Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Wiener model identification and nonlinear model predictive control of a pH neutralization process based on Laguerre filters and least squares support vector machines*

Qing-chao WANG[1], Jian-zhong ZHANG†‡[1,2]

([1]*School of Energy Science and Engineering, Harbin Institute of Technology, Harbin 150001, China*)

([2]*SINOPEC Safety Engineering Institute, Qingdao 266071, China*)

†E-mail: zjzhit1984@163.com

Received Dec. 21, 2009; Revision accepted Mar. 24, 2010; Crosschecked Dec. 6, 2010

Abstract: This paper deals with Wiener model based predictive control of a pH neutralization process. The dynamic linear block of the Wiener model is parameterized using Laguerre filters while the nonlinear block is constructed using least squares support vector machines (LSSVM). Input-output data from the first principle model of the pH neutralization process are used for the Wiener model identification. Simulation results show that the proposed Wiener model has higher prediction accuracy than Laguerre-support vector regression (SVR) Wiener models, Laguerre-polynomial Wiener models, and linear Laguerre models. The identified Wiener model is used here for nonlinear model predictive control (NMPC) of the pH neutralization process. The set-point tracking performance of the proposed NMPC is compared with those of the Laguerre-SVR Wiener model based NMPC, Laguerre-polynomial Wiener model based NMPC, and linear model predictive control (LMPC). Validation results show that the proposed NMPC outperforms the other three controllers.

Key words: Wiener model, Nonlinear model predictive control (NMPC), pH neutralization process, Laguerre filters, Least squares support vector machines (LSSVM)

doi:10.1631/jzus.C0910779 **Document code:** A **CLC number:** TP273

1 Introduction

The control of pH neutralization processes plays a significant role in a variety of industries, including wastewater treatment, pharmaceuticals, biotechnology, and chemical processing. It is often difficult to achieve a high performance and robust pH control due to their time-varying and severe nonlinear characteristics (Henson and Seborg, 1994). Hence, pH control is often considered a benchmark for new models and control strategies.

A number of control strategies and algorithms have been suggested in the literature for the control of pH processes. Many of these techniques involve adaptive control and model predictive control (MPC). Henson and Seborg (1994) proposed an adaptive nonlinear output feedback control scheme for a pH neutralization process, developed by combining an input-output linearizing controller with an open-loop nonlinear observer. Loh *et al.* (1995) presented a direct model reference adaptive neural network control scheme for pH control. Lakshmi Narayanan *et al.* (1997) developed an adaptive nonlinear internal model controller based on a strong acid equivalent technique and adaptive mechanism. Böling *et al.* (2007) proposed a multi-model adaptive proportional-integral-derivative (PID) controller based on a set of simple linear dynamic models for a pH neutralization process.

MPC has become a standard control technique in the process industries over the past three decades due

‡ Corresponding author

* Project (No. 60574022) supported by the National Natural Science Foundation of China

to its intrinsic capability of dealing with constraints and with multivariable systems. As an alternative control method for pH processes, nonlinear model predictive control (NMPC) has been studied by several researchers. Numerous NMPC algorithms reported are based on the empirical models of pH neutralization processes, including NARMAX models (Pröll and Karim, 1994), ARIMAX models (Altinten, 2007), neural networks (Akesson *et al.*, 2005), support vector machines (Bao *et al.*, 2007), Hammerstein models (Fruzzetti *et al.*, 1997), and Wiener models (Norquay *et al.*, 1998). In particular, Wiener models, which consist of a linear time invariant (LTI) subsystem cascaded with a static nonlinear subsystem, have a special capability of facilitating the application of NMPC. There are many successful applications of Wiener model based NMPC algorithms for pH neutralization processes in the literature. Gómez and Baeyens (2004) presented a subspace-based black-box Wiener type model for a pH neutralization process, the identified model being used as the internal model in an NMPC controller design. Saha *et al.* (2004) used quadratic polynomial (QP) functions as well as artificial neural networks (ANN) to construct a nonlinear output map and Laguerre filters to represent LTI components of the Wiener model. The identified models were further used to formulate MPC schemes. Mahmoodi *et al.* (2009) also employed Laguerre filters and polynomial functions as the linear and nonlinear blocks of the Wiener model, respectively. The so-called Wiener-Laguerre model was used to evaluate identification and NMPC of a pH neutralization process.

Wiener models have an advantage in that the complexity of system dynamics is contained in the linear element, whereas the complexity of nonlinearity is only in the static element. This advantage reduces the difficulty in controller design because a divide-and-conquer strategy, using linear and nonlinear system theories, can be applied to deal with the two subsystems separately (Hsu and Wang, 2009). Many techniques have been proposed for the identification of Wiener models. These techniques differ mainly in the way the dynamic linear and static nonlinear blocks are represented. Reported structures of the dynamic linear block include the auto regressive with exogenous inputs (ARX) model (Norquay *et al.*, 1998), step-response model (Norquay *et al.*, 1998; 1999), state-space model (Lovera *et al.*, 2000; Cervantes *et al.*, 2003; Shafiee *et al.*, 2008), finite impulse response (FIR) model (Kalafatis *et al.*, 1995), and Laguerre filters (Saha *et al.*, 2004; Mahmoodi *et al.*, 2009). The nonlinear block is represented by polynomial functions (Saha *et al.*, 2004; Mahmoodi *et al.*, 2009), piecewise linear functions (Cervantes *et al.*, 2003; Shafiee *et al.*, 2008), neural networks (Fang and Chow, 2000; Aadaleesan *et al.*, 2008; Arefi *et al.*, 2008), support vector machines (Tötterman and Toivonen, 2009), etc.

Orthonormal filter networks such as Laguerre filters (Zervos and Dumont, 1988; Wahlberg, 1991) and Kautz filters (Wahlberg, 1994) have received considerable attention in linear model approximation and identification. Laguerre models are insensitive to the choice of the sampling rate, and they can approximate linear systems with a lower model order compared with the traditional ARX models (Wahlberg, 1991). Another advantage of Laguerre approximation is that no explicit knowledge is required about the system time constant and time delay for model development (Saha *et al.*, 2004). Wiener (1958) originally introduced orthonormal Laguerre functions as the linear dynamic part of the Laguerre-Wiener model, while the nonlinearity was captured using Hermite polynomials. The use of Laguerre filters in combination with polynomial functions, neural networks, and support vector regression in constructing Wiener-type models can be found in Aadaleesan *et al.* (2008), Mahmoodi *et al.* (2009), and Tötterman and Toivonen (2009).

