

Article

Just-in-Time Correntropy Soft Sensor with Noisy Data for Industrial Silicon Content Prediction

Kun Chen ¹, Yu Liang ², Zengliang Gao ² and Yi Liu ^{2,*}

¹ Department of Electrical and Information Engineering, Shaoxing University, Shaoxing 312000, China; kchen@usx.edu.cn

² Institute of Process Equipment and Control Engineering, Zhejiang University of Technology, Hangzhou 310014, China; yuliang@zjut.edu.cn (Y.L.); zlgao@zjut.edu.cn (Z.G.)

* Correspondence: yliuzju@zjut.edu.cn; Tel.: +86-571-88871060; Fax: +86-571-88320842

Received: 28 May 2017; Accepted: 27 July 2017; Published: 8 August 2017

Abstract: Development of accurate data-driven quality prediction models for industrial blast furnaces encounters several challenges mainly because the collected data are nonlinear, non-Gaussian, and uneven distributed. A just-in-time correntropy-based local soft sensing approach is presented to predict the silicon content in this work. Without cumbersome efforts for outlier detection, a correntropy support vector regression (CSVR) modeling framework is proposed to deal with the soft sensor development and outlier detection simultaneously. Moreover, with a continuous updating database and a clustering strategy, a just-in-time CSVR (JCSVR) method is developed. Consequently, more accurate prediction and efficient implementations of JCSVR can be achieved. Better prediction performance of JCSVR is validated on the online silicon content prediction, compared with traditional soft sensors.

Keywords: soft sensor; industrial blast furnace; silicon content; local learning; support vector regression; outlier detection

1. Introduction

The silicon content is an important index of the thermal state in industrial blast furnaces. It must be controlled appropriately to facilitate stable production. To achieve this goal, the accurate state of silicon content in hot metal should be known online. Previously, there was extensive research on the kinetic and thermodynamic behaviors occurring in the blast furnace iron-making process [1–6]. Nevertheless, an accurate mechanism model suitable for industrial applications is still not available. Nowadays, process data can be easily measured in industrial blast furnaces. To explore and utilize useful information hidden in the data, several empirical soft-sensing approaches were applied to online predict the silicon content. Existing popular data-driven methods include artificial neural networks [7–14], multivariate regression [14,15], time series analysis [16–19], fuzzy systems [20], subspace identification [21], support vector regression (SVR) and least squares SVR (LSSVR) [22–25], and multi-scale and multiple models [26–30]. These data-driven empirical models for short-term silicon content prediction can be constructed in a quick way [31–33].

Actually, the development of a good soft sensor model is easy. This is mainly because the modeling data are noisy and often contain unwanted outliers. They may come from instrument degradation, transmission problems, etc. Generally, a good soft sensor model is dependent on the high quality of modeling data. Different kinds of noise should be considered when training artificial neural networks and other data-driven models [34–36]. Without enough attention, a soft sensor model trained with outliers and inappropriate data may tend to be over-fitting and, thus, lead to unreliable prediction. For practical use, a reliable prediction model should be constructed by reducing the negative effect of outliers. Generally, obvious outliers can be deleted by most of traditional outlier

detection methods [37–40]. However, it is not easy to detect those inconspicuous outliers mainly because they may be masked by their adjacent data. In our opinion, the soft sensor development and outlier detection should be integrated into a unified framework rather than be separated into two tasks.

For industrial blast furnace iron-making processes, only using a global/fixed model is not possible to describe the complex characteristics. Additionally, it is difficult to update the global models quickly when the process dynamics are changing [41]. Nurkkala et al. [30] proposed multiple autoregressive vector models to describe complex systems. To construct the local models automatically, several just-in-time learning methods were utilized for nonlinear process modeling problems [42–44]. Different from most traditional soft sensors, just-in-time-based models are built in a lazy learning manner when the query sample is required to be predicted. Consequently, the advantage is that the prediction for the query sample can be optimized locally, which might increase the prediction performance. For the silicon content prediction, Liu and Gao [25] utilized the just-in-time LSSVR (JLSSVR) modeling method to better describe process nonlinearity directly. Unfortunately, data samples utilized for construction of a JLSSVR model are assigned with the same weight regardless of their different effects. In such a situation, the negative effect of outliers may not be removed.

A novel online local model is developed for reliable prediction of the industrial silicon content. To handle the noisy data with non-Gaussian and uneven distributions, a just-in-time correntropy SVR (JCSVR) soft sensor is proposed. Compared with traditional soft sensors, the proposed JCSVR method is more reliable and practical in two ways. First, by reduction of the outliers' negative effect, more accurate prediction of the silicon content can be obtained. Second, the reliability of the database can be improved gradually. These two properties make the JCSVR method better for long-term utilization.

The remainder of this work is structured thusly: The correntropy SVR (CSVR) soft sensing approach is formulated in Section 2. In Section 3, the clustering-based JCSVR local modeling method is proposed. Additionally, the database maintenance is implemented. In Section 4, the JCSVR method is applied to online silicon content prediction. Finally, a conclusion is drawn in Section 5.

