Journal of Zhejiang University SCIENCE
ISSN 1009-3095
http://www.zju.edu.cn/jzus
E-mail: jzus@zju.edu.cn

New predictive control algorithms based on Least Squares Support Vector Machines*

LIU Bin (刘 斌)†[1,2], SU Hong-ye (苏宏业)[1], CHU Jian (褚 健)[1]

([1]National Laboratory of Industrial Control Technology, Institute of Advanced Process Control, Zhejiang University, Hangzhou 310027, China)

([2]School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China)

†E-mail: liubin0130@163.com

Received Feb. 15, 2004; revision accepted June 15, 2004

Abstract: Used for industrial process with different degree of nonlinearity, the two predictive control algorithms presented in this paper are based on Least Squares Support Vector Machines (LS-SVM) model. For the weakly nonlinear system, the system model is built by using LS-SVM with linear kernel function, and then the obtained linear LS-SVM model is transformed into linear input-output relation of the controlled system. However, for the strongly nonlinear system, the off-line model of the controlled system is built by using LS-SVM with Radial Basis Function (RBF) kernel. The obtained nonlinear LS-SVM model is linearized at each sampling instant of system running, after which the on-line linear input-output model of the system is built. Based on the obtained linear input-output model, the Generalized Predictive Control (GPC) algorithm is employed to implement predictive control for the controlled plant in both algorithms. The simulation results after the presented algorithms were implemented in two different industrial processes model; respectively revealed the effectiveness and merit of both algorithms.

Key words: Least Squares Support Vector Machines, Linear kernel function, RBF kernel function, Generalized predictive control
doi:10.1631/jzus.2005.A0440 **Document code:** A **CLC number:** TP273

INTRODUCTION

Model predictive control (MPC), based on predictive model and receding horizon optimization, has become an attractive feedback control strategy, because it has found successful applications, especially in the process industry. For this kind of control strategy, the predictive model is a crucial component because the essence of MPC is to optimize the forecast of process behavior (Rawlings, 2000), and the forecast is accomplished with the predictive model. If the controlled plant is linear or weakly nonlinear, it can be fitted by linear predictive model effectively. But, if the plant has strongly nonlinear characteristics and operates over large region in variable space, the nonlinear predictive model must be used to approximate the system dynamics.

The learning algorithms for traditional modeling approaches, including classical neural networks, fuzzy modeling, etc. (Babuška and Verbruggen, 2003), are almost all based on the expectation risk minimization principle. These kinds of algorithms often lead to the problem of overfitting (Zhang, 2000). Simply speaking, for a given learning task with a given finite amount of training data, lesser training error may result in poorer generalization performance. Vapnik (1998) presented based on the statistical learning and structural risk minimization principle the Support Vector Machines (SVM), which can give attention to both the expectation risk and the generalization performance and can be used to approximate

* Project supported by the National Outstanding Youth Science Foundation of China (No. 60025308) and the Teach and Research Award Program for Outstanding Young Teachers in Higher Education Institutions of MOE, China

nonlinear functions. Suykens and Vandewalle (1999) presented LS-SVM method, in which the objective function includes an additional sum squared error term. Using the equality constraint, LS-SVM does not solve quadratic programming existing in the standard SVM. In this paper, the LS-SVM methods with linear kernel and Radial Basis Function (RBF) kernel are employed to build the model of controlled systems with different degree of nonlinearity.

LS-SVM

SVM is one of the methods by which the statistical learning theory can be introduced to practical application. It has its own advantages in solving the pattern recognition problem with small samples, nonlinearity, and higher dimension. SVM can be easily introduced into learning problem such as function estimation.

Suykens and Vandewalle (1999) presented the LS-SVM approach, in which the following function is used to approximate the unknown function,

$$y(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x}) + b \tag{1}$$

where, $\boldsymbol{x}\in\mathrm{R}^n$, $y\in\mathrm{R}$, $\varphi(\cdot):\mathrm{R}^n \to \mathrm{R}^{n_h}$ is a nonlinear function which maps the input space into a higher dimension feature space.

