Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Adaptive Predictive Control Strategy Using Wavenet Based Plant Modeling

Yul Y. Nazaruddin and Ferdian Cahyadi

Instrumentation and Control Research Group, Department of Engineering Physics,
Institut Teknologi Bandung, Bandung 40132, Indonesia
(Tel. +62-22-2504424; e-mail: yul@tf.itb.ac.id)

Abstract: An alternative control strategy for nonlinear processes which is an integration between the generalized predictive control and wavelet network (wavenet) approach is proposed in this paper. The wavenet is used for modeling the process as it has learning capability from the numerical data obtained from the measurements and subsequently used as process model in the generalized predictive control scheme. The process model is represented in a Nonlinear AutoRegressive with eXogeneous variables (NARX) model. The modeling process is implemented on-line at each control action and this allows the control to be done adaptively. The proposed adaptive control scheme with its wavenet based modeling is applied to control an ammonia stripper which is basically an aqueous ammonia binary distillation column of a fertilizer plant in Gresik, East Java, Indonesia. The results show how the proposed control scheme has satisfactorily tracking capability as well as performance with respect to changes in process dynamics.

1. INTRODUCTION

Generalized Predictive Control (GPC), which is considered as universal method for model-based predictive control, is proven to be successful in handling various kind of processes and has also been successfully applied in various industries. GPC can be used either to control a simple plant with little prior knowledge or a complex plant such as non-minimum phase, open-loop unstable and a process having variable dead-time (see e.g. Clark *et al.*, 1987, Garcia *et al.*, 1989, Rawlings, 2000). A very critical step toward the success of the implementation of GPC is the availability of a reliable process model as an accurate plant model is necessary to derive a set of future plant output close to its corresponding reference signal sequence. As most processes in industry have nonlinear behavior, then the modeling process is even more difficult (Nazaruddin and Sudarto, 2007).

Nowadays, the application of neural network can be found in many areas, such as signal processing, pattern recognition and system identification (Pham and Liu, 1995). The main advantage of the neural network is its capability for the self-learning from the previous data. The neural network can be used as function for nonlinear tool fitting to develop a model from input-output data or to provide mapping between input and output space. However, there are some difficulties in implementing the neural network i.e. in the mathematical analysis and construction methods where the network representation is not single so that its use is inefficient, network convergence is not guaranteed, or the training procedure depends on the initial condition, and possibility to be trapped into local minima. Besides, the setting of the network structure (number of layers and neuron in each layer) is difficult due to unavailability of supporting theory and uneasiness in the network initialization which usually randomly chosen. The shortcoming of neural network approach has initiated an alternative idea to develop a new approach which is an integration between wavelet theory and neural network (wavelet network or wavenet) where the activation function sigmoid is replaced by wavelet basis function localized in space and frequency domain (Zhang and Benveniste, 1992). Wavenet itself is a feedforward neural network using wavelet basis function as an activation function where the wavenet parameters namely dilation, translation and weighting are optimized during the learning phase and the gradient method is used for the optimization of the parameters.

In practice, to determine the number of neuron (wavelet in network) and the network initialization becomes the major problems. Good initialization of wavenet is of important to obtain faster convergence of the algorithm (Oussar and Dreyfus, 2000). This is the different and the advantages of the wavenet compared to neural network. Moreover, the task to determine the number of wavelet in the network is critical as it will reduce the model order. Here, the Final Prediction Error (FPE) method will be applied.

In this paper, an alternative control strategy for nonlinear processes which is an integration between the generalized predictive control and wavelet network (wavenet) approach will be presented. The modeling will be performed based on a set of input-output data obtained from real-time measurements of an ammonia stripper which is basically an aqueous ammonia binary distillation column of a fertilizer plant in Gresik, East Java, Indonesia. The control strategy is then performed using the identified model.