2. CSVR-Based Soft Sensor Model

In this section, how to integrate the maximization correntropy criterion [45,46] and SVR into a CSVR-based unified framework is formulated. The CSVR model f is to fit a dataset $\{\mathbf{S}\} = \{\mathbf{X}, \mathbf{Y}\}$, where $\{\mathbf{X}\} = \{\mathbf{x}_i\}_{i=1}^N$ and $\{\mathbf{Y}\} = \{y_i\}_{i=1}^N$ are N input and output data samples, respectively. The relationship is formulated as [45]:

$$y_i = f(\mathbf{x}_i; \boldsymbol{\theta}) + e_i = f(\mathbf{x}_i; \mathbf{w}, b) + e_i = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i) + b + e_i, i = 1, \dots, N \quad (1)$$

where $f(\cdot)$ is the model; e_i is the noise item of the i th sample; \mathbf{x}_i is an input vector composed of several online-measured variables. The model parameter vector and the bias are \mathbf{w} and b , respectively, and $\boldsymbol{\theta} = [\mathbf{w}^T, b]^T$. The CSVR model is solved using the optimization problem below [45]:

$$\begin{cases} \min J(\mathbf{w}, b, \rho) = \frac{\gamma}{2} \sum_{i=1}^N \rho(e_i) e_i^2 + \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{s.t. } y_i - \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i) - b - e_i = 0, i = 1, \dots, N \end{cases} \quad (2)$$

where γ ($\gamma > 0$) is the regularization parameter determining the trade-off between the approximation accuracy and the model's complexity. Several approaches [45] are available for selection of

the kernel width σ of the related items $\rho(e_i) = \frac{\exp\left(-\frac{e_i^2}{2\sigma^2}\right)}{\sigma^3 \sqrt{2\pi}}$. Here, it is simply adopted as $\sigma = \frac{\max|e_i|}{2\sqrt{2}}, i = 1, \dots, N$ [46].

According to Liu and Chen [46], a two-step iterative algorithm is adopted to obtain the solution of above problem in Equation (2). Finally, a CSVSR soft sensor model is established. For a test sample \mathbf{x}_t , its prediction \hat{y}_t is formulated below:

$$\hat{y}_t = f(\mathbf{w}, b; \mathbf{x}_t) = \sum_{i=1}^N \alpha_i \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_t) \rangle + b = \boldsymbol{\alpha}^T \mathbf{k}_t + b \quad (3)$$

where $\mathbf{k}_t = [k_{t1}, \dots, k_{tN}]^T \in R^{N \times 1}$, $k_{ti} = \langle \phi(\mathbf{x}_t), \phi(\mathbf{x}_i) \rangle$, $\forall i = 1, \dots, N$ is a kernel column.

For an established CSVSR model, the corresponding weight of a training sample \mathbf{x}_i is $\rho(e_i) = \frac{\exp\left(-\frac{e_i^2}{2\sigma^2}\right)}{\sigma^3 \sqrt{2\pi}}$. Using the weights, the uncertainty of the training data can be quantified. Generally, the outliers are only a small portion of all data, and they can be automatically assigned with relatively smaller weights [46]. Consequently, those candidate outliers can be identified using a simple criterion in Equation (4):

$$\rho(e_i) < \bar{\rho} \quad (4)$$

where $\bar{\rho}$ is a cutoff value and it can be chosen as a small one less than 1 after simply normalizing all the weights $\rho(e_i)$, $i = 1, \dots, N$ into the range of $[0, 1]$.

In summary, the candidate outliers can be detected simultaneously using the weights of an established CSVSR model. Interestingly, although the outliers are temporarily not removed out, they cannot degrade the prediction performance of CSVSR due to their relatively small weights [45,46]. At a glance, the CSVSR method is similar with some weighted SVR methods, e.g., in [47,48]. However, most weighted SVR methods are heuristic [47,48]. For complex industrial data, it is difficult to design suitable weighted strategies. Unlike those heuristic schemes, a reliable CSVSR model can be constructed more directly for noisy data.

3. Correntropy-Based Local Modeling and Prediction Method

3.1. JCSVSR-Based Local Model

In this section, the JCSVSR modeling method for online prediction of a query sample \mathbf{x}_q is described. First, search similar samples in the database \mathbf{S} as a similar set \mathbf{S}_q using some defined similarity criteria. Second, establish a JCSVSR model $f_{\text{JCSVSR}}(\mathbf{x}_q)$ with \mathbf{S}_q . Third, obtain \hat{y}_q for \mathbf{x}_q online. With the same implementations, a new JCSVSR model can be constructed for another query sample.

As a common similarity, the Euclidean-distance-based similarity index (SI) is defined below [42]:

$$SI_{qi} = \exp(-d_{qi}) = \exp(-\|\mathbf{x}_i - \mathbf{x}_q\|), i = 1, \dots, N \quad (5)$$

where d_{qi} denotes the similarity between \mathbf{x}_q and \mathbf{x}_i in the historical set. Obviously, $0 \leq SI_{qi} \leq 1$. When SI_{qi} approaches to 1, \mathbf{x}_q and \mathbf{x}_i are almost the same. Other similarity criteria (e.g., correlation-based similarity) [41,43,44] can also be utilized to search similar samples.

To select a suitable dataset \mathbf{S}_q with n_q similar samples, the n_{\max} most similar samples can be ranked using the SI criterion in Equation (5). Correspondingly, a cumulative similarity factor (CSF) CSF_{qn} is defined below [44]:

$$CSF_{qn} = \frac{\sum_{i=1}^{n_q} SI_{qi}}{\sum_{i=1}^{n_{\max}} SI_{qi}}, n_q \leq n_{\max} \quad (6)$$

where CSF_{qn} denotes the cumulative similarity of n_q most similar samples of all n_{\max} samples. The CSF index can determine the most similar samples simply. For example, $CSF_{qn} = 0.85$ means 85% of the similar samples are chosen [44]. Using the similarity criterion, a similar dataset \mathbf{S}_q is utilized to construct the JCSVSR model.