Given training data, LS-SVM defines an optimization problem as follows,

$$\min_{w,b,e} J(\boldsymbol{w},\boldsymbol{e}) = \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + \gamma\frac{1}{2}\sum_{k=1}^{N} e_k^2,\ (\gamma > 0) \tag{2}$$

subject to the equality constraints

$$y_k = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x}_k) + b + e_k, \quad k = 1,\dots,N.$$

To solve this optimization problem, one defines the following Lagrange function,

$$L(\boldsymbol{w},b,\boldsymbol{e};\boldsymbol{\alpha}) = J(\boldsymbol{w},\boldsymbol{e}) - \sum_{k=1}^{N}\alpha_k\{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x}_k) + b + e_k - y_k\} \tag{3}$$

where, $\boldsymbol{\alpha} = \{\alpha_k\}_{k=1}^N$ is the Lagrange multiplier set. Calculating the partial derivatives of $L(\boldsymbol{w},b,\boldsymbol{e};\boldsymbol{\alpha})$ with respect to $\boldsymbol{w}$, b, $\boldsymbol{e}$, $\boldsymbol{\alpha}$, one gets the optimal condition for Eq.(2) as

$$\begin{cases} \dfrac{\partial L}{\partial \boldsymbol{w}} = 0 \to \boldsymbol{w} = \sum_{k=1}^{N}\alpha_k\boldsymbol{\varphi}(\boldsymbol{x}_k) \\ \dfrac{\partial L}{\partial b} = 0 \to \sum_{k=1}^{N}\alpha_k = 0 \\ \dfrac{\partial L}{\partial e_k} = 0 \to \alpha_k = re_k \qquad k = 1,\dots,N \\ \dfrac{\partial L}{\partial \alpha_k} = 0 \to \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x}_k) + b + e_k - y_k = 0 \quad k = 1,\dots,N \end{cases} \tag{4}$$

Expressing e_k and $\boldsymbol{w}$ with α_k and b, one can transform the above equality into

$$\begin{bmatrix} 0 & \mathbf{1}^{\mathrm{T}} \\ \mathbf{1} & \boldsymbol{\Omega} + \gamma^{-1}\boldsymbol{I} \end{bmatrix}\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}, \tag{5}$$

where $\boldsymbol{y}=[y_1, \dots, y_N]^{\mathrm{T}}$, $\mathbf{1}=[1, \dots, 1]^{\mathrm{T}}$, $\boldsymbol{\alpha}=[\alpha_1, \dots, \alpha_N]^{\mathrm{T}}$, and $\boldsymbol{\Omega}$ is a square matrix in which the element located on kth column and lth row is

$$\boldsymbol{\Omega}_{kl} = \boldsymbol{\varphi}(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x}_l) = K(\boldsymbol{x}_k, \boldsymbol{x}_l) \quad k,l = 1,\dots,N.$$

Choosing $\gamma>0$, ensures that the matrix

$$\boldsymbol{\Phi} = \begin{bmatrix} 0 & \mathbf{1}^{\mathrm{T}} \\ \mathbf{1} & \boldsymbol{\Omega} + \gamma^{-1}\boldsymbol{I} \end{bmatrix}$$

is invertible. Then we have the analytical solution of $\boldsymbol{\alpha}$ and b

$$\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \boldsymbol{\Phi}^{-1}\begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}. \tag{6}$$

Substituting the obtained b and $\boldsymbol{\alpha}$ into Eq.(4), we get

$$y(\boldsymbol{x}) = \sum_{k=1}^{N}\alpha_k K(\boldsymbol{x}, \boldsymbol{x}_k) + b \tag{7}$$

where $K(\boldsymbol{x}, \boldsymbol{x}_k)$ is the Kernel function, which can be any symmetric function satisfying Mercer's condition

(Smola, 1996).

