

Research Article

Fuzzy Pruning Based LS-SVM Modeling Development for a Fermentation Process

Weili Xiong,^{1,2} Wei Zhang,² Dengfeng Liu,² and Baoguo Xu²

¹ Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, China

² School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China

Correspondence should be addressed to Weili Xiong; greenpre@163.com

Received 16 December 2013; Revised 14 January 2014; Accepted 14 January 2014; Published 27 February 2014

Academic Editor: Shen Yin

Copyright © 2014 Weili Xiong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to the complexity and uncertainty of microbial fermentation processes, data coming from the plants often contain some outliers. However, these data may be treated as the normal support vectors, which always deteriorate the performance of soft sensor modeling. Since the outliers also contaminate the correlation structure of the least square support vector machine (LS-SVM), the fuzzy pruning method is provided to deal with the problem. Furthermore, by assigning different fuzzy membership scores to data samples, the sensitivity of the model to the outliers can be reduced greatly. The effectiveness and efficiency of the proposed approach are demonstrated through two numerical examples as well as a simulator case of penicillin fermentation process.

1. Introduction

For the limitation of advanced measurement techniques, some important process variables in biochemical industrial processes, such as product composition, product concentration, and biomass concentration, are difficult or impossible to measure online. However, these variables are very important for the products quality and the result of the whole reaction process. A soft sensor model is always needed to construct between variables which are easy to measure online and one which is difficult to measure. Then a value of an objective variable can be inferred by this model. The approaches and corresponding applications of soft sensors have been discussed in some literature [1–4]. For example, partial least squares (PLS) and principal component analysis (PCA) [5, 6] are the most popular projection based soft sensor modeling methods for modeling and prediction. However, a drawback of these models is their linear nature. If it is known that the relation between the easy-to-measure and the difficult-to-measure variables is nonlinear, then a nonlinear modeling method should be used. In last decades, data-based soft sensor modeling approaches have been intensively studied, such as nonlinear partial least squares (NPLS), nonlinear principal component analysis (NPCA), artificial neural networks (ANNs), and support vector machine (SVM) [7–10]. Although the NPCA is a well-established and powerful

algorithm, it has several drawbacks. One of them is that the principal components describe very well the input space but do not reflect the relation between the input and the output data space. A solution to this drawback is given by the NPLS method. NPLS models are appropriate to study the behavior of the process. Unfortunately, sometimes the algorithm of NPLS is available only for specific nonlinear relationships. To break through the limitation of NPLS, ANN is adopted to solve the complexity and highly nonlinear problem in the case of the sample data tending to infinity. The disadvantage of ANNs is that during their learning they are prone to get stuck in local minima, which can result in suboptimal performance. Meanwhile, SVM has been demonstrated to work very well for a wide spectrum of applications under the limited training data samples, so it is not surprising that it has also been successfully applied as soft sensor.

Support vector machine (SVM) proposed by Vapnik [11, 12], which is based on statistical learning theory, obtains the optimal classification of the sample data through a quadratic programming. So it can balance the risk of learning algorithm and promotion of the extension ability. As a sophisticated soft sensor modeling method, SVM has a lot of advantages in solving small sample data and nonlinear and high dimensional pattern recognition and has been applied to the fermentation process successfully [13, 14]. Least squares

support vector machine (LS-SVM) proposed by Suykens and Vandewalle [15] is an extension of the standard SVM. It can solve linear equations with faster solution speed and figure out the robustness, sparseness, and large-scale computing problems. However, all training data are treated as the normal support vector which loses the sparseness of SVM [16–19]. In this paper, the effective work addressed in Section 3 could improve the performance of the standard LS-SVM effectively.

