A nonlinear system modelling approach to industrial cane sugar crystallization

Guancheng Lu, Yanmei Meng*, Jian Chen, Zhihong Tang, Xiaochun Wang, Xian Yu

College of Mechanical Engineering, Guangxi University, No. 100, Daxue Road, Nanning, China

Received 1 July 2014, www.cmnt.lv

Abstract

Cane sugar crystallization is a non-linear process where multiple control parameters are involved, which makes it rather difficult to reveal its internal mechanism by mechanism modelling. Derived from variants of standard support vector machine method, an online control system modelling method based on multi-input and multi-output proximal least square support vector machine is proposed to be applied in sugar crystallization process. This method takes multiple process control parameters as the input and output of machine learning algorithm, through which the inherent law between key and auxiliary parameters in the sugar crystallization process is established. The ultimate goal is to control the sugar crystallization process automatically. The experimental results show that the accuracy rate of the model output is 95%.

Keywords: multi-input and multi-output proximal least square SVM, sugar crystallization control system, machine learning, nonlinear modelling

1 Introduction

Nowadays techniques like soft sensor based on computer and sensing technology, process identification and nonlinear process control are widely applied in modern industrial field. Among them, process identification and soft sensor based on computer sensing technology is a method that establishes the mathematic relationship between leading variables and auxiliary variables, where the leading variables are those hard to be measured or even can't be measured and the auxiliary variables are those relevant and easy to be measured. Then the evaluation of leading variables is implemented by auxiliary variables. An organic combination of computer and industrial process knowledge has been achieved by soft sensor technology, which makes variables measurement easier and more accessible by replacing hardware with software and taking auxiliary variables to evaluate the leading ones.

Cane sugar crystallization is a nonlinear and complex process as well as hard to be modelled. The key parameters in that control process field like super saturation and Brix are not stable enough for long term measurement. The accuracy of measuring sensors would be compromised once their surfaces are covered with scaling sugar in crystallization process, which stands in the way of automatic controlling in Chinese sugar industry field. However, machine learning methods could identify the inherent law between the input and output in control process through learning and modelling the relevant data, without knowing exactly what kind of mechanism the process is. It means that the machine learning methods can approach to the actual process closely.

Support Vector Machine (SVM) is a new machine learning technique based on minimization principle, with features of solid theoretical foundation and basis [1-3]. SVM which features with good generalization ability can still obtain forecast value under some extreme conditions, like nonlinearity, small sample and high-dimension data [4-6]. SVM is generally applied in industrial fields of iron and steel production, biopharming and food processing for parameter forecasting, with trend of penetrating into other industrial fields.

2 Measurement and control system model for sugar crystallization process

Crystallization is the most significant phase of sugar boiling, and its mechanism is complex and hard to be modelled. The key control parameters like massecuite Brix and supersaturation can be obtained through measuring and analysing the following external parameters: temperature, vacuum degree, steam pressure etc. Since those external parameters determine the massecuite Brix and supersaturation, there should exist some certain law between them. Even though the inherent law can be established by analysing the mechanism of sugar crystallization process, it is still difficult to figure it out from the perspective of the mechanism modelling since multiple parameters are involved and the process is nonlinear.

Since the internal mechanism of sugar crystallization process is hard to be modelled based on traditional theory, this paper applies machine learning to cave the implicit relation among the key control parameters in sugar

* ***Corresponding author*** e-mail: gxu_mengyun@163.com

crystallization process. After that, the measurement and control system model for sugar crystallization process has been built, shown as Figure 1.

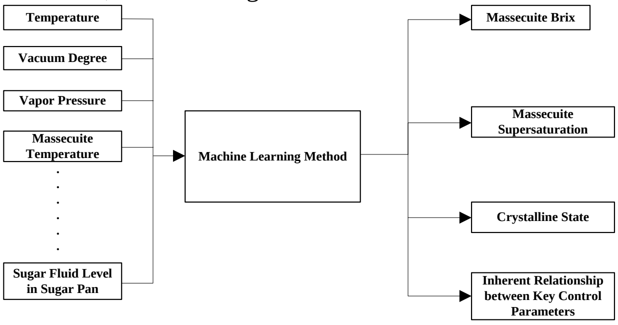


FIGURE 1 Control modelling scheme of cane sugar crystallization process

3 Base modelling building

In this paper, machine-learning method was adopted to extract the inherent law between the multiple key parameters and the auxiliary parameters in the nonlinear cane sugar crystallization process, which aimed to controlling the process automatically. The multiple control parameters in that process were seen as the input and output of the machine learning method.

