International Journal of Innovative Computing, Information and Control
Volume **12**, Number **6**, December **2016**

ICIC International ©2016 ISSN 1349-4198
pp. 2041–2052

SOFT SENSOR FOR DISTILLATION PROCESS BASED ON THE DYNAMIC MOVING WINDOW LEAST SQUARES SUPPORT VECTOR MACHINE ALGORITHM

Qi Li[1,*], Wenya Liu[1], Minghao Li[1] and Wei Ba[2]

[1]School of Control Science and Engineering
Dalian University of Technology
No. 2, Linggong Road, Ganjingzi District, Dalian 116024, P. R. China
*Corresponding author: qili@dlut.edu.cn

[2]The Key Laboratory for Underwater Scientific Test and Control Technology
Dalian Scientific Test and Control Technology Institute
Dalian 116013, P. R. China

Received May 2016; revised September 2016

Abstract. *Due to the fact that time-varying factors will affect the prediction performance of a soft sensor for the distillation processes, a dynamic soft sensor modeling method based on the moving window least squares support vector machine algorithm is proposed in this paper. By using a moving window, a least squares support vector machine (LS-SVM) regression with the incremental and decremental algorithm is developed to implement on-line dynamic soft sensing of distillation compositions. In this approach, the LS-SVM regression algorithm is used to build the composition soft sensor and the moving window strategy is used to update the dynamic model. A sensitivity matrix analysis method is presented to select the most suitable secondary variables to be used as the soft sensors inputs. Simulation results demonstrate the effectiveness of the proposed method.*
Keywords: Soft sensor, Distillation column, Dynamic modelling, Time-varying, Moving window

1. Introduction.

The composition of distillation products is a key quality specification in various oil refineries, whereas not many hardware sensors can be used for on-line measurement of distillation compositions. Since hardware sensors such as gas chromatographs and NIR (Near-InfraRed) typically have significant time lags and require high investment and maintenance cost, the use of economic temperature measurement on the column trays to infer distillation compositions is widely practiced in industrial applications [1,2]. Generally, the tray temperature control is based on the assumption that each of the product compositions can meet the desired specification when the corresponding tray temperature is kept the set-point value. However, in a multi-component column it is often difficult to keep a product composition at the desired value by using temperature control, because the tray temperature does not correspond exactly to the composition [3-5]. An alternative solution is to construct soft sensors to circumvent this issue.

Many approaches have been proposed to build the soft sensor models. We may summarize these methods into two categories. One method is to build the soft sensors from the mechanism model of the distillation column on-line. However, it is often difficult in refineries, due to the complexity of industrial distillation processes. Physical modeling can be very time-consuming and significant parameters are generally unknown. The other method is to adopt the empirical model or the data-driven soft sensor.

The first generation of data-driven soft sensors relied on offline modeling using the recorded historical data. There are many algorithms which have been proposed to build the soft sensor in this way, including using multivariate regression analysis [6-8], artificial neural networks [9], support vector machine regression [10-12]. Due to the strong correlation among tray temperature measurements of the distillation column, composition soft sensors based on principal component analysis (PCA) and partial least squares (PLS) regression have been widely used. However, large samples are needed in these methods and the composition soft sensors are insensitive to measurement errors. Moreover, due to the nonlinear model of composition soft sensors, many composition soft sensors based on artificial neural networks have been proposed and successfully applied in industrial processes [10]. However, there are no guarantees of avoidance of local minima, and the overfitting phenomenon and the number of hidden units in general neural networks are usually difficult to choose. As the industrial distillation column always has the characteristic of nonlinear and time-varying, the second generation of data-driven soft sensors relied on online modeling or dynamic modeling methods which use the new data to update the offline model. These dynamic modeling methods have been proposed to build the soft sensor in this way, including the moving window method, the recursive method, and the distance-based just-in-time method [13-20]. There are no dynamic model methods having high-predictive ability in all process states, and the prediction accuracy of each dynamic model method depends on a process state. Kaneko and Funatsu discussed characteristics of dynamic models and confirmed the discussion results through the numerical simulation data and real industrial data analyses [16].

Due to the fact that time-varying factors will affect the prediction performance of a soft sensor for the distillation processes, we present an LS-SVM with moving window method (MWLS-SVM) as a dynamic version of the LS-SVM in this paper, which allows the use of moving window to update the LS-SVM model. LS-SVM is a classical machine learning algorithm which has good performance of generalization for the nonlinear regression. By using the moving window to update the LS-SVM soft sensor model, the time-varying factors which affect a soft sensor to quality prediction can be overcome directly. In this paper, the sensitivity matrix analysis method is used to select the most suitable secondary variables as the soft sensors inputs. The simulation results demonstrate the effectiveness of the proposed method.