Although the polynomial function is often selected to represent the nonlinear component of Wiener models because of its simplicity, it does suffer from definite limitations, particularly in the case of processes with highly nonlinear gains (Norquay *et al.*, 1999). There are many more powerful nonlinear modeling approaches such as piecewise linear functions and neural networks which have been introduced in Wiener model construction. In recent years, the support vector machine (SVM) has become a popular approach in the field of classification and regressions (Vapnik, 1998). It has several appealing properties for black-box model identification. Tötterman and Toivonen (2009) applied SVM to identify

Wiener models, where the linear subsystem is represented by Laguerre filters.

In this paper, Laguerre filters are combined with least squares support vector machines (LSSVM) to construct a Wiener model, which is an improved version of the support vector regression (SVR) based Wiener model (Tötterman and Toivonen, 2009). It is shown that the parameters of LSSVM can be obtained by solving a set of linear equations instead of the quadratic programming problem in the SVM case, and this can reduce the computational load. The developed Laguerre-LSSVM Wiener model is used to model a pH neutralization process. Based on the identified Wiener model, a nonlinear model predictive control is implemented, and the set-point tracking performance is compared with those of linear model predictive control (LMPC), Laguerre-polynomial Wiener model based NMPC, and Laguerre-SVR Wiener model based NMPC.

2 Laguerre-LSSVM Wiener model identification

In this section, a Laguerre-LSSVM Wiener model is introduced. The state-space representation of the Laguerre filter network is used to describe the dynamic linear subsystem, and the LSSVM model is used to describe the static nonlinearity. Laguerre filter models can represent any linear dynamical system efficiently, and the LSSVM model can approximate arbitrary functions (even in the mildly nonlinear system), with adequate accuracy, by a proper choice of its parameters. The proposed Wiener model (Fig. 1) can be described as

$$\begin{cases} \boldsymbol{L}(k+1)=\boldsymbol{A}\boldsymbol{L}(k)+\boldsymbol{B}u(k), \\ y(k)=f(\boldsymbol{L}(k)), \end{cases} \tag{1}$$

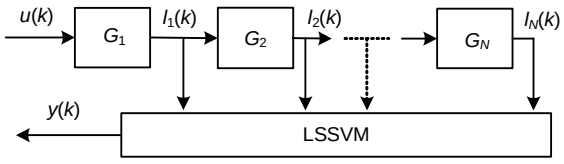

Fig. 1 Laguerre-LSSVM Wiener model

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are coefficient matrices, $u(k)$ is the scalar input signal, $y(k)$ is the scalar output signal, the Laguerre filter state vector $\boldsymbol{L}(k)=[l_1(k), l_2(k), \ldots, l_N(k)]^{\mathrm{T}}$ acts as the inputs of LSSVM, and $f(\cdot)$ denotes the LSSVM nonlinear mapping.

The linear-nonlinear (L-N) approach for Wiener model identification is used in this study. Note that Laguerre filters and the LSSVM model are estimated in separate steps. First, the linear block is identified by selecting proper Laguerre filter parameters, and then the static nonlinearity is estimated according to the input-output dataset $\left\{\boldsymbol{L}_{\mathrm{t}}(k), y_{\mathrm{t}}(k)\right\}_{k=1}^{M}$. Details are introduced in the following.

2.1 Dynamic linear subsystem

Considering a single-input single-output (SISO) linear system modeled using a Laguerre filter, the model output can be expressed as (Saha *et al.*, 2004; Aadaleesan *et al.*, 2008)

$$\hat{y}(z)=\left(\sum_{i=1}^{N} c_i L_i(z)\right)u(z), \tag{2}$$

where $L_i(z)$, which is the z-domain representation of the $l_i(k)$ in Eq. (1), denotes the ith order Laguerre filter. $L_i(z)$ is given as

$$L_i(z)=\sqrt{(1-\alpha^2)T}\,\frac{(1-\alpha z)^{i-1}}{(z-\alpha)^i}, \tag{3}$$

where $\alpha=\mathrm{e}^{-pT}$, p is the time scale parameter of the Laguerre filter, N is the number of Laguerre filters, and T is the sampling interval. Defining an N dimensional state vector $\boldsymbol{L}(k)=[l_1(k), l_2(k), \ldots, l_N(k)]^{\mathrm{T}}$, a discrete state space representation proposed by Zervos and Dumont (1988) is

$$\boldsymbol{L}(k+1)=\boldsymbol{A}\boldsymbol{L}(k)+\boldsymbol{B}u(k), \tag{4}$$

where $u(k)$ is the input. Defining

$$\tau_1=\mathrm{e}^{-pT},\quad \tau_2=T+\frac{2}{P}(\mathrm{e}^{-pT}-1),$$

$$\tau_3=-T\,\mathrm{e}^{-pT}-\frac{2}{P}(\mathrm{e}^{-pT}-1),\quad \tau_4=\frac{\sqrt{2p}}{p}(1-\tau_1),$$

then

$$\boldsymbol{A}=\begin{bmatrix} \tau_1 & 0 & \cdots & 0 \\ \dfrac{-\tau_1\tau_2-\tau_3}{T} & \tau_1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \dfrac{(-1)^{N-1}\tau_2^{N-2}(\tau_1\tau_2+\tau_3)}{T^{N-1}} & \cdots & \dfrac{-\tau_1\tau_2-\tau_3}{T} & \tau_1 \end{bmatrix}. \quad (5)$$

$$\boldsymbol{B}=[\tau_4, (-\tau_2/T)\tau_4, ..., (-\tau_2/T)^{N-1}\tau_4]. \quad (6)$$

When $p>0$, the above state space system is stable, observable, and controllable (Zervos and Dumont, 1988). For linear systems, Wahlberg (1991) suggested that the value of p should be chosen close to the inverse of the system dominant time constant in order to obtain a fast convergence rate.

For the linear model given by Eq. (2), the output of the process can be approximated by the weighted sum of the outputs of the Laguerre filter states:

$$\hat{y}(k)=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{L}(k), \quad (7)$$

where $\boldsymbol{C}^{\mathrm{T}}=[c_1, c_2, \ldots, c_N]$ are Laguerre filter coefficients.

2.2 Static nonlinear subsystem

To construct a Wiener type model using Laguerre filters, a nonlinear mapping is substituted into Eq. (7), and the model output is formulated as

$$\hat{y}(k)=f(\boldsymbol{L}(k)), \quad (8)$$

where $f(\cdot)$: $\mathbb{R}^N \to \mathbb{R}$ denotes a memoryless nonlinear mapping.