The input vector $x(x \ni R^d)$ in d dimensions was defined as the auxiliary parameter in the process; whereas the output vector $Y(Y \ni R^n)$ in n dimensions was defined as the key parameter. Measurement and control system model for cane sugar crystallization process was built based on multi-input multi-output proximal least square support vector machine (MIMO PLS-SVM) [7, 8]. The object function for optimizing the model satisfies the following equation:

$$\min_{w_j, e_{i,j}} J^{(n)}\left(\mathbf{w}_j, e_{i,j}\right) = \frac{1}{2}\sum_{j=1}^{n}\left(\mathbf{w}_j^T\mathbf{w}_j + b_j^2\right) + \frac{c}{2}\sum_{j=1}^{n}\sum_{i=1}^{m} e_{i,j}^2, c<0,$$
$$s.t\begin{cases} y_{i,1} = \mathbf{w}_1^T\Phi_1\left(\mathbf{x}_i\right) + b_1 + e_{i,1}, \quad i = 1,2,...,\mathrm{m} \\ y_{i,2} = \mathbf{w}_2^T\Phi_2\left(\mathbf{x}_i\right) + b_2 + e_{i,2}, \; i = 1,2,...,\mathrm{m} \\ \ldots \\ y_{i,n} = \mathbf{w}_n^T\Phi_n\left(\mathbf{x}_i\right) + b_n + e_{i,n}, \; i = 1,2,...,\mathrm{m} \end{cases} \quad ,(1)$$

where b_j is the j-dimensional offset vector; w_j is the j-dimensional weight vector; x is the high dimensional transformation function; $e_{i,j}$ is the quantitative error in i^{th} row and j^{th} column; c is the penalty factor of irrelevant samples. The following Lagrange function is deduced from the object function:

$$L^{(n)}\left(\mathbf{w}, b_j, e_{i,j}, a_{i,j}\right) = J^{(n)} - \sum_{j=1}^{n}\sum_{i=1}^{m} a_{i,j}\left\{\mathbf{w}_j^T\Phi_j\left(\mathbf{x}_i\right) + b_j + e_{i,j} - y_{i,j}\right\}, \quad (2)$$

The conditions for optimality are the following equations:

$$\begin{cases} \dfrac{\partial L}{\partial \mathbf{w}_j} = \mathbf{w}_j - \sum_{i=1}^{m} a_{i,j}\Phi_j\left(\mathbf{x}_i\right) = 0 \to \mathbf{w}_j = \sum_{i=1}^{m} a_{i,j}\Phi_j\left(\mathbf{x}_i\right) \\ \dfrac{\partial L}{\partial b_j} = b_j - a_{i,j} = 0 \to b_j = a_{i,j} \\ \dfrac{\partial L}{\partial e_{i,j}} = ce_{i,j} = a_j = 0 \to e_{i,j} = \dfrac{a_{i,j}}{c} \\ \dfrac{\partial L}{\partial a_{i,j}} = 0 - \left\{\mathbf{w}_j^T\Phi_j\left(\mathbf{x}_i\right) + b_j + e_{i,j} - y_{i,j}\right\} = 0 \to \mathbf{w}_j^T\Phi_j\left(\mathbf{x}_i\right) + b_j + e_{i,j} - y_{i,j} = 0 \\ i = 1,2,...,m \\ j = 1,2,...\mathrm{n} \end{cases} . \quad (3)$$

Linear equations system of eliminating the transition variables in Equation (3):

$$\begin{bmatrix} \Phi\left(\mathbf{x}_1\right)\Phi\left(\mathbf{x}_1\right) + 1 + \dfrac{1}{c} & \cdots & \Phi\left(\mathbf{x}_1\right)\Phi\left(\mathbf{x}_m\right) + 1 \\ \vdots & \ddots & \vdots \\ \Phi\left(\mathbf{x}_m\right)\Phi\left(\mathbf{x}_1\right) + 1 & \cdots & \Phi\left(\mathbf{x}_m\right)\Phi\left(\mathbf{x}_m\right) + 1 + \dfrac{1}{c} \end{bmatrix} \times \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} = Y, \quad (4)$$