Penicillin fermentation process is a typical biochemical reaction process with the features of nonlinearity and dynamic, which is caused by the factors such as genetic variation of somatic cell, microbial sensitivity to environment changing, and instability of raw material and seed quality that bring about serious nonlinearity and uncertainty [20]. For this process, key variables are concentration of the biomass, product, and substrate which are difficult to measure directly. However, some other auxiliary variables are easy to measure. So we choose aeration rate, dissolved oxygen concentration, agitator power, and others as auxiliary variables and the concentration of penicillin as the quality variable in this process. The next step is to construct the inferred model between the auxiliary variables and the quality variable. Outliers are commonly encountered in penicillin fermentation process which may be treated as the normal support vector and always has a bad influence on the precision of the soft sensor model. So applying the idea of fuzzy pruning for LS-SVM algorithm to cut off these outliers and reduce the number of support vectors will improve the sparseness and precision of the original LS-SVM model. Also assigning different fuzzy membership scores to sample data, the sensitivity to the outliers is reduced and the accuracy of the model is further improved as well. Finally, the LS-SVM and fuzzy pruning based LS-SVM soft sensor models for the penicillin fermentation process are constructed based on the optimal parameters obtained by using particle swarm optimization algorithm [21, 22]. Thus a soft sensor model with higher prediction precision and better generalization capability for penicillin fermentation process is completed.

The remainder of this paper is organized as follows. Section 2 begins with the revisit of LS-SVM algorithm and lays out the mathematical formulations. Detailed descriptions of improved LS-SVM based on fuzzy pruning algorithm are provided in Section 3. Two numerical simulation examples are illustrated in Section 4 which aims to demonstrate the effectiveness of the proposed method in developing soft sensors. Thereafter, a soft sensor application for the penicillin fermentation process using the proposed approach is presented in Section 5. Section 6 draws conclusions based on the results obtained in this paper.

2. The LS-SVM Revisit

Given the training data $\{T = (x_i, y_i), i = 1, 2, 3, \dots, l\}$, $x_i \in R^n$ and $y_i \in R$ denote the input patterns and one-dimension output data, respectively. Similar to the standard SVM, LS-SVM nonlinear regression is mapping the data to a higher dimension space F by using a nonlinear function $\phi(x)$ and

constructing an optimal linear regression function in the higher dimension space:

$$y(x) = \omega^T \cdot \phi(x) + b. \quad (1)$$

Here ω is the weight value and b is the threshold.

The main difference between LS-SVM and SVM is that LS-SVM adopts the equality constraints instead of inequality constraints, and empirical risk is the deviation of the quadratic rather than one square deviation. By introducing the Kernel function σ and the penalty factor C , one considers the following optimization problem:

$$\begin{aligned} \min_{\omega, b, e} \quad & J(\omega, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^l e_i^2 \\ \text{s.t.} \quad & y_i = \omega^T \phi(x_i) + b + e_i, \quad i = 1, 2, \dots, l. \end{aligned} \quad (2)$$

To solve the optimization problem, the constrained optimization problem should be converted to unconstrained optimization problem first. By introducing Lagrange multiplier α_i , we obtain the following Lagrange function as follows:

$$L(\omega, b, e, \alpha) = J(\omega, e) - \sum_{i=1}^l \alpha_i \{ \omega^T \phi(x_i) + b + e_i - y_i \}. \quad (3)$$

Then according to the Mercer condition, the specific form of the nonlinear mapping does not need to be known a priori. Suppose the kernel function takes the form $k(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$; this optimization problem could be changed into several linear equations. Based on the conditions of Karush-Kuhn-Tucker, calculating the partial derivative of $L(\omega, b, e_i, \alpha_i)$ with respect to ω , b , e_i , and α_i , respectively, and setting to zero yield

$$\begin{aligned} \frac{\partial L}{\partial \omega} = 0, \quad & \omega = \sum_{i=1}^l \alpha_i \phi(x_i), \\ \frac{\partial L}{\partial b} = 0, \quad & \sum_{i=1}^l \alpha_i = 0, \\ \frac{\partial L}{\partial e_i} = 0, \quad & \alpha_i = C e_i, \end{aligned} \quad (4)$$

$$\frac{\partial L}{\partial \alpha_i} = 0, \quad \omega^T \phi(x_i) + b + e_i - y_i = 0.$$