A polynomial function described by Eq. (9) was introduced by Mahmoodi *et al*. (2009) to represent the nonlinear subsystem:

$$\hat{y}(z)=h_0+\sum_{i=1}^{N}h_i l_i(k)+\sum_{i=1}^{N}\sum_{j=i}^{N}h_{ij}l_i(k)l_j(k)+\cdots+\sum_{i=1}^{N}\sum_{j=i}^{N}\cdots\sum_{k=r}^{N}\sum_{m=k}^{N}h_{ij\ldots km}l_i(k)l_j(k)\ldots l_k(k)l_m(k). \quad (9)$$

Eq. (9) can be rewritten as $\hat{y}(z)=\boldsymbol{H}^{\mathrm{T}}\boldsymbol{\psi}(k)$, where $\boldsymbol{H}=[h_0, h_1, \ldots, h_R]^{\mathrm{T}}$, and regression vector $\boldsymbol{\psi}(k)$ is defined by $\boldsymbol{\psi}(k)=[1, l_1, ..., l_N, l_1^2, l_1l_2, ..., l_N^2, l_1^3, l_1l_2l_3, ..., l_N^3, ..., l_N^M]^{\mathrm{T}}$.

A drawback of the polynomial function is that the number of parameters to be identified, $R=\sum_{i=1}^{n}\binom{N+i-1}{i}$, will be excessively large if either polynomial order n or Laguerre filter number N increases, and this will cause a dimensional disaster.

In this study, an LSSVM model is used to identify the static nonlinearity of the Wiener model. In this approach, the nonlinear block is represented as

$$\hat{y}(k)=f(\boldsymbol{L}(k))=\boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{L}(k))+b, \quad (10)$$

where the weight vector $\boldsymbol{w}\in\mathbb{R}^{nh}$, $\varphi(\cdot)$: $\mathbb{R}^N\to\mathbb{R}^{nh}$ is a nonlinear mapping from the input space to the feature space, and b is the bias term.

The identification problem of the proposed Wiener model is based on the plant input-output dataset $\{u_{\mathrm{t}}(k), y_{\mathrm{t}}(k)\}_{k=1}^{M}$. For clarity, subscript 't' is used to distinguish the training input-output data from general inputs and outputs. Note that the Laguerre filter state vector $\boldsymbol{L}_{\mathrm{t}}(k)$ can be computed according to Eq. (4); hence, the LSSVM can be identified according to the data sequence $\{\boldsymbol{L}_{\mathrm{t}}(k), y_{\mathrm{t}}(k)\}_{k=1}^{M}$.

The approximation error of sample k is defined as $e_k=y_{\mathrm{t}}(k)-\hat{y}_{\mathrm{t}}(\boldsymbol{L}(k))$. In LSSVM regression, the cost function is formulated as (Suykens and Vandewalle, 1999; Goethals *et al.*, 2005)

$$\min\{\boldsymbol{w}, b, e\}: J(\boldsymbol{w}, e)=\frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}+\frac{\gamma}{2}\sum_{k=1}^{M}e_k^2 \quad (11)$$

$$\text{subject to } y_{\mathrm{t}}(k)=\boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{L}_{\mathrm{t}}(k))+b+e_k, \quad (12)$$

where $k=1, 2, \ldots, M$, and the scalar γ is a regularization constant to govern the smoothness of the solution and data fitting. To solve the constrained optimization problem, the Lagrangian is

$$\xi(\boldsymbol{w}, b, e; \alpha)=J(\boldsymbol{w}, e)-\sum_{k=1}^{M}\alpha_k\left(\boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{L}_{\mathrm{t}}(k))+b+e_k-y_{\mathrm{t}}(k)\right), \quad (13)$$

where $\alpha_k\in\mathbb{R}$ ($k=1, 2, \ldots, M$) are Lagrange multipliers. The conditions for optimality are given as

$$\frac{\partial \xi}{\partial \boldsymbol{w}}=0 \rightarrow \boldsymbol{w}=\sum_{k=1}^{M} \alpha_k \varphi(\boldsymbol{L}_{\mathrm{t}}(k)), \tag{14}$$

$$\frac{\partial \xi}{\partial b}=0 \rightarrow \sum_{k=1}^{M} \alpha_k=0, \tag{15}$$

$$\frac{\partial \xi}{\partial e_k}=0 \rightarrow \alpha_k=\gamma e_k, \tag{16}$$

$$\frac{\partial \xi}{\partial \alpha_k}=0 \rightarrow y_{\mathrm{t}}(k)=\boldsymbol{w}^{\mathrm{T}} \varphi(\boldsymbol{L}_{\mathrm{t}}(k))+b+e_k. \tag{17}$$

Elimination of e_k and $\boldsymbol{w}$ through substitution results in the following set of linear equations:

$$\begin{pmatrix} \boldsymbol{0} & \boldsymbol{1}_M^{\mathrm{T}} \\ \boldsymbol{1}_M & \boldsymbol{\Omega}+\gamma^{-1}\boldsymbol{I}_M \end{pmatrix} \begin{pmatrix} b \\ \boldsymbol{\alpha} \end{pmatrix}=\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{Y}_{\mathrm{t}} \end{pmatrix}, \tag{18}$$

where $\Omega_{ij}=K(\boldsymbol{L}_{\mathrm{t}}(i), \boldsymbol{L}_{\mathrm{t}}(j))=\varphi(\boldsymbol{L}_{\mathrm{t}}(i)^{\mathrm{T}}\boldsymbol{L}_{\mathrm{t}}(j))$, $i, j=1, 2, \ldots, M$ with K being the positive definite kernel function. $\boldsymbol{Y}_{\mathrm{t}}=[y_{\mathrm{t}}(1), y_{\mathrm{t}}(2), \ldots, y_{\mathrm{t}}(M)]^{\mathrm{T}}$, $\boldsymbol{\alpha}=[\alpha_1, \alpha_2, ..., \alpha_M]^{\mathrm{T}}$, $\boldsymbol{1}_M=[1, 1, ..., 1]^{\mathrm{T}}$. To calculate the elements in Eq. (18), the mapping $\varphi(\cdot)$ from the input space to the feature space does not need to be defined explicitly. Only its inner product in the form of a positive kernel is needed (de Kruif and de Vries, 2003). In this work, a radial basis function (RBF) kernel described by Eq. (19) is adopted as a kernel function due to its good performance under a general smoothness assumption.

$$K(\boldsymbol{L}_{\mathrm{t}}(i), \boldsymbol{L}_{\mathrm{t}}(j))=\exp(-\left\|\boldsymbol{L}_{\mathrm{t}}(i)-\boldsymbol{L}_{\mathrm{t}}(j)\right\|_2^2/\sigma^2), \tag{19}$$

where $\sigma>0$ denotes the bandwidth of the kernel.