LS-SVM MODELING AND GPC CONTROL

Generally, the following input-output model

$$y=f(\boldsymbol{x}) \tag{8}$$

can be employed to denote the controlled system characteristics. Where, $\boldsymbol{x}=[x(1), x(2), \ldots, x(nu+ny)]$ denotes the regression vector including the past input-output data of the system. $f(\cdot)$, a linear or nonlinear function, is used to fit the system characteristics. nu and ny denote input and output order of the system respectively. Input-output data of the system are collected and constitute the training dataset $\{\boldsymbol{x}_k, y_k\}_{k=1}^{N}$. Here, $\boldsymbol{x}_k$ is the regression vector at different sampling instant and y_k is the system output corresponding to $\boldsymbol{x}_k$.

Modeling of weakly nonlinear system

For the system with weak nonlinearity, we use the LS-SVM with linear kernel and get the model of the controlled system as follows,

$$y(\boldsymbol{x})=\sum_{k=1}^{n}\alpha_k(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{x})+b \tag{9}$$

Apparently, Eq.(9) is a linear LS-SVM model which can be transformed into a linear input-output relation of the controlled system. Given the following regression vector at current instant

$$\begin{aligned}\boldsymbol{x}&=[x(1),\ldots,x(nu+ny)]\\&=[u(t-1),\ldots,u(t-nu),y(t-1),\ldots,y(t-ny)],\end{aligned} \tag{10}$$

we can get the system output at next sampling time:

$$\begin{aligned}y(t)&=\sum_{k=1}^{N}\alpha_k\{\boldsymbol{x}_k^{\mathrm{T}}[u(t-1),\ldots,u(t-nu),\\&\qquad y(t-1),\ldots,y(t-ny)]^{\mathrm{T}}\}+b\\&=\sum_{k=1}^{N}\alpha_k\{x_{k,1}u(t-1)+\ldots+x_{k,nu}u(t-nu),\\&\qquad x_{k,nu+1}y(t-1),\ldots,x_{k,nu+ny}y(t-ny)]^{\mathrm{T}}\}+b\\&=\sum_{k=1}^{N}[\alpha_k(x_{k,1}+\ldots+x_{k,nu}z^{-nu+1})]u(t-1)\\&\quad+\sum_{k=1}^{N}\left[\alpha_k(x_{k,nu+1}z^{-1}+\ldots+x_{k,nu+ny}z^{-ny})\right]y(t)+b.\end{aligned} \tag{11}$$

From Eq.(11), we have the linear input-output equation

$$A(z^{-1})y(t)=B(z^{-1})u(t-1)+b\,. \tag{12}$$

where

$$A(z^{-1})=1-\sum_{k=1}^{N}\left[\alpha_k(x_{k,nu+1}z^{-1}+\ldots+x_{k,nu+ny}z^{-ny})\right],$$

$$B(z^{-1})=\sum_{k=1}^{N}\left[\alpha_k(x_{k,1}+x_{k,2}z^{-1}+\ldots+x_{k,nu}z^{-nu+1})\right].$$

Modeling of strongly nonlinear system

For the system with strong nonlinearity, we use the LS-SVM with RBF kernel and get the off-line model of the controlled system as follows,

$$y(\boldsymbol{x})=\sum_{k=1}^{N}\alpha_k\exp\{-\|\boldsymbol{x}-\boldsymbol{x}_k\|_2^2/\sigma^2\}+b\,. \tag{13}$$

Eq.(13) is a nonlinear LS-SVM model of the controlled system. To avoid solving the nonlinear programming problem resulting from the nonlinear predictive model, Eq.(13) is linearized at each sampling period of system running. Let data regression vector $\boldsymbol{x}$ at current instant be the same as Eq.(10), and let $\boldsymbol{x}_0=\boldsymbol{x}$ for the sake of simple expression. Linearizing Eq.(13) at $\boldsymbol{x}_0$, we have