3.2. Implementations of the Proposed Method

In this section, the JCSVR-based online modeling method is enhanced for a relative long-term utilization. Generally, there are some outliers in the initial training dataset. In the offline modeling stage, the CSVR method is first applied to the initial training dataset. After this preprocessing step, some outliers can be identified. Additionally, to make the JCSVR method more efficient in computation, the training data are clustered into several groups. This can divide the whole dataset into several subsets. The data in each subset show similar characteristics. Consequently, for online prediction of a query sample, only its similar data are searched. This can improve the computation efficiency. The step-by-step procedures of the JCSVR-based modeling method for online silicon content prediction are summarized below:

- Step 1.** Collect the process input and output data, i.e., $\{\mathbf{S}\} = \{\mathbf{X}, \mathbf{Y}\}$, for training of the CSVR model.
- Step 2.** Train a CSVR model using the common cross-validation training strategy [46]. The weights $\rho(e_i)$, $i = 1, \dots, N$ can be obtained simultaneously. Then normalize all the weights $\rho(e_i)$, $i = 1, \dots, N$ into the range of $[0, 1]$. Using Equation (4) to identify the outliers and assign them into a outlier set $\mathbf{S}_{outlier}$. The relative clean dataset can be denoted as $\mathbf{S}_{normal} = \{\mathbf{X}_{normal}, \mathbf{Y}_{normal}\}$.
- Step 3.** Applying a simple fuzzy c-means (FCM) clustering approach [49] to \mathbf{S}_{normal} , the training samples are clustered into l sub-classes, denoted as $\{\mathbf{S}_{normal,1}, \mathbf{S}_{normal,2}, \dots, \mathbf{S}_{normal,l}\}$. For \mathbf{X}_{normal} , each sub-class has a center denoted as $\{\mathbf{c}_{normal,1}, \mathbf{c}_{normal,2}, \dots, \mathbf{c}_{normal,l}\}$.
- Step 4.** For online prediction of a new input measurement \mathbf{x}_q , judge which center of the sub-classes $\{\mathbf{c}_{normal,1}, \mathbf{c}_{normal,2}, \dots, \mathbf{c}_{normal,l}\}$ is its nearest one. If $\mathbf{c}_{normal,j}$ is the nearest to \mathbf{x}_q , only search the similar set \mathbf{S}_q in $\mathbf{S}_{normal,j}$ using the similarity criterion (Equations (5) and (6)). A JCSVR model for \mathbf{x}_q can be online constructed and the prediction \hat{y}_q is obtained.
- Step 5.** If new training data $\mathbf{S}_{new} = \{\mathbf{X}_{new}, \mathbf{Y}_{new}\}$ are available, combine these data into \mathbf{S} (i.e., $\mathbf{S} = \mathbf{S}_{new} \cup \mathbf{S}$) and go to step 1. Otherwise, go to step 4 and repeat the same procedure for online prediction of another new input \mathbf{x}_{q+1} .

The main implemented steps of the JCSVR-based soft sensor modeling and prediction are summarized in Figure 1. For industrial data, candidate outliers are simply identified without considerable efforts. Additionally, step 2 and step 3 can be implemented offline. This can improve the computation efficiency for the online JCSVR modeling method. Consequently, the proposed JCSVR-based local method can provide a relative long-term utilization for the silicon content prediction.

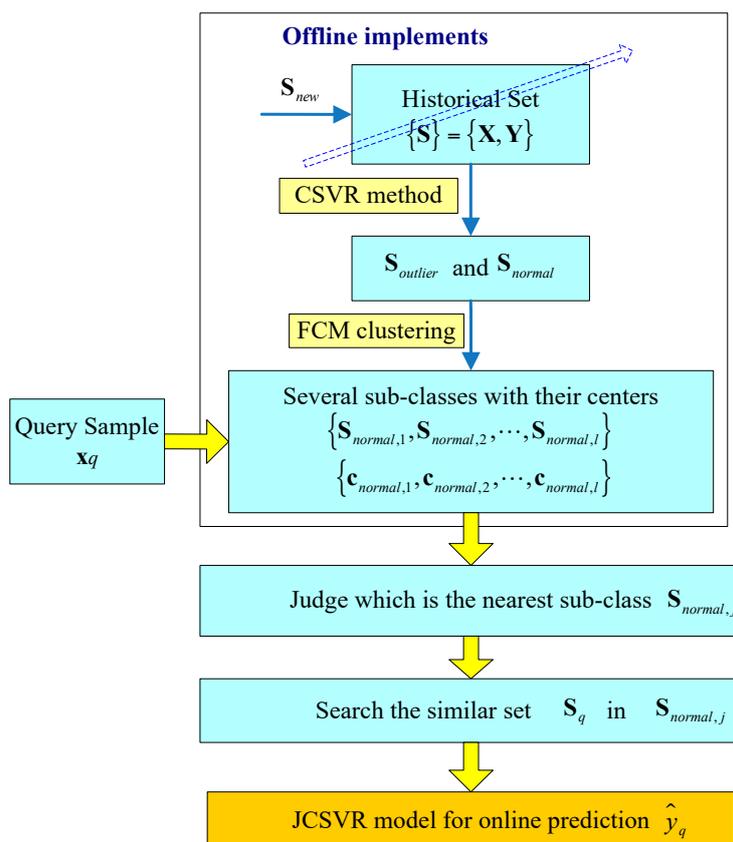


Figure 1. Main implemented steps of the proposed soft sensor modeling method integrating both of offline implements and online JCSVR method.

4. Industrial Silicon Content Prediction

The presented JCSVR-based local modeling method is applied to online prediction of the silicon content in an industrial blast furnace in China. The input variables correlated with the silicon content include the blast temperature, the blast volume, the gas permeability, the top pressure, the top temperature, the ore/coke ratio, and the pulverized coal injection [21,22,24]. The sampling time of most of these input variables is 1 min. Additionally, the time difference between the silicon content and input variables is selected according to expert experience [31]. For example, the time difference between the silicon content and the top pressure is about 2 h. The silicon content is analyzed offline and infrequently. Consequently, the soft sensor is constructed using the online measured variables.