The remainder of this paper is organized as follows. In Section 2, the first-principles dynamic model of distillation column and data collection are discussed. In Section 3, the LS-SVM regression algorithm with moving window method is described. In Section 4, the performances of the MWLS-SVM method are discussed. Finally, summary and discussion are given in Section 5.

2. Distillation Process Description.

2.1. First-principles model of distillation column.

A continuous distillation column is chosen as the research object which is well known for being used in composition soft sensor performance studies [4,5,21]. The distillation column consists of 41 theoretical stages including a total condenser and a reboiler. The components in the feed flow are set as $N_c = 2$, where N_c is the number of components [4,5,21]. The schematic diagram of the continuous distillation column is shown in Figure 1.

A first-principles dynamic model is considered as the representation of the continuous distillation column.

Overall material balance:

$$dM_i/dt = d\left(M_{iL} + M_{iV}\right)/dt = L_{i+1} + V_{i-1} - L_i - V_i \tag{1}$$

where i is the index for stages and M_i is the liquid holdup on stage i, L_i [kmol/min] is the values for the liquid flow, and V_i [kmol/min] is the values for the vapor flow.

Component material balances:

$$dN_{ij}/dt = L_{i+1}x_{i+1,j} + V_{i-1}y_{i-1,j} - L_i x_{i,j} - V_i y_{i,j} \quad (2)$$

where $N_{ij} = M_{Li}x_{ij} + M_{Vi}y_{ij}$ and j is the index for components. x, y are the liquid and vapor phase mole fractions, respectively.

Vapor-liquid equilibria:

$$y_i = \alpha * x_i / (1 + (\alpha - 1) * x_i) \quad (3)$$

where α is the relative volatility between light and heavy component.

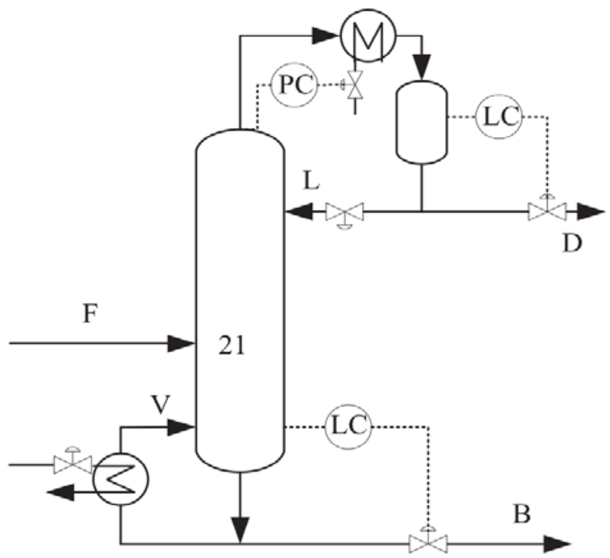

FIGURE 1. Schematic control configuration of distillation column

Unlike other models used for similar studies, this model considers varying molar holdups on each tray, that is, it includes liquid flow dynamics.

$$L_i = L_{0i} + (M_i - M_{0i})/\tau_L + (V_{i-1} - V_{0,i-1})\,\lambda \quad (4)$$

where L_{0i} [kmol/min] and M_{0i} [kmol] are the nominal values for the liquid flow and holdup on stage. The energy balances are not included in the dynamic model; therefore, the vapor flow rate is constant inside the column. Other assumptions are ideal trays, well-mixed capacities, boiling feed, total condensation with no subcooling, negligible heat losses and constant pressure operation.

2.2. **Dynamic simulation conditions and process data collection.** In order to get more various data from the dynamic simulation experiment, we use the LV-configuration in Figure 1 and the indirect temperature control loop is not included in this experiment. The data sets needed to develop the composition soft sensors of the distillation column were generated by running the first-principles dynamic model under these operating conditions.

Random perturbations within ±5% of the steady-state value were added to the feed composition during simulations. In addition to these random disturbances, the total feed flow rate changes stepwise by ±1%, while the fluctuation of the total flow rate is restricted within ±2% of its steady-state value. Measurement noises of the distribution $N(0°\text{C}, 0.1°\text{C})$ were added to the tray temperatures measurements. Simulated data for validating composition soft sensors are obtained under almost the same conditions as described above. The differences are the seeds of the random signals. The sampling period of tray temperature variables is set at 1 s. As the composition sampling time is often longer than the sampling time of inputs, the sampling period of the compositions is set at 30 s. The total simulation time is set at 5 h. By this way, we collected 601 sampling data altogether.

