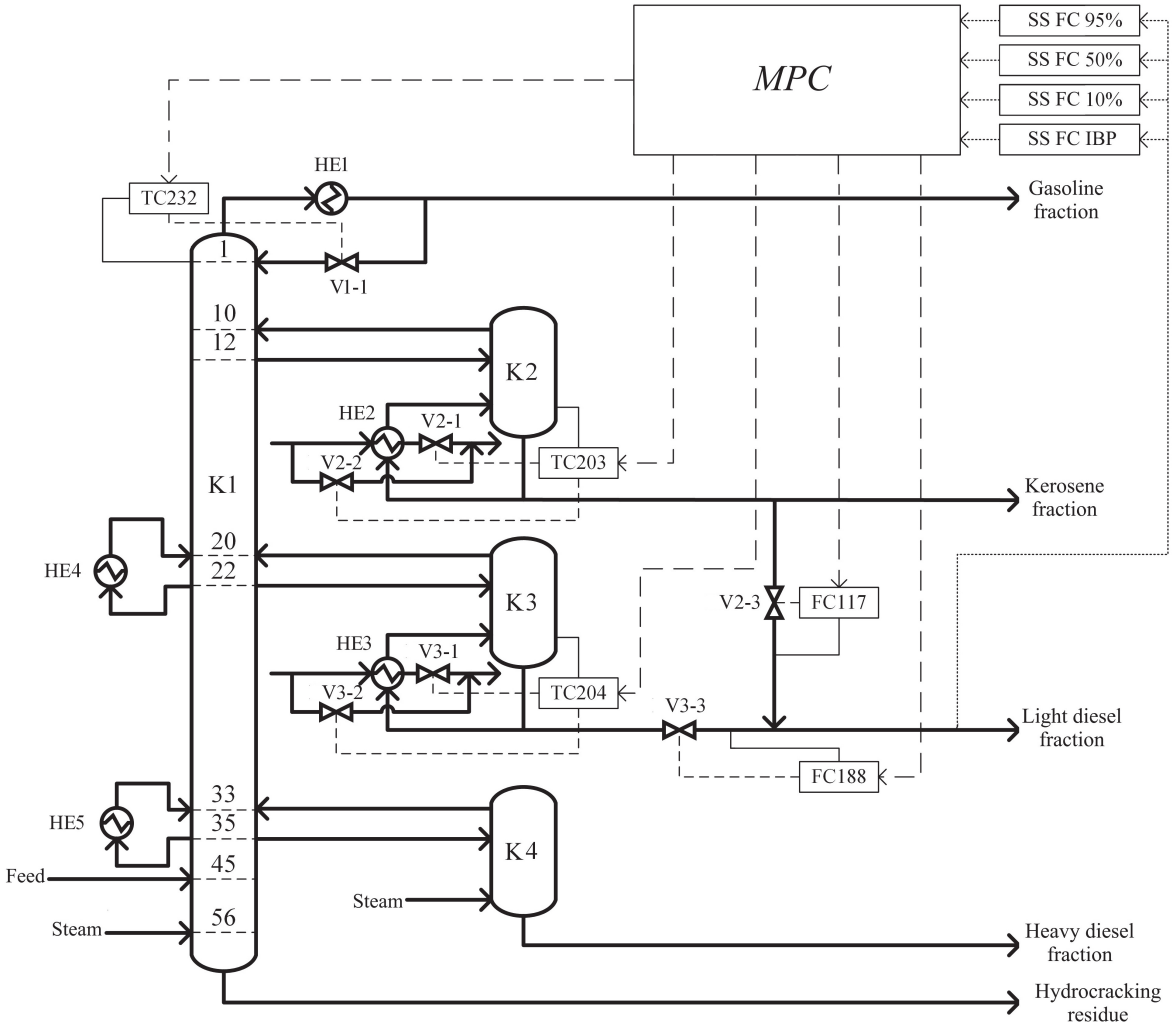

Fig. 1. Industrial plant scheme and control system.

column; K2 – kerosene fraction stripping column; K3 – diesel fraction stripping column; K4 – heavy diesel fraction stripping column; HE1 – gasoline fraction vapor condenser; HE2 – reboiler of column K2, HE3 – reboiler of column K3, HE4 – cooler of upper circulation reflux; HE5 – cooler of lower circulation reflux; V1-1 – valve on bypass of HE1; V2-1 – valve on outlet of HE2; V2-2 – valve on bypass of HE2; V2-3 – valve on flow line of KF for mixing with LDF; V3-1 – valve on outlet of HE3; V3-2 – valve on bypass of HE3; V3-3 – valve on flow line of LDF for mixing with KF. Table 1 shows a fragment of matrix of gain factors of process object. Table 2 shows the input variables of the fractional composition (FC) of the LDF used in industrial conditions: FC initial boiling point (IBP), 10, 50, 95%.

Table 1. Fragment of the gain factors matrix of an industrial plant

SS	TC232 (K1 top temp.)	TC203 (K2 temp.)	TC204 (K3 temp.)	FC117 (KF rate for mixing with LDF)	FC188 (LDF rate for mixing with KF)
FC IBP	2,2	2,4	2,7	0,5	
FC 10%	1,1	1,2	2,3	0,35	
FC 50%				0,2	1,2
FC 90%					1,7

Table 2. MOSS input variables

No.	Designation	Description	Unit of meas.
1	P131	K1 top pressure	MPa
2	P124	K1 bot pressure	MPa
3	FC105	K1 reflux rate	m^3/h
4	FC106	K1 top pumparound stream	m^3/h
5	FC107	K1 bottom pumparound stream	m^3/h
6	FC116	KF product rate	m^3/h
7	FC117	KF rate for mixing with LDF	m^3/h
8	FC188	LDF rate for mixing with KF	m^3/h
9	TC232	K1 top temperature	°C
10	T233	Return upper product temperature from K2 to K1	°C
11	T234	Return upper product temperature from K3 to K1	°C
12	T235	Return upper product temperature from K4 to K1	°C
13	T238	Sidestream temperature from K1 to K2	°C
14	T239	Sidestream temperature from K1 to K3	°C
15	T240	Sidestream temperature from K1 to K4	°C
16	T242	Temperature above the feedstock input zone in K1	°C
17	T247	Temperature at the outlet of HE2 into K2	°C
18	T248	Temperature at the outlet from K2 to HE2	°C
19	T253	Temperature at the outlet from HE3 to K3	°C
20	T254	Temperature at the outlet from K3 to HE3	°C
21	TC275	Temperature of bottom pumparound to K1	°C
22	TC279	Temperature of top pumparound to K1	°C
23	FY098	K1 feed rate	m^3/h

The problem of estimating the parameters of the MOSS and constructing a model of correlated error predictor of the SS is considered. In general, the MOSS model with a correlated error predictor