To simplify the equations, we can get a compressed matrix equation:

$$\begin{pmatrix} 0 & e^T \\ e & Q + \frac{1}{C} I \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}, \quad (5)$$

where $Q_{ij} = \phi(x_i) \cdot \phi(x_j) = k(x_i, x_j)$, $e = (1, 1, 1 \dots 1)^T$, $y = (y_1, y_2 \dots y_l)^T$, $\alpha = (\alpha_1, \alpha_2 \dots \alpha_l)^T$, $i = 1, 2, 3 \dots l$, $C > 0$ denotes the penalty factor, and I denotes the identity matrix. Solving the matrix equation (5), eventually the function of least squares vector machines is estimated as

$$y(x) = \sum_{i=1}^l \alpha_i k(x, x_i) + b. \quad (6)$$

3. Improved LS-SVM with Fuzzy Pruning Algorithm

3.1. *The Idea of Fuzzy Pruning Algorithm.* Compared with SVM, the computational load of LS-SVM is reduced greatly. However, LS-SVM loses its sparseness because all training data are treated as support vectors even the outliers which always have a bad influence on the precision of the soft sensor model. In this paper, aiming to minimize effects of the outliers as well as the antidisturbance ability of sampling data [23, 24], fuzzy pruning approach is employed to handle the problem. The number of the support vectors is reduced which improves the sparseness of LS-SVM and model accuracy as well. Furthermore, the sensitivity to outliers of the proposed algorithm can be reduced through the fuzzy membership score assigned to the data samples.

The absolute value of Lagrange multiplier determines the importance of data in the training process which means the higher the absolute value, the greater the influence degree. The absolute value of Lagrange multiplier of outliers is often higher than that of the normal data. Based on this situation, the data which have the higher absolute value of Lagrange multiplier will be cut off according to certain proportion (e.g., 5%). When these data are cut off, the impact of outlier data is minimized, and the model sparseness and accuracy are improved simultaneously.

Since Lagrange multiplier plays an important role in constructing model, a fuzzy membership score is introduced to adjust the weight of data for modeling. Fuzzy membership value is defined as

$$s_i = (1 - \delta) \frac{|\alpha_i| - |\alpha|_{\min}}{|\alpha|_{\max} - |\alpha|_{\min}}, \quad 0 \leq \delta < 1, \quad (7)$$

where s_i is the fuzzy membership score and α_i is the Lagrange multiplier of the i th sample data. Meanwhile, δ need to be given an appropriate value between 0 and 1.

It is noticed that the fuzzy membership score is near to zero when Lagrange multiplier is very small. So the corresponding sampling data may play no role in modeling, which means a part of sample data can be cut off according to the absolute value of Lagrange multiplier that is very small. As a result, the sparseness of the proposed LS-SVM algorithm is further improved.

3.2. *Description of Fuzzy Pruning Based LS-SVM Algorithm.* Adding fuzzy membership score s_i to error e_i , the new quadratic programming problem is expressed as follows:

$$\begin{aligned} \min_{\omega, b, e} \quad & J(\omega, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^l s_i e_i^2 \\ \text{s.t.} \quad & y_i = \omega^T \Phi(x_i) + b + e_i, \quad i = 1, 2, \dots, l. \end{aligned} \quad (8)$$

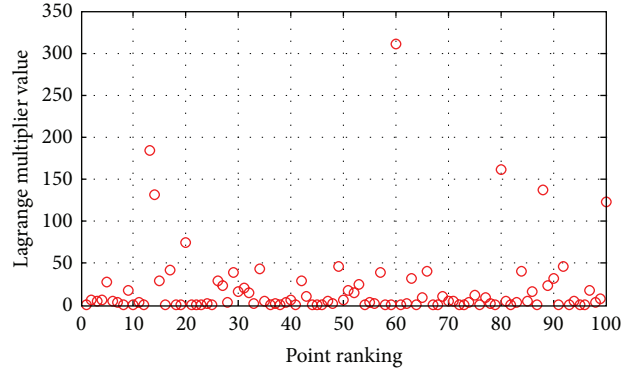


FIGURE 1: Lagrange multiplier value.