2. PREDICTIVE CONTROL LAW

Suppose that a future set-point or reference sequence $[w(t+j); j=1,2,\cdots]$ is available. In most cases $w(t+j)$ will be constant w equal to current set-point $w(t)$, though sometimes (as batch process control or robotic) future variations in $w(t+j)$ would be known. The objective of the predictive control law is then to derive the future plant outputs $y(t+j)$ close to $w(t+j)$ in some sense, bearing in mind that the control activity required to do so. This is done using a receding-horizon approach for which at each sample-instant t :

(1). the future set-point sequence $w(t+j)$ is calculated;

(2). the prediction model is used to generate a set of predicted outputs $\hat{y}(t+j\mid t)$ with corresponding predicted errors $e(t+j)=w(t+j)-\hat{y}(t+j\mid t)$, by noting that $\hat{y}(t+j\mid t)$ for $j>k$ depends in part on the future control signals $u(t+j\mid t)$ which are to be determined;

(3). an appropriate quadratic function of the future error and control is minimized, provide a suggested sequence of future controls $u(t+j\mid t)$;

The quadratic function, also known as the cost function, is defined as

$$J=\sum_{j=N_1}^{N_2}\left[\hat{y}(t+j)-w(t+j)\right]^2+\sum_{J=1}^{N_U}\lambda(j)\left[\Delta u(t+j-1)\right]^2 \tag{1}$$

where N_2, N_1, N_u and $\lambda(j)$ is the horizon maximum $(1\le N_2)$, the horizon minimum $(1\le N_1\le N_2)$, the control horizon $(1\le N_u\le N_2)$ and a control-weighting sequence, respectively.

3. WAVENET BASED MODELING

A. Identification Using Wavenet

In the last few years, several methods have been elaborated for the identification of linear/nonlinear systems using neural network approach. The identification procedure is based on a set of input-output data of experiments so that a sequence of training data (x^n, y_m^n) can be developed, where $x^n=\{x_1^n, x_2^n,\ldots,x_{Ni}^n\}$ is input and y_m^n is output vector. The main problem of identification is to obtain the relation between previous data input $[x^{t-1}, y^{t-1}]$ and the output $y(t)$, which can be written as

$$y(t)=f(x^{t-1},y^{t-1})+e(t) \tag{2}$$

where $e(t)$ is noise signal which means that the output $y(t)$ is not exact with the previous output. Assuming that the function f can be written as

$$f(x^{t-1},y^{t-1},\theta)=f(\varphi(t),\theta) \tag{3}$$

where

$$\varphi(t)=[y(t-1)\ \cdots\ y(t-n_a)\ x(t-1)\ \cdots\ x(t-n_b)]^T \tag{4}$$

with $d=n_a+n_b$ and $\varphi(t)\in R^d$. Using a sequence of observation data as training data, which is written as

$$Z^N=\{y(t),\varphi(t),k=1,2,\ldots N\} \tag{5}$$

then the following relation can be established

$$\hat{y}(t)=\hat{f}_N(\varphi(t),\theta) \tag{6}$$

where $\hat{f}_N$ is an unknown nonlinear function. In fact, the model $\hat{f}_N(\varphi(t))$ can be determined using several methods. As a possible approximation, the nonlinear function $\hat{f}_N$ can be written as an expansion of basis function, which takes the form

$$\hat{f}_N(\varphi(t),\theta)=\sum_{k=1}^{K}w_k h_k(\varphi) \tag{7}$$

where h_k is a basis function. An example of local basis function is a wavelet basis function. Also, nonlinear function $\hat{f}_N$ will be determined using wavenet (wavelet network) approach, where θ is wavenet parameter

B. *Wavelet Function and Network* (*Wavenet*)

The term wavelet means a little wave, which has a minimum oscillation and a fast decay to zero, in both the positive and negative directions, of its amplitude. This property is analogous to an admissibility condition of a function that is required for the wavelet transform. To approximate a signal, then sets of wavelets are used. The objective is then to find a set of daughter wavelets, which are generated by a dilated (scaled or compressed) and translated (shifted) original wavelets or mother wavelets.