(MOSS with a CEP) has the following form:

$$\hat{\boldsymbol{Y}}_t^* = K\boldsymbol{X}_t + \tilde{\boldsymbol{b}} + \Phi\boldsymbol{E}_t^{past}, \tag{1}$$

where $\hat{\boldsymbol{Y}}_t^* = \begin{pmatrix}\hat{y}_{1,t}^* & \dots & \hat{y}_{j,t}^* & \dots & \hat{y}_{N,t}^*\end{pmatrix}^{\mathrm{T}}$ is a vector of estimated values of quality indicators; $K = \begin{pmatrix} k_{1,1} & \dots & k_{1,j} & \dots & k_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ k_{h,1} & \dots & k_{h,j} & \dots & k_{h,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ k_{N,1} & \dots & k_{N,j} & \dots & k_{N,m} \end{pmatrix}$ is a MOSS coefficient matrix; $\boldsymbol{X}_t = \begin{pmatrix}x_{1,t} & \dots & x_{j,t} & \dots & x_{m,t}\end{pmatrix}^{\mathrm{T}}$ is MOSS input variables vector; $\tilde{\boldsymbol{b}} = \begin{pmatrix}b_1 & \dots & b_h & \dots & b_N\end{pmatrix}^{\mathrm{T}}$ is a MOSS free term vector; m is a number of MOSS input variables; N is a number of MOSS output variables; $\Phi = \begin{pmatrix} \phi_{1,1|1} & \cdots & \phi_{1,1|\alpha} & \cdots & \phi_{1,N|1} & \cdots & \phi_{1,N|\alpha} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ \phi_{N,1|1} & \cdots & \phi_{N,1|\alpha} & \cdots & \phi_{N,N|1} & \cdots & \phi_{N,N|\alpha} \end{pmatrix}$ is a coefficient matrix of MOSS correlated error predictor; $\phi_{i,j|\alpha}$ is αth MOSS correlated error predictor coefficient for considering of the influence between the jth and ith components of the multivariable error series; $\boldsymbol{E}_t^{past} = \begin{pmatrix}e_{1,t-1} & \dots & e_{1,t-\alpha} & \dots & e_{N,t-1} & \dots & e_{N,t-\alpha}\end{pmatrix}^{\mathrm{T}}$ is an error predictor input vector; α is an error predictor order; $e_{i,t} = y_{i,t} - \left(b_i + \left[K\right]_i \boldsymbol{X}_t\right)$ is a MOSS estimation error of the ith quality indicator at time t; $\left[\cdot\right]_i$ is an ith matrix row.

MOSS with a preliminary whitening correlated error predictor (i.e. with passage through a filter that equalizes the spectral density) of an error series using the AR model (MOSS with a WCEP) is described by the expression:

$$\hat{\boldsymbol{Y}}_t^{**} = K\boldsymbol{X}_t + \tilde{\boldsymbol{b}} + W\boldsymbol{E}_t^{past} + \tilde{\Phi}\boldsymbol{V}_t^{past}, \tag{2}$$

where $\boldsymbol{E}_t^{past} = \begin{pmatrix}e_{1,t-1} & \dots & e_{1,t-q} & \dots & e_{N,t-1} & \dots & e_{N,t-q}\end{pmatrix}^{\mathrm{T}}$ is a MOSS error vector; $W = \begin{pmatrix} w_{1,1|1} & \dots & w_{1,1|q} & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & w_{2,2|1} & \dots & w_{2,2|q} & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & w_{N,N|1} & \dots & w_{N,N|q} \end{pmatrix}$ is a whitening predictor coefficient matrix with order q; $w_{i,j|q}$ is a qth whitening predictor coefficient for considering the influence of the jth element on the ith multidimensional error series; $\tilde{\Phi} = \begin{pmatrix} \tilde{\phi}_{1,1|1} & \cdots & \tilde{\phi}_{1,1|\tilde{\alpha}} & \cdots & \tilde{\phi}_{1,N|1} & \cdots & \tilde{\phi}_{1,N|\tilde{\alpha}} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\phi}_{N,1|1} & \cdots & \tilde{\phi}_{N,1|\tilde{\alpha}} & \cdots & \tilde{\phi}_{N,N|1} & \cdots & \tilde{\phi}_{N,N|\tilde{\alpha}} \end{pmatrix}$ is a whitening correlated error predictor coefficient matrix with order $\tilde{\alpha}$; $\boldsymbol{V}_t^{past} = \begin{pmatrix}v_{1,t-1} & \dots & v_{1,t-\tilde{\alpha}} & \dots & v_{N,t-1} & \dots & v_{N,t-\tilde{\alpha}}\end{pmatrix}^{\mathrm{T}}$ is an error predictor input vector; $v_{i,t} = e_{i,t} - \left[W\right]_i \boldsymbol{E}_t^{past}$ is a whitening error for ith quality indicator MOSS error.

The required parameters matrix θ_1 for (1) and θ_2 for (2) have the following form:

$$\theta_1 = \begin{pmatrix} K & \tilde{\boldsymbol{b}} & \Phi \end{pmatrix}, \tag{3}$$

$$\theta_2 = \begin{pmatrix} K & \tilde{\boldsymbol{b}} & W & \tilde{\Phi} \end{pmatrix}. \tag{4}$$

The determination of estimates of unknown parameters θ_z $(z = 1, 2)$ is carried out by solving an optimization problem for a given quality functional J:

$$\hat{\theta}_z = \underset{\theta_z \in \mathbb{R}}{\operatorname{argmin}} J(\theta_z). \tag{5}$$

The MOSS mean square error is considered as the criterion J:

$$J = \frac{1}{n} \sum_{i=1}^{n} \left|\left| \boldsymbol{Y}_i - \hat{\boldsymbol{Y}}_i(\theta_z) \right|\right|^2, \tag{6}$$

where $\boldsymbol{Y}_i$ and $\hat{\boldsymbol{Y}}_i$ are the vectors of measured and estimated values of the target product quality indicators, respectively; n is an observations number in a sample; $||\cdot||$ is an Euclidean norm.

3. DEVELOPING A MULTI-OUTPUT SOFT SENSOR

The diagram of the proposed multi-output soft sensor with a predictor as part of the control system is shown in Fig. 2, where MPC is a model predictive controller; $G(U)$ is an industrial plant; U is the control actions vector; $\hat{\boldsymbol{Y}}_t$ is a SS estimate (no consideration of error predictor); $\boldsymbol{e}_t$ is a vector of SS errors; $\boldsymbol{v}_t$ is a vector of whitening errors; $\hat{\boldsymbol{e}}_t$ is an error prediction; $\hat{\boldsymbol{v}}_t$ is a whitening error prediction; B is the back shift operator. The main advantage of the proposed multi-output soft sensor lies in accounting for the dynamic influence of correlated errors in evaluating the quality indicators of the target product of a complex fractionating column. It should be noted that the proposed predictor is an add-on to any model within the multi-output soft sensor and can be used in combination with other models, such as neural network models that form the basic component of the yield estimation $\hat{\boldsymbol{Y}}_t$.