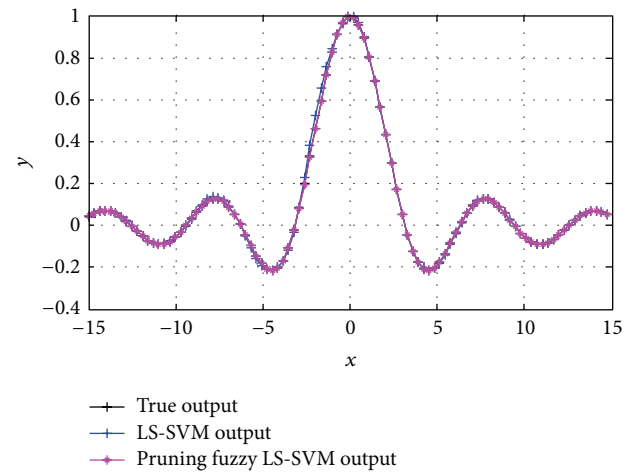


FIGURE 2: Prediction output of one-dimension function.

Since the direct optimization is not tractable, Lagrange method is introduced to convert it to become an unconstrained optimization problem. Therefore, the Lagrange function can be obtained as

$$L(\omega, b, e, \alpha) = J(\omega, e) - \sum_{i=1}^l \alpha_i \{ \omega^T \phi(x_i) + b + e_i - y_i \}. \quad (9)$$

The optimization requires the computation of the derivative of $L(\cdot)$ with respect to ω , b , e_i , and α_i , respectively. Thereafter, a set of linear equations are obtained and can be simplified as

$$\begin{pmatrix} 0 & e^T \\ e & Q + M_n \end{pmatrix} \begin{pmatrix} b^* \\ \alpha^* \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}, \quad (10)$$

where $Q_{ij} = (\phi(x_i) \cdot \phi(x_j)) = k(x_i, x_j)$, $e = (1, 1, 1 \dots 1)^T$, $y = (y_1, y_2 \dots y_l)^T$, $\alpha^* = (\alpha_1, \alpha_2 \dots \alpha_l)^T$, $M_n = \text{diag}\{1/Cs_1, \dots, 1/Cs_l\}$, and $C > 0$ denotes the penalty factor.

Eventually, the fuzzy pruning based LS-SVM function takes the form as follows:

$$y(x) = \sum_{i=1}^l \alpha_i^* k(x, x_i) + b^*. \quad (11)$$

3.3. *The Modeling Steps Based on Fuzzy Pruning LS-SVM.* The proposed LS-SVM algorithm based on fuzzy pruning technique can be summarized as follows.

- (1) Based on the training data set $\{x_i, y_i\}_{i=1}^l$, we can calculate the Lagrange multiplier α_i .
- (2) Choose a suitable δ ; the fuzzy membership scores s_i of training data are obtained from (7).
- (3) Build a new data set $\{x_i, y_i, s_i\}_{i=1}^l$, and train the new data set under the scheme of fuzzy pruning LS-SVM algorithm again; then we can get the new α_i^* .
- (4) Sort the Lagrange multiplier α_i , and cut off the data taking larger Lagrange multiplier according to certain proportion (e.g., 5%).
- (5) Then the fuzzy pruning based LS-SVM algorithm is applied to train the current data set. If the fitting performance degrades, the training procedure is done. Otherwise, switch to (4).

4. Two Numerical Simulations

4.1. *One-Dimension Function.* The effectiveness and efficiency of handing the outliers through the proposed approach are evaluated through two numerical functions. All the simulation experiments are run on a 2.8 GH CPU with 1024 MB RAM PC using Matlab 7.11.

Consider one-dimension function defined as follows:

$$y = \frac{\sin x}{x}, \quad -15 < x < 15. \quad (12)$$

100 data are generated in $[-15, 15]$ randomly as the training data set. To test the performance of detecting outliers, 30% disturbance is added to the 20th, 40th, 60th, 80th, and 100th data sample, respectively. And another 100 data are collected for evaluation.