According to the principle of SVM, kernel function is introduced as follows:

$$\mathbf{K} = \begin{bmatrix} K\left(\mathbf{x}_1, \mathbf{x}_1\right) & \cdots & K\left(\mathbf{x}_1, \mathbf{x}_m\right) \\ \vdots & \ddots & \vdots \\ K\left(\mathbf{x}_m, \mathbf{x}_1\right) & \cdots & K\left(\mathbf{x}_m, \mathbf{x}_m\right) \end{bmatrix}, \quad (5)$$

$$a = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}, \quad (6)$$

$$E=\begin{bmatrix}1 & \cdots & 1\\ \vdots & \ddots & \vdots\\ 1 & \cdots & 1\end{bmatrix}, \tag{7}$$

$$I=\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}. \tag{8}$$

Consequently, the linear equations system can be written as matrix form:

$$(\mathbf{K}+\mathbf{E}+c\mathbf{I})\mathbf{a}=\mathbf{Y}. \tag{9}$$

Left Matrix in Equations (9) is named as model identification matrix as follows:

$$\mathbf{H}=\begin{bmatrix}K(\mathbf{x}_1,\mathbf{x}_1)+1+\frac{1}{c} & \cdots & K(\mathbf{x}_1,\mathbf{x}_m)+1\\ \vdots & \ddots & \vdots\\ K(\mathbf{x}_m,\mathbf{x}_1)+1 & \cdots & K(\mathbf{x}_m,\mathbf{x}_m)+1+\frac{1}{c}\end{bmatrix}. \tag{10}$$

Introduced by the positive definiteness of the kernel function, we can get the following relationship:

$$K(\mathbf{x}_i,\mathbf{x}_j)=(\Phi(\mathbf{x}_i)\Phi(\mathbf{x}_j))=(\Phi(\mathbf{x}_j)\Phi(\mathbf{x}_i))=K(\mathbf{x}_j,\mathbf{x}_i),$$

which suggest matrix H is symmetric. Assuming z is a nonzero vector, do the operation as follows:

$$\mathbf{z}^T\mathbf{H}\mathbf{z}=\mathbf{z}^T\mathbf{K}\mathbf{z}+\mathbf{z}^T\mathbf{E}\mathbf{z}+c\mathbf{z}^T\mathbf{I}\mathbf{z}=$$
$$\left\|\sum_{i=1}^{l}\lambda_i\Phi(\mathbf{x}_i)\right\|^2+\left(\sum_{i=1}^{l}z_i\right)^2+\frac{1}{c}\sum_{i=1}^{l}z_i^2>0. \tag{11}$$

Since matrix H is symmetric and positive, Equation (9) has a unique solution due to the presence of its inverse matrix.

Known from the object function, the measurement and control model built on MIMO PLS-SVM can feed the quantified correlativity among multiple variables in the cane sugar crystallization process back to the output variables efficiently and synchronously. Mutual adjustment between those outputs variables are made on condition of the identical input variables, which leads to coordination of the whole control objects. In a certain sense, the whole measurement and control process is optimized.

4 Online measurement and control model building of cane sugar crystallization process

According to Equations (9) and (10), the solution of Equation (9) is determined by the inverse matrix of the symmetric and positive H. However, solving the inverse matrix directly is of high computation cost and time-consuming, which is incompatible with the online measurement and control model. Obviously, matrix inversion approach available for online measurement and control condition is highly demanded.

At time t, matrix H is defined as follows:

$$\mathbf{H}_t=\begin{bmatrix}K(\mathbf{x}_1,\mathbf{x}_1)+1+\frac{1}{c} & \cdots & K(\mathbf{x}_1,\mathbf{x}_t)+1\\ \vdots & \ddots & \vdots\\ K(\mathbf{x}_t,\mathbf{x}_1)+1 & \cdots & K(\mathbf{x}_t,\mathbf{x}_t)+1+\frac{1}{c}\end{bmatrix}. \tag{12}$$