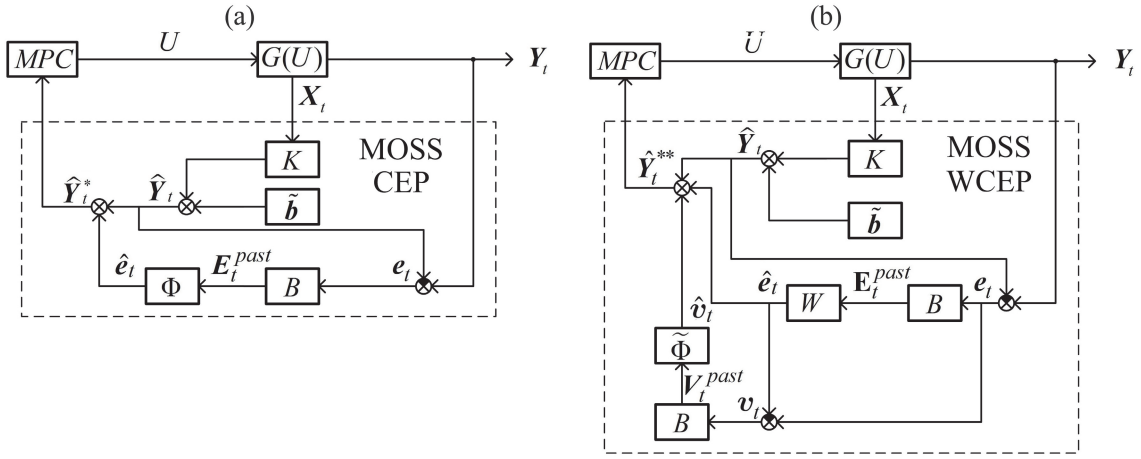

Fig. 2. Structural scheme of the control system with MOSS with a predictor of correlated errors: (a) without preliminary whitening; (b) with preliminary whitening.

In expressions (1) and (2) the first two terms $(K\boldsymbol{X}_t + \tilde{\boldsymbol{b}})$ represent the multiple linear regression (MLR) model, which is widely used in industrial settings [27].

The search for a solution to the optimization problem (5) consists of several stages. At the first stage, the first two blocks of the matrix θ_z ($\hat{K}$ and $\hat{\tilde{\boldsymbol{b}}}$), which are responsible for the model within the MOSS, are determined:

$$\{\hat{K}, \hat{\tilde{\boldsymbol{b}}}\} = \underset{K, \tilde{\boldsymbol{b}} \in \mathbb{R}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} \left|\left| \boldsymbol{Y}_i - \left(K\boldsymbol{X}_i + \tilde{\boldsymbol{b}}\right) \right|\right|^2. \tag{7}$$

At the second stage, the MOSS error predictor is developed, i.e. the third block of matrices θ_z is determined. At this stage, for the multi-output soft sensor with a correlated error predictor (CEP), $\hat{\Phi}$ is determined taking into account the cross-correlations of the MOSS errors:

$$\hat{\Phi} = \underset{\Phi\in\mathbb{R}}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n}\left|\left|\boldsymbol{e}_i - \Phi\boldsymbol{E}_i^{past}\right|\right|^2. \tag{8}$$

For the MOSS with a whitening correlated error predictor (WCEP), $\hat{W}$ is determined, i.e. the parameters of the "whitening" predictor are determined without taking into account the cross-correlations of the MOSS errors:

$$\hat{W} = \underset{W\in\mathbb{R}}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n}\left|\left|\boldsymbol{e}_i - W\boldsymbol{E}_i^{past}\right|\right|^2. \tag{9}$$

Since the interdependence of the error series is not accounted for in the MOSS with a WCEP in the previous stage, a third stage becomes necessary, during which the optimal parameters of the "whitening" error predictor $\hat{\tilde{\Phi}}$ are determined, taking their cross-correlation into account:

$$\hat{\tilde{\Phi}} = \underset{\tilde{\Phi}\in\mathbb{R}}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n}\left|\left|\boldsymbol{v}_i - \tilde{\Phi}\boldsymbol{V}_i^{past}\right|\right|^2. \tag{10}$$

Well-known autoregressive (AR) models can be used as a model of the "whitening" predictor W. Typically, single-output AR models make it possible to significantly increase the SS accuracy. However, this approach has its drawbacks, namely, it does not account for the correlation of SS errors for interconnected quality indicators.

The MOSS with a CEP described in (1) is further considered for the case $N = 2$, and its parameters θ_1 are found using (7) and (8). In this paper, the CEP for the MOSS is implemented based on the following VAR model:

$$\begin{pmatrix}\hat{e}_{1,t} & \hat{e}_{2,t}\end{pmatrix}^{\mathrm{T}} = \Omega\boldsymbol{E}_t^{(q)}, \tag{11}$$

where $\Omega = \begin{pmatrix}\omega_{1,1|1} & \dots & \omega_{1,1|j} & \dots & \omega_{1,1|q} & \omega_{1,2|1} & \dots & \omega_{1,2|j} & \dots & \omega_{1,2|q}\\ \omega_{2,1|1} & \dots & \omega_{2,1|j} & \dots & \omega_{2,1|q} & \omega_{2,2|1} & \dots & \omega_{2,2|j} & \dots & \omega_{2,2|q}\end{pmatrix}$ is the coefficients matrix of the VAR model; $\boldsymbol{E}_t^{(q)} = \begin{pmatrix}e_{1,t-1} & \dots & e_{1,t-j} & \dots & e_{1,t-q} & e_{2,t-1} & \dots & e_{2,t-j} & \dots & e_{2,t-q}\end{pmatrix}^{\mathrm{T}}$ is the input vector of the VAR model; q is VAR model order.

This paper also considers another method for implementing the CEP using an ARDL model [26] of the following form:

$$\hat{e}_{1|t} = \boldsymbol{a}_1 E_{1|t}^{endo(1)} + \boldsymbol{l}_1 E_{2|t}^{exo(1)},\ \hat{e}_{2|t} = \boldsymbol{a}_1 E_{2|t}^{endo(2)} + \boldsymbol{l}_2 E_{1|t}^{exo(2)}, \tag{12}$$

where $\boldsymbol{a}_1 = \begin{pmatrix}a_{1,1} & \dots & a_{1,j} & \dots & a_{1,A_1}\end{pmatrix}$ and $\boldsymbol{a}_2 = \begin{pmatrix}a_{2,1} & \dots & a_{2,j} & \dots & a_{2,A_2}\end{pmatrix}$ are the regression coefficients of the endogenous component of the model error series $\hat{e}_1$ and $\hat{e}_2$; $\boldsymbol{l}_1 = \begin{pmatrix}l_{1,1} & \dots & l_{1,j} & \dots & l_{1,L_1}\end{pmatrix}$ and $\boldsymbol{l}_2 = \begin{pmatrix}l_{2,1} & \dots & l_{2,j} & \dots & l_{2,L_2}\end{pmatrix}$ are the regression coefficients of the exogenous component of the model error series $\hat{e}_1$ and $\hat{e}_2$; $E_{1|t}^{endo(1)} = \begin{pmatrix}e_{1,t-1} & \dots & e_{1,t-j} & \dots & e_{1,t-A_1}\end{pmatrix}^{\mathrm{T}}$ and $E_{2|t}^{endo(2)} = \begin{pmatrix}e_{2,t-1} & \dots & e_{2,t-j} & \dots & e_{2,t-A_2}\end{pmatrix}^{\mathrm{T}}$ are the input vectors of the endogenous component of the model error series $\hat{e}_1$ and $\hat{e}_2$; $E_{2|t}^{exo(1)} = \begin{pmatrix}e_{2,t-d_1} & \dots & e_{2,t-d_1-j} & \dots & e_{2,t-d_1-L_1+1}\end{pmatrix}^{\mathrm{T}}$ and

$E_{1|t}^{exo(2)} = \begin{pmatrix} e_{1,t-d_2} & \dots & e_{1,t-d_2-j} & \dots & e_{1,t-d_2-L_2+1} \end{pmatrix}^{\mathrm{T}}$ are the input vectors of the exogenous component of the model error series $\hat{e}_1$ and $\hat{e}_2$; A_1 and A_2 are the regression orders of the endogenous component of the model error series $\hat{e}_1$ and $\hat{e}_2$; d_1 and d_2 are the lags of the exogenous component of the model error series $\hat{e}_1$ and $\hat{e}_2$; L_1 and L_2 are the regression orders of the exogenous component of the model error series $\hat{e}_1$ and $\hat{e}_2$.