The wavelet theory has an associated transform, i.e. it is an operation that transforms a function by integrating it with modified version of some kernel function called mother wavelet and the modified version is called daughter wavelet. A function which performs as mother wavelet must be admissible. For a given function $h(t)$, the admissibility condition for wavelet function is

$$\int_{-\infty}^{+\infty}h(t)dt=0 \tag{8}$$

The wavelet transform of a function $f\in L^2(R)$ with respect to a given admissible mother wavelet $h(t)$ is defined as

$$W_f(a,b)=\int_{-\infty}^{+\infty}f(t)h^*{}_{a,b}(t)dt \tag{9}$$

where * denotes the complex conjugate, although most wavelets are real. The daughter wavelets are constructed from single mother wavelet *h(t)* by dilation and translation, or

$$h_{a,b}(t)=\frac{1}{\sqrt{a}}h\left(\frac{t-b}{a}\right) \tag{10}$$

where $a>0$ and b is the dilation and translation factor, respectively.

Wavelet function can be classified in 2 categories, i.e. orthogonal wavelet and wavelet frame. An orthogonal wavelet is a wavelet where the associated wavelet transform is orthogonal. Wavelet frame is constructed from simple operation of translation and dilation of mother wavelet function. Efficient nonlinear identification can be implemented if the wavelet family constitutes a frame (Zhang, 1997). A wavelet function of type POLYnomials

WindOwed with Gaussian (POLYWOG) is applied in this work, which is a derivative of a Gaussian function and satisfies the admissible condition of mother wavelet.

C. Determination of the Number of Wavelet

To obtain a wavenet model of minimal order is an important task. For this purpose, a part of the model data is used to approximate the model order. There are several methods which are commonly used, and one of an ordinary method to determine the number of wavelet is Akaike Final Prediction Error (FPE), which has been implemented (Nazaruddin and Yuliati, 2006). FPE will minimize the following equation

$$J_{FPE}(\hat{f}) = \frac{1 + n_d / n}{1 - n_d / n} \frac{1}{2n} \sum_{n=1}^{n} \left(\hat{f}(x_n) - y_n \right)^2 \quad (11)$$

where n_d is the number of parameter in the estimator and d is input dimension. For wavelet network, then the number of parameters is

$$n_d = K(d+2) \quad (12)$$

where K is the number of wavelet in the network. In (12), number of parameters in the network is weight, dilation and translation.

D. Wavenet Algorithm

To approximate an arbitrarily nonlinear function, there are two processes in the algorithm for neural network adaptive wavelet namely network construction and error minimization. A wavelet network (wavenet) as a class of feedforward neural network with wavelet as an activation function was introduced in (Zhang and Benveniste, 1992).

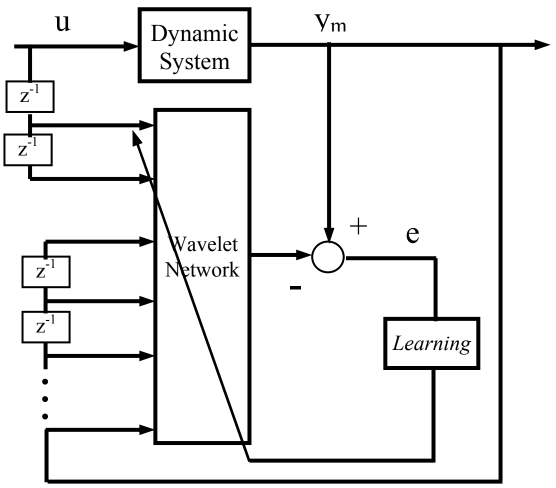


Fig. 1. Serial-parallel wavenet model for system identification.

It seems that there is similarity between the inverse of discrete wavelet transformation with a single hidden layer in the neural network. Using Nonlinear AutoRegressive with eXogenous (NARX) structure as a model approximation, then the configuration of the network which is used for systems identification is wavenet serial-parallel model, which structure is shown in Fig. 1.

The wavenet architecture approximates a signal $y(t)$ with linear combination of a set of daughter wavelet (10) which is formed with dilation a and translation b of the mother wavelet. Assuming that the output function of the network satisfies the admissibility condition and the network could approximate the target, i.e. time-frequency region which is effectively covered by their K windows, then the approximate signal of the network $\hat{y}(t)$ can be written as

$$\hat{y}(t) = u(t) \sum_{k=1}^{K} w_k h_{a_k, b_k}(t) \quad (13)$$

with K is the number of wavelet and w_k is the weight coefficient. Fig. 2 illustrates the structure of adaptive wavelet network (Lekutai, 1997).

