International Journal of Computer Science & Applications © 2007 Technomathematics Research Foundation

Vol. 4 Issue 3, pp 119-133

Using Artificial Neural networks for the modelling Of a distillation column

Yahya CHETOUANI
Yahya.Chetouani@univ-rouen.fr

Université de Rouen, Département Génie Chimique, Rue Lavoisier, 76821 Mont Saint Aignan Cedex, France.

Abstract

The main aim of this paper is to establish a reliable model both for the steady-state and unsteady-state regimes of a nonlinear process. The use of this model should reflect the true behavior of the process under its normal operating conditions and allow distinguishing a normal mode from an abnormal one. In order to obtain this reliable model for the process dynamics, the neural black-box identification by means of a NARMAX model has been chosen in this study. The modelling study shows the choice and the performance of the neural network in the training and test phases. Also an analysis of the inputs choice, time delay, hidden neurons and their influence on the behavior of the neural estimator is carried out. Three statistical criteria are used for the validation of the experimental data. The model is implemented by training a Multi-Layer Perceptron Artificial Neural Network (MLP-ANN) with input-output experimental data. After describing the system architecture a realistic and complex application as a distillation column is presented in order to illustrate the proposed ideas concerning the dynamics modelling. Satisfactory agreement between identified and experimental data is found and results show that the neural model successfully predicts the evolution of the product composition.

Keywords: Reliability, modelling, neural network, NARMAX, distillation column

1 Introduction

Process development and continuous request for productivity led to an increasing complexity of industrial units. In chemical industries, it is absolutely necessary to control the process and any drift or anomaly must be detected as soon as possible in order to prevent risks and accidents [1]. The intrinsic highly nonlinear behavior in the industrial process, especially when a chemical reaction is used, poses a major problem for the formulation of good predictions and the design of reliable control systems [2]. Due to the relevant number of degree of freedom, to the nonlinear coupling of different phenomena and to the processes complexity, the mathematical modelling of the process is computationally heavy and may produce an unsatisfactory correspondence between experimental and simulated data. Similar problems arise also from the uncertainty for the parameters of the process, such as the reaction rate, activation energy, reaction enthalpy, heat transfer coefficient, and their unpredictable variations. In fact, note that most of the chemical and thermo-physical variables both strongly depend and influence instantaneously the temperature of the reaction mass [3, 4].

One way of addressing this problem is the use of a reliable model for the on-line prediction of the system dynamic evolution. There are essentially two approaches by which nonlinear models can be developed for a distillation column; from first principles by using the process knowledge or empirically from input/output data. The advantages and disadvantages of each approach are well known. In industrial practice, it is not always possible in general to obtain accurate first principles models for high-purity distillation columns. Most industrial columns are used to separate multi-component mixtures whose constituent elements are often not known completely; the fundamental thermodynamics of multi-component vapor-liquid equilibria, the physical property data, and other essential constitutive relations required for the successful development of a

first principles model are not always available. And even when such knowledge is available, the resulting models usually occur in the form of a very large system of coupled nonlinear ordinary differential equations, and may therefore not always be the most convenient for controller or fault detection (FD) design. On the other hand, when fundamental process knowledge is unavailable or incomplete, or when the resulting model may not be particularly suitable for FD applications, the input/output models identified from plant data may be more useful. Engell et al. [5] used a semi-batch reactive distillation process. A comparison was carried out between conventional control structures and model-based predictive control by using a neural net plant model. Nanayakkara et al. [6] presented a novel neural network to control an ammonia refrigerant evaporator. The objective is to control evaporator heat flow rate and secondary fluid outlet temperature while keeping the degree of refrigerant superheat at the evaporator outlet by manipulating refrigerant and evaporator secondary fluid flow rates. Owing to their inherent nature to model and learn 'complexities', ANNs have found wide applications in various areas of chemical engineering and related fields [7, 8].

The purpose of this study is to obtain a powerful model of reference allowing to reproduce the dynamics of a complex process as a distillation column. This reliable model enables to reproduce the process dynamics under different operating conditions (steady-state or/and unsteady-state). The present study focuses on the development, and implementation of a NARMAX neural model for the forecasting of the distillation column dynamics. Experiments were performed in a distillation column and experimental data were used both to define and to validate the model. The performance of this neural model was then evaluated using the performance criteria. Results show that the NARMAX neural model is representative for the dynamic behavior of this nonlinear process. The modelling procedure, the experimental set-up and prediction results are described in the following sections.