At time $t+1$, the new data described as input vector x_{t+1} is inserted to the online model, then H can be written as follows:

$$\mathbf{H}_{t+1}=\begin{bmatrix}K(\mathbf{x}_1,\mathbf{x}_1)+1+\frac{1}{c} & \cdots & K(\mathbf{x}_1,\mathbf{x}_t)+1 & K(\mathbf{x}_1,\mathbf{x}_{t+1})+1\\ \vdots & \ddots & \vdots & \vdots\\ K(\mathbf{x}_t,\mathbf{x}_1)+1 & \cdots & K(\mathbf{x}_t,\mathbf{x}_t)+1+\frac{1}{c} & K(\mathbf{x}_t,\mathbf{x}_{t+1})+1\\ K(\mathbf{x}_{t+1},\mathbf{x}_1)+1 & \cdots & K(\mathbf{x}_{t+1},\mathbf{x}_t)+1 & K(\mathbf{x}_{t+1},\mathbf{x}_{t+1})+1+\frac{1}{c}\end{bmatrix}. \tag{13}$$

Assuming:

$$\mathbf{h}_{new}=\begin{bmatrix}K(\mathbf{x}_{t+1},\mathbf{x}_1)+1\\ \vdots\\ K(\mathbf{x}_{t+1},\mathbf{x}_t)+1\end{bmatrix}, \tag{14}$$

$$h_{new}=K(\mathbf{x}_{t+1},\mathbf{x}_{t+1})+1+\frac{1}{c}. \tag{15}$$

According to the inversion of block matrix when new data is inserted to the online model, the inversion of H_{t+1} is obtained using the inverse of H_t at time t, which simplifies the computation process massively [9]. At time $t+1$, the solution of linear equations system of the online model is defined as a_{t+1}; the newly added data is regarded as input vector y_{t+1}; Y_{t+1} denotes the output matrix; Consequently a_{t+1} can be rewritten by Equation (17), where the needed computing time is decreased largely. Equations system of online measurement and control model is assumed as $H_{t+1}a_{t+1}=Y_{t+1}$, and the solution deduction is as follows:

$$\mathbf{H}_{t+1}^{-1}=\begin{bmatrix}\mathbf{H}_t & \mathbf{h}_{new}\\ \mathbf{h}_{new}^T & h_{new}\end{bmatrix}^{-1}=\begin{bmatrix}\mathbf{H}_t^{-1}+\mathbf{H}_t^{-1}\mathbf{h}_{new}^T\mathbf{H}_t^{-1}\left(h_{new}-\mathbf{h}_{new}^T\mathbf{H}_t^{-1}\mathbf{h}_{new}^T\right)^{-1} & \mathbf{H}_t^{-1}\mathbf{h}_{new}\left(h_{new}-\mathbf{h}_{new}^T\mathbf{H}_t^{-1}\mathbf{h}_{new}^T\right)^{-1}\\ \mathbf{h}_{new}^T\mathbf{H}_t^{-1}\left(h_{new}-\mathbf{h}_{new}^T\mathbf{H}_t^{-1}\mathbf{h}_{new}^T\right)^{-1} & \left(h_{new}-\mathbf{h}_{new}^T\mathbf{H}_t^{-1}\mathbf{h}_{new}^T\right)^{-1}\end{bmatrix}=$$
$$\begin{bmatrix}\mathbf{H}_t & 0\\ 0 & 0\end{bmatrix}+\begin{bmatrix}\mathbf{H}_t^{-1}\mathbf{h}_{new}\\ -1\end{bmatrix}\begin{bmatrix}\mathbf{h}_{new}^T\mathbf{H}_t^{-1} & -1\end{bmatrix}\frac{1}{h_{new}-\mathbf{h}_{new}^T\mathbf{H}_t^{-1}\mathbf{h}_{new}^T}, \tag{16}$$

$$\mathbf{a}=\mathbf{H}_{t+1}^{-1}\mathbf{Y}_{t+1}=\left(\begin{bmatrix}\mathbf{H}_t & 0\\ 0 & 0\end{bmatrix}+\begin{bmatrix}\mathbf{H}_t^{-1}\mathbf{h}_{new}\\ -1\end{bmatrix}\begin{bmatrix}\mathbf{h}_{new}^T\mathbf{H}_t^{-1} & -1\end{bmatrix}\frac{1}{h_{new}-\mathbf{h}_{new}^T\mathbf{H}_t^{-1}\mathbf{h}_{new}^T}\right)\mathbf{Y}_{t+1}=$$
$$\begin{bmatrix}\mathbf{H}_t^{-1}\mathbf{Y}_t\\ 0\end{bmatrix}+\frac{1}{h_{new}-\mathbf{h}_{new}^T\mathbf{H}_t^{-1}\mathbf{h}_{new}^T}\left(\begin{bmatrix}\mathbf{h}_{new}^T\mathbf{H}_t^{-1} & -1\end{bmatrix}\mathbf{Y}_t-\mathbf{y}_{t+1}\right)\begin{bmatrix}\mathbf{H}_t^{-1}\mathbf{h}_{new}\\ -1\end{bmatrix}, \tag{17}$$