The ARDL model is a more flexible tool for accounting for the correlation of SS errors in comparison with the VAR model, since it makes it possible to set different regression orders for endogenous and exogenous components, as well as the lag value for the exogenous component. The proposed approach to developing a SS includes the possibility of preliminary whitening (2) of the SS error series (correction of the SS error using the AR model) before applying the VAR or ARDL models (MOSS with a WCEP). The parameters of the MOSS with a WCEP are found using (7), (9), (10). In this case, the VAR and ARDL models are used to forecast not the SS errors but the whitening errors:

$$v_{1,t} = e_{1,t} - \hat{e}_{1,t},\ v_{2,t} = e_{2,t} - \hat{e}_{2,t}. \tag{13}$$

Therefore, the VAR (11) and ARDL (12) models take the form (14) and (15), respectively:

$$\begin{pmatrix} \hat{v}_{1,t} & \hat{v}_{2,t} \end{pmatrix}^{\mathrm{T}} = \Omega \boldsymbol{V}_t^{(q)}, \tag{14}$$

where $\boldsymbol{V}_t^{(q)} = \begin{pmatrix} v_{1,t-1} & \dots & v_{1,t-j} & \dots & v_{1,t-q} & v_{2,t-1} & \dots & v_{2,t-j} & \dots & v_{2,t-q} \end{pmatrix}^{\mathrm{T}}$ is a VAR model input vector.

$$\hat{v}_{1|t} = \boldsymbol{a}_1 V_{1|t}^{endo(1)} + \boldsymbol{l}_1 V_{2|t}^{exo(1)},\ \hat{e}_{2|t} = \boldsymbol{a}_2 V_{2|t}^{endo(2)} + \boldsymbol{l}_2 V_{1|t}^{exo(2)}, \tag{15}$$

where $V_{1|t}^{endo(1)} = \begin{pmatrix} v_{1,t-1} & \dots & v_{1,t-j} & \dots & v_{1,t-A_1} \end{pmatrix}^{\mathrm{T}}$ and
$V_{2|t}^{endo(2)} = \begin{pmatrix} v_{2,t-1} & \dots & v_{2,t-j} & \dots & v_{2,t-A_2} \end{pmatrix}^{\mathrm{T}}$ are the input vectors of the endogenous component of the model error series $\hat{v}_1$ and $\hat{v}_2$;
$V_{2|t}^{exo(1)} = \begin{pmatrix} v_{2,t-d_1} & \dots & v_{2,t-d_1-j} & \dots & v_{2,t-d_1-L_1+1} \end{pmatrix}^{\mathrm{T}}$ and
$V_{1|t}^{exo(2)} = \begin{pmatrix} v_{1,t-d_2} & \dots & v_{1,t-d_2-j} & \dots & v_{1,t-d_2-L_2+1} \end{pmatrix}^{\mathrm{T}}$ are the input vectors of the exogenous component of the model error series $\hat{v}_1$ and $\hat{v}_2$.

As a rule, the order of an AR model for a time series is determined by visual analysis of its correlogram or autospectrum. Despite the wide prevalence of this approach, in the present work, a numerical method is used to determine the AR model order, based on the analysis of the partial autocorrelation estimate $\hat{\psi}_k$, determined by the Yule-Walker equation:

$$\hat{\Psi}^{(k)} = P_k^{-1} \rho^{(k)}, \tag{16}$$

where k is the lag; $\hat{\Psi}^{(k)} = \begin{pmatrix} \hat{\psi}_1 & \dots & \hat{\psi}_j & \dots & \hat{\psi}_k \end{pmatrix}^{\mathrm{T}}$ is an AR model coefficients vector with order k; $\rho^{(k)} = \begin{pmatrix} r_1 & \dots & r_j & \dots & r_k \end{pmatrix}^{\mathrm{T}}$ is an autocorrelation vector of estimates before lag k; $P_k = \begin{pmatrix} 1 & r_1 & r_2 & \dots & r_{k-1} \\ r_1 & 1 & r_1 & \dots & r_{k-2} \\ r_2 & r_1 & 1 & \dots & r_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{k-1} & r_{k-2} & r_{k-3} & \dots & 1 \end{pmatrix}$ is the Toeplitz matrix based on $\rho^{(k)}$. The equation (16) links the coefficients of the AR model and the values of the autocorrelation estimate [28]. The use of

numerical methods allows for the automation of the AR model order determination process and the implementation of structural predictor adaptation. The method for determining the AR model order based on the analysis of the estimate $\hat{\psi}_k$ is as follows. The value of the partial autocorrelation estimate at lag k is set to zero if $\left|\hat{\psi}_k\right| < 2\hat{\sigma}_{\hat{\psi}_k}$, where $\hat{\sigma}_{\hat{\psi}_k}$ is the standard error of the partial autocorrelation estimates, defined as $\hat{\sigma}_{\hat{\psi}_k} = \frac{1}{n}$. Then, the maximum lag value k for which $\hat{\psi}_k \neq 0$ is used as the order p.

Partial autocorrelation is also used to determine the order q of the VAR model. Using the approach described above, the orders of the AR models for each error series included in the vector model are determined, after which the smallest of the previously obtained AR model orders is chosen as the vector model order q. This approach helps to avoid including redundant variables in the model.

To determine the values of A and L for the ARDL model, the method from [29] can be used, which is based on identifying the instability of the figures of the model parameter estimates as the model structure becomes more complex. A sequential increase in the model order is performed as long as the following condition is met:

$$\sum_{i=0}^{q-1} \operatorname{sgn} b_i^{(q-1)} \operatorname{sgn} b_i^{(q)} = q - 1, \tag{17}$$

where q is order of model: $\hat{x}_{t+1} = b_0^{(q)} x_t + b_1^{(q)} x_{t-1} + \ldots + b_q^{(q)} x_{t-q} = \sum_{i=0}^{q} b_i^{(q)} x_{t-i}$. The orders A and L of the ARDL model are determined in a similar manner. The methods presented above for determining the orders of AR, VAR, and ARDL models are used to implement the structural adaptation of the error predictors in online-mode, i.e., at each discrete time step.

The coefficients of the AR model are determined using the Yule-Walker equation:

$$\mathbf{W}_{YW} = P_\tau^{-1} \rho^{(\tau)}, \tag{18}$$

where $\mathbf{W}_{YW} = \begin{pmatrix} w_1 & \ldots & w_j & \ldots & w_p \end{pmatrix}^{\mathrm{T}}$ is the coefficients vector of AR model order p; $\rho^{(\tau)} = \begin{pmatrix} r_1 & \ldots & r_j & \ldots & r_\tau \end{pmatrix}^{\mathrm{T}}$ is the vector of autocorrelation estimates before the lag $\tau = p$; P_τ is the Toeplitz matrix based on $\rho^{(\tau)}$ determined as in the (16). The tasks (8) and (10) for finding the parameters of the VAR and ARDL models are solved using the least squares method.