The output of the LSSVM model for a new input $\boldsymbol{L}^*$ can be evaluated as

$$f(\boldsymbol{L}^*)=\sum_{k=1}^{M} \alpha_k K(\boldsymbol{L}_{\mathrm{t}}(k), \boldsymbol{L}^*)+b, \tag{20}$$

where α_k and b are the solutions of Eq. (18).

3 Wiener model based nonlinear model predictive control formulation

In this section, the Laguerre-LSSVM Wiener model based NMPC is developed. MPC refers to a class of computer control algorithms that predict and optimize the future behavior of a plant by using the process dynamic model. At each sampling interval, the MPC algorithm computes a sequence of manipulated variables over a control horizon M_{C} to optimize the future plant behavior. The optimization is performed by minimizing a criterion function based on a desired output trajectory over a prediction horizon P. The first value in the optimal sequence is injected into the plant, and the entire optimization is repeated at subsequent control intervals (Gómez and Baeyens, 2004).

Based on the Laguerre-LSSVM Wiener model, P-step-ahead prediction can be written as

$$\begin{aligned} \boldsymbol{L}(k+i|k) &= \boldsymbol{A}\boldsymbol{L}(k+i-1|k)+\boldsymbol{B}u(k+i-1|k) \\ &= \boldsymbol{A}^i\boldsymbol{L}(k|k)+\sum_{j=1}^{i}\boldsymbol{A}^{i-j}\boldsymbol{B}u(k+j-1|k), \quad i=1, 2, ..., P, \end{aligned} \tag{21}$$

$$\hat{y}(k+i|k)=f(\boldsymbol{L}(k+i|k)), \tag{22}$$

where $u(k+M_{\mathrm{C}}-1+j|k)=u(k+M_{\mathrm{C}}-1|k)$, $j=1, 2, \ldots, P-M_{\mathrm{C}}+1$.

In order to compensate the plant-model mismatch and account for unmeasured disturbance, a mismatch correction term is incorporated in the prediction model as follows (Saha *et al.*, 2004; Mahmoodi *et al.*, 2009):

$$y_{\mathrm{c}}(k+i|k)=\hat{y}(k+i|k)+d(k|k), \quad i=1,2,...,P, \tag{23}$$

$$d(k|k)=y(k)-\hat{y}(k|k-1), \tag{24}$$

where $y(k)$ represents the measured plant output at the kth instant, and $\hat{y}(k|k-1)$ represents the measured plant output at the kth instant.

The control goal of the nonlinear process is that the model output can reach the set point quickly and smoothly. Hence, the plant output should follow a reference trajectory as

$$\begin{cases} y_{\mathrm{r}}(k+i)=\alpha_{\mathrm{ref}}^i y_{\mathrm{r}}(k)+(1-\alpha_{\mathrm{ref}}^i)y_{\mathrm{sp}}, \ i=1, 2, ..., P, \\ y_{\mathrm{r}}(k)=y(k), \end{cases} \tag{25}$$

where $y_{\mathrm{r}}(k+i)$ is the reference trajectory, y_{sp} is the set point, and α_{ref} is a parameter ($0\leq\alpha_{\mathrm{ref}}\leq1$).

The purpose of the MPC controller is to compute future control moves, which will minimize a criterion function based on the desired output trajectory over the prediction horizon subject to constraints on the output and control variables. A commonly used criterion function for the constrained MPC is

$$\min_{u(k|k),\dots,u(k+M_C-1|k)}\left(\sum_{i=1}^{P}\left\|e(k+i|k)\right\|_{\boldsymbol{Q}}^2+\sum_{i=1}^{M_C}\left\|u(k+i|k)\right\|_{\boldsymbol{R}}^2+\sum_{i=1}^{M_C}\left\|\Delta u(k+i|k)\right\|_{\boldsymbol{S}}^2\right), \tag{26}$$

subject to

$$\begin{cases}u_{\min}\le u(k+i|k)\le u_{\max},\\ \Delta u_{\min}\le \Delta u(k+i|k)\le \Delta u_{\max},\\ y_{\min}\le y(k+i|k)\le y_{\max},\end{cases} \tag{27}$$

where

$$e(k+i|k)=y_r(k+i|k)-y_c(k+i|k). \tag{28}$$

$$\Delta u(k+i|k)=u(k+i|k)-u(k+i-1|k). \tag{29}$$

Here, $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{S}$ are positive semi-definite weight matrices of the prediction error (e), magnitude of the input (u), and the change in the input (Δu).

In this study, the resulting nonlinear programming problem is solved using the successive quadratic programming (SQP) algorithm. $\boldsymbol{Q}$, $\boldsymbol{R}$, $\boldsymbol{S}$, as well as the control horizon M_C and prediction horizon P are design parameters that must be tuned to provide the controller with a satisfactory performance.

4 Case study

4.1 Dynamic linear subsystem

Schematic diagram of the UCSB pH neutralization process is shown in Fig. 2 (Henson and Seborg, 1994). The process has three inlet streams: an acid stream q_1 (HNO_3), a buffer stream q_2 ($NaHCO_3$), and a base stream q_3 (NaOH with trace amounts of $NaHCO_3$). The liquid level (h) and effluent pH (pH) are measured variables. Usually, the objective is to control the pH by manipulating the base flow rate, while the flow rates of buffer and acid streams represent unmeasured disturbances (Pottmann and Seborg, 1997).

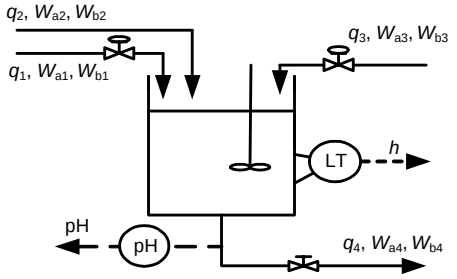

Fig. 2 Schematic diagram of the UCSB pH neutralization process

The process has three inlet streams: an acid stream q_1 (HNO_3), a buffer stream q_2 ($NaHCO_3$), and a base stream q_3 (NaOH with trace amounts of $NaHCO_3$). The liquid level (h) and effluent pH (pH) are measured variables. LT: liquid transmitter

The dynamic model of the pH neutralization process is described by three nonlinear ordinary equations and one nonlinear algebraic equation. More details can be found in Henson and Seborg (1994).

$$\frac{\mathrm{d}h}{\mathrm{d}t}=\frac{1}{A}(q_1+q_2+q_3-C_v h^{0.5}), \tag{30}$$

$$\frac{\mathrm{d}W_{a4}}{\mathrm{d}t}=\frac{1}{Ah}[(W_{a1}-W_{a4})q_1+(W_{a2}-W_{a4})q_2+(W_{a3}-W_{a4})q_3], \tag{31}$$

$$\frac{\mathrm{d}W_{b4}}{\mathrm{d}t}=\frac{1}{Ah}[(W_{b1}-W_{b4})q_1+(W_{b2}-W_{b4})q_2+(W_{b3}-W_{b4})q_3], \tag{32}$$

$$\frac{W_{b4}(1+2\times10^{\mathrm{pH}-\mathrm{pK}_2})}{1+10^{\mathrm{pK}_1-\mathrm{pH}}+10^{\mathrm{pH}-\mathrm{pK}_2}}+W_{a4}+10^{\mathrm{pH}-14}=10^{-\mathrm{pH}}. \tag{33}$$

Here, W_{a4} and W_{b4} are the reaction invariants of the effluent stream, and pK_1 and pK_2 are the first and second disassociation constants of the weak acid H_2CO_3, respectively. The nominal operating conditions are listed in Table 1.