After simply removing obvious outliers using the 3-sigma criterion, a set of 440 data samples is investigated. The historical set consists of 240 data. The rest 200 data points are for testing. It should be noted that the data are noisy and still contain some inconspicuous outliers. The normal probability of two input variables, including the top pressure and the top temperature, is shown in Figure 2a,b, respectively. The distribution results indicate that the process variables violate the Gaussian distribution denoting by the red lines in Figure 2a,b, respectively. The other process variables not plotted here are also non-Gaussian distribution. Additionally, as illustrated in Figure 3, several input variables exhibit the nonlinear relationship, and the data in different operating areas are distributed irregularly.

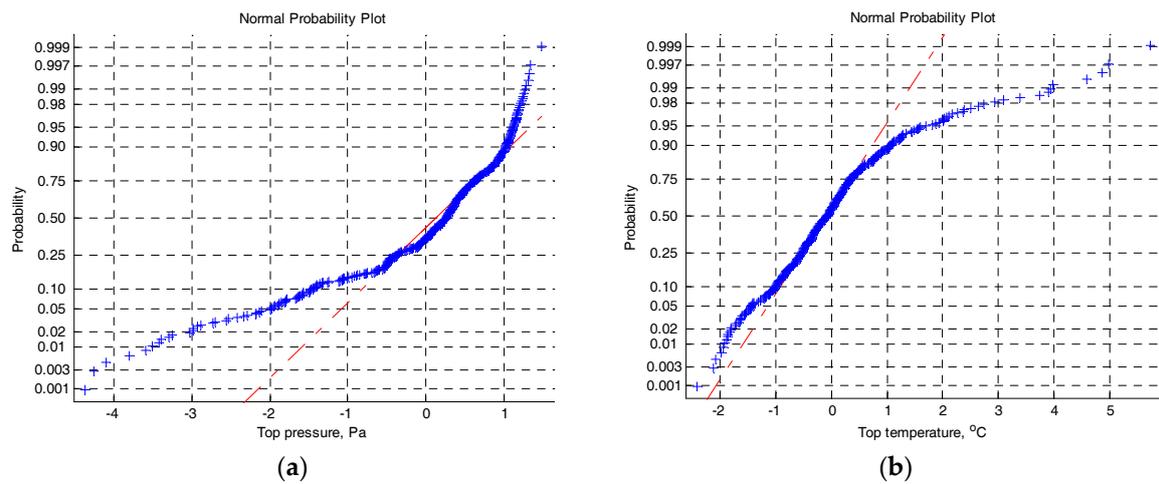


Figure 2. (a) The normal probability plot of the top pressure variable in the training set; and (b) the normal probability plot of the top temperature variable in the training set.

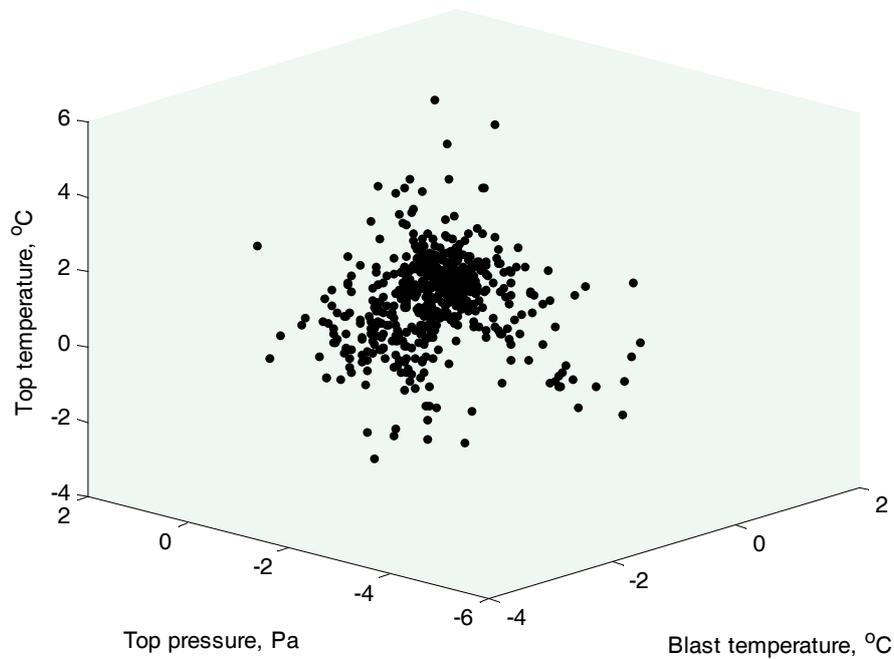


Figure 3. The spatial relationship of several process input variables in the training set.

To show the advantage of JCSVR, it is compared with three SVR-based soft sensors, including JLSSVR [25], CSVN [46], and LSSVR [47]. To evaluate the prediction performance of different models, three indices of the root-mean-square error (RMSE), relative RMSE (simply noted as RE), and the hit rate (HR) [21–31] are utilized and defined below, respectively:

$$\text{RMSE} = \sqrt{\sum_{q=1}^{N_{\text{tst}}} \left(\frac{y_q - \hat{y}_q}{N_{\text{tst}}} \right)^2} \quad (7)$$

$$\text{RE} = \sqrt{\sum_{q=1}^{N_{\text{tst}}} \left(\frac{y_q - \hat{y}_q}{y_q} \right)^2 / N_{\text{tst}}} \quad (8)$$

$$\begin{cases} \text{HR} = \frac{\sum_{q=1}^{N_{\text{tst}}} H_q}{N_{\text{tst}}} \times 100\% \\ \text{where } H_q = \begin{cases} 1, & |\hat{y}_q - y_q| \leq 0.1 \\ 0, & \text{else} \end{cases} \end{cases} \quad (9)$$

where y_q and \hat{y}_q are the actual value and the predicted value, respectively, and N_{tst} is the number of test data point.