The data of the online model would increase rapidly along with time flowing and data updating, which compromises the operation speed. Aimed to solve this problem, the data amount needed processing must be limited without compromising the identification precision of the model. When the data amount reaches to a certain point, space for processing the newly data would be squeezed out by deleting the old data. In the meantime, the inverse matrix and solution of the equations system of the model would be updated. Sliding window techniques are used to clear the old data out the online model. Preprocessing is made as follows:

$$\mathbf{h}_{old}=\begin{bmatrix}K(\mathbf{x}_2,\mathbf{x}_1)+1\\ \vdots\\ K(\mathbf{x}_{t+1},\mathbf{x}_1)+1\end{bmatrix}, \tag{18}$$

$$h_{old}=K(\mathbf{x}_1,\mathbf{x}_1)+1+\frac{1}{c}. \tag{19}$$

Then we have:

$$\mathbf{H}_{t+1}^{-1}=\begin{bmatrix}h_{old} & \mathbf{h}_{old}\\ \mathbf{h}_{old}^T & \bar{\mathbf{H}}_{t+1}\end{bmatrix}^{-1}=\begin{bmatrix}\left(h_{old}-\mathbf{h}_{old}\bar{\mathbf{H}}_{t+1}^{-1}\mathbf{h}_{old}^T\right)^{-1} & -\mathbf{h}_{old}^T\bar{\mathbf{H}}_{t+1}^{-1}\left(h_{old}-\mathbf{h}_{old}^T\bar{\mathbf{H}}_{t+1}^{-1}\mathbf{h}_{p;d}^T\right)^{-1}\\ -\bar{\mathbf{H}}_{t+1}^{-1}\mathbf{h}_{old}^T\left(h_{old}-\mathbf{h}_{old}^T\bar{\mathbf{H}}_{t+1}^{-1}\mathbf{h}_{old}^T\right)^{-1} & \bar{\mathbf{H}}_{t+1}^{-1}+\bar{\mathbf{H}}_{t+1}^{-1}\mathbf{h}_{old}^T\mathbf{h}_{old}\bar{\mathbf{H}}_{t+1}^{-1}\left(h_{old}-\mathbf{h}_{old}\bar{\mathbf{H}}_{t+1}^{-1}\mathbf{h}_{old}^T\right)^{-1}\end{bmatrix} \tag{20}$$

$$\mathbf{H}_{t+1}^{-1}=\begin{bmatrix}h_{inverse\_old} & \mathbf{h}_{inverse\_old}\\ \mathbf{h}_{inverse\_old}^T & \mathbf{H}_{inverse\_new}\end{bmatrix}. \tag{21}$$

Compare Equation (20) with (21), the inverse matrix of cane sugar measurement and control model is as follows, when the old data is clear at time t+1:

$$\bar{\mathbf{H}}_{t+1}^{-1}=\mathbf{H}_{inverse\_new}-\frac{\mathbf{h}_{inverse\_old}^T\mathbf{h}_{inverse\_old}}{h_{inverse\_old}}, \tag{22}$$

Equations (23) and (24) determine the solution of the model, whose deductive process is as follows:

$$\mathbf{a}_{t+1}=\mathbf{H}_{t+1}^{-1}\mathbf{Y}_{t+1},$$

$$\bar{\mathbf{a}}_{t+1}=\bar{\mathbf{H}}_{t+1}^{-1}\bar{\mathbf{Y}}_{t+1}\Rightarrow$$

$$\bar{\mathbf{a}}_{t+1}=\left(\mathbf{H}_{inverse\_new}-\frac{\mathbf{h}_{inverse\_old}^T\mathbf{h}_{inverse\_old}}{h_{inverse\_old}}\right)\times\bar{\mathbf{Y}}_{t+1}\Rightarrow$$