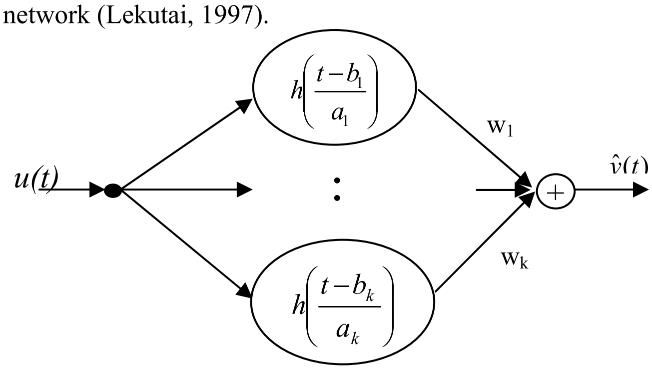


Fig. 2. Structure of adaptive wavelet network

The neural network parameters w_k, b_k, a_k can be optimized using Least Mean Square (LMS) by minimizing the cost function as a function of energy at time t, with

$$e(t) = y(t) - \hat{y}(t) \quad (14)$$

denotes the error as time variant function, where $y(t)$ is the expected output/target. The objective function is defined as

$$E = \frac{1}{2} \sum_{t=1}^{T} e^2(t) \quad (15)$$

and to minimize E, the steepest descent method is applied, which needs the gradient $\frac{\partial E}{\partial w_k}, \frac{\partial E}{\partial a_k}, \frac{\partial E}{\partial b_k}$ to update the parameters w_k, b_k, a_k for each wavelet, where the gradient of E to each network parameter is given by equations

$$\frac{\partial E}{\partial w_k} = -\sum_{t=1}^{T} e(t) h(\tau) \quad (16)$$

$$\frac{\partial E}{\partial b_k} = -\sum_{t=1}^{T} e(t) w_k \frac{\partial h(\tau)}{\partial b_k} \quad (17)$$

$$\frac{\partial E}{\partial a_k} = -\sum_{t=1}^{T} e(t) w_k \tau \frac{\partial h(\tau)}{\partial b_k} = \tau \frac{\partial E}{\partial b_k} \quad (18)$$

where

$$\tau = \frac{t - b_k}{a_k} \quad (19)$$

Consequently, each coefficient w, b, a of the network is updated using the following relation

$$\underline{w}(n+1) = \underline{w}(n) + \mu_w \Delta \underline{w} \quad (20)$$

$$\underline{b}(n+1) = \underline{b}(n) + \mu_b \Delta \underline{b} \quad (21)$$

$$\underline{a}(n+1) = \underline{a}(n) + \mu_a \Delta \underline{a} \quad (22)$$

where μ is the learning rate.

4. FORMULATION OF CONTROL SCHEME

The main idea of the proposed control algorithm is to apply predictive control strategy using nonlinear process model, which is obtained from wavenet based modeling approach. Modeling process is done at each control action, and consequently this allows the control to be done adaptively. The basic structure of the proposed control scheme can be seen in Figure 3. The whole algorithm can be broken up into two main steps, modelling and control. These two steps are explained below.

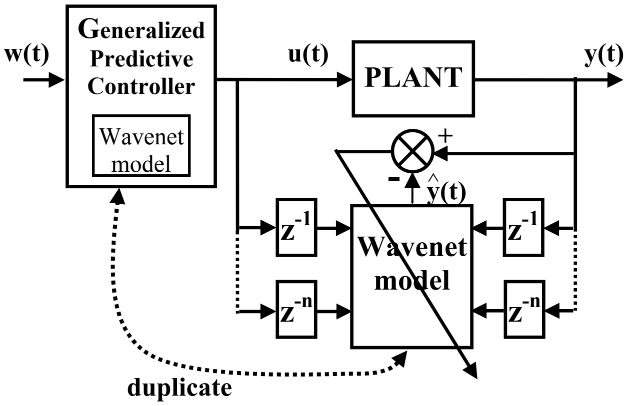


Fig. 3. The generalized predictive adaptive control scheme with wavenet based modelling