3. Moving Window LS-SVM Algorithm.

3.1. LS-SVM algorithm.

LS-SVM regression algorithm is an improved SVM algorithm based on structural risk minimization principle and has shown powerful ability in modeling with nonlinear, limited samples, high dimension and so on. The LS-SVM regression algorithm is described as follows [11,12].

Let training sample data:

$$D = \{(x_i, y_i)|i = 1, 2, \cdots, n\}, \quad x_i \in R^n,\ y_i \in R \tag{5}$$

where x_i is input data, and y_i is output data. The optimization problem in weight w space can be described as:

$$\begin{cases} \min\limits_{w,b,e} J(w,e) = \dfrac{1}{2}w^T w + \dfrac{1}{2}\gamma \sum\limits_{k=1}^{n} e_i^2 \\ \text{s.t.} \quad y_i = w^T \phi(x_i) + b + e_i \end{cases} \tag{6}$$

in which $\phi(\cdot) : R_n \longrightarrow R^{n_h}$ is kernel space mapping function, weight vector $w \in R^{n_h}$, error variable $e_i \in R$, b is the bias, losing J is the sum of SSE error and regularized variable, and γ is adjusted constant. The kernel mapping function is to extract characters from origin space, and map the sample in origin space to a vector in the high dimension characters space.

According to optimization (6), it defines Lagrange function as:

$$L(w,b,e,a) = J(w,e) - \sum_{i=1}^{n} a_i \left[w^T \phi(x_i) + b + e_i - y_i\right] \tag{7}$$

in which Lagrange factor (viz. support vector) $a_i \in R$. We can do optimization computation on the above equation:

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \longrightarrow w = \sum\limits_{i=1}^{n} a_i \phi(x_i) \\ \dfrac{\partial L}{\partial b} = 0 \longrightarrow \sum\limits_{i=1}^{n} a_i = 0 \\ \dfrac{\partial L}{\partial e_i} = 0 \longrightarrow a_i = \gamma e_i \\ \dfrac{\partial L}{\partial a_i} = 0 \longrightarrow w^T \phi(x_i) + b + e_i - y_i = 0 \end{cases} \tag{8}$$

Eliminating w, e, we can get the matrix equation:

$$\begin{bmatrix} 0 & l^T \\ l & \Omega + \frac{1}{\gamma} I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{9}$$

where $l = [1; 1; \cdots; 1] \in R^n$; $\Omega_{ij} = \phi(x_i)^T \phi(x_j)$.

According to Mercer condition, there exists mapping function ϕ and kernel function $K(\cdot,\cdot)$, satisfying:

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \tag{10}$$

The function approximation of LS-SVM is:

$$y(x) = \sum_{i=1}^{n} a_i K(x, x_i) + b \tag{11}$$

There are some typical kernel functions:

1) Polynomial kernels:

$$K_{poly}(x, x_i) = (x \cdot x_i + 1)^q \tag{12}$$

2) Radial basis function kernels:

$$K_{rbf}(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2p^2}\right) \tag{13}$$

3) Sigmoid kernels:

$$K_{sig}(x, x_i) = \tanh(v(x, x_i) + c) \tag{14}$$

where q is degree of polynomial function; p is the width parameter of the radial basis function; v is scale, and c is offset.

3.2. **MWLS-SVM algorithm.** For dynamic modelling of nonlinear and time-varying processes, whenever a new sample arrives, the trained LS-SVM model is updated by incrementing the training set with the new sample and thereby should discard the old sample in the training set. To decrease the computational complexity, an incremental algorithm is employed to increment the training set and a decremental algorithm to discard the old sample in the paper. Unlike other moving window algorithms [14-16], the incremental and decremental algorithms employed for MWLS-SVM in this paper are discussed in the following.

Incremental algorithm:

Let (x_{n+1}, y_{n+1}) be a new data pair, if the trained LS-SVM model computation errors satisfied this condition, we update the trained LS-SVM model.

$$|y_{n+1} - y_{LS\text{-}SVM}(n+1)| \geq \varepsilon \tag{15}$$

where ε is the max error which we can accept. If the max error is greater than ε, then the incremental algorithm updates the trained LS-SVM model by adding the new data (x_{n+1}, y_{n+1}) to the train set. That is to say the moving window is updating the train set if the model output is not very accurate. If the LS-SVM model output can predict very accurately and can overcome the time-varying factors, we do not update the trained LS-SVM model.

Decremental algorithm:

With the incremental algorithm, the oldest date information is removed from the training data sets in traditional decremental algorithm [13-15]. The drawback of traditional decremental algorithm is that the similarity data in the training data is not considered. Here, a novel decremental algorithm is described. We consider the training data sets maintain a constant number of data with the moving windows width. The similarity of the data x_q with the data set D in the moving windows is described as [16,17]:

$$s_{qi} = \rho\|x_i - x_q\| + (1 - \rho)\cos(\theta_{qi}) \tag{16}$$

where ρ is the weight parameter and is constrained between 0 and 1, and θ_{qi} is the angle. We calculate the similarity of the data, and exclude the minimal similar data.