2 Modelling by input-output approach

In chemical systems, parameter variations and uncertainty play a fundamental role on the system dynamics and are very difficult to be accurately modeled [2]. Modelling strategies of various kinds by means of input-output measurements are commonly used in many situations in which it is not necessary to achieve a deep mathematical knowledge of the system under study, but it is sufficient to predict the system evolution [9, 10]. This is often the case in control applications, where satisfactory predictions of the system that are to be controlled and sufficient robustness to parameter uncertainty are the only requirements.

To model as a stochastic process, Auto-Regressive Moving Average (ARMA) models should be used [10, 11]. In earlier studies, it was difficult to implement ARMA modelling in real-time because of limitations in the speed of computers. Other studies [12, 13] have demonstrated that ARMA modelling, if equipped with an appropriate recursive estimator, can perform real time machining error modelling and control with good accuracy and a minimal computational load. However, the above studies were based on linear systems, which were represented by time-invariant or slowly changing time-varying difference equations. It was found by Chen et al. [14] that a wide class of nonlinear systems was best represented by a polynomial NARMAX model, and an application of a nonlinear identification methodology to an industrial process was reported in [15].

In order to provide a closer approximation to the dynamic behavior of the distillation process is some situations, a nonlinear NARMAX model [16, 17] is employed in this study, which is identified by means of Artificial Neural Networks. The NARMAX model was obtained by using a MLP-ANN [14, 18] to describe accurately the process behavior. The nonlinear model of a finite dimensional system [19] with order (n_y, n_u, n_b) and scalar variables y, u and b are defined by:

$$y(k) = \Phi(y(k-1),...,y(k-n_y),\ u(k-n_u)_{1\le n_u\le m},\ b(k-1),...,b(k-n_b)) \qquad (1)$$

where *y(k)* is the Auto-Regressive (AR) variable or system output; *u(k)* is the eXogenous (X) variable or system input; *b(k)* is the Moving Average (MA) variable or white noise i.e. $E(b(t)) = 0$ and $E(b(t)b(t-k)) = d_k s^2$. d_k is the Kronecker delta defined, as follows $d_k = 0$ if $k \neq 0$ and $d_k = 1$ if $k = 0$. s^2 represents the process variance. n_y, n_u and n_b are the AR, MA and X orders, respectively. n_u and m are

respectively the time delay and the maximum time delay. Φ is a neural nonlinear function.

2.1. Structure of the ANN used in this study

For engineering purposes, the neural network can be thought of as a black box model which accepts inputs, processes them and produces outputs according to some nonlinear transfer function [20]. The feed-forward MLP-ANN is used in this study and consists in a large number of highly connected nonlinear simple neurons. The structure is based on a result by Cybenko [21] who proved that a neural network with one hidden layer of sigmoid or hyperbolic tangent units and an output layer of linear units is capable of approximating any continuous function. In this study, we used the tan–sigmoid transfer function on the hidden layer and a linear transfer function on the output layer. The input data are propagated through the network to the output layer to obtain the network output. In this study, a back-propagation training function for feed-forward networks using momentum and adaptive learning rate techniques is used [22, 23]. This version of back-propagation learns quickly compared to classical algorithm and minimizes the chance that the network parameters will become stuck in a high error minimum

2.2 Calculation of the NN output

Though the applicability of neural networks to solve several nonlinear complex problems has been amply demonstrated, the time taken to train neural networks off-line can be quite excessive [10]. The following steps explain the calculation of the ANNs output based on the input vector.

1. Assign $\hat{w}^T(k)$ to the input vector $x^{\mathrm{T}}(k)$ and apply it to the input units where $\hat{w}^T(k)$ is the regression vector given by the following equation:

$$\hat{w}^T(t) = [y(t-1),...,y(t-n_y),\quad u(t-1-n_k),...,u(t-n_u-n_k)] \tag{2}$$

2. Calculate the input to the hidden layer units:

$$net_j^h(k) = \sum_{i=1}^{p} W_{ji}^h(k)x_i(k) + b_j^h \tag{3}$$

where p is the number of input nodes of the network, i.e. $p = n_y + n_u + n_b$; j is the jth hidden unit; W_{ji}^h is the connection weight between ith input unit and jth hidden unit; b_j^h is the bias term of the jth hidden unit.
3. Calculate the output from a node in the hidden layer:

$$z_j = f_j^h(net_j^h(k)) \tag{4}$$

where f_j^h is the tan-sigmoid transfer function.
4. Calculate the input to the output nodes:

$$net_l^q(k) = \sum_{j=1}^{h} W_{lj}^q(k)z_j(k) \tag{5}$$

where l is the lth output unit; $W_{lj}^q(k)$ is the connection weight between jth hidden unit and lth output unit.
5. Calculate the outputs from the output nodes:

$$\hat{v}_l(k) = f_l^q(net_l^q(k)) \tag{6}$$

where f_l^q is the linear activation function defined by:

$$f_l^q(net_l^q(k)) = net_l^q(k) \tag{7}$$

2.3 Back-propagation training algorithm

The error function E is defined as:

$$E = \frac{1}{2}\sum_{l=1}^{q}(v_l(k) - \hat{v}_l(k))^2 \tag{8}$$

where q is the number of output units and $v_l(k)$ is the lth element of the output vector of the network. Within each time interval from k to k+1, the back-propagation (BP) algorithm tries to minimize the error for the output value as defined by E by adjusting the weights of the network connections, i.e. W_{ji}^h and W_{lj}^q. The BP algorithm uses the following procedure (Eqs. 9, 10, 11, 12):

$$W_{ji}^h(k+1) = W_{ji}^h(k) + a\,DW_{ji}^h(k) - h\frac{\partial E}{\partial W_{ji}^h(k)} \tag{9}$$

$$W_{lj}^q(k+1) = W_{lj}^q(k) + a\,DW_{lj}^q(k) - h\frac{\partial E}{\partial W_{lj}^q(k)} \tag{10}$$

where ? and a are the learning rate and the momentum factor, respectively; ? W_{ji}^h and ? W_{lj}^q are the amounts of the previous weight changes; $?E/?W_{ji}^h(k)$ and $?E/?W_{lj}^q(k)$ are given by:

$$\frac{\partial E}{\partial W_{ji}^h(k)} = -\left[z_j(k)(1 - z_j(k))x_i(k)\right]\sum_{l=1}^{q}\left[(v_l(k) - \hat{v}_l(k))\hat{v}_l(k)W_{lj}^h(k)\right] \tag{11}$$

$$\frac{\partial E}{\partial W_{lj}^q(k)} = -(v_l(k) - \hat{v}_l(k))z_j(k) \tag{12}$$

The implementation of the ANN for forecasting is as follows:

1. Initialize the weights using small random values and set the learning rate and momentum factor for the ANN.
2. Apply the input vector given by Eq. 2 to the input units.
3. Calculate the forecast value of the error using the data available at (k- 1)th sample (Eqs. 2, 3, 4, 5, 6, 7).
4. Calculate the error between the forecast value and the measured value.
5. Propagate the error backwards to update the weights (Eqs. 9, 10, 11, 12).
6. Go back to step 2.

For weights initialization, the Nguyen-widrow initialization method [24] is best suited for the use with the sigmoid/linear network which is often used for function approximation.

3 Results and discussion

3.1 Experimental device: distillation column

The feed tank (fig. 1) contains a mixture to be separated (Toluene-Methylcyclohexane) with a mass composition at 23 % in methylcyclohexane. The product is introduced through the optimal feed tray so that the light components are volatilized, while the heavy part goes down again with the reflux in the column reboiler. Feed preheating system is constituted by three elements of 250 w each one. In addition it has a low liquid level switch in order to avoid the running if the level is excessively low. The reciprocating feed pump is constituted by a membrane allowing firstly the suction of the mixture and the discharge towards the tank with a flow capacity F = 4.32 $L.h^{-1}$. The column has also a reboiler of 2 liters hold-up capacity, an immersion

heater of a power Q_b = 3.3 kw and of a level liquid switch sensor which allows the automatic stop of heating if the level is insufficient. The stirring of the mixture in the reboiler is ensured by the boiling mixture. The internal packing is made of Multiknit stainless 316L which enhances the mass transfer between the vapor and liquid phases. In order to approach the adiabatic conditions, a heat-insulating made of glass wool is laid around the column. A condenser is placed at the column overhead in order to condense the entire vapor coming out from the column. The cooling medium used in exchangers is water. The heat-transfer area of the total overhead condenser is 0.08 m^2. Moreover the reflux timer (R_t) is constituted by an electromagnetic valve and allows to control the reflux ratio (R_r). It is monitored by the overhead product temperature (T_d). When the required distillate temperature (T_d) is attained, the reflux timer opens. In the opposite case, it remains closed. Supervision control system allows to modify the parameters and to follow their evolution such as the pressure drop (*DP*), the flow or the temperatures at different points of the distillation column. The unit has twelve sensors which measure continuously the temperature throughout the column.