Table 1 Nominal operating parameters of the system

Parameter	Value	Parameter	Value
A (cm^2)	207	W_{b1} (mol/L)	0
C_v (ml/(cm·s))	8.75	W_{b2} (mol/L)	3×10^{-2}
pK_1	6.35	W_{b3} (mol/L)	5×10^{-5}
pK_2	10.25	q_1 (ml/s)	16.6
W_{a1} (mol/L)	3×10^{-3}	q_2 (ml/s)	0.55
W_{a2} (mol/L)	-3×10^{-2}	q_3 (ml/s)	15.55
W_{a3} (mol/L)	3.05×10^{-3}	h (cm)	14.0

4.2 Wiener model identification of the process

To construct the Laguerre-LSSVM Wiener model, input-output data of the pH neutralization process should be collected. A mechanistic model of the pH process is used to generate training and testing data. The buffer stream (q_2) acts as the model input, while the effluent pH (pH) is the model output.

Traditionally, pseudo-random binary sequences (PRBS) are used as the exciting signal to the plant to generate process input-output data. This is sufficient for the identification of linear models because only frequency of the test signal needs to be considered. For the identification of Wiener models, binary test signals applied on the process will not be possible to detect the nonlinearity. Therefore, multi-amplitude or continuous amplitude signals should be used to excite the nonlinearity (Zhu, 2001). GMN, which is a multiple-level generalization of generalized binary noise (GBN) (Tulleken, 1990), is used as the excitation signal to the pH neutralization process in this work. The amplitude of GMN has a uniform distribution over the operating range (0, 30), and switching time is T_{sw}=50. In order to determine the robustness of the identified models with respect to measurement noise, Gaussian white noise with zero mean is added in the model output. The signal-to-noise ratio (SNR) as defined by Eq. (34) is SNR=20.

$$\mathrm{SNR}=10\lg(\sigma^2_{\mathrm{signal}}/\sigma^2_{\mathrm{noise}}), \tag{34}$$

where $\sigma^2_{\mathrm{signal}}$ and $\sigma^2_{\mathrm{noise}}$ are the variances of signal and noise, respectively.

A set of 4000 input-output samples are collected. The first 3000 data are used to identify the Wiener model and the remaining 1000 data are used for model validation. Fig. 3 shows the GMN test signal, and Fig. 4 shows the corresponding model outputs.

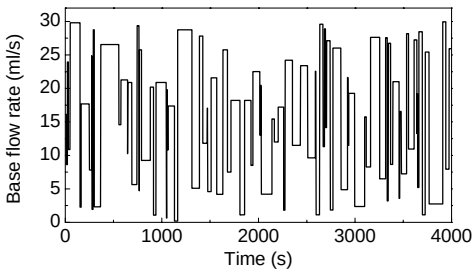

Fig. 3 Generalized multiple-level noise (GMN) test signal

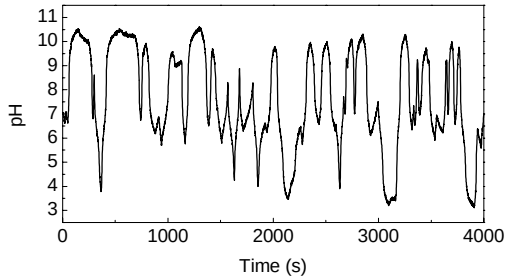

Fig. 4 Output data for model identification

There are four parameters which affect the estimation precision and thus need to be tuned: p and N in Eqs. (2) and (3), which decide the dynamic performance of the Wiener model; γ in Eq. (18), which determines the trade-off between minimizing training errors and model complexity; and, σ in Eq. (19), which influences the number of initial eigenvectors and the fitting level of the static model. As the parameters are problem-specific, the cross-validation technique (Tötterman and Toivonen, 2009) is used in this work. Selection of parameters is based on the model performance on independent test data which are not used for identification. Test results are compared according to the variance accounted for (VAF) criterion (Mahmoodi *et al.*, 2009), which is described by

$$\mathrm{VAF}=\max\left(1-\frac{\mathrm{var}\{y-\hat{y}\}}{\mathrm{var}\{y\}},0\right)\times 100\%, \tag{35}$$

where y denotes the real output sequence, $\hat{y}$ denotes the model output sequence, and var{·} denotes the variance of the signal.

Tables 2 and 3 show the testing results of various Wiener model parameters. Table 2 shows the VAF of testing data with fixed LSSVM parameters γ=1, σ=200, and various Laguerre filter parameters p and N, while Table 3 shows the testing results of various LSSVM parameters with fixed Laguerre filter parameters p=0.013 and N=4. It is seen that the cross-validation approach gives Laguerre-LSSVM Wiener model parameters p=0.013, N=4, γ=1, and σ=200, and the corresponding VAF is 97.2875. Validation results of the estimated Wiener model are depicted in Fig. 5.

Linear Laguerre models and three-order polynomial function based Wiener models are constructed

for comparison. Different orders (N) and time scale parameters (p) of Laguerre filters are tested. Tables 4 and 5 show the testing results.

For the linear Laguerre models, the highest VAF value 96.1677 is obtained when N=4 and p=0.009. The highest VAF value of Laguerre-polynomial Wiener models is 97.0521 when N=2 and p=0.011.