$$\bar{\mathbf{a}}_{t+1}=\mathbf{H}_{inverse\_new}\bar{\mathbf{Y}}_{t+1}+\mathbf{h}_{inverse\_old}^T\mathbf{Y}_{old}-\mathbf{h}_{inverse\_old}^T\mathbf{Y}_{old}-\frac{\mathbf{h}_{inverse\_old}^T\mathbf{H}_{inverse\_old}\mathbf{h}_{inverse\_old}^T\bar{\mathbf{Y}}_{t+1}}{h_{inverse\_old}}\Rightarrow$$

$$\bar{\mathbf{a}}_{t+1}=\mathbf{H}_{inverse\_new}\bar{\mathbf{Y}}_{t+1}+\mathbf{h}_{inverse\_old}^T\mathbf{Y}_{old}\text{ - }\mathbf{h}_{inverse\_old}^T\frac{h_{inverse\_old}\mathbf{Y}_{old}+\mathbf{h}_{inverse\_old}\bar{\mathbf{Y}}_{t+1}}{h_{inverse\_old}},$$

$$\Rightarrow\begin{bmatrix}\mathbf{a}_{old}\\ \hat{\mathbf{a}}_{t+1}\end{bmatrix}=\begin{bmatrix}h_{inverse\_old} & \mathbf{h}_{inverse\_old}\\ \mathbf{h}_{inverse\_old}^t & \mathbf{H}_{inverse\_new}\end{bmatrix}\times\begin{bmatrix}\mathbf{Y}_{old}\\ \bar{\mathbf{Y}}_{t+1}\end{bmatrix},$$

$$\mathbf{a}_{old}=h_{inverse\_old}\mathbf{Y}_{old}+\mathbf{h}_{inverse\_old}\bar{\mathbf{Y}}_{t+1}, \tag{23}$$

$$\hat{\mathbf{a}}_{t+1}=\mathbf{h}_{inverse\_old}^T\mathbf{Y}_{old}+\mathbf{H}_{inverse\_old}\bar{\mathbf{Y}}_{t+1}. \tag{24}$$

When the old data is replaced by newly data at time t+1, the solution of model is defined as $\bar{\mathbf{a}}_{t+1}$ while the output vector is $\bar{\mathbf{Y}}_{t+1}$. Then the solution of model can be determined by Equation (25), through which the computing process is simplified and large-scale matrix operation is avoided.

$$\bar{\mathbf{a}}_{t+1} = \hat{\mathbf{a}}_{t+1} - \frac{\mathbf{h}^T_{inverse\_old}\mathbf{a}_{old}}{\mathbf{h}_{inverse\_old}} . \quad (25)$$

The online identification process is as follows:

1) Acquire a little sample data from the control or detection process. Optimize the kernel function parameter and penalty factor of the model, aiming to obtain the optimal parameters.

2) The parameters of model are offline trained in order to build the offline model.

3) The input data of model is received and identified online.

4) The actual control or measurement process determines if new data is inserted to the data set of model [10]. If new data is inserted in, add the data to the data set, otherwise the new data is abandoned and jump to step (3).

5) Set the size of data set, once its size is larger than the limited one, delete the original data and jump to step (3).

If a table is too long to fit onto one page, the table number and headings should be repeated on the next page before the table is continued.

Alternatively, the table can be spread over two consecutive pages (first on even-numbered, then on odd-numbered page).

For a wide table you can use 1-column section (Table 1), for a small standard table 2-column section is used (Table 2).

5 Simulation and experimental analysis

This paper takes the data provided by a local cane sugar production factory as experiment object. Data selected from multiple cane sugar crystallization phases is defined as experiment sample set to make sure that the experiment can cover the whole range of sugar crystallization process and data for machine learning is sufficient. The sample set has 212 samples, and parts of them are shown as Table 1. It is a common sense that the massecuite Brix and supersaturation are two vital factors in the cane sugar crystallization process. Massecuite temperature, steam pressure, steam temperature, vacuum degree and sorts of these are defined as input vectors, whereas massecuite Brix and supersaturation are defined as output vectors. The measurement and control model of the cane sugar crystallization process can synchronously identify the massecuite Brix and supersaturation online. It can also compare the results with the measured Brix.