The effect of a CSVR model is first investigated. After training, the main results, including the weighted terms $\rho(e_i)$, of a CSVR model are shown in Figure 4. Using the correntropy-based strategy, the outliers can be assigned with smaller weights different from most normal samples. As a result, the bad influence of outliers can be reduced. Here, the cut-off parameter is selected as $\bar{\rho} = 0.7$. As shown in the bottom subplot of Figure 4, some candidate outliers can be identified directly. Finally, altogether 44 candidate outliers are chosen from all 240 training data. About 20% ($44/240 = 18.3\%$) abnormal data, this indicates that the training data are noisy and contain several inconspicuous outliers. If the negative effect of these outliers are not removed, the prediction performance of established soft sensors cannot be good.

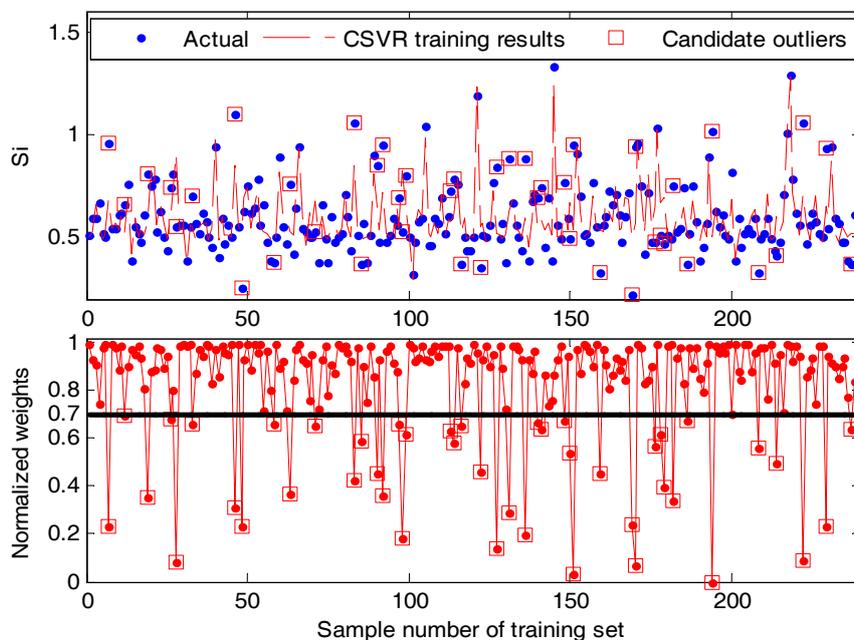


Figure 4. The trained CSVR model for fitting data with the normalized weights $\rho(e_i)$. As an affiliated product, those data with $\rho(e_i) < \bar{\rho} = 0.7$ can be simply identified as candidate outliers.

For comparison, the performance indices of the CSVR and LSSVR methods for the training data are listed in Table 1. The fitting results of both CSVR and LSSVR methods are not good. One main reason is that the data are noisy, non-Gaussian, and unevenly distributed. If the model fits all noisy training data, especially for the outliers, the over-fitting problem occurs. It can be noticed that the traditional LSSVR model cannot provide more information about the training data, treating all the training data equally. Different from LSSVR, the CSVR model can distinguish outliers from normal data and assign the training data with suitable weights.

The dynamics may change in an industrial blast furnace. In such a situation, a fixed soft sensor model may be not accurate for future data [25,30]. Here, the proposed JCSVr-based method is compared with a recent local modeling method, named JLSSVR [25]. For the test data, the online prediction results and corresponding absolute prediction errors ($|y_q - \hat{y}_q|$) of JCSVr and JLSSVR methods are shown in Figures 5 and 6, respectively. To show the result more clear, only the first 70

testing samples are plotted. As aforementioned, clean data are needed for online construction of a good local model. With some unwanted outliers, the prediction performance of a local model may be unreliable. Therefore, the prediction results shown in Figures 5 and 6 indicate that JCSVR is superior to JLSSVR for industrial data-driven modeling problems with noisy data.

Table 1. Fitting results comparison of CSVR and LSSVR soft sensor models for the training set.

Soft Sensor Model	RMSE	RE (%)	HR (%)
CSVR [46]	0.116	21.49	73.33
LSSVR [24,47]	0.122	23.04	66.25

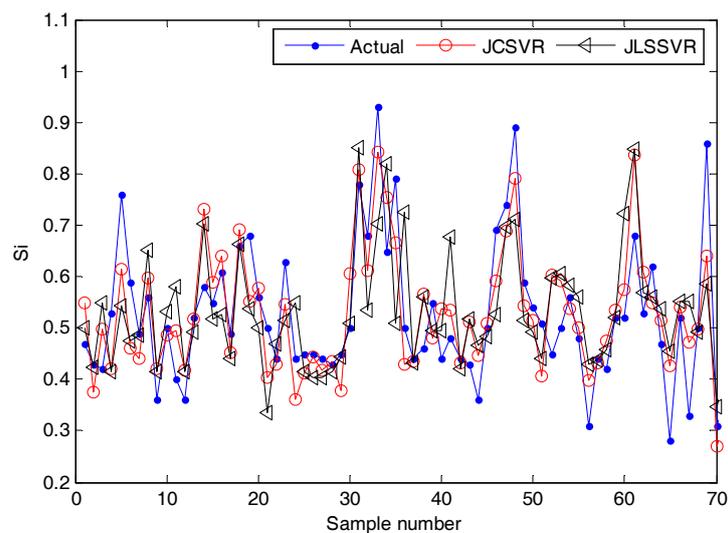


Figure 5. Comparison results of online silicon content prediction using JCSVR and JLSSVR models (test set).