Table 2 Testing results of various Laguerre filter parameters (Laguerre-LSSVM Wiener model)

N	VAF*	p	VAF**
2	97.1941	0.007	97.0801
3	96.8608	0.009	97.0461
4	**97.2785**	0.011	97.2509
5	97.1093	0.013	**97.2785**
6	95.4080	0.015	97.0969

* p=0.013, γ=1, σ=200; ** N=4, γ=1, σ=200

Table 3 Testing results of various LSSVM parameters

γ	VAF*	σ	VAF**
0.1	94.9601	160	95.9706
0.5	96.9407	180	96.9872
1	**97.2785**	200	**97.2875**
5	96.5022	220	97.1074
10	95.1206	240	96.6820

* p=0.013, N=4, σ=200; ** p=0.013, N=4, γ=1

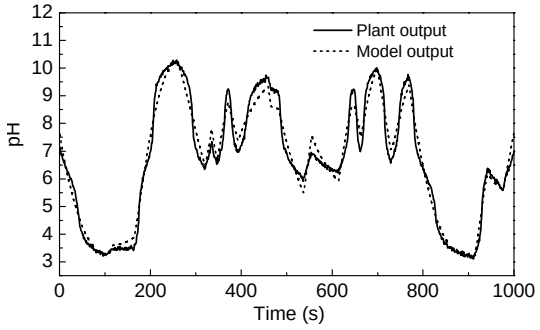

Fig. 5 True and estimated output for the Laguerre-LSSVM Wiener model

Table 4 Testing results of linear Laguerre filters

p	VAF			
	N=2	N=3	N=4	N=5
0.007	94.9032	95.9931	96.1494	96.1425
0.009	96.0576	96.1482	**96.1677**	95.9079
0.011	96.1608	96.1659	96.0419	95.6370
0.013	96.1061	96.1274	95.8276	95.6403

Table 5 Testing results of Laguerre-Wiener models

p	VAF			
	N=2	N=3	N=4	N=5
0.007	95.2014	96.3930	96.8106	84.6421
0.009	97.0033	95.4103	95.6267	85.6034
0.011	**97.0521**	95.5994	95.6931	85.1963
0.013	96.9390	96.8880	96.0204	83.8282

The ε-SVR based Wiener model is also used for comparison. Details of the SVR based Wiener model can be found in Tötterman and Toivonen (2009). Beside the linear Laguerre parameters, the support vector method requires the selection of design parameter ε in the loss function, C in the structural risk function, and RBF kernel band width σ. The threshold ε affects the model accuracy and the number of support vectors. The parameter C determines the trade-off between the empirical risk and the complexity penalty in the structural risk function.

The Laguerre-SVR Wiener model parameters are selected using cross-validation techniques. To acquire a high prediction accuracy, ε is assumed equal to a very small value (ε=1E-6). Training and testing data are scaled to [−1, 1]. Then the four parameters p, N, C, and σ are adjusted. Tables 6 and 7 show the testing results. It is seen that the Wiener model parameters obtained are p=0.011, N=3, C=1, and σ=0.9, and the corresponding VAF is 97.1520.

Table 6 Testing results of various Laguerre filter parameters (Laguerre-SVR Wiener model)

N	VAF*	p	VAF**
2	97.0659	0.007	96.4894
3	**97.1520**	0.009	96.4836
4	94.6221	0.011	**97.1520**
5	95.1430	0.013	97.0252
6	95.3828	0.015	96.5285

* p=0.011, C=1, σ=0.9; ** N=3, C=1, σ=0.9

Table 7 Testing results of various SVR parameters

C	VAF*	σ	VAF**
0.1	96.3985	0.8	95.9723
0.5	97.1295	0.9	**97.1520**
1	**97.1520**	1	97.0954
5	96.9898	1.1	97.0040
10	96.7016	1.2	96.8398

* p=0.011, N=3, σ=0.9; ** p=0.011, N=3, C=1

In this simulation, the numbers of support vectors are all 3000 because of the selection of a small ε. To reduce the model complexity, ε should be adjusted. Table 8 shows the testing results of various ε, showing that a small ε gives high prediction accuracy at the expense of an increased number of support vectors.

Table 8 Testing results of various ε

ε	VAF	Number of support vectors
1E-6	97.1520	3000
0.1	97.0355	2205
0.3	96.6228	1374
0.5	95.3731	479

Note that the SVR has an attractive property of sparseness; i.e., many elements in the QP solution vector are equal to zero. This reduces the Wiener model complexity. The application of SVR in large datasets, however, has been limited by the time and memory consumption optimization. This disadvantage has been overcome by LSSVM, which solves linear equations instead of a QP problem. Simulation results indicate that the proposed Laguerre-LSSVM Wiener model has higher prediction accuracy than the other three models, and that the identified model can represent the dynamic behavior of the pH neutralization process efficiently.

4.3 Model predictive control simulation results

The NMPC based on the Laguerre-LSSVM Wiener model is applied to the pH neutralization process. The objective is to control the pH by manipulating the base flow rate. The system sample time is chosen as Δt=10 s. The prediction horizon is chosen as P=4, while the number of control horizons is chosen as M_C=2. The weighting matrices $\boldsymbol{Q}$=10$\boldsymbol{I}$, $\boldsymbol{R}$=$\boldsymbol{0}$, and $\boldsymbol{S}$=4$\boldsymbol{I}$ are evaluated through simulation, trying to solve the trade-off between the smooth control action and fast response. A lower limit of 0 ml/s and an upper limit of 30 ml/s are chosen for manipulated variable restrictions to guarantee that the operating region of the process is always inside the identified region. Also, the lower and upper limits of the change in the input are chosen as ±1 ml/s.

For comparison, linear Laguerre model based MPC, Laguerre-polynomial Wiener model based NMPC, and Laguerre-SVR Wiener model based NMPC are developed. All of the control schemes have the same set of parameters as the proposed NMPC algorithm. Set-point tracking behaviors of the four controllers are depicted in Figs. 6a–6c. The corresponding base flow rates are shown in Fig. 7. In Table 9, mean absolute errors (MAE), mean squared errors (MSE), and sum squared errors (SSE) of tracking errors for the three controllers are compared. As can been seen, the proposed NMPC performs better than the other three controllers.