$$\begin{aligned}y(\boldsymbol{x})&=y(\boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{x}_0}+\frac{\partial y}{\partial x(1)}\bigg|_{\boldsymbol{x}=\boldsymbol{x}_0}[x(1)-x_0(1)]+\ldots\\&\quad+\frac{\partial y}{\partial x(nu+ny)}\bigg|_{\boldsymbol{x}=\boldsymbol{x}_0}[x(nu+ny)-x_0(nu+ny)]\\&=y(\boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{x}_0}-\frac{\partial y}{\partial x(1)}\bigg|_{\boldsymbol{x}=\boldsymbol{x}_0}x_0(1)-\ldots\\&\quad-\frac{\partial y}{\partial x(nu+ny)}\bigg|_{\boldsymbol{x}=\boldsymbol{x}_0}x_0(nu+ny)\\&\quad+\frac{\partial y}{\partial x(1)}\bigg|_{\boldsymbol{x}=\boldsymbol{x}_0}x(1)+\ldots+\frac{\partial y}{\partial x(nu+ny)}\bigg|_{\boldsymbol{x}=\boldsymbol{x}_0}x(nu+ny)\end{aligned}$$

$$= p + b_1x(1) + ... + b_{nu}x(nu)$$
$$-a_1x(nu+1) - ... - a_{ny}x(nu+ny)$$
$$= p + b_1u(k-1) + ... + b_{nu}u(k-nu)$$
$$-a_1y(k-1) - ... - a_{ny}y(k-ny). \quad (14)$$

where, p is constant at current sampling period and depends only on the current sampling point. Namely,

$$A(z^{-1})y(k) = B(z^{-1})u(k-1) + p\,. \quad (15)$$

where

$A(z^{-1})=1+a_1z^{-1}+\ldots+a_{ny}z^{-ny}$

$B(z^{-1})=b_1+b_2z^{-1}+\ldots+b_{nu}z^{-nu+1}$.

GPC of the controlled system

We have obtained the approximate linear input-output model of weakly nonlinear system in Eq.(12) and the linear online model of strongly nonlinear system in Eq.(15). Both the equations can be unified into the following equation

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + v(t) \quad (16)$$

where, $v(t)$ is the error and disturbance resulting from one fitting current system characteristics with Eq.(16). We decompose $v(t)$ as follows

$$v(t)=v_{dc}+v_{ac}(t),$$

where, v_{dc} is the direct-current component independent of time. In the weakly nonlinear case, v_{dc} includes the bias term b in Eq.(12), while in the strongly nonlinear case, v_{dc} includes p in Eq.(14). The amplitude of v_{dc} is equal to the mean of $v(t)$. v_{ac} is the AC component whose mean is zero. Modeling v_{ac} with $w(t)/\Delta$, we can transform the input-output equation at current time into

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + v_{dc} + w(t)/\Delta \quad (17)$$

where, $w(t)$ is the disturbance with zero mean, $\Delta=1-z^{-1}$ is the difference operator. Δ is introduced into the system model in order to provide integral action and therefore eliminate steady-state offsets.

In order to forecast the future system output based on the past input-output data and the future system input, we introduce the following Diophantine equations,

$$1 = E_j(z^{-1})A(z^{-1})\Delta + z^{-j}F_j(z^{-1}) \quad (18)$$

$$E_j(z^{-1})B(z^{-1}) = G_j(z^{-1}) + z^{-j}H_j(z^{-1}) \quad (19)$$

Multiplying both sides of Eq.(13) by $E_j(z^{-1})\Delta z^j$, we have

$$E_j(z^{-1})A(z^{-1})\Delta y(t+j) = E_j(z^{-1})B(z^{-1})\Delta u(t+j-1)$$
$$+E_j(z^{-1})(\Delta v_{dc} + w(t+j)) \quad (20)$$

v_{dc} is independent of time, so $\Delta v_{dc}=0$. Namely, the system performance should not be influenced by the direct-current component of model error. Then the above equation is simplified as

$$E_j(z^{-1})A(z^{-1})\Delta y(t+j)$$
$$= E_j(z^{-1})B(z^{-1})\Delta u(t+j-1) + E_j(z^{-1})w(t+j) \quad (21)$$