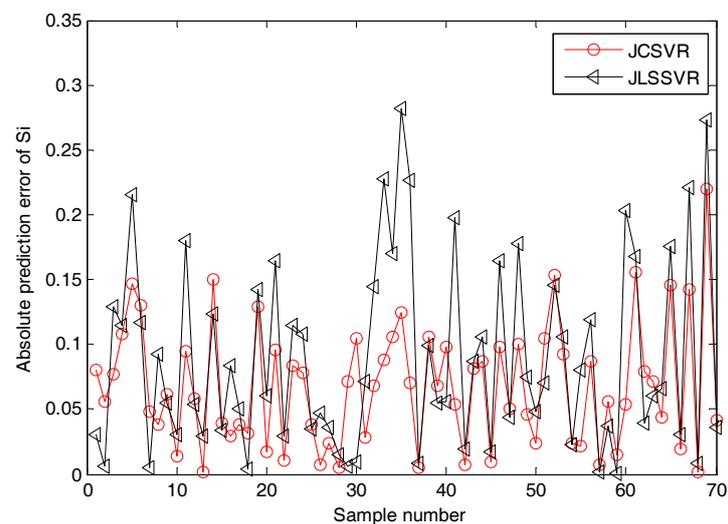


Figure 6. Comparison results of online silicon content prediction error using JCSVR and JLSSVR models (test set).

The main properties of the JCSVR, JLSSVR, CSVR, and LSSVR approaches are summarized in Table 2. Briefly, the outlier identification and local modeling are integrated into the JCSVR method. Detailed values about online silicon content prediction comparisons of four methods are listed in Table 3. It shows that the JCSVR method, for the test set, achieves the best prediction performance.

Additionally, local models are generally more accurate than their global ones. For example, JCSVR shows better prediction performance than only using a CSV model.

Table 2. Brief description of four different prediction models.

Prediction Model	Brief Description			
	Local and Unfixed	Outlier Detection	Main Pros	Main Cons
JCSVR	Yes	Yes	More suitable for noisy and uneven distributed data	Model selection is relatively difficult
JLSSVR [25]	Yes	No	Suitable for online modeling of nonlinear processes	Not robust for outliers
CSV [46]	No	Yes	Suitable for construction of a global model with noisy data	Prediction accuracy for some local regions may not be enough
LSSVR [24,47]	No	No	A simple nonlinear modeling method	Not robust for outliers and relatively inaccurate for some local regions

Table 3. Comparison results of online silicon content prediction for the test set using four different models.

Prediction Model	RMSE	RE (%)	HR (%)
JCSVR	0.079	17.70	80.5
JLSSVR [25]	0.105	22.51	64.5
CSV [46]	0.123	28.16	61.5
LSSVR [24,47]	0.141	31.29	52.5

To show the relative prediction errors (i.e., $\frac{y_q - \hat{y}_q}{y_q}$) of four methods, their corresponding box plots are shown in Figure 7. The band inside the box is the median value, and the box edges denote the first and third quartiles. A few outliers are shown individually. Among four methods, JCSVR exhibits the narrowest distribution. The median value of JCSVR is nearest to 0. These results imply that JCSVR has the best prediction performance. One main reason is that the database is maintained continually. In contrast, without maintenance of the database, JLSSVR and LSSVR become unreliable and not suitable for long-term prediction. This is a common problem of traditional soft sensor models utilized in industrial processes. Based on all the prediction results and analysis, JCSVR is the most suitable one among all of the methods.

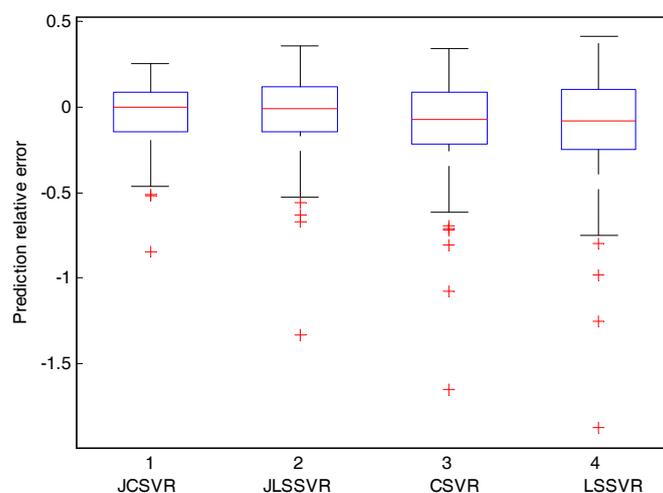


Figure 7. Comparison results of silicon content prediction relative error distribution with JCSVR, JLSSVR, CSV, and LSSVR models (test set).

5. Conclusions

This work has proposed a correntropy-based local soft sensor modeling method for silicon content prediction when the collected data contain uncertainties. Its main distinguished characteristics are summarized. First, the soft sensor and outlier detection can be integrated into a CSVR modeling framework. By simply removing the candidate outliers, the updated historical data are more reliable for construction of local models. Second, by incorporating the database update into the clustering-based JCSVR method, better prediction performance can be achieved. Consequently, the proposed method can reduce the effect of outliers. Compared with several methods, better silicon content prediction results of JCSVR are obtained. There are still several interesting research directions worth investigating. First, other forms of correntropy can be adopted to adapt to the uncertainty of sensor data. Second, development of efficient feature extraction method for noisy data is interesting. Third, how to incorporate process knowledge to further improve the prediction accuracy is important and challenging.

Acknowledgments: The authors would like to gratefully acknowledge the National Natural Science Foundation of China (grant no. 61640312) and Foundation of Key Laboratory of Advanced Process Control for Light Industry (Jiangnan University), Ministry of Education, China (grant no. APCLI1603) for their financial support.