By this way, the minimal similar data information is removed from the training data pairs. In the decremental algorithm, the training data sets retain more difference and useful information for training and building the model, and the model discards historical information as quickly as possible.

The incremental and decremental algorithms for updating the LS-SVM, presented above, make online learning for the LS-SVM possible. This algorithm is so straightforward and decreases the computational complexity.

4. Simulation Experiments.

The performances of the proposed method with the other methods are compared using a simulation case study and a distillation process case study.

4.1. Simulation case study.

To verify the effectiveness of the proposed method, we firstly analyzed various types of simulation time-varying data. The relationships between and change from moment to moment for one simulation data set and the relationships have strong nonlinearity for the other simulation data sets. The compared models are as follows:

PLS: Non-adaptive PLS model;

SVR: Non-adaptive support vector machine regression model;

LS-SVM: Non-adaptive least squares-support vector machine regression;

MWPLS: PLS model with moving window method;

OSVR: Updated SVR model.

MWPLS and OSVR models were updated with newly obtained data. Each model including the time variable was also used in the case studies. The hyperparameters for the PLS, SVR, LS-SVM, OSVR, MWPLS and MWLS-SVM models were selected with the five-fold cross-validation.

The number of X variables was set as two. First, x_1 and x_2 of uniform pseudorandom numbers whose ranges were from 0 to 10 were prepared. Then, y was set as follows [16,17]:

$$y = [x_1 \quad x_2][1 \quad b_2]^T + N(0, 0.1) \tag{17}$$

where b_2 means the magnitude of contribution of x_2 to y and $N(0, 0.1)$ is random numbers from normal distribution given a standard deviation of 0.1 and a mean of 0. We set b_2 as follows:

$$b_2 = 3\sin(0.01\pi t) + 1 \tag{18}$$

$$b_2 = 3\sin(0.02\pi t) + 1 \tag{19}$$

where t was set as 1, 2, $\cdots$, 200, and the number of data was 200. The first 100 data were used for training and the next 100 data were the test data.

The window size, which means the number of data for the model construction, was set as 20 and 50 for the MWPLS, OSVR and MWLS-SVM models. When the window size was 20, the number of the last training data was 80, and when the window size was 50, the number of the last training data was 50.

The static and dynamic model is evaluated on the basis of root mean squared error ($RMSE$) and the determination coefficient (r_p^2) of prediction data sets. The r_p^2 is the determination coefficient for test data and represents the prediction accuracy of each model. The higher r_p^2 value of a model is, the more predictive accuracy the model has. The $RMSE$ for dynamic model is calculated by applying the model to the validation data in the moving windows width w:

$$RMSE = \left(\frac{1}{N} \sum_{n=1}^{N} (x(n) - \hat{x}(n))^2 \right)^{\frac{1}{2}} \tag{20}$$

where x is a measurement of product composition, $\hat{x}$ is its prediction value from the soft sensor, and N is the number of measurements.

Table 1 shows the prediction results for test data when only $x1$ and $x2$ were used and the time variable t was not used as X variables. From Table 1, we can see that those static models could not adapt to the time-varying process characteristics. The prediction results of the dynamic models were better than those of static models. As t was not added to X variables ($x1$ and $x2$), any models could not accurately adapt to the time-varying change of the relationships between X and y. But yet, the prediction results were improved compared with other methods by using the MWLS-SVM method, especially when the moving windows width was set as 20, its prediction result was better than others.

The prediction results are shown in Table 2 when t was added to X variables ($x1$ and $x2$). In PLS, SVR, LS-SVM modeling, some of the prediction results were improved, compared with the results without the time variable t, but yet, the improvements were not enough. For the MWPLS models, the predictive accuracy was drastically decreased by increasing the window size from 20 to 50. The nonlinear and time-dependent changes in process characteristics were linearly-approximated in MWPLS modeling. By increasing the window size, the difference between the actual nonlinear relationship and the linearly

TABLE 1. The prediction results without the time variable

Model	$b_2 : E_q(18)$		$b_2 : E_q(19)$	
	r_p^2	$RMSE$	r_p^2	$RMSE$
PLS	−10.6	19	0.092	11
LS-SVM	−10.71	22.05	−0.083	13.655
SVR	−9.9	19	0.067	11
MWPLS (20)	0.546	3.8	0.574	7.3
MWPLS (50)	−1.3	8.7	−0.44	13
OSVR (20)	0.498	4.1	0.536	7.6
OSVR (50)	−1.1	8.3	−0.20	12
MWLS-SVM (20)	0.628	2.86	0.752	3.574
MWLS-SVM (50)	0.672	4.715	0.559	5.454