The process model will be represented in a NARX model which is in the form

$$y(t) = F\big(y(t-1), \cdots, y(t-n), u(t-d-1), \cdots, u(t-d-n)\big) + e(t) \quad (23)$$

Here *y(t)* and *u(t)* are the sampled process output and input at time instant *t* respectively, *e(t)* is the equation error, *n* denotes the order of the process, *d* represents the process dead time as an integer number of samples and *F(.)* is an unknown nonlinear function to be identified. With modeling process using wavenet approach, the function *F(.)* is obtained, so that the following relation is obtained

$$\hat{y}(t) = F\big(y(t-1), \cdots, y(t-n), u(t-1), \cdots\big) \quad (24)$$

where $\hat{y}$ denotes the output prediction of wavenet based model. The model obtained is further used to predict the process outputs $\hat{y}(t+N_1)$,..., $\hat{y}(t+N_2)$ when any sets of inputs $u(t),...,u(t+N_U)$ is applied to process, where N_1, N_2, N_U are minimum, maximum and control horizon of the generalized predictive control scheme. Next, the task is to find a set of control signals that minimizes the cost function in eq. (1). This can be done by giving any arbitrary control signal to the process, then using the steepest descent method, the control signal will be calculated iteratively to find desired control signal. Mathematically it can be written as

$$u(t)_{i+1} = u(t)_i - \eta_u \frac{\partial^+ J}{\partial u(t)} \quad (25)$$

where η_u is the learning rate. Then, using chain rule combined with *backpropagation error* method, the derivative of cost function $\frac{\partial^+ J}{\partial u(t)}$ is calculated.

5. IMPLEMENTATION AND RESULTS

5.1 Plant Description

Ammonia stripper, a subsystem of ammonia plant plays an important role in the ammonia production, especially when the energy conservation is important. Basically, ammonia stripper is a binary distillation column. It separates the feed flow, namely aqueous ammonia into two products, ammonia (distillate) and water (bottom product). The feed flow comes from an ammonia scrubber, which has 14.1 %wt ammonia. For energy conservation reason, its bottom product is recycled back to the ammonia scrubber (Kellog,1992). However, there is a limitation that the maximum ammonia content in the bottom product shall be 0.14 %wt. Otherwise, significant problem in the next process, namely hydrogen recovery unit, will occur. From this point of view, the success of separation in the ammonia stripper plays an important role in the ammonia production.

The ammonia stripper consists of a re-boiler, a condenser and a reflux accumulator. Presently, due to the oversized valve and other dynamic factors, the ammonia stripper is difficult to be controlled especially during start-up operation (IKPT, 1992) Further, a multivariable distillation column presents a number of challenging problem both for system identification and control due to the nonlinear and ill-conditioned nature. These two characteristics cause difficulties in identification and control design of the distillation column. For that reasons, this work is concentrated on how to implement the adaptive predictive control of the ammonia stripper based on nonlinear identification technique, i.e. using wavenet based modeling.

In the fertilizer plant, two inferential measurements of product concentrations, namely temperature measurement for top product composition, and a conductivity measurement for its bottom composition, have been installed. Because of any limitation and much complexity in the plant operation, data for system identification have been obtained from a number of single-input multi-output experiments. Only two control loops were considered for the implementation of control strategy including system identification, i.e. reflux loop and steam loop, as can be seen in Fig. 4. Due to the high risk of its plant operation, the reflux loop experiment has been done in a closed-loop manner, meanwhile another one in an open-loop condition.

In this investigation, implementation of control strategy is performed by assuming that the system is represented by *Single Input Multi Output* (SIMO) model. Data for plant modeling was obtained from valve opening of reflux flow as manipulated variable, and top product temperature as well as bottom conductivity as output. Meanwhile, steam flow to the boiler remained constant at 815Kg/H $\pm$ 0.5%. Data records were perfomed using sampling time of 1 minute due to the available facilities in the existing Distributed Control System

(DCS) in the plant. Because the control valve available in the system was calibrated and tested properly, it could be assumed that the control output value was directly related to the % valve opening position.