It can be seen from Figure 1 that the outliers have the higher value of Lagrange multiplier as mentioned above. Using PSO algorithm (w keeps linear decline from 1.2 to 0.4, population size is 20, and maximum number of iterations of the population is 200) to optimize kernel parameter σ and the penalty factor C , then the LS-SVM and fuzzy pruning LS-SVM models are constructed to predict and compare (Figures 2 and 3). Figure 3 is the 45-degree line comparison between different measurements. If two measurements agree with the true outputs, then all data points will fall into the black 45-degree line. The blue circles denote the LS-SVM measurements and the pink asterisks denote the model predictions of fuzzy pruning LS-SVM. We can see that the estimation with the fuzzy pruning LS-SVM fits the black line better and thus provides a superior performance compared to the LS-SVM observation.

The detailed results such as the maximum absolute error (Max EE), the mean absolute error (Mean EE), and root mean square error (RMSE) are calculated and listed in Table 1. The RMSE decreased from 1.21% to 0.052%, which indicates the fuzzy pruning LS-SVM has higher prediction performance and better antidisturbance.

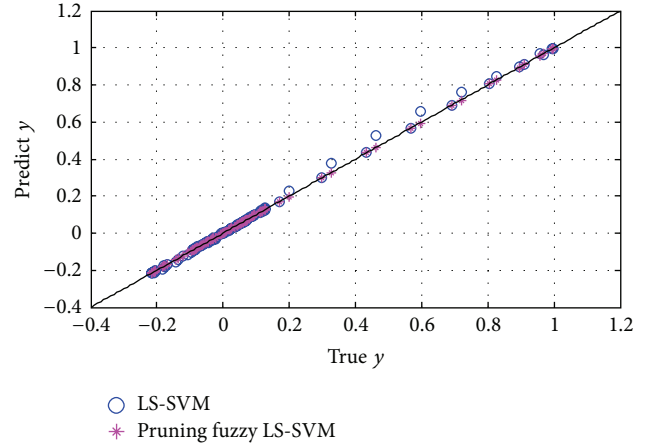


FIGURE 3: 45-degree comparison of the two soft sensors.

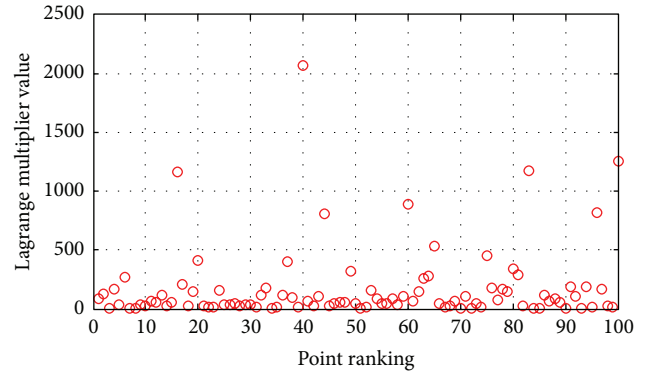


FIGURE 4: Lagrange multiplier value.

4.2. *Two-Dimension Function.* A two-dimension function is described as

$$z = \sin x \cos y, \quad x \geq -\pi, \quad y \leq \pi. \quad (13)$$

100 data are generated randomly in the range of $[-\pi, \pi]$, which makes up a training data set. Then the 20th, 40th, 60th, 80th, and 100th data points are added with 30% disturbance separately and the performance is tested by using another different 100 data. As is shown in Figure 4, Lagrange multiplier value of data points that corrupted by some disturbance always has the higher value. Compared results are shown in Figure 5. From Table 2, prediction accuracy of fuzzy pruning LS-SVM is much higher than LS-SVM, which indicates the five outliers have been detected and cut off effectively using the proposed method.

5. An Experiment Simulation

The Pensim simulator provides a simulation of a fed-batch fermentation process for penicillin production. The main component of the process is a fermenter, where the biological reaction takes place. It fully considers the most factors influencing the penicillin fermentation process, such as PH, aeration rate, substrate feed rate, carbon dioxide, and

TABLE 1: One-dimension function predicted results.