TABLE 2. The prediction results with the time variable

Model	$b_2 : E_q(18)$		$b_2 : E_q(19)$	
	r_p^2	$RMSE$	r_p^2	$RMSE$
PLS	−10.6	19	−6.4	31
LS-SVM	0.276	5.484	−1.11	19.07
SVR	−1.2	8.4	−5.9	20
MWPLS (20)	0.848	2.2	0.848	4.4
MWPLS (50)	0.147	5.3	0.239	9.8
OSVR (20)	0.996	0.35	0.994	0.85
OSVR (50)	0.995	0.42	0.993	0.93
MWLS-SVM (20)	0.999	0.066	0.999	0.089
MWLS-SVM (50)	0.999	0.113	0.999	0.124

approximated relationship would be large. The significant improvement of the predictive ability for the MWLS-SVM model of the window size 20 was confirmed. The proposed MWLS-SVM model could appropriately deal with the time-varying change of the relationship between X and y by updating the LS-SVM models and by adding the time variable to X variables.

4.2. **Distillation process case study.** The structure of the dynamic MWLS-SVM prediction model is taking as this form:

$$\begin{aligned} y(k) &= f(X(k), \theta) \\ X(k) &= [x_1(k), x_1(k-1), \cdots, x_1(k-d) \\ &\quad x_2(k), x_2(k-1), \cdots, x_2(k-d) \\ &\quad \vdots \\ &\quad x_n(k), x_n(k-1), \cdots, x_n(k-d)] \end{aligned} \tag{21}$$

where $f(\cdot)$ is the nonlinear model, and θ is the model's parameter.

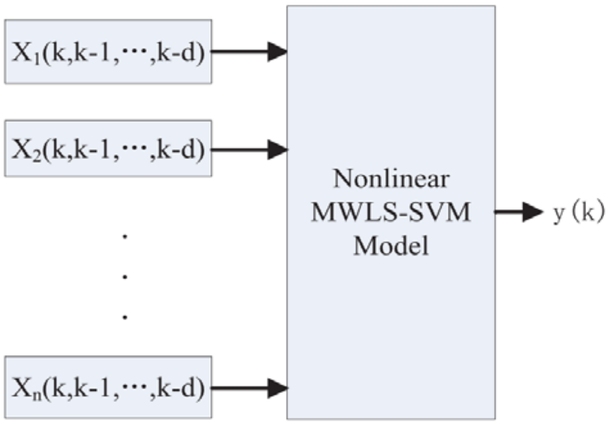

FIGURE 2. Nonlinear model structure with multi-point input

For optimal selection of the input variables, the sensitivity matrix analysis method is presented to select the most suitable secondary variables to be used as the soft sensor's inputs [21]. We define this cumulative frequency index, because the components in the feed flow $N_c = 2$ are considered in this paper, the temperatures on 21, 20, 22, 19 trays are chosen as the sensitive variables to the composition soft sensor's inputs.

The radial basic function kernel is chosen as the kernel function for all soft sensor models. The optimal selection of parameters is very important to the soft sensor's performance. In order to obtain the optimal parameters for the soft sensor models, the five-fold cross-validation technique is used to choose the optimal parameters. As a result, the optimal parameters range of the soft sensor models are defined as $p \in [0.5, 5]$, $\gamma \in [1, 10000]$, and in order to decrease the computational complexity, the number of validation data takes the last 10% data in the training data sets, and the max error which we can accept ε is defined as $\varepsilon = 0.0005$. The training data set which we select for the static LS-SVM soft sensor is $N = 300$. So, the moving windows width for dynamic MWLS-SVM soft sensor is also set as $w = 300$.

The simulation results are shown in Figures 3-6. In Figure 3 and Figure 4, the dashed dots with blue color represent the true distillation compositions while the red solid lines represent the results of dynamic soft sensor predicting. The statistical results are summarized in Table 3. The simulation results show that the MWLS-SVM soft sensor has better performance of generalization than the static LS-SVM soft sensor.