Accordingly, the well-known method (Clarke, 1989) is used to obtain the multistep prediction of $y(t)$, which can be denoted as the following expression in the form of vector

$$\boldsymbol{y}^{\mathrm{o}} = \boldsymbol{G}\boldsymbol{u} + \boldsymbol{F}\boldsymbol{y}(t) + \boldsymbol{H}\Delta\boldsymbol{u}(t-1)\,.$$

Finally, the control law is constructed as follows

$$\boldsymbol{u} = (\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G} + \lambda\boldsymbol{I})^{-1}\boldsymbol{G}^{\mathrm{T}}[\boldsymbol{y}_r - \boldsymbol{F}\boldsymbol{y}(t) - \boldsymbol{H}\Delta\boldsymbol{u}(t-1)]\,.$$

All the variables in the above expression have the same definition as in Clarke's paper (Clarke *et al.*, 1989).

SIMULATION

In this section, we introduce two nonlinear industrial process models–the pulp washing process and the pH neutralization process. The former possesses weak nonlinearity and the latter strong nonlinearity.

Example 1 In the pulp washing process (Yang *et al.*, 1997), Dilution Factor (*DF*) is a crucial index for pulp quality. The physical model of this process is detailed as follows:

$$y(k) = a_1(k)y(k-1) + a_2(k)y(k-2) + b_0(k)u(k-1) + b_1(k)u(k-2)$$

with

$a_1(k)$=0.15+0.002sin0.02k,
$a_2(k)$=0.2+0.003cos0.02k,
$b_0(k)$=1.2, $b_1(k)$=0.5.

In this model, the input variable is the wash flow u, and the output is DF. Apparently, this process can be regarded as a typical system with weak nonlinearity.

Using the LS-SVM with linear kernel function, we can build the approximate linear model of the pulp washing process. A group of white noise signals are used as the input of the system to produce training dataset $\{\boldsymbol{x}_k, y_k\}_{k=1}^{100}$. We choose the order of the output and input variable as nu=1 and ny=1 respectively. In addition, for giving attention to both the precision and generalization capacity of the model, we choose the parameter γ in Eq.(2) as 10. In order to test the approximation performance of the linear LS-SVM model, another group of white noise signals are employed as input to produce test dataset, in which the amplitude of system output is in the interval (0, 70). We denote the actual system output as 'Y', and the model output as 'Ym'. The error between actual system and model output is shown in Fig.1.

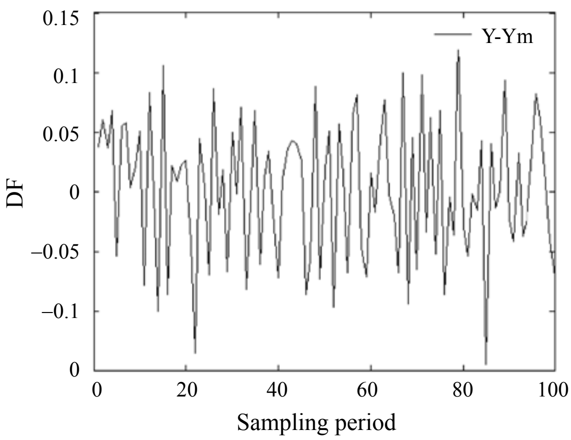

Fig.1 The error between actual system and model output

The obtained linear LS-SVM model is transformed into approximate linear input-output relation of the controlled system, and then the GPC is used to implement the predictive control. Let predictive horizon be P=5, control horizon be M=3, then the obtained tracking curve is shown in Fig.2. In Fig.2, 'Yr' is the reference trajectory, and 'Y' is the system output.