Model	C	σ	Max EE	Mean EE	RMSE
LS-SVM	4616.2	1.5118	0.0640	0.0079	0.0121
Fuzzy pruning LS-SVM	4616.2	1.5118	0.0028	0.0002	5.2456e(-4)

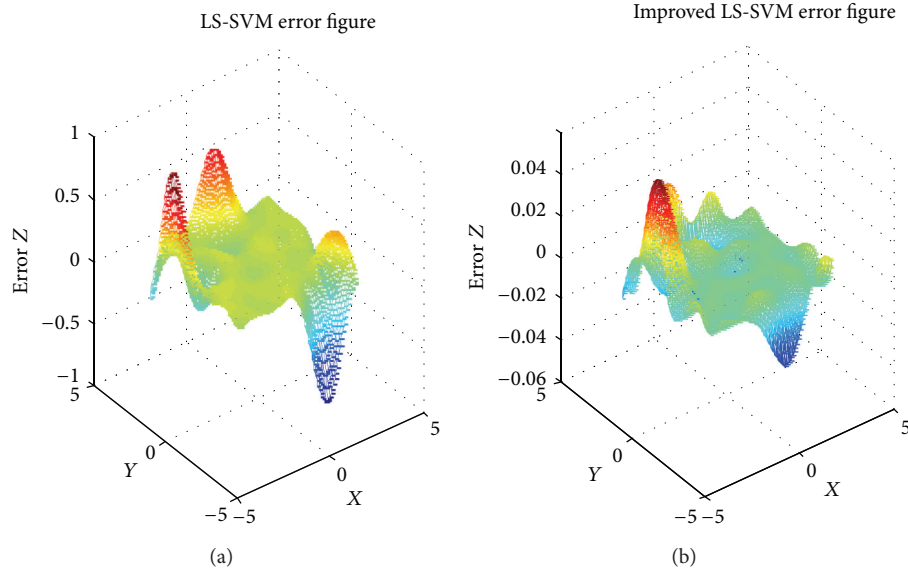


FIGURE 5: Prediction error of two-dimension function.

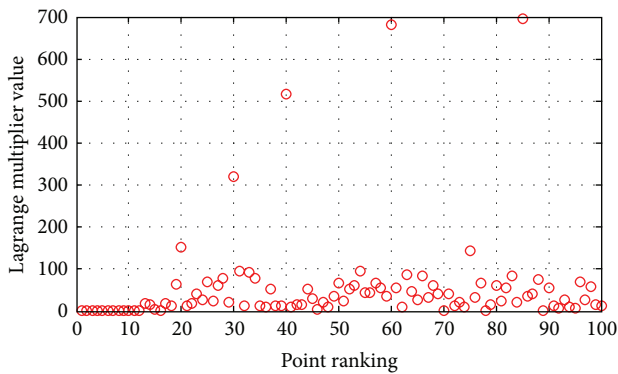


FIGURE 6: Lagrange multiplier value.

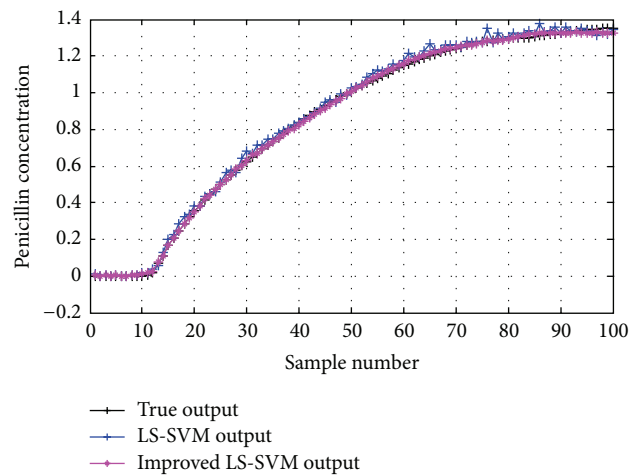


FIGURE 7: Penicillin concentration prediction.

penicillin production. The practicability and validity of the platform have been fully verified [25–27] and it has been a benchmark problem for modeling and diagnosis detection.

In this paper Pensim simulation platform is used to generate the original 100 training data. Then 30% disturbance is added to the 20th, 30th, 40th, 60th, and 85th, respectively, and another 100 data are used as test data to verify the constructed model. The simulation results are shown in Figures 7 and 8.