TABLE 1 Experimental data

Sample number	Brix (°Bx)	Supersaturation (%)	Massecuite temperature (°C)	Vacuum degree (kPa)	Steam temperature (°C)	Steam pressure (MPa)
1	81.75	1.18	58.91	82.56	108.9	0.068
2	79.18	1.10	64.36	84.56	109.2	0.065
3	81.69	1.21	57.85	81.36	107.5	0.036
4	79.91	1.09	59.81	82.32	108.2	0.03
5	83.97	1.25	58.06	80.4	108.2	0.052
6	82.39	1.19	59.67	83.4	106.5	0.037
7	81.62	1.14	56.78	83.04	106.9	0.038
…	…	…	…	…	…	…
212	78.37	1.06	67.65	83.16	109.8	0.063

Gaussian radial basis function is selected as the kernel function of the model talked about above. The experiment program is compiled In VC9.0 environment. Parameters and penalty factors of the kernel function are optimized using particle swarm optimization (PSO) algorithm and leave-one-out cross validation [11]. Assume the optimization range of parameters was from 0.01 to 1000, and the penalty factors' range was from 1 to 10000. 162 samples are selected randomly from the sample set as training set and the rest of sample set is regarded as testing set (50 samples). Iterative algebra of PSO is 150 and population size of it is 100. Emulation results are shown in Table 2. Aimed to verify the experiment results, the difference between the measured data and output data from the model has been studied, which is shown in Figures 2 and 3 shows the accuracy rate of the results; experimental errors were demonstrated as Figure (4).

TABLE 2 Results

Sample number	Brix (°Bx)	Supersaturation (%)	Massecuite temperature (°C)	Vacuum degree (kPa)	Error rate of Brix (%)	Error rate of supersaturation (%)
1	78.17	1.06	78.69	1.09	0.67	2.83
2	78.12	1.05	78.09	1.07	0.04	1.90
3	85.1	1.26	84.78	1.28	0.37	1.59
4	78.07	1.06	79.40	1.09	1.70	2.83
5	79.07	1.07	80.03	1.09	1.22	1.87
6	83.26	1.21	81.51	1.2	2.10	0.83
7	82.81	1.2	81.90	1.21	1.10	0.83
…	…	…	…	…	…	…
50	78.91	1.1	79.06	1.12	0.19	1.82

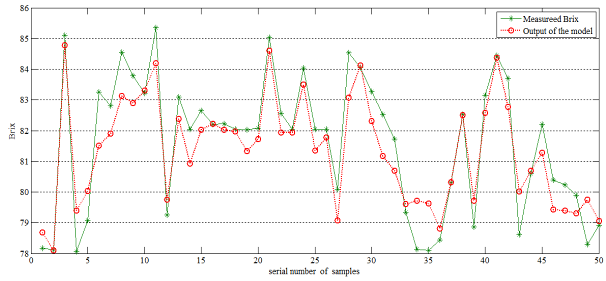


FIGURE 2 Comparative results of Brix

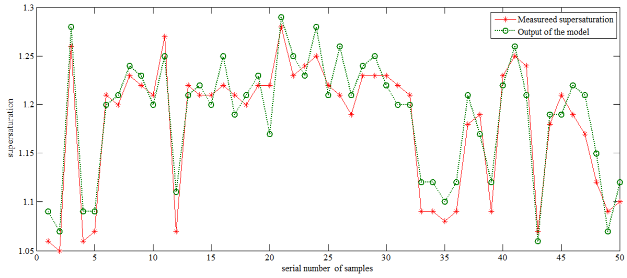


FIGURE 3 Comparative results of supersturation

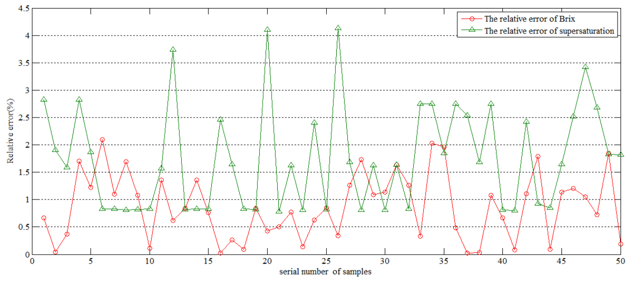


FIGURE 4 The relative error of Brix and supersaturation

6 Conclusion

Modelling with MIMO PLS-SVM can still perceive the predictive value of high accuracy when numbered samples are trained, by which challenges of nonlinearity, numbered samples and high dimension are easily dissolved. The inherent relationship between the multiple key variables and auxiliary variables is established with application of machine learning. The auxiliary variables are defined as the input of machine learning algorithm, while the key variables are defined as the output. Experimental results demonstrated that the accuracy rate of output was 95%. In an overall sense, the model discussed above is of great utility value in process identification, soft sensor and nonlinear control field.