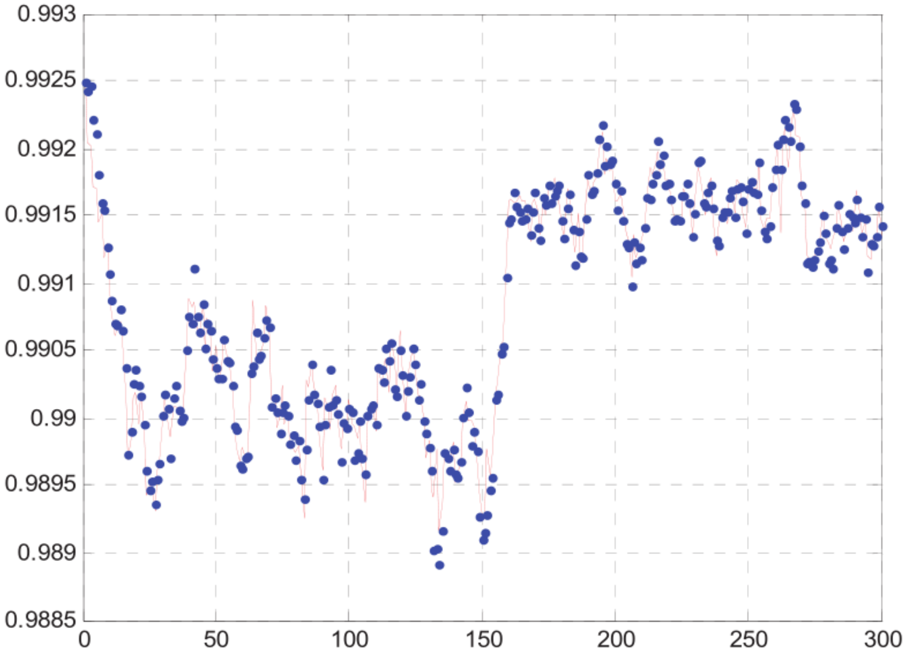

FIGURE 3. Prediction with the moving windows width $w = 300$

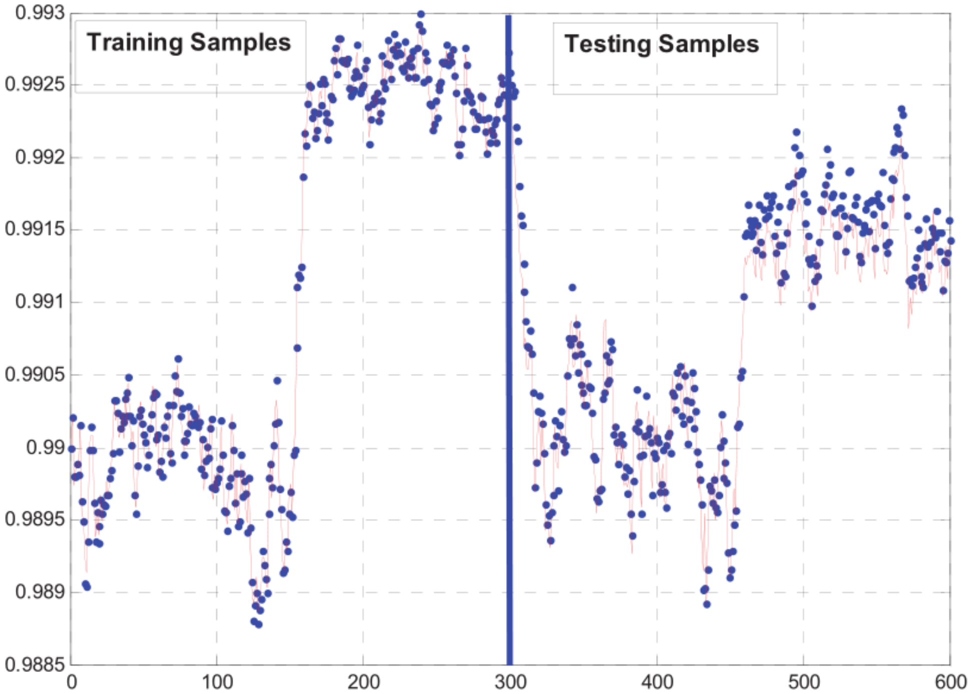

FIGURE 4. Prediction with static LS-SVM

In order to develop the dynamic MWLS-SVM composition soft sensor, we should define the moving windows width w. In this section, we also analyze the performance of the dynamic composition soft sensor with different moving windows width. In the simulation experiments, we defined $w = 50, 100, 200, 300$. The statistical results of the MWLS-SVM soft sensors in different data sets are summarized in Table 4.