To further exhibit the difference of the two methods, the indexes of Max EE, Mean EE, and RMSE of each method are also calculated and listed in Table 3.

Compared to LS-SVM, the proposed approach makes RMSE decrease from 2.44% to 0.97%, which indicates the fuzzy pruning LS-SVM has better prediction performance.

Lagrange multiplier values according to each data point are shown in Figure 6, and we can easily find out the outliers obviously have much bigger Lagrange multiplier. Figure 8 is the 45-degree line comparison between two different soft sensors. Clearly, the fuzzy pruning based LS-SVM exhibits the better capability of approximating the true process. It has effectively handled the disturbance of the outliers so that their impact on modeling is minimized to lowest.

TABLE 2: Two-dimension function predicted results.

Model	C	σ	Max EE	Mean EE	RMSE
LS-SVM	9528.2	1.4557	0.0475	0.0146	0.0207
Pruning fuzzy LS-SVM	9528.2	1.4557	0.0415	0.0029	0.0066

TABLE 3: The predicted concentration of penicillin.

Model	C	σ	Max EE	Mean EE	RMSE
LS-SVM	2131.3	2.5448	0.0678	0.0191	0.0244
Pruning fuzzy LS-SVM	2131.3	2.5448	0.0261	0.0075	0.0097

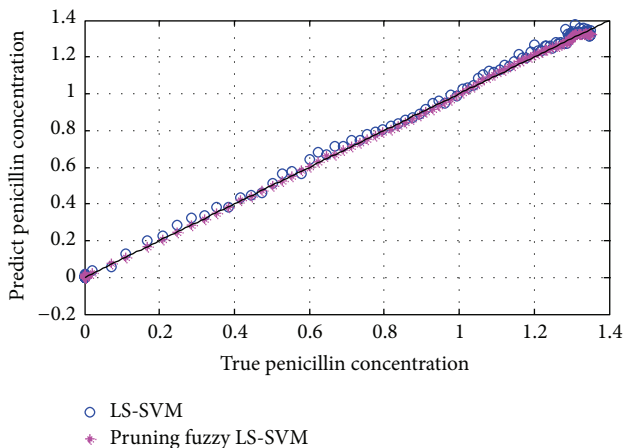


FIGURE 8: 45-degree comparison of the two soft sensors.

6. Conclusions

A novel LS-SVM method based on fuzzy pruning technique is investigated in this paper. Pruning algorithm is applied to cut off the outliers. Therefore the number of support vectors is reduced which improves the sparseness and accuracy of LS-SVM algorithm. On the other hand, assigning different fuzzy membership score to each of the sample data makes those sample data that play a small role in soft sensor modeling not participate in the construction of the model. Furthermore, the sensitivity to the outliers of the proposed algorithm can be reduced through the fuzzy membership score. The simulation examples demonstrated that the proposed method can effectively handle the outliers and achieved satisfied performance of modeling and prediction.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank the financial support by the National Natural Science Foundation of China (nos. 21206053, 21276111, and 61273131) and partial support by the 111 Project (B12018)