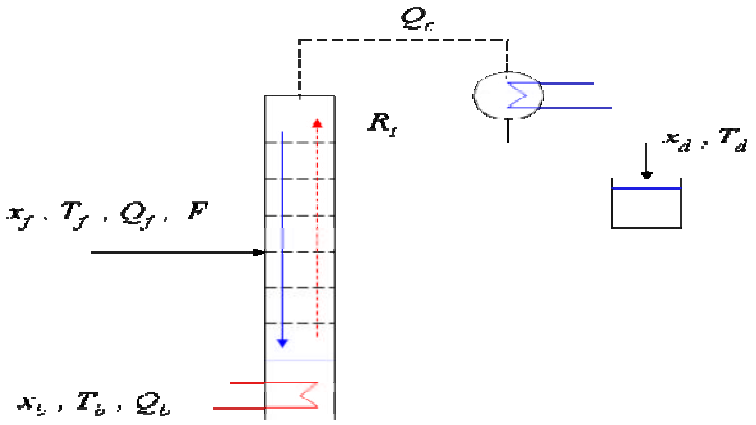

Figure 1: Experimental device: Distillation column.

3.2 Relative importance–determination of the time delay n_u

The neural network weight matrix can be used to assess the relative importance of the different input variables upon the output variables of the process [7]. Garson proposed an equation based on partitioning of connection weights [25]. In the present study, another strategy of the determination of the relevant input variables is defined. In order to determine the relevance of the input variables and the time delay n_u between each input variable and the output variables; two main parameters for a modelling which takes the physical nature of the process and the interactions between variables into account, each input variable is modified and its influence on the output variables is observed. It has to be bear in mind that this analysis of the relative importance is based on a long-term experimental work but it takes the physical nature into account as well. For this particular reason, the distillation operates in the temperature range ($T_{d\,min}$ = 101.3°C, $T_{d\,max}$ =103.5°C) i.e. ($x_{d\,min}$ = 0.49, $x_{d\,max}$ = 0.84) from the isobar diagram of the toluene-methylcyclohexane by using the Wilson thermo-dynamic model for vapor-liquid equilibrium (VLE). This interval is selected in order to obtain a mixture rich in methylcyclohexane. Moreover the bubble and dew temperatures will have to be rather far from each other in order to facilitate the separation process. The set of the most important measurable variables of the process are (R_t ,Q_b ,*DP* ,Q_f ,T_f , F ,Q_c) as input and (T_d ,T_b) as input.
It is important to notice that only three temperature measurements amongst all were retained. Amongst the remaining variables, an experimental analysis of the process is carried out in order to observe the influence of each input variable on the output variables. This analysis aims at reducing more the system. This study is determined for a nominal steady-state regime of the column defined in fig. 2.