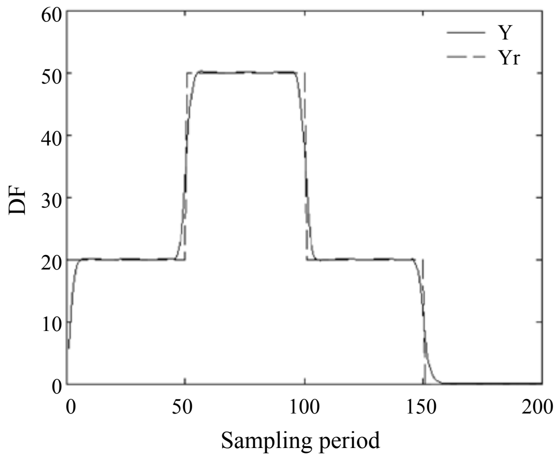

Fig.2 Tracking curve of the pulp washing process

The above simulation results illustrate the effectiveness of the predictive control algorithm based on LS-SVM with linear kernel function on weakly nonlinear system.

Example 2 We introduce a pH neutralization process (Nie *et al*., 1996) as a typical system with strong nonlinearity. In this process, an acetic acid (weak acid) is neutralized by a strong base NaOH in water. The physical model of this process consists of two parts: the linear dynamic model and the nonlinear static model. The dynamic model is given by:

$$V\frac{\mathrm{d}w_\mathrm{a}}{\mathrm{d}t} = F_1C_\mathrm{a} - (F_1 + F_2)w_\mathrm{a}$$

$$V\frac{\mathrm{d}w_\mathrm{b}}{\mathrm{d}t} = F_2C_\mathrm{b} - (F_1 + F_2)w_\mathrm{b}.$$

In this model, F_1 and F_2 are flow rate of acid and the strong base, of which w_a and w_b are the concentrations respectively. The static model is given by

$$w_b + 10^{-pH} - 10^{pH-14} - \frac{w_a}{1+10^{pK_a-pH}} = 0.$$

Consider the SISO system: the input of the pH neutralization process is NaOH flow rate F_2 in the inlet, and the output is pH in the outlet of the container. The acid flow rate F_1 in the inlet is a constant. The other parameter values used in the model are coefficients related to the specified system (Nie *et al*., 1996). The nonlinearity exhibited in the weak acid and strong base system is severe, especially near pH=9.

Using the LS-SVM with RBF kernel, we can build the off-line model of the pH neutralization process. The same as in Example 1, a group of white noise signals is used as the input signal of the system to produce training dataset $\{\boldsymbol{x_k}, y_k\}_{k=1}^{497}$. We choose the order of the output and input variable as nu=3 and ny=3 respectively. The parameter γ is chosen as γ=5.

In order to test the approximation performance of the off-line model, a signal composed of four kinds of sine wave with different frequency is employed as test input to produce test dataset. For the different values of σ (generally, $1<\sigma\leq 10$), we can build different models using training data. By means of model validation based on the test data, the parameter σ is chosen as 1.6. The comparison between the actual system and model output is given in Fig.3. In Fig.3, 'Y' is the actual system output and 'Ym' is the model output.

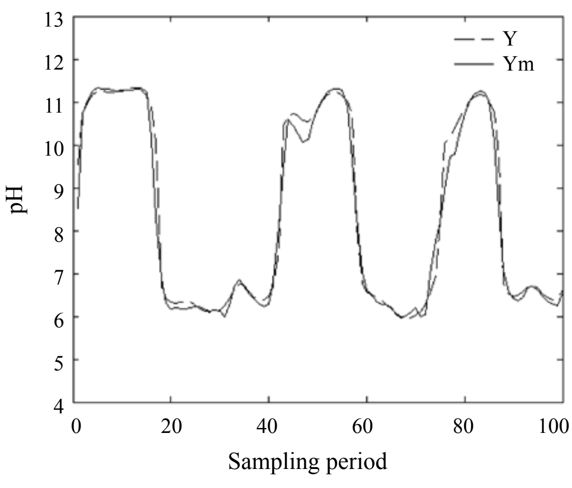

Fig.3 The comparison between the actual system and model output

The off-line model is linearized on-line and the GPC strategy is used to compute the control action at each sampling point. Let predictive horizon be P=10, control horizon be M=5, and the sampling time be T_s=0.5 min. We get the tracking curve of the closed-loop system in Fig.4, in which 'Yr' is the reference trajectory and 'Y' is the output of the practical system. From Fig.4, we can see that the presented predictive algorithm based on LS-SVM with RBF kernel can control the given strongly nonlinear system rapidly and stably.