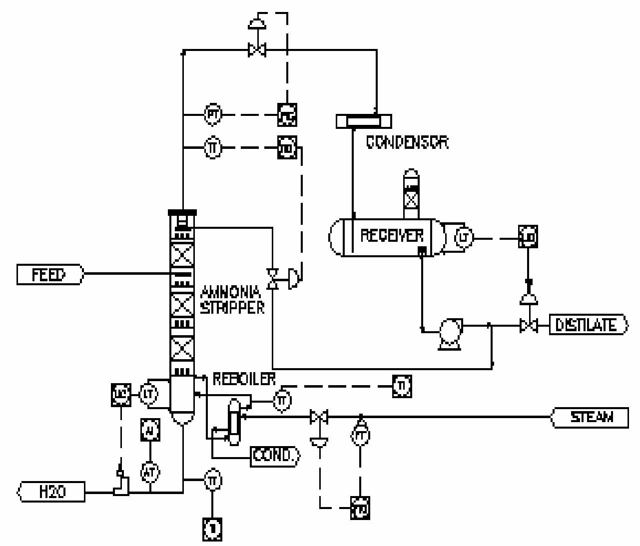


Fig. 4. Plant schematic

In the experiment, 1000 data have been collected for each corresponding inputs and outputs. The first 500 data will be used for identification, and the rest for model validation. Prior to the modeling, those experiments data have been analyzed to verify any linear or nonlinear relationship and possibility of system time delay. Coherence test was used for every pair of input-output data and the results showed that almost all input-output relations indicated nonlinear behaviour.

5.2 Identification Results and Validation

For validation of the obtained models, beside a direct obeservation by comparing the plot of measured signals and the output of model based on the estimated parameters, a quantitative criteria of Root Mean Square Error (RMSE), defined as

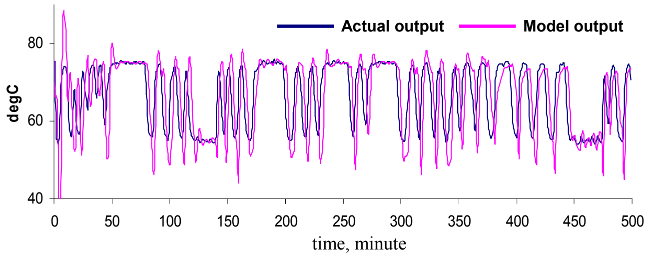
$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y(t)-w(t))^2} \tag{25}$$

where n is the number of data, was used for a measure of a good model fitting

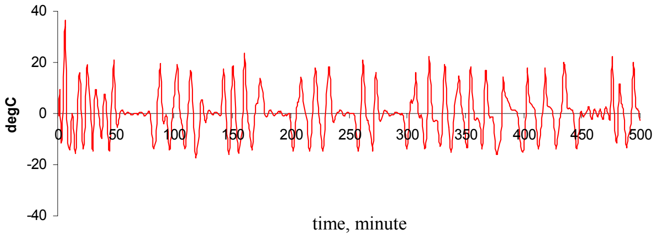
For training process and modeling purpose, the equation of the estimator / identifier has been chosen as

$$\hat{y} = \Psi(y_p(t-1), y_p(t-2), x(t-1), x(t-2), \theta) \tag{26}$$

where x is a input network and θ is the wavenet parameter i.e. weight, dilation, translation. The learning rate constant for each parameter network was chosen as 0.5 and the number of wavelet was equal to 11. The results of modeling using the first 500 data applying the obtained identified model for Top Product Temperature as a function of Reflux Flow, and its corresponding error, which is the different between identified model output and measured signal, using wavenet method are shown in Fig. 5. The result of RMSE value from the modeling was 8.491.