4. EXPERIMENTAL VALIDATION AND DISCUSSION OF RESULTS

The proposed approach for developing a multi-output soft sensor with a predictor was tested as follows. To implement the error predictors that account for the correlation of soft sensor errors for the fractional composition (FC), pairs are considered: the first pair is the initial boiling point of fractional composition (IBP FC) and 10% FC; the second pair is 50% FC and 95% FC. Figures 3 and 4 show the presence of a dynamic relationship between the selected pairs of soft sensor error time series using smoothed squared coherency spectral estimate smoothed sample estimates [30] of the squared coherence spectrum :

$$\overline{\mathrm{K}}_{12}^2(\lambda) = \frac{\overline{A}_{12}^2(\lambda)}{\overline{C}_{11}(\lambda)\overline{C}_{22}(\lambda)}, \quad 0 \leqslant \lambda \leqslant F, \tag{19}$$

where $\overline{A}_{12}(\lambda) = \sqrt{\overline{L}_{12}^2(\lambda) + \overline{Q}_{12}^2(\lambda)}$, $0 < \lambda < F$ is the smoothed cross amplitude spectral estimate; $\overline{L}_{12}(\lambda) = 2\left(\tilde{l}_{12}(0) + 2\sum\limits_{\kappa=1}^{L-1} \tilde{l}_{12}(\kappa)\beta(\kappa)\cos\frac{\pi\kappa\lambda}{F}\right)$, $0 \leqslant \lambda \leqslant F$ and $\overline{Q}_{12}(\lambda) = 4\sum\limits_{\kappa=1}^{L-1} \tilde{q}_{12}(\kappa)\beta(\kappa)\sin\frac{\pi\kappa\lambda}{F}$,

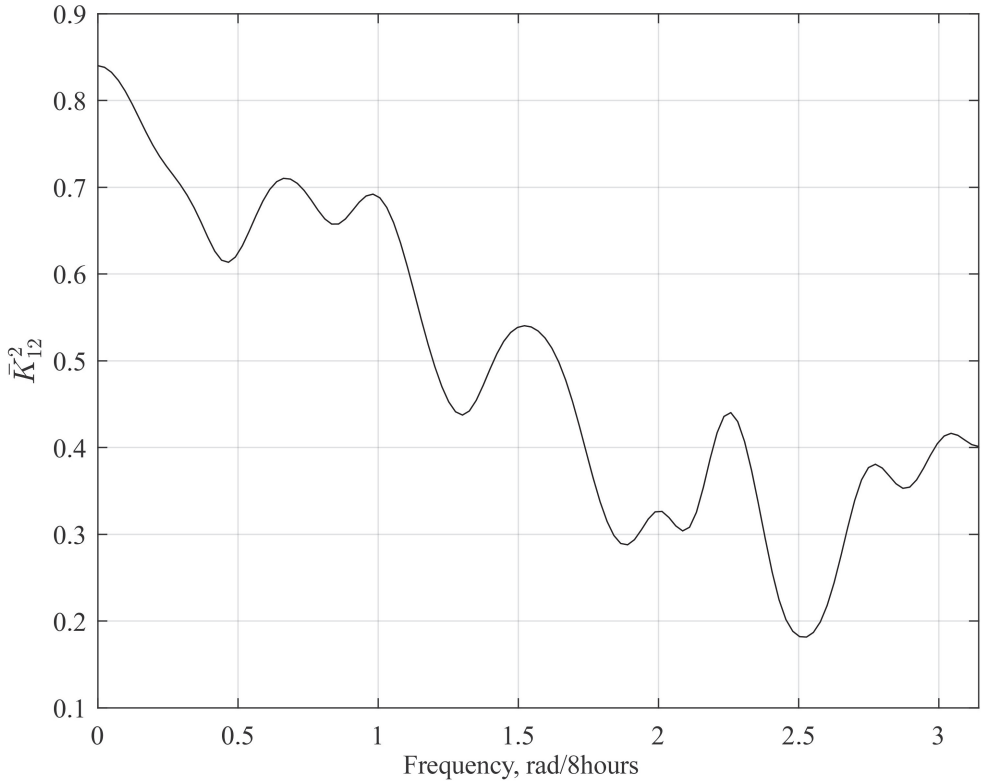

Fig. 3. Smoothed sample estimate of the quadrature coherency spectrum of SS IBP FC and FC 10% errors on the training sample.

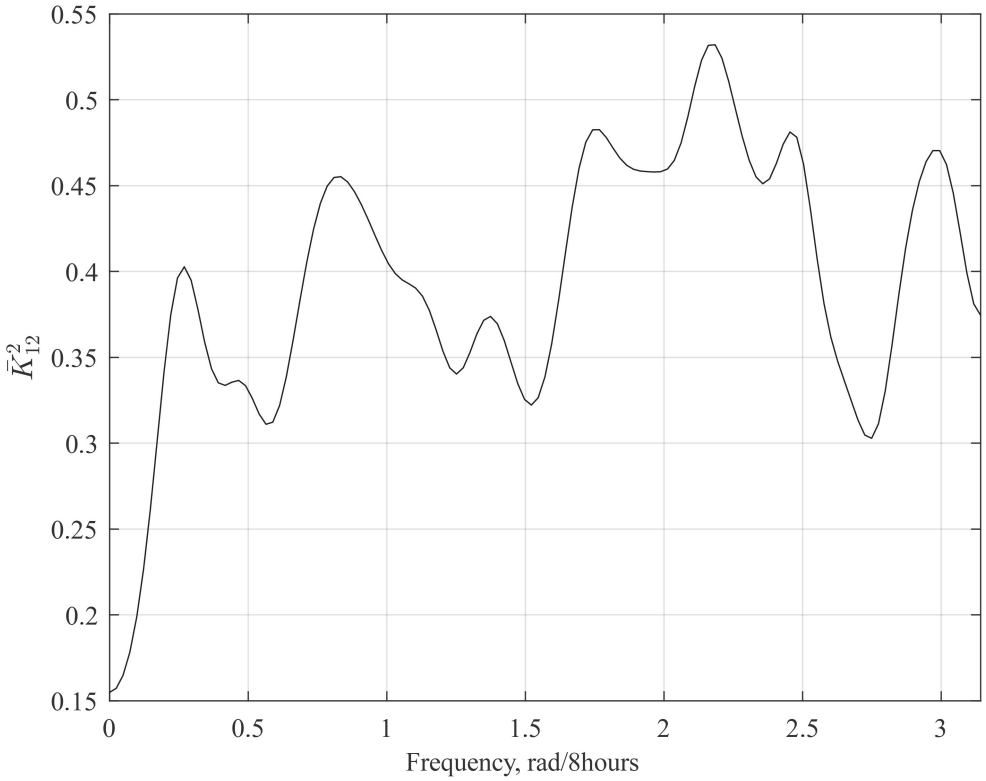

Fig. 4. Smoothed sample estimate of the quadrature coherency spectrum of SS FC 50% and FC 95% errors on the training sample.