Author Contributions: K. Chen and Y. Liu conceived and designed the experiments; Y. Liang performed the experiments; K. Chen and Y. Liu analyzed the data; Y. Liu contributed analysis tools; K. Chen, Z. Gao and Y. Liu wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CSF	Cumulative Similarity Factor
CSVR	Correntropy Support Vector Regression
HR	Hit Rate
JCSVR	Just-in-time Correntropy Support Vector Regression
JLSSVR	Just-in-time Least Squares Support Vector Regression
LSSVR	Least Squares Support Vector Regression
RE	Relative root-mean-square Error
RMSE	Root Mean Square Error
SI	Similarity Index
SVR	Support Vector Regression

References

1. Wijk, O.; Torssell, K.; Bl, X.; Torssell, K.; Wijk, O. Prediction of the blast furnace process by a mathematical model. *ISIJ Int.* **1992**, *32*, 481–488.
2. Sugawara, K.; Morimoto, K.; Sugawara, T.; Dranoff, J.S. Dynamic behavior of iron forms in rapid reduction of carbon-coated iron ore. *AIChE J.* **1999**, *45*, 574–580. [[CrossRef](#)]
3. Radhakrishnan, V.R.; Ram, K.M. Mathematical model for predictive control of the bell-less top charging system of a blast furnace. *J. Process. Control.* **2001**, *11*, 565–586. [[CrossRef](#)]
4. Nogami, H.; Chu, M.; Yagi, J.I. Multi-dimensional transient mathematical simulator of blast furnace process based on multi-fluid and kinetic theories. *Comput. Chem. Eng.* **2005**, *29*, 2438–2448. [[CrossRef](#)]
5. Nishioka, K.; Maeda, T.; Shimizu, M. A three-dimensional mathematical modelling of drainage behavior in blast furnace hearth. *ISIJ Int.* **2005**, *45*, 669–676. [[CrossRef](#)]
6. Ueda, S.; Natsui, S.; Nogami, H.; Jun-Ichiro, Y.; Ariyama, T. Recent progress and future perspective on mathematical modeling of blast furnace. *ISIJ Int.* **2010**, *50*, 914–923. [[CrossRef](#)]
7. Singh, H.; Sridhar, N.V.; Deo, B. Artificial neural nets for prediction of silicon content of blast furnace hot metal. *Steel Res. Int.* **1996**, *67*, 521–526. [[CrossRef](#)]

8. Radhakrishnan, V.R.; Mohamed, A.R. Neural networks for the identification and control of blast furnace hot metal quality. *J. Process. Control.* **2000**, *10*, 509–524. [[CrossRef](#)]
9. Chen, J. A predictive system for blast furnaces by integrating a neural network with qualitative analysis. *Eng. Appl. Artif. Intell.* **2001**, *14*, 77–85. [[CrossRef](#)]
10. Jimenez, J.; Mochon, J.; Sainz, D.A.J.; Obeso, F. Blast furnace hot metal temperature prediction through neural networks-based models. *ISIJ Int.* **2007**, *44*, 573–580. [[CrossRef](#)]
11. Saxen, H.; Pettersson, F. Nonlinear prediction of the hot metal silicon content in the blast furnace. *ISIJ Int.* **2007**, *47*, 1732–1737. [[CrossRef](#)]
12. Pettersson, F.; Chakraborti, N.; Saxén, H. A genetic algorithms based multi-objective neural net applied to noisy blast furnace data. *Appl. Soft Comput.* **2007**, *7*, 387–397. [[CrossRef](#)]
13. Nurkkala, A.; Pettersson, F.; Saxén, H. Nonlinear modeling method applied to prediction of hot metal silicon in the ironmaking blast furnace. *Ind. Eng. Chem. Res.* **2011**, *50*, 9236–9248. [[CrossRef](#)]
14. Hao, X.; Shen, F.; Du, G.; Shen, Y.; Xie, Z. A blast furnace prediction model combining neural network with partial least square regression. *Steel Res. Int.* **2005**, *76*, 694–699. [[CrossRef](#)]
15. Bhattacharya, T. Prediction of silicon content in blast furnace hot metal using partial least squares (PLS). *ISIJ Int.* **2005**, *45*, 1943–1945. [[CrossRef](#)]
16. And, M.W.; Saxén, H. On the development of predictive models with applications to a metallurgical process. *Ind. Eng. Chem. Res.* **2000**, *39*, 982–988.
17. Gao, C.; Zhou, Z.; Chen, J. Assessing the predictability for blast furnace system through nonlinear time series analysis. *Ind. Eng. Chem. Res.* **2008**, *47*, 3037–3045. [[CrossRef](#)]
18. Waller, M.; Saxen, H. Application of nonlinear time series analysis to the prediction of silicon content of pig iron. *ISIJ Int.* **2002**, *42*, 316–318. [[CrossRef](#)]
19. Miyano, T.; Kimoto, S.; Shibuta, H.; Nakashima, K.; Ikenaga, Y.; Aihara, K. Time series analysis and prediction on complex dynamical behavior observed in a blast furnace. *Phys. D: Nonlinear Phenom.* **2000**, *135*, 305–330. [[CrossRef](#)]
20. Martin, R.D.; Obeso, F.; Mochon, J.; Barea, R.; Jimenez, J. Hot metal temperature prediction in blast furnace using advanced model based on fuzzy logic tools. *Ironmak. Steelmak.* **2007**, *34*, 241–247. [[CrossRef](#)]
21. Zeng, J.S.; Gao, C.H. Improvement of identification of blast furnace ironmaking process by outlier detection and missing value imputation. *J. Process. Control.* **2009**, *19*, 1519–1528. [[CrossRef](#)]
22. Jian, L.; Gao, C.; Xia, Z. A sliding-window smooth support vector regression model for nonlinear blast furnace system. *Steel Res. Int.* **2011**, *82*, 169–179. [[CrossRef](#)]
23. Gao, C.; Jian, L.; Luo, S. Modeling of the thermal state change of blast furnace hearth with support vector machines. *IEEE Trans. Ind. Electron.* **2012**, *59*, 1134–1145. [[CrossRef](#)]
24. Jian, L.; Shen, S.; Song, Y. Improving the solution of least squares support vector machines with application to a blast furnace system. *J. Appl. Math.* **2012**, *2012*, 12. [[CrossRef](#)]
25. Liu, Y.; Gao, Z. Enhanced just-in-time modelling for online quality prediction in bf ironmaking. *Ironmak. Steelmak.* **2015**, *42*, 321–330. [[CrossRef](#)]
26. Gao, C.; Chen, J.; Zeng, J.; Liu, X.; Sun, Y. A chaos-based iterated multistep predictor for blast furnace ironmaking process. *AIChE J.* **2009**, *55*, 947–962. [[CrossRef](#)]
27. Luo, S.; Gao, C.; Zeng, J.; Huang, J. Blast furnace system modeling by multivariate phase space reconstruction and neural networks. *Asian J. Control.* **2013**, *15*, 553–561. [[CrossRef](#)]
28. Gao, C.; Zeng, J.; Zhou, Z. Identification of multiscale nature and multiple dynamics of the blast furnace system from operating data. *AIChE J.* **2011**, *57*, 3448–3458. [[CrossRef](#)]
29. Chu, Y.; Gao, C. Data-based multiscale modeling for blast furnace system. *AIChE J.* **2014**, *60*, 2197–2210. [[CrossRef](#)]
30. Nurkkala, A.; Pettersson, F.; Saxén, H. A study of blast furnace dynamics using multiple autoregressive vector models. *ISIJ Int.* **2012**, *52*, 1763–1770. [[CrossRef](#)]
31. Saxén, H.; Gao, C.; Gao, Z. Data-driven time discrete models for dynamic prediction of the hot metal silicon content in the blast furnace—A review. *IEEE Trans. Ind. Inform.* **2013**, *9*, 2213–2225. [[CrossRef](#)]
32. Kano, M.; Nakagawa, Y. Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry. *Comput. Chem. Eng.* **2008**, *32*, 12–24. [[CrossRef](#)]
33. Kadlec, P.; Gabrys, B.; Strandt, S. Data-driven soft sensors in the process industry. *Comput. Chem. Eng.* **2009**, *33*, 795–814. [[CrossRef](#)]