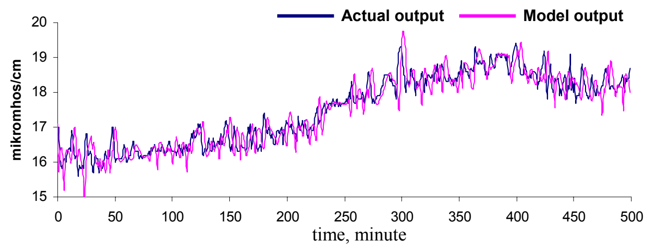


(a).

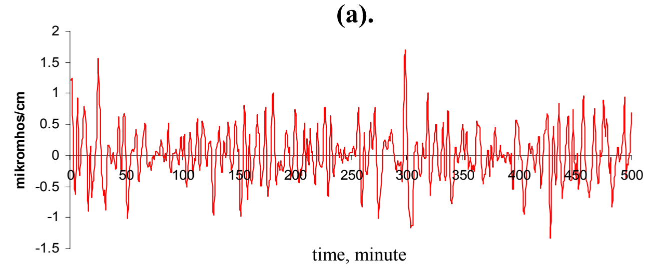


(b).

Fig. 5. (a). Validation result of the identified model using wavenet (Top Product Temperature vs. Reflux Flow) and (b). its corresponding error



(a).



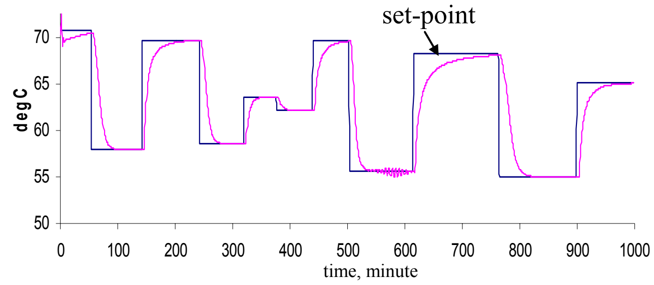
(b).

Fig. 6. (a). Validation result of the identified model using wavenet (Bottom Conductivity vs. Reflux Flow) and (b). its corresponding error

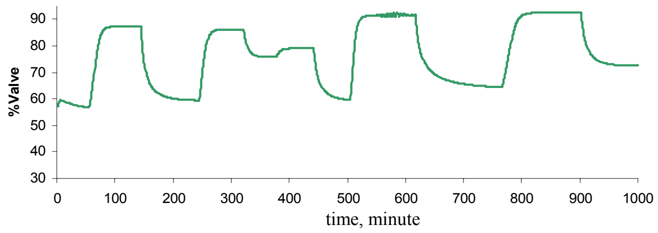
For the case of modeling the bottom conductivity as a function of valve opening of reflux flow, the number of wavelet was equal to 10. The results of modeling using the first 500 data applying the obtained identified model, and its corresponding error, is shown in Fig. 6. with the RMSE value was 0.406. Above results revealed that satisfactory model matching were obtained which means that the models have captured the real basic features of the nonlinear dynamics of the ammonia stripper unit

5.2. *Results of Control*

Experiments were conducted to observe the performance of the proposed control scheme with respect to the set-point changes and to plant dynamics changes, especially for the Top Product Temperature loop. Fig. 7 shows the results of control implementation and its corresponding control signal. The output temperature of the ammonia stripper can track the set-point quite satisfactorily with RMSE value of 3.34.



(a).

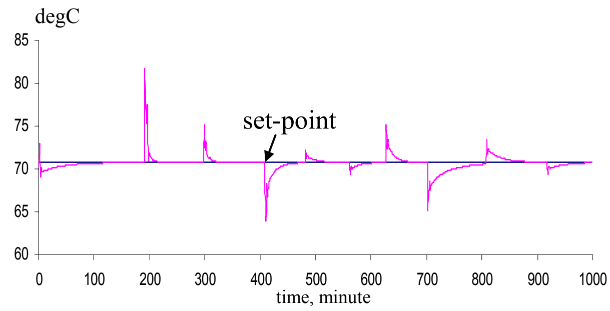


(b).