and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

- [1] S. Yin, H. Luo, and S. Ding, "Real-time implementation of fault-tolerant control systems with performance optimization," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 2402–2411, 2014.
- [2] Q.-D. Yang, F.-L. Wang, and Y.-Q. Chang, "Soft sensor of biomass based on improved BP neural network," *Control and Decision*, vol. 23, no. 8, pp. 869–878, 2008.
- [3] G. Liu, D. Zhou, H. Xu, and C. Mei, "Soft sensor modeling using SVM in fermentation process," *Chinese Journal of Scientific Instrument*, vol. 30, no. 6, pp. 1228–1232, 2009.
- [4] L. Huang, Y. Sun, X. Ji, Y. Huang, and B. Wang, "Soft sensor of lysine fermentation based on tPSO-BPNN," *Chinese Journal of Scientific Instrument*, vol. 31, no. 10, pp. 2317–2321, 2010.
- [5] S. J. Qin, "Recursive PLS algorithms for adaptive data modeling," *Computers and Chemical Engineering*, vol. 22, no. 4-5, pp. 503–514, 1998.
- [6] W. Li, H. H. Yue, S. Valle-Cervantes, and S. J. Qin, "Recursive PCA for adaptive process monitoring," *Journal of Process Control*, vol. 10, no. 5, pp. 471–486, 2000.
- [7] S. J. Qin and T. J. McAvoy, "Nonlinear PLS modeling using neural networks," *Computers and Chemical Engineering*, vol. 16, no. 4, pp. 379–391, 1992.
- [8] D. Dong and T. J. McAvoy, "Nonlinear principal component analysis: based on principal curves and neural networks," *Computers and Chemical Engineering*, vol. 20, no. 1, pp. 65–78, 1996.
- [9] J. C. B. Gonzaga, L. A. C. Meleiro, C. Kiang, and R. Maciel Filho, "ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process," *Computers and Chemical Engineering*, vol. 33, no. 1, pp. 43–49, 2009.
- [10] S. Yin, G. Wang, and H. Karimi, "Data-driven design of robust fault detection system for wind turbines," *Mechatronics*, 2013.
- [11] V. N. Vapnik, *The Nature Statistical Learning Theory*, Springer, New York, NY, USA, 1999.
- [12] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, 1999.
- [13] G. Liu, D. Zhou, H. Xu, and C. Mei, "Microbial fermentation process soft sensors modeling research based on the SVM," *Chinese Journal of Scientific Instrument*, vol. 30, no. 6, pp. 1228–1232, 2009.
- [14] X.-J. Gao, P. Wang, C.-Z. Sun, J.-Q. Yi, Y.-T. Zhang, and H.-Q. Zhang, "Modeling for Penicillin fermentation process based on

- support vector machine," *Journal of System Simulation*, vol. 18, no. 7, pp. 2052–2055, 2006.
- [15] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [16] X. Wang, J. Chen, C. Liu, and F. Pan, "Hybrid modeling of penicillin fermentation process based on least square support vector machine," *Chemical Engineering Research and Design*, vol. 88, no. 4, pp. 415–420, 2010.
- [17] L. LI, H. SU, and J. CHU, "Modeling of isomerization of C8 aromatics by online least squares support vector machine," *Chinese Journal of Chemical Engineering*, vol. 17, no. 3, pp. 437–444, 2009.
- [18] L. Li, K. Song, and Y. Zhao, "Modeling of ARA fermentation based on affinity propagation clustering," *CIESC Journal*, vol. 62, no. 8, pp. 2116–2121, 2011.
- [19] J. W. Cao and H. W. Ma, *Microbial Engineering*, Science press, Beijing, China, 2002.
- [20] G. Guo, S. Z. Li, and K. L. Chan, "Support vector machines for face recognition," *Image and Vision Computing*, vol. 19, no. 9-10, pp. 631–638, 2001.
- [21] R.-Q. Chen and J.-S. Yu, "Soft sensor modeling based on particle swarm optimization and least squares support vector machines," *Journal of System Simulation*, vol. 19, no. 22, pp. 5307–5310, 2007.
- [22] L. Huang, Y. Sun, X. Ji, Y. Huang, and T. Du, "Soft sensor modeling of fermentation process based on the combination of CPSO and LSSVM," *Chinese Journal of Scientific Instrument*, vol. 32, no. 9, pp. 2066–2070, 2011.
- [23] X. Zhang, "Using class-center vectors to build support vector machines," in *Proceedings of the 9th IEEE Workshop on Neural Networks for Signal Processing (NNSP '99)*, pp. 3–11, August 1999.
- [24] C.-F. Lin and S.-D. Wang, "Fuzzy support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 464–471, 2002.
- [25] S. Yin, S. Ding, A. Haghani, and H. Hao, "Data-driven monitoring for stochastic systems and its application on batch process," *International Journal of Systems Science*, vol. 44, no. 7, pp. 1366–1376, 2013.
- [26] Y. Liu and H.-Q. Wang, "Pensim simulator and its application in penicillin fermentation process," *Journal of System Simulation*, vol. 18, no. 12, pp. 3524–3527, 2006.
- [27] S. Yin, S. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.