The simulation results show that the estimated outputs of soft sensor based on MWLS-SVM model to the top composition match real values of the top composition and follow the varying trend of the top composition very well. In Table 3, it also shows that the

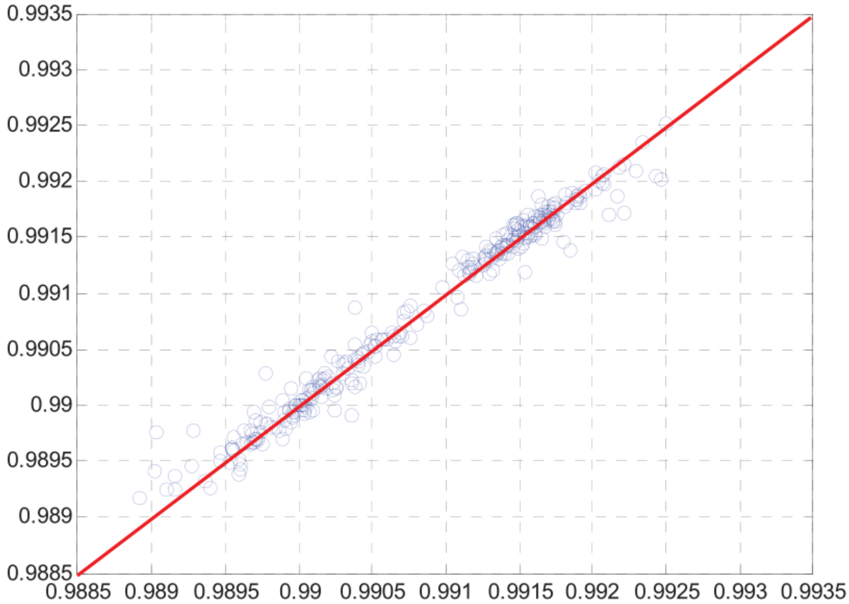

FIGURE 5. The values estimated by the dynamic model

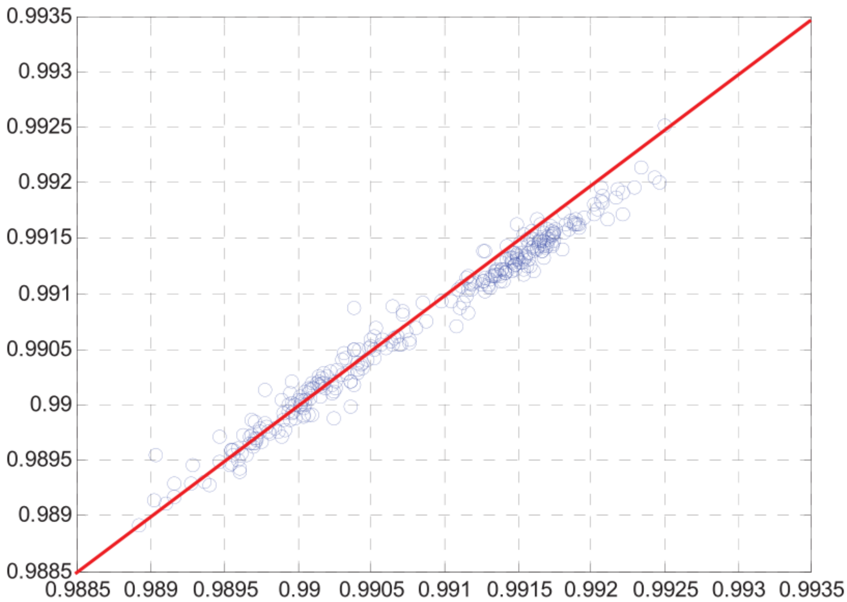

FIGURE 6. The values estimated by the static model

TABLE 3. Comparison of dynamic and static soft sensors

Method	$RMSE$ (10^{-4})
LS-SVM $N = 300$	1.7991
MWLS-SVM $w = 300$	1.3270

dynamic MWLS-SVM soft sensor has good performance of generalization. From Figure 3 and Figure 5, we find that the dynamic MWLS-SVM soft sensor model has a good generalization performance in estimation of the top composition with $w = 300$. Compared with the dynamic model with $w = 50, 100, 200$, statistical results in Table 4 show that

TABLE 4. Comparison of dynamic soft sensors

Dynamic Model	$RMSE$ (10^{-4})
MWLS-SVM $w = 50$	3.2973
MWLS-SVM $w = 100$	3.5329
MWLS-SVM $w = 200$	1.8743
MWLS-SVM $w = 300$	1.3270

the MWLS-SVM soft sensor with the moving windows width $w = 300$ predicts more accurately in the time-varying data. The simulation results show that the long moving windows width covers more information and predicts more accurately in this model, but we find that the training of the dynamic model will consume more time than other models in our experiments.

5. **Conclusions.** This paper does research on on-line dynamic soft sensor for composition estimator in the distillation column based on an MWLS-SVM and sensitivity matrix analysis method. Through the optimal selection of the second variables and building the dynamic MWLS-SVM composition soft sensor, the simulation result shows that this method is efficient.

With the development of composition soft sensors, it is possible to implement inferential control of quality variables of the distillation process on-line. In summary, successful installation of the composition soft sensors in an existing refinery can ensure better product quality control with higher productivity. Future work will concern on how to reduce the computation time and improve the performance of the dynamic moving windows algorithm.