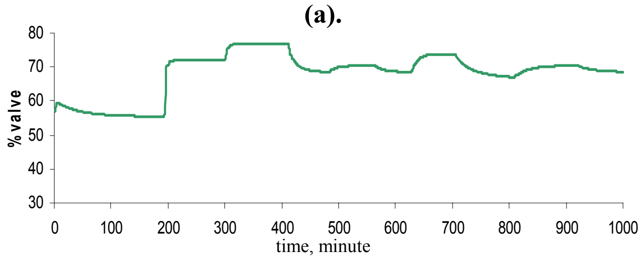
Fig. 7. (a). Response of top product temperature to the set-point changes and (b). its corresponding control signal (% opening of the valve) using the adaptive predictive controller

Parameters of predictive adaptive controller were set after extensive trial and error procedure, and give optimum results as follows :

- Minimum horizon = N_1 =1
- Maximum horizon = N_2 = 10
- Control horizon = N_u = 1
- Control-weighting signal = = λ = 0.00001



(a).



(b).

Fig. 8. (a). Response of top product temperature plant to sudden changes of plant process dynamics and (b). its corresponding control signal (% opening of the valve) using the adaptive predictive controller

Objective of the next experiment was to observe the adaptive capability of the control strategy. For this purposes, several changes of process dynamic were applied to the process with a fixed set-point of top product temperature which was 71°C. This was done by changing the parameters of plant model from 5 to 20 percent. Fig. 8 shows the results of control implementation and its corresponding control signal. As can be observed, the control scheme is capable to tracks the set-points after several changes of process dynamics.

6. CONCLUSIONS

Satisfactory control performance has been shown by the proposed adaptive predictive control strategy using wavenet based modeling, especially its application to the ammonia stripper, which is basically an aqueous ammonia binary distillation column of a fertilizer plant. This plant has nonlinear characteristics. Further application of the proposed control strategy using various methods of identification, especially for nonlinear processes in the ammonia plant is currently under study.

REFERENCES

Clark, D., C. Mohtadi, and P. Toffs (1987). Generalized predictive control. Part I: The basic algorithm, *Automatica,* **Vol. 23**, No. 1, pp. 137-148

Garcia, C.E., D.M. Prett, and M. Morari (1989). Model Predictive Control - Theory and Practice - A Survey. *Automatica*, **Vol. 25**, No. 3, pp. 335-348

Rawlings, J. (2000). Tutorial overview of model predictive control *IEEE Control Systems Magazine,* **Vol. 20**, No. 1, pp. 38-52

Nazaruddin, Y.Y. and Sudarto (2007), Neuro-Fuzzy and Wavelet Network (Wavenet) Based Modeling of an Ammonia Converter Unit, *Proceed. the 4th Int. Conference on Computational Intelligence, Robotics and Autonomous System (CIRAS'2007)*, Palmerston North, New Zealand

Pham, D.T. and X. Liu (1995). *Neural Networks for Identification, Prediction and Control*, Springer Verlag, London, 1995

Zhang, Q., and A. Benveniste. (1992). Wavelet networks, *IEEE Trans. on Neural Networks*, 3(6), pp. 889-898

Oussar, Y. and G. Dreyfus (2000). Initialization by selection for wavelet network training, *Neurocomputing*, **Vol. 34**, 2000, pp. 131-143

Zhang, Q. (1997). Using wavelet network in nonparametric estimate, *IEEE Trans. on Neural Networks*, 8(2), pp. 227-236

Nazaruddin, Y.Y. and Yuliati (2006). Wavenet Based Modeling of Vehicle Suspension System, *Proceed. of 32nd Annual Conference of the IEEE Industrial Electronics Society*, Paris, France, 2006

Lekutai, G. (1997). *Adaptive Self-Tuning Neuro Wavelet Network Controllers*, Ph.D. dissertation, Virginia Polytechnic Institute and State University, Virginia

Kellog, MW. (1992). *Operating Instruction Manual For 1350 MTPD Ammonia Unit*, PT. Petrokimia Gresik.

IKPT, PT. (1992). *Specification of Air Operated Control Valve for 1350 MTPD Ammonia*, PT. Petrokimia Gresik