$1 \leqslant \lambda \leqslant F-1$ are the smoothed co- and quadrature spectral estimates, respectively; $\tilde{l}_{12}(\kappa) = 0.5\,(c_{12}(\kappa) + c_{12}(-\kappa)),\ 0 \leqslant \kappa \leqslant L-1$ are the even cross covariance function estimates; $\tilde{q}_{12}(\kappa) = 0.5\,(c_{12}(\kappa) - c_{12}(-\kappa)),\ 0 \leqslant \kappa \leqslant L-1$ are the odd cross covariance function estimates; $c_{12}(\kappa) = \frac{1}{n}\sum\limits_{t=1}^{n-\kappa}(e_{1,t} - \overline{e}_1)(e_{2,t+\kappa} - \overline{e}_2),\ c_{12}(-\kappa) = \frac{1}{n}\sum\limits_{t=1}^{n-\kappa}(e_{1,t+\kappa} - \overline{e}_1)(e_{2,t} - \overline{e}_2),\ 0 \leqslant \kappa \leqslant L-1$ is the cross covariance function estimate; $\overline{C}_{11}(\lambda) = 2\left(c_{11}(0) + 2\sum\limits_{\kappa=1}^{L-1} c_{11}(\kappa)\,\beta(\kappa)\cos\frac{\pi\kappa\lambda}{F}\right), 0 \leqslant \lambda \leqslant F$ is the smoothed spectral estimate; $\beta(\kappa) = \begin{cases} 1 - \frac{|\kappa|}{M}, & |\kappa| \leqslant M \\ 0, & |\kappa| > M \end{cases}$ is the Bartlett window; $c_{11}(\kappa) = \frac{1}{n}\sum\limits_{t=1}^{n-\kappa}(e_{1,t} - \overline{e}_1)(e_{1,t+\kappa} - \overline{e}_1), 0 \leqslant \kappa \leqslant L-1$ is the autocovariance function estimate; $\overline{e}_1 = \frac{1}{n}\sum\limits_{t=1}^{n} e_{1,t}$ is mean of the SS error; L is number of lags of used covariation functions; F is the maximum frequency of calculating spectral estimates; estimates $\overline{C}_{22}(\lambda)$ and $c_{22}(\kappa)$ are calculated similarly as $\overline{C}_{11}(\lambda)$ and $c_{11}(\kappa)$.

The amplitude values of the squared coherence spectrum shown in Fig. 3 are greater than 0.6 in the low-frequency region, indicating a close relationship between the soft sensor errors for IBP FC and 10% FC and low autocorrelation of the error series for these indicators. The graph in Fig. 4 reflects a relationship between the soft sensor errors for 50% FC and 95% FC, but there is significant autocorrelation of the soft sensor errors for one of the indicators in the low-frequency region, which is expressed by a low amplitude value of the smoothed sample estimate of the squared coherence spectrum, around 0.15.

The proposed MOSS is compared with an error predictor based on an AR model (which does not account for SS error correlation and is widely used in industry) through the procedure of bias update (BU) for the MLR model within the SS. The proposed approach is implemented for both adaptive (AdaptSS) and non-adaptive SS. The adaptation is implemented within the framework of the moving window (MW) method. Versions of the error predictors with structural adaptation (SA) and parametric adaptation (PA) are also implemented. PA is understood as the recalculation of the predictor parameters on the shifted MW without changing their order, which is determined by the methods described in the previous section during the first MW iteration. SA is understood as the calculation of the predictor parameters along with the determination of their order at each shift of the MW.

The total available sample contains 1382 observations. It is divided into a training sample with 967 observations and a test sample with 415 observations. For adaptive soft sensors and predictors, the size of the SW is the same as the size of the training sample for non-adaptive SS.

Tables 3 and 4 present the test results for non-adaptive and adaptive soft sensors, respectively. The coefficient of determination R_j^2 and the mean absolute error (MAE_j) were used as accuracy criteria for the jth quality indicator:

$$R_j^2 = 1 - \frac{\sum_{i=1}^{n}(y_{j,i} - \hat{y}_{j,i})^2}{\sum_{i=1}^{n}\left(y_{j,i} - \overline{y}_j\right)^2}, \tag{20}$$

$$\mathrm{MAE}_j = \frac{1}{n}\sum_{i=1}^{n}|y_{j,i} - \hat{y}_{j,i}|, \tag{21}$$

where $y_{j,i}$ – laboratory value of jth quality indicator; $\hat{y}_{j,i}$ – estimate of jth quality indicator; $\overline{y}_j = \frac{1}{n}\sum_{i=1}^{n} y_{j,i}$.

It should be noted that the model parameters were found according to criterion (6). Criteria (20) and (21) are used for comparative analysis due to their widespread use in monitoring and automation systems for industrial processes.

Table 3. Testing results for non-adaptive SS

Type of SS	Pair with correlated errors				Pair with correlated errors			
	IBP FC		10% FC		50% FC		95% FC	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
Industrial approaches								
SS MLR	0.250	13.590	0.506	10.177	0.824	4.735	0.274	7.302
SS BU	0.667	**9.023**	0.776	**6.808**	0.905	**3.293**	0.601	**5.105**
The proposed method in comparison with the AR model (non-adaptive)								
SS AR	0.872	5.108	0.851	4.974	0.928	2.606	0.375	6.418
SS ARDL	0.868	5.180	0.909	**3.731**	0.927	2.679	0.707	**4.047**
SS VAR	0.868	5.184	0.909	3.732	0.928	2.678	0.707	4.048
SS AR ARDL	0.873	**5.063**	0.839	5.166	0.928	**2.606**	0.397	6.298
SS AR VAR	0.872	5.108	0.851	4.975	0.928	2.607	0.375	6.416
The proposed method in comparison with the AR model (parametric adaptation)								
SS AR PA	0.887	4.558	0.829	5.373	0.933	2.443	0.372	6.378
SS ARDL PA	0.883	4.654	0.914	3.510	0.933	2.473	0.715	3.910
SS VAR PA	0.883	4.647	0.914	**3.507**	0.933	2.472	0.715	**3.909**
SS AR PA ARDL PA	0.888	**4.522**	0.819	5.552	0.933	2.450	0.395	6.269
SS AR PA VAR PA	0.887	4.551	0.829	5.375	0.933	**2.441**	0.373	6.375
The proposed method in comparison with the AR model (structural adaptation)								
SS AR SA	0.885	4.652	0.828	5.366	0.933	2.479	0.372	6.378
SS ARDL SA	0.883	4.654	0.914	3.510	0.933	2.473	0.715	3.910
SS VAR SA	0.883	4.607	0.914	**3.485**	0.933	2.472	0.715	3.909
SS AR SA ARDL SA	0.887	**4.565**	0.817	5.592	0.932	**2.469**	0.395	6.274
SS AR SA VAR SA	0.885	4.645	0.828	5.369	0.933	2.478	0.373	6.375