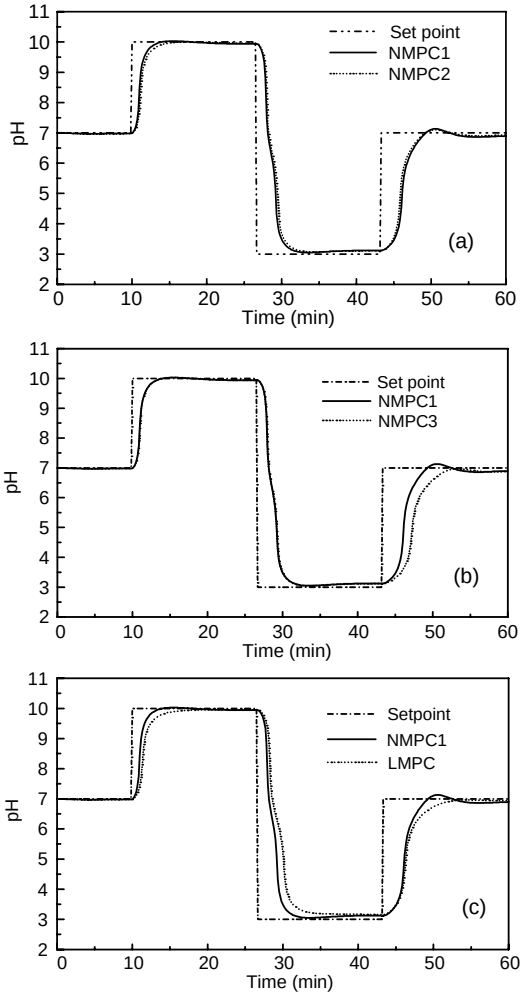

Fig. 6 Set-point tracking performances of NMPC1 and NMPC2 (a), NMPC1 and NMPC3 (b), and NMPC1 and LMPC (c)

NMPC1: the Laguerre-LSSVM Wiener model based nonlinear model predictive control; NMPC2: the Laguerre-SVR Wiener model based nonlinear model predictive control; NMPC3: the Laguerre-polynomial Wiener model based nonlinear model predictive control; LMPC: linear model predictive control

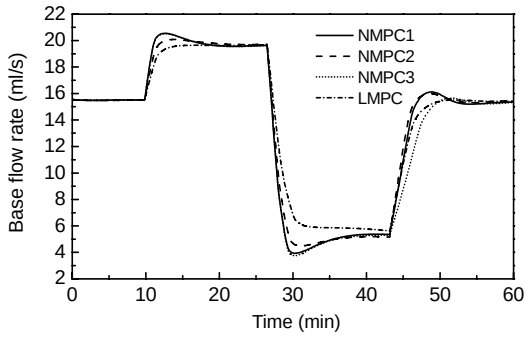

Fig. 7 Manipulated variables for the different models

NMPC1: the Laguerre-LSSVM Wiener model based nonlinear model predictive control; NMPC2: the Laguerre-SVR Wiener model based nonlinear model predictive control; NMPC3: the Laguerre-polynomial Wiener model based nonlinear model predictive control; LMPC: linear model predictive control

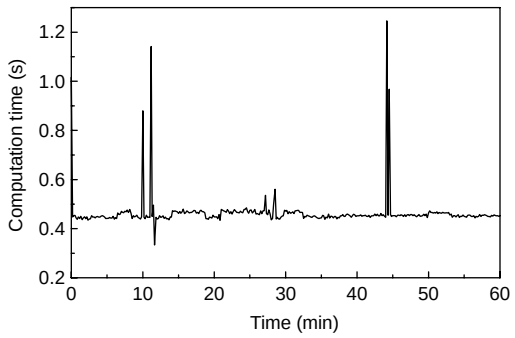

Fig. 8 Computation time of the Laguerre-LSSVM Wiener model based nonlinear model predictive control (NMPC1)

Table 9 Performance comparison of four controllers

Scheme	MAE	SSE	MSE
NMPC1	0.5466	725.9188	2.0109
NMPC2	0.5569	753.3591	2.0869
NMPC3	0.6220	815.6396	2.2594
LMPC	0.6836	889.4822	2.4639

MAE: mean absolute errors; SSE: sum squared errors; MSE: mean squared errors. NMPC1: the Laguerre-LSSVM Wiener model based nonlinear model predictive control; NMPC2: the Laguerre-SVR Wiener model based nonlinear model predictive control; NMPC3: the Laguerre-polynomial Wiener model based nonlinear model predictive control; LMPC: linear model predictive control

Note that the proposed Laguerre-LSSVM Wiener model has a drawback in that the number of terms (M) in Eq. (20) for model prediction equals the number of training data (M=3000 in this study), while the other three Wiener models can make predictions with many fewer parameters. Thus, a large number of training data will increase the model complexity and the computation load, especially in the application of MPC. In this study, all the set-point tracking experiments are performed using Matlab 7.0 on a computer with a Pentium III processor, 512 MB RAM. The computation time required for generating the control signal in the NMPC1 simulation is shown in Fig. 8.

As can be seen, the average computation time for optimization is 0.4641 s, and the maximum computation time for optimization is 1.2459 s, which is sufficiently below the chosen sampling time 10 s. For a system that has a smaller sampling time T_s or large prediction horizon P, however, the number of training data needs to be pruned to reduce the online computation load of NMPC.

5 Conclusions

A novel Laguerre-LSSVM Wiener type model is proposed for pH neutralization process identification and nonlinear model predictive control. Laguerre filter network is used to describe the dynamic linear subsystem of the Wiener model, while the LSSVM model is used to represent the static nonlinearity. Input-output data of the process are used for Wiener model identification, and the simulation results show that the identified Wiener model performs better than the Laguerre-SVR Wiener model, Laguerre-polynomial Wiener model, and linear Laguerre model. The identified Wiener model is used as the internal model to develop a nonlinear model predictive control algorithm for the pH neutralization process, and the successive quadratic programming algorithm is used to solve the resulting nonlinear programming problem. The set-point tracking performance of the proposed NMPC is compared with those of the Laguerre-SVR Wiener model based NMPC, Laguerre-polynomial Wiener model based NMPC, and LMPC. Simulation results show that the proposed NMPC outperforms the other three controllers.

References

Aadaleesan, P., Miglan, N., Sharma, R., Saha, P., 2008. Nonlinear system identification using Wiener type Laguerre-wavelet network model. *Chem. Eng. Sci.*, **63**(15):3932-3941. [doi:10.1016/j.ces.2008.04.043]

Akesson, B.M., Toivonen, H.T., Waller, J.B., Nyström, R.H., 2005. Neural network approximation of a nonlinear model predictive controller applied to a pH neutralization process. *Comput. Chem. Eng.*, **29**(2):323-335. [doi:10.