Acknowledgement

This research is supported by National Natural Science Foundation of China (No. 51065004), Guangxi Natural Science Foundation (No. 2011GXNSFA018168) and Nanning Scientific and Technological Project (No. 20131079).

References

[1] Suykens J A K, Lukas L, Vandewalle J 2000 Sparse approximation using least square vector machine *Proceedings of the ISCAS 2000 Geneva The 2000 IEEE International Symposium on Circuits and Systems* **2** 757-60

[2] Suykens J A K, Lukas L, Vandewalle J 2002 Weighted least squares support vector machines: Robustness and sparse approximation *Neurocomput* **48**(1-4) 85-105

[3] Mountrakis G, Im J, Ogole C 2011 Support vector machines in remote sensing: A review *ISPRS Journal of Photogrammetry and Remote Sensing* **66**(3) 247-59

[4] Chang C C, Lin C 2011 LIBSVM: A Library for Support Vector Machine *ACM Transactions on Intelligent Systems and Technology* **2**(3) Article No. 27

[5] Kim H S, Sohn S Y 2010 Support vector machines for default prediction of SMEs based on technology credit *European Journal of Operational Research* **201** 838-46

[6] Kamandar M, Ghassemian H 2013 Linear Feature Extraction for Hyperspectral Images Based on Information Theoretic Learning *IEEE Geoscience and Remote Sensing Letters* **10**(4) 702-6

[7] Charrada A, Samet A 2012 SIMO-OFDM Channel Estimation based on Nonliear Complex LS-SVM *International Journal of Computer Applications* **43**(3) 1-7

[8] Sun Y, Todorovis S, Goodison S 2010 Local Learning Based Feature Selection for High Dimensional Data Analysis *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(9) 1610-26

[9] Jiao L C, Bo L F, Wang L 2007 Fast Sparse Approximation for Least Squares Support Vector Machine *IEEE Transactions on Neural Networks* **18**(3) 685-697

[10] Olivieri M, Clinton J 2012 An almost fair comparisonbetween Earthworm and SeisComp3 *Seismol Res Lett* **83**(4) 720-7

[11] Bo L, Wang L, Jiao L C 2006 Feature scaling for kernel fisher discriminant analysis using leave-one-out cross validation *Neural Comput* **18**(4) 961-78

Authors

Guancheng Lu, born in 1984, China

Current position, grades: teacher and engineer at Guangxi University.
University studies: Master degree in mechanical and electronic engineering, Guangxi University, China, 2010.
Scientific interest: measurement and control system modelling, automatic control implementation for mechanical engineering.
Publications: 3 Patents, 6 papers.

Yanmei Meng, born in 1963, China

Current position, grades: PhD, professor, vice dean of the mechanism and engineering college in Guangxi University. China.
University studies: PhD degree in Mechanical Manufacturing and Automation in Hefei University of Technology.
Scientific interest: Complex mechatronic product design and optimization, intelligent detection and control technology.
Publications: 11 Patents, 60 Papers.

Jian Chen, born in 1989, China

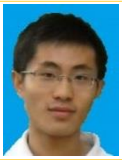
Current position, grades: postgraduate student in Guangxi University.
University studies: Bachelor degree in Mechatronic Engineering in 2012.
Scientific interest: model predictive control, PLC implementation.
Publications: 1 paper.

Zhihong Tang, born in 1964, China

Current position, grades: senior engineer in college of Mechanism engineering of Guangxi University.
University studies: Bachelor degree in Mechatronic Engineering.

Xiaochun Wang, born in 1954, China.

Current position, grades: associate professor at the college of Mechanism engineering of Guangxi University.
University studies: Bachelor degree in Mechanical manufacturing in 1982.
Scientific interest: advanced manufacturing technology, industrial engineering.
Publications: 20 Papers.
Experience: 20 years in Guangxi University.



Xian Yu, born in 1989, China

Current position, grades: postgraduate student in Guangxi University.
University studies: bachelor degree in Mechatronic Engineering in 20.12.
Scientific interest: artificial intelligent control, model optimization and data mining, first-principle and hybrid modeling of sugar crystallization.
Publications: 1 Patent, 2 Papers.