34. Zur, R.M.; Jiang, Y.; Pesce, L.L.; Drukker, K. Noise injection for training artificial neural networks: A comparison with weight decay and early stopping. *Med. Phys.* **2009**, *36*, 4810–4818. [[CrossRef](#)] [[PubMed](#)]
35. Arbabi, V.; Pouran, B.; Campoli, G.; Weinans, H.; Zadpoor, A.A. Determination of the mechanical and physical properties of cartilage by coupling poroelastic-based finite element models of indentation with artificial neural networks. *J. Biomech.* **2016**, *49*, 631–637. [[CrossRef](#)] [[PubMed](#)]
36. Arbabi, V.; Pouran, B.; Weinans, H.; Zadpoor, A.A. Combined inverse-forward artificial neural networks for fast and accurate estimation of the diffusion coefficients of cartilage based on multi-physics models. *J. Biomech.* **2016**, *49*, 2799–2805. [[CrossRef](#)] [[PubMed](#)]
37. Liu, H.; Shah, S.; Jiang, W. Online outlier detection and data cleaning. *Comput. Chem. Eng.* **2004**, *28*, 1635–1647. [[CrossRef](#)]
38. Pearson, R.K. Outliers in process modeling and identification. *IEEE Trans. Control. Syst. Technol.* **2002**, *10*, 55–63. [[CrossRef](#)]
39. Chiang, L.H.; Pell, R.J.; Seasholtz, M.B. Exploring process data with the use of robust outlier detection algorithms. *J. Process. Control.* **2003**, *13*, 437–449. [[CrossRef](#)]
40. Hodge, V.J.; Austin, J. A survey of outlier detection methodologies. *Artif. Intell. Rev.* **2004**, *22*, 85–126. [[CrossRef](#)]
41. Fujiwara, K.; Kano, M.; Hasebe, S.; Takinami, A. Soft-sensor development using correlation-based just-in-time modeling. *AIChE J.* **2009**, *55*, 1754–1765. [[CrossRef](#)]
42. Bontempi, G.; Birattari, M.; Bersini, H. Lazy learning for local modelling and control design. *Int. J. Control.* **1999**, *72*, 643–658. [[CrossRef](#)]
43. Cheng, C.; Chiu, M.S. A new data-based methodology for nonlinear process modeling. *Chem. Eng. Sci.* **2004**, *59*, 2801–2810. [[CrossRef](#)]
44. Liu, Y.; Chen, J. Integrated soft sensor using just-in-time support vector regression and probabilistic analysis for quality prediction of multi-grade processes. *J. Process. Control.* **2013**, *23*, 793–804. [[CrossRef](#)]
45. Liu, W.; Pokharel, P.P.; Principe, J.C. Correntropy: Properties and applications in non-gaussian signal processing. *IEEE Trans. Signal. Process.* **2007**, *55*, 5286–5298. [[CrossRef](#)]
46. Liu, Y.; Chen, J. Correntropy kernel learning for nonlinear system identification with outliers. *Ind. Eng. Chem. Res.* **2014**, *53*, 5248–5260. [[CrossRef](#)]
47. Suykens, J.A.K.; Van Gestel, T.; De Brabanter, J.; De Moor, B.; Vandewalle, J. *Least Squares Support Vector Machines*; World Scientific: Singapore, 2002.
48. Wen, W.; Hao, Z.; Yang, X. A heuristic weight-setting strategy and iteratively updating algorithm for weighted least-squares support vector regression. *Neurocomputing* **2008**, *71*, 3096–3103. [[CrossRef](#)]
49. Miyamoto, S.; Ichihashi, H.; Honda, K. *Algorithms for Fuzzy Clustering: Methods in c-Means Clustering with Applications*; Springer: New York, NY, USA, 2008.