1016/j.compchemeng.2004.09.023]

Altinten, A., 2007. Generalized predictive control applied to a pH neutralization process. *Comput. Chem. Eng.*, **31**(10): 1199-1204. [doi:10.1016/j.compchemeng.2006.10.005]

Arefi, M.M., Montazeri, A., Poshtan, J., Jahed-Motlagh, M.R., 2008. Wiener-neural identification and predictive control of a more realistic plug-flow tubular reactor. *Chem. Eng. J.*, **138**(1-3):274-282. [doi:10.1016/j.cej.2007.05.044]

Bao, Z.J., Pi, D.Y., Sun, Y.X., 2007. Nonlinear model predictive control based on support vector machine with multi-kernel. *Chin. J. Chem. Eng.*, **15**(5):691-697. [doi:10.1016/S1004-9541(07)60147-5]

Böling, J.M., Seborg, D.E., Hespanha, J.P., 2007. Multi-model adaptive control of a simulated pH neutralization process. *Control Eng. Pract.*, **15**(6):663-672. [doi:10.1016/j.conengprac.2006.11.008]

Cervantes, A.L., Agamennoni, O.E., Figueroa, J.L., 2003. A nonlinear model predictive control system based on Wiener piecewise linear models. *J. Process Control*, **13**(7):655-666. [doi:10.1016/S0959-1524(02)00121-X]

de Kruif, B.J., de Vries, T.J.A., 2003. Pruning error minimization in least squares support vector machines. *IEEE Trans. Neur. Networks*, **14**(3):696-702. [doi:10.1109/TNN.2003.810597]

Fang, Y., Chow, T.W.S., 2000. Orthogonal wavelet neural networks applying to identification of Wiener model. *IEEE Trans. Circ. Syst.-I: Fundam. Theory Appl.*, **47**(4): 591-593. [doi:10.1109/81.841863]

Fruzzetti, K.P., Palazoglu, A., McDonald, K.A., 1997. Nonlinear model predictive control using Hammerstein models. *J. Process Control*, **7**(1):31-41. [doi:10.1016/S0959-1524(97)80001-B]

Goethals, I., Pelckmans, K., Suykens, J.A.K., de Moor, B., 2005. Identification of MIMO Hammerstein models using least squares support vector machines. *Automatica*, **41**(7): 1263-1272. [doi:10.1016/j.automatica.2005.02.002]

Gómez, J.C., Baeyens, E., 2004. Wiener model identification and predictive control of a pH neutralisation process. *IEE Proc.-Control Theory Appl.*, **151**(3):329-338. [doi:10.1049/ip-cta:20040438]

Henson, M.A., Seborg, D.E., 1994. Adaptive nonlinear control of a pH neutralization process. *IEEE Trans. Control Syst. Technol.*, **2**(3):169-182. [doi:10.1109/87.317975]

Hsu, Y.L., Wang, J.S., 2009. A Wiener-type recurrent neural network and its control strategy for nonlinear dynamic applications. *J. Process Control*, **19**(6):942-953. [doi:10.1016/j.jprocont.2008.12.002]

Kalafatis, A., Arifin, N., Wang, L., Cluett, W.R., 1995. A new approach to the identification of pH processes based on the Wiener model. *Chem. Eng. Sci.*, **50**(23):3693-3701. [doi:10.1016/0009-2509(95)00214-P]

Lakshmi Narayanan, N.R., Krishnaswamy, P.R., Rangaiah, G.P., 1997. An adaptive internal model control strategy for pH neutralization. *Chem. Eng. Sci.*, **52**(18):3067-3074. [doi:10.1016/S0009-2509(97)00130-9]

Loh, A.P., Looi, K.O., Fong, K.F., 1995. Neural network modeling and control strategies for a pH process. *J. Process Control*, **5**(6):355-362. [doi:10.1016/0959-1524(95)00005-B]

Lovera, M., Gustafsson, T., Verhaegen, M., 2000. Recursive subspace identification of linear and non-linear Wiener state-space models. *Automatica*, **36**(11):1639-1650. [doi:10.1016/S0005-1098(00)00103-5]

Mahmoodi, S., Poshtan, J., Jahed-Motlagh, M.R., Montazeri, A., 2009. Nonlinear model predictive control of a pH neutralization process based on Wiener-Laguerre model. *Chem. Eng. J.*, **146**(3):328-337. [doi:10.1016/j.cej.2008.06.010]

Norquay, S.J., Palazoglu, A., Romagnoli, J.A., 1998. Model predictive control based on Wiener models. *Chem. Eng. Sci.*, **53**(1):75-84. [doi:10.1016/S0009-2509(97)00195-4]

Norquay, S.J., Palazoglu, A., Romagnoli, J.A., 1999. Application of Wiener model predictive control (WMPC) to a pH neutralization experiment. *IEEE Trans. Control Syst. Technol.*, **7**(4):437-445. [doi:10.1109/87.772159]

Pottmann, M., Seborg, D.E., 1997. A nonlinear predictive control strategy based on radial basis function models. *Comput. Chem. Eng.*, **21**(9):965-980. [doi:10.1016/S0098-1354(96)00340-7]

Pröll, T., Karim, M.N., 1994. Model predictive pH control using real-time NARX approach. *AIChE J.*, **40**(2):269-282. [doi:10.1002/aic.690400207]

Saha, P., Krishnan, S.H., Rao, V.S.R., Patwardhan, S.C., 2004. Modeling and predictive control of MIMO nonlinear systems using Wiener-Laguerre models. *Chem. Eng. Commun.*, **191**(8):1083-1119. [doi:10.1080/00986440490276452]

Shafiee, G., Arefi, M.M., Jahed-Motlagh, M.R., Jalali, A.A., 2008. Nonlinear predictive control of a polymerization reactor based on piecewise linear Wiener model. *Chem. Eng. J.*, **143**(1-3):282-292. [doi:10.1016/j.cej.2008.05.013]

Suykens, J.A.K., Vandewalle, J., 1999. Least squares support vector machine classifiers. *Neur. Process. Lett.*, **9**(3):293-300. [doi:10.1023/A:1018628609742]

Tötterman, S., Toivonen, H.T., 2009. Support vector method for identification of Wiener models. *J. Process Control*, **19**(7):1174-1181. [doi:10.1016/j.jprocont.2009.03.003]

Tulleken, H.J.A.F., 1990. Generalized binary noise test-signal concept for improved identification experiment design. *Automatica*, **26**(1):37-49. [doi:10.1016/0005-1098(90)90156-C]

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

Wahlberg, B., 1991. System identification using Laguerre models. *IEEE Trans. Autom. Control*, **36**(5):551-562. [doi:10.1109/9.76361]

Wahlberg, B., 1994. System identification using Kautz models. *IEEE Trans. Autom. Control*, **39**(6):1276-1282. [doi:10.1109/9.293196]

Wiener, N., 1958. Nonlinear Problems in Random Theory. Technology Press, MIT and Wiley, New York.

Zervos, C.C., Dumont, G.A., 1988. Deterministic adaptive control based on Laguerre series representation. *Int. J. Control*, **48**(6):2333-2359. [doi:10.1080/00207178808906334]

Zhu, Y., 2001. Multivariable System Identification for Process Control. Elsevier Science Ltd., Oxford.