Acknowledgment. The authors gratefully acknowledge the financial support of the National Natural Science Foundation of China (61403058), the PetroChina Innovation Foundation (2014D-5006-0601) and the Fundamental Research Funds for the Central Universities of China in the completion of this work.

REFERENCES

[1] M. Kano, K. Miyazaki, S. Hasebe and I. Hashimoto, Inferential control system of distillation compositions using dynamic partial least squares regression, *Journal of Process Control*, vol.10, pp.157-166, 2000.
[2] M. Kano, K. Miyazaki, S. Hasebe and I. Hashimoto, Inferential control of distillation compositions: Selection of model and control configuration, *Control Engineering Practice*, vol.11, pp.927-933, 2003.
[3] B. Joseph and C. B. Brosilow, Inferential control of processes: Part I. Steady state analysis and design, *AIChE Journal*, vol.24, pp.485-492, 1978.
[4] T. Mejdell and S. Skogestad, Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression, *Industrial and Engineering Chemistry Research*, vol.30, pp.2543-2555, 1991.
[5] T. Mejdell and S. Skogestad, Composition estimator in a pilot-plant distillation column using multiple temperatures, *Industrial and Engineering Chemistry Research*, vol.30, pp.2555-2564, 1991.
[6] E. Zamprogna, M. Barolo and D. E. Seborg, Estimating product composition profiles in batch distillation via partial least squares regression, *Control Engineering Practice*, vol.12, pp.917-929, 2004.
[7] E. Zamprogna, M. Barolo and D. E. Seborg, Optimal selection of soft sensor inputs for batch distillation columns using principal component analysis, *Journal of Process Control*, vol.15, pp.39-52, 2005.
[8] J. Zhang, Inferential feedback control of distillation composition based on PCR and PLS models, *Proc. of the American Control Conference*, pp.1196-1201, 2001.

[9] S. Bhartiya and J. R. Whiteley, Development of inferential measurements using neural networks, *ISA Transactions*, vol.40, pp.307-323, 2001.
[10] W. W. Yan, H. H. Shao and X. F. Wang, Soft sensing modeling based on support vector machine and Bayesian model selection, *Computers and Chemical Engineering*, vol.28, pp.1489-1498, 2004.
[11] V. Vapnik, *The Nature of Statistical Learning Theory*, 2nd Edition, Springer, Berlin, 1999.
[12] Q. Li, Q. Du, W. Ba and C. Shao, Multiple-input multiple-output soft sensors based on KPCA and MKLS-SVM for quality prediction in atmospheric distillation column, *International Journal of Innovative Computing, Information and Control*, vol.8, no.12, pp.8215-8230, 2012.
[13] H. S. Tang, S. T. Xue, R. Chen and T. Sato, Online weighted LSSVM for hysteretic structural system identification, *Eng. Struct.*, vol.28, pp.1728-1735, 2006.
[14] W.-W. Yan, J.-L. Chang and H.-H. Shao, Least square SVM regression method based on sliding time window and its simulation, *Shanghai Jiaotong Daxue Xuebao*, vol.38, no.4, pp.524-526, 2004.
[15] S. Tatinati, Y. Wang, G. Shafiq and K. C. Veluvolu, Online LS-SVM based multi-step prediction of physiological tremor for surgical robotics, *Proc. of the 35th Annual International Conference of the IEEE EMBS*, pp.6043-6046, 2013.
[16] H. Kaneko and K. Funatsu, Adaptive soft sensor model using online support vector regression with the time variable and discussion on appropriate hyperparameter settings and window size, *Comput. Chem. Eng.*, vol.58, pp.288-297, 2013.
[17] H. Kaneko and K. Funatsu, Application of online support vector regression for soft sensors, *AIChE Journal*, vol.60, pp.600-612, 2014.
[18] Z. Cheng, An intelligent method of change-point detection based on LS-SVM algorithm, *HKIE Transactions*, vol.20, no.3, pp.141-147, 2013.
[19] Q. Li, L. Xing, W. Liu and W. Ba, Adaptive soft sensor based on a moving window just-in-time learning LS-SVM for distillation processes, *IFAC PapersOnline*, vol.48, no.28, pp.51-56, 2015.
[20] C. Shang, X. Huang, J. A. K. Suykens and D. Huang, Enhancing dynamic soft sensors based on DPLS: A temporal smoothness regularization approach, *Journal of Process Control*, vol.28, pp.17-26, 2015.
[21] Q. Li and C. Shao, Soft sensing modelling based on optimal selection of secondary variables and its application, *International Journal of Automation and Computing*, vol.6, no.4, pp.379-384, 2009.