Table 4. Testing results for adaptive SS

Type of SS	Pair with correlated errors				Pair with correlated errors			
	IBP FC		10% FC		50% FC		95% FC	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
Industrial approaches								
AdaptSS MLR	0.549	10.642	0.659	8.522	0.892	3.401	0.494	5.579
AdaptSS BU	0.771	**7.504**	0.828	**5.899**	0.930	**2.566**	0.677	**4.251**
The proposed method in comparison with the AR model (non-adaptive)								
AdaptSS AR	0.877	4.904	0.885	4.131	0.938	2.279	0.545	5.328
AdaptSS ARDL	0.877	4.935	0.912	**3.694**	0.939	2.268	0.729	**3.657**
AdaptSS VAR	0.877	4.938	0.912	3.694	0.939	**2.267**	0.729	3.659
AdaptSS AR ARDL	0.879	**4.861**	0.877	4.268	0.938	2.272	0.560	5.225
AdaptSS AR VAR	0.877	4.905	0.885	4.131	0.938	2.280	0.546	5.325
The proposed method in comparison with the AR model (parametric adaptation)								
AdaptSS AR PA	0.879	4.857	0.885	4.137	0.939	2.247	0.545	5.334
AdaptSS ARDL PA	0.879	4.886	0.913	3.671	0.939	2.248	0.728	3.661
AdaptSS VAR PA	0.879	4.883	0.913	**3.669**	0.939	2.248	0.728	**3.660**
AdaptSS AR PA ARDL PA	0.880	**4.810**	0.877	4.286	0.938	**2.245**	0.561	5.217
AdaptSS AR PA VAR PA	0.879	4.854	0.885	4.137	0.939	2.246	0.545	5.332
The proposed method in comparison with the AR model (structural adaptation)								
AdaptSS AR SA	0.876	4.932	0.885	4.175	0.938	2.247	0.545	5.330
AdaptSS ARDL SA	0.879	4.886	0.913	3.671	0.939	2.248	0.728	3.661
AdaptSS VAR SA	0.882	**4.802**	0.915	**3.647**	0.939	2.248	0.728	**3.660**
AdaptSS AR SA ARDL SA	0.878	4.849	0.876	4.347	0.939	2.249	0.561	5.219
AdaptSS AR SA VAR SA	0.876	4.929	0.885	4.176	0.939	**2.246**	0.545	5.329

Based on Tables 3 and 4, for IBP FC estimation, the best results in terms of accuracy criteria are achieved using the MOSS with a WCEP of type (2). For the 10% FC estimation, the MOSS with a CEP of type (1) shows the best results. For the 95% FC, the MOSS with a CEP of type (1) also shows the best results. A high degree of similarity is noted in the results of the MOSS with a CEP of type (1) based on VAR and ARDL models, and their significant advantage over the SS with an error predictor based on an AR model is also noted.

According to Fig. 4 and Tables 3, 4, the minor effect from using the CEP for the 50% FC SS (in comparison with the CEP for other SSs) is associated with an insufficiently strong dependence of the 50% FC SS errors on the 95% FC SS errors, as well as the presence of significant serial correlation in the 50% FC SS errors.

Tables 5 and 6 show the percentage change in accuracy criteria for the best result relative to the baseline methods (SS, SS BU, SS AR, SS AR PA, SS AR SA), which do not account for error cross-correlation for non-adaptive and adaptive SS, respectively. According to Tables 5 and 6, the lowest mean absolute error (MAE) is most often achieved using a CEP based on a VAR model with parametric or structural adaptation (VAR PA, VAR SA are highlighted in bold in the tables).

Table 5. Change of the accuracy criteria for non-adaptive SS

Type of SS	Pair with correlated errors				Pair with correlated errors			
	IBP FC		10% FC		50% FC		95% FC	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
Basic method								
SS MLR	71.83%	66.73%	44.62%	65.76%	11.67%	48.41%	61.69%	46.47%
SS BU	24.93%	49.88%	15.10%	48.81%	3.08%	25.81%	15.92%	23.44%
SS AR	1.82%	11.47%	6.95%	29.94%	0.53%	6.26%	47.55%	39.10%
SS AR PA	0.11%	0.79%	9.32%	35.15%	0.00%	0.00%	47.90%	38.72%
SS AR SA	0.36%	2.79%	9.39%	35.07%	0.08%	1.48%	47.90%	38.72%
Best result								
	SS AR PA ARDL PA		**SS VAR SA**		SS AR PA		**SS VAR PA**	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
	0.888	4.522	0.914	3.485	0.933	2.443	0.715	3.909

Table 6. Change of the accuracy criteria for adaptive SS

Type of SS	Pair with correlated errors				Pair with correlated errors			
	IBP FC		10% FC		50% FC		95% FC	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
Basic method								
AdaptSS MLR	37.71%	54.88%	27.97%	57.20%	4.96%	33.99%	32.14%	34.39%
AdaptSS BU	12.53%	36.00%	9.43%	38.17%	0.92%	12.53%	6.99%	13.90%
AdaptSS AR	0.53%	2.08%	3.30%	11.70%	0.03%	1.51%	25.04%	31.30%
AdaptSS AR PA	0.34%	1.12%	3.25%	11.84%	−0.04%	0.10%	25.12%	31.38%
AdaptSS AR SA	0.61%	2.64%	3.21%	12.65%	−0.01%	0.10%	25.09%	31.33%
Best result								
	AdaptSS VAR SA		**AdaptSS VAR SA**		AdaptSS AR PA ARDL PA		**AdaptSS VAR PA**	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
	0.882	4.802	0.915	3.647	0.938	2.245	0.728	3.660

The results presented in Tables 3–6 make it possible to conclude that the proposed method (SS/AdaptSS VAR, ARDL, AR VAR, AR ARDL) is effective compared to the existing methods considered. For non-adaptive and adaptive SS, the reduction in MAE compared to BU was on average 37 and 25.1%, respectively. For FC IBP – 49.9 and 36%, for FC 10% – 48.8 and 38.2%, for FC 50% – 25.8 and 12.5%, for FC 90% – 23.4 and 13.9%. Compared to the AR-based error predictor, the MAE reduction for FC IBP was 11.5 and 2.1%, for FC 10% – 29.9 and 11.7%, for FC 50% – 6.3 and 1.5%, for FC 90% – 39.1 and 31.3%, averaging 21.7 and 11.7% for non-adaptive and adaptive SS, respectively.

5. CONCLUSION

The problem of developing a MOSS with a correlated error predictor has been solved. The effectiveness of the proposed approach for developing a MOSS with a predictor has been demonstrated for use in a quality control system for the fractional composition indicators of the light diesel fraction from a complex fractionating column in a hydrocracking unit. The comparison of the proposed MOSS with the BU and the AR model-based error predictor without considering error cross-correlation showed a reduction in MAE by an average of 29.3 and 21% and an increase in the coefficient of determination by 14 and 7% for non-adaptive and adaptive soft sensors, respectively.

FUNDING

The research was carried out within the state assignment of IACP FEB RAS on theme FW-2025-0002 (implementation of the proposed method and testing on experimental data) and theme FWFW-2021-0003 (method for constructing a multidimensional SS with an error predictor).

REFERENCES

1. Lawrence, N.P., Damarla, S.K., Kim, J.W., Tulsyan, A., Amjad, F., Wang, K., Chachuat, B., Lee, J.M., Huang, B., and Gopaluni, R.B., Machine Learning for Industrial Sensing and Control: A Survey and Practical Perspective, *Control Engineering Practice*, 2024, vol. 145, 105841. https://doi.org/10.1016/j.conengprac.2024.105841

2. Bakhtadze, N.N., Virtual Analyzers: Identification Approach, *Autom. Remote Control*, 2004, vol. 65, no. 11, pp. 1691–1709. https://doi.org/10.1023/B:AURC.0000047885.52816.c7

3. Logunov, P.L., Shamanin, M.V., Kneller, D.V., Setin, S.P., and Shunderyuk, M.M., Advanced Process Control: From a PID loop up to refinery-wide optimization, *Autom. Remote Control*, 2020, vol. 81, pp. 1929–1943. https://doi.org/10.1134/S0005117920100100

4. Shvartser, V.I., Determination of Rectification Column Characteristics on the Basis of Normal Operation Data by Potential Function Method, *Autom. Remote Control*, 1968, no. 6, pp. 142–144.