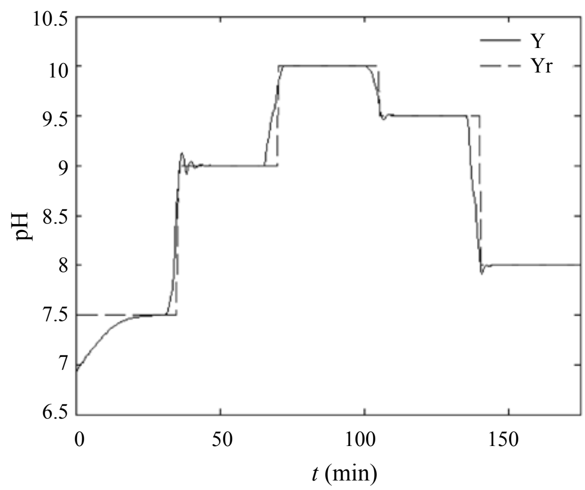

Fig.4 Tracking curve of the pH neutralization process

From Fig.3, we can see that the off-line model does not fit the actual system characteristic very accurately, which can be explained by the fewer training data. Even so, from Fig.4, we know that the GPC strategy can effectively control the plant on the basis of given predictive model, which is attributable to the robustness and the capability of eliminating steady-state error for GPC. In fact, the better approximation effect can be obtained by using more training data, but the fewer training data and the corresponding model can be used here to prove the robustness of the predictive algorithm.

CONCLUSION

Two predictive control algorithms based on LS-SVM are put forward in this paper. LS-SVM is a modeling approach based on structural risk minimization principle. It can give attention to both the expectation risk and the generalization performance. Its modeling process has analytical solution formula and less indeterminate parameters. For the system with weak nonlinearity, the LS-SVM with linear kernel function is used to build the approximate linear LS-SVM model for the controlled system, and then the linear LS-SVM model is transformed into a linear input-output model. For the system with strong nonlinearity, the LS-SVM with RBF kernel function is used to build the off-line nonlinear model for the controlled system. To avoid the need for the nonlinear programming problem to be resolved at each sampling period, we linearize the off-line model online at each sampling point. For both classes of nonlinear system with different degree of nonlinearity, we find a uniform linear input-output expression, and then employ GPC to implement the predictive control strategy. The results of the experiment showed that the presented algorithms are effective.

References

Babuška, R., Verbruggen, H.B., 2003. Neuro-fuzzy methods for nonlinear system identification. *Annual Reviews in Control*, **27**(1):73-85.

Clarke, D.W., Mohtadi, C., Tuffs, P.S., 1987. Generalized predictive control–Part I. The basic algorithm. *Automatic*, **23**(2):137-148.

Nie, J.H., Loh, A.P., Hang, C.C., 1996. Modeling pH neutralization processes using fuzzy-neutral approaches. *Fuzzy Sets and Systems*, **78**:5-22.

Rawlings, J.B., 2000. Tutorial overview of model predictive control. *Control Systems Magazines, IEEE*, **20**(3):38-52.

Smola, A.J., 1996. Regression Estimation with Support Vector Learning Machines. Master's Thesis, Technische Universität München.

Suykens, J.A.K., Vandewalle, J., 1999. Least Squares Support Vector Machine classifiers. *Neural Processing Letters*, **9**(3):293-300.

Vapnik, V., 1998. Statistical Learning Theory. John Wiley, New York, U.S.A.

Yang, C.J., Sun, Y.X., Bao, B.L., 1997. A simplified GPC algorithm for pulp washing process. *Mechanical & Electrical Engineering Magazine*, **24**(6):9-12.

Zhang, X.G., 2000. Introduction to statistical learning theory and Support Vector Machines. *ACTA AUTOMATICA SINICA*, **26**(1):32-42.