5. Mejdell, T. and Skogestad, S., Estimation of Distillation Compositions from Multiple Temperature Measurements Using Partial-Least-Squares Regression, *Industrial & Engineering Chemistry Research*, 1991, vol. 30, no. 12, pp. 2543–2555. https://doi.org/10.1021/ie00060a007

6. Torgashov, A. and Skogestad, S., The Use of First Principles Model for Evaluation of Adaptive Soft Sensor for Multicomponent Distillation Unit, *Chemical Engineering Research and Design*, 2019, vol. 151, pp. 70–78. https://doi.org/10.1016/j.cherd.2019.08.017

7. Ferreira, J., Pedemonte, M., and Torres, A.I., Development of a Machine Learning-Based Soft Sensor for an Oil Refinery's Distillation Column, *Computers & Chemical Engineering*, 2022, vol. 161, 107756. https://doi.org/10.1016/j.compchemeng.2022.107756

8. Chereshko, A.A., Soft Sensors Based on Digital Models, *Autom. Remote Control*, 2023, vol. 84, no. 7, pp. 788–796. https://doi.org/10.1134/s0005117923070044

9. Wang, K., Shang, C., and Liu, L., Dynamic Soft Sensor Development Based on Convolutional Neural Networks, *Industrial & Engineering Chemistry Research*, 2019, vol. 58, no. 26, pp. 11521–11531. https://doi.org/10.1021/acs.iecr.9b02513

10. Xibilia, M.G., Latino, M., and Marinkovic, Z., Soft Sensors Based on Deep Neural Networks for Applications in Security and Safety, *IEEE Transactions on Instrumentation and Measurement*, 2020, vol. 69, no. 10, pp. 7869–7876. https://doi.org/10.1109/TIM.2020.2984465

11. Xuefeng, Y., Hybrid Artificial Neural Network Based on BP-PLSR and Its Application in Development of Soft Sensors, *Chemometrics and Intelligent Laboratory Systems*, 2010, vol. 103, no. 2, pp. 152–159. https://doi.org/10.1016/j.chemolab.2010.07.002

12. Li, L., Li, N., and Wang, X., Multi-Output Soft Sensor Modeling Approach for Penicillin Fermentation Process Based on Features of Big Data, *Expert Systems with Applications*, 2023, vol. 213, pp. 119208. https://doi.org/10.1016/j.eswa.2022.119208

13. Mastelini, S.M., Da Costa, V.G.T., and Santana, E.J., Multi-Output Tree Chaining: An Interpretative Modelling and Lightweight Multi-Target Approach, *J. Signal Proc. Syst.*, 2019, vol. 91, no. 2, pp. 191–215. https://doi.org/10.1007/s11265-018-1376-5

14. Xu, S., An, X., and Qiao, X., Multi-Output Least-Squares Support Vector Regression Machines, *Pattern Recognition Letters*, 2013, vol. 34, no. 9, pp. 1078–1084. https://doi.org/10.1016/j.patrec.2013.01.015

15. Grbić, R., Slišković, D., and Kadlec, P., Adaptive Soft Sensor for Online Prediction and Process Monitoring Based on a Mixture of Gaussian Process Models, *Computers & Chemical Engineering*, 2013, vol. 58, pp. 84–97. https://doi.org/10.1016/j.compchemeng.2013.06.014

16. Bakirov, R., Gabrys, B., and Fay, D., Multiple Adaptive Mechanisms for Data-Driven Soft Sensors, *Computers & Chemical Engineering*, 2017, vol. 96, pp. 42–54. https://doi.org/10.1016/j.compchemeng.2016.08.017

17. Saptoro, A., State of the Art in the Development of Adaptive Soft Sensors Based on Just-in-Time Models, *Procedia Chemistry*, 2014, vol. 9, pp. 226–234. https://doi.org/10.1016/j.proche.2014.05.027

18. Pan, B., Jin, H., Wang, L., Qian, B., Chen, X., Huang, S., and Li, J. Just-in-Time Learning Based Soft Sensor with Variable Selection and Weighting Optimized by Evolutionary Optimization for Quality Prediction of Nonlinear Processes, *Chemical Engineering Research and Design*, 2019, vol. 144, pp. 285–299. https://doi.org/10.1016/j.cherd.2019.02.004

19. Fujiwara, K., Kano, M., Hasebe, S., and Takinami, A., Soft-Sensor Development Using Correlation-Based Just-in-Time Modeling, *AIChE J.*, 2009, vol. 55, no. 7, pp. 1754–1765.

20. Sangoi, E., Sanseverinatti, C.I., Clementi, L.A., and Vega, J.R., A Bayesian Bias Updating Procedure for Automatic Adaptation of Soft Sensors, *Computers & Chemical Engineering*, 2021, vol. 147, pp. 107250. https://doi.org/10.1016/j.compchemeng.2021.107250

21. Xiong, W., Zhang, W., Xu, B., and Huang, B., JITL Based MWGPR Soft Sensor for Multi-Mode Process with Dual-Updating Strategy, *Computers & Chemical Engineering*, 2016, vol. 90, pp. 260–267. https://doi.org/10.1016/j.compchemeng.2016.04.033

22. Klimchenko, V., Torgashov, A., Shardt, Y.A.W., and Yang, F., Multi-Output Soft Sensor with a Multivariate Filter That Predicts Errors Applied to an Industrial Reactive Distillation Process, *Mathematics*, 2021, vol. 9, no. 16, p. 1947. https://doi.org/10.3390/math9161947

23. Klimchenko, V.V., Snegirev, O.Yu., Shevlyagina, S.A., and Torgashov, A.Yu., Developing an Adaptive Soft Sensor Using a Predictive Filter for a Nonstationary Process, *Autom. Remote Control*, 2022, vol. 83, pp. 1984–1994. https://doi.org/10.1134/S0005117922120104

24. Torgashov, A., Snegirev, O., Klimchenko, V., and Yang, F., Development of a Multivariate Quality Estimator for an Industrial Fractionator in the Presence of Cross-Correlation of Output Variables, *IFAC-PapersOnLine*, 2023, vol. 56, no. 2, pp. 7160–7165. https://doi.org/10.1016/j.ifacol.2023.10.596

25. Hamilton, J.D., *Time Series Analysis*, Princeton: Princeton University Press, 1994.

26. Pesaran, M.H., *An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis*, Department of Applied Economics, University of Cambridge, Cambridge, 1995, vol. 9514, pp. 371–413.

27. Yeo, W.S., Saptoro, A., Kumar, P., and Kano, M., Just-in-Time Based Soft Sensors for Process Industries: A Status Report and Recommendations, *J. Process Control*, 2023, vol. 128, 103025. https://doi.org/10.1016/j.jprocont.2023.103025

28. Box, G.E.P., Jenkins, G.M., Reinsel, G.C., and Ljung G.M., *Time series analysis: forecasting and control fifth ed.*, Hoboken, New Jersey: Wiley, 2015.

29. Alt, F.L., Distributed Lags, *Econometrica*, 1942, vol. 10, no. 2, pp. 113–128. https://doi.org/10.2307/1905784

30. Jenkins, G.M. and Watts, D.G., *Spectral Analysis and Its Applications*, San Francisco: Holden-Day, 1968.

This paper was recommended for publication by N.N. Bakhtadze, a member of the Editorial Board