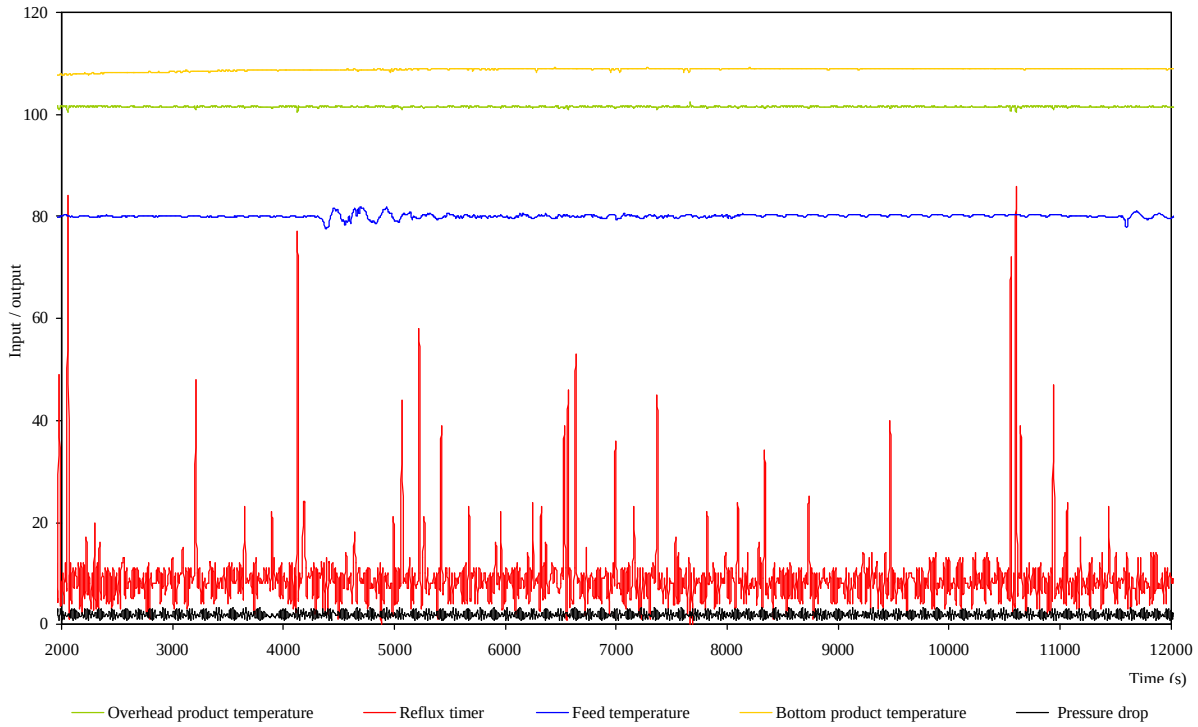

Figure 2: Nominal steady-state behavior of the studied column.

In order to define this nominal mode, all the regulation systems of the column are put in a closed-loop configuration (R_t, Q_b, DP, Q_f, T_d, T_b, …). When the steady-state regime is achieved, the column was operated for approximately four hours in order to collect a sufficient number of data on the steady-state behavior (T_d = 102 °C). The data is taken every 11 s. Table 1 gives the values of the measurable variables found on average from the nominal steady-state regime. The analysis of the relative importance is based on a long-term and an accurate experimental work. As a result of that the analysis realized for the study of the reflux timer and the heating power are shown in this paper. The other analyses were defined in the same way i.e. only one input variable is modified while the other input variables are maintained constants in order to observe the evolution of the output variables due to this particular input variation.

R_t	Q_b	T_f	T_d	F	Q_c
14 %	45 %	80 °C	102 °C	50 %	250 L.h^{-1}

Table 1: The operating conditions of the nominal steady-state regime.

3.2.1 Influence of the reflux timer

The reflux timer allows to route part of the vapor condensed in the column, therefore the distillate temperature (T_d) decreases as the reflux increases. It also allows to recover the distillate, which causes a temperature (T_d) increase. In this experiment, it is important to highlight the importance the opening percentage of the reflux timer, in other words, the reflux ratio is modified from a minimal value to a maximum one in order to reveal clearly in fig. 3 the influence of this input variable. At the same time, the mean time delay (if it exists) is calculated between this input variable and the output variables of the system.

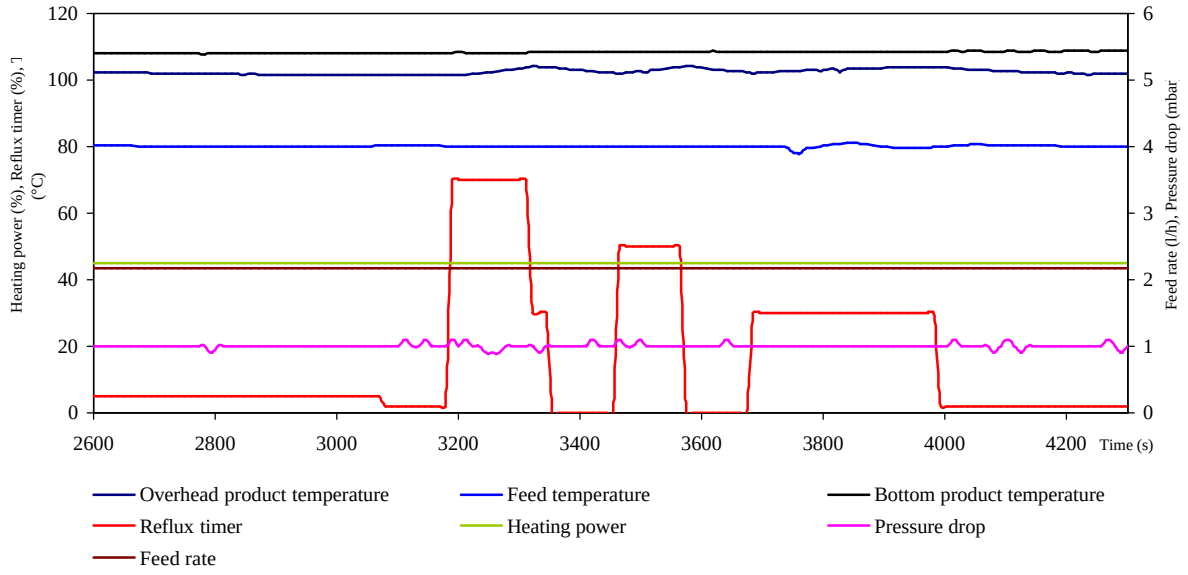

Figure 3: Relative importance of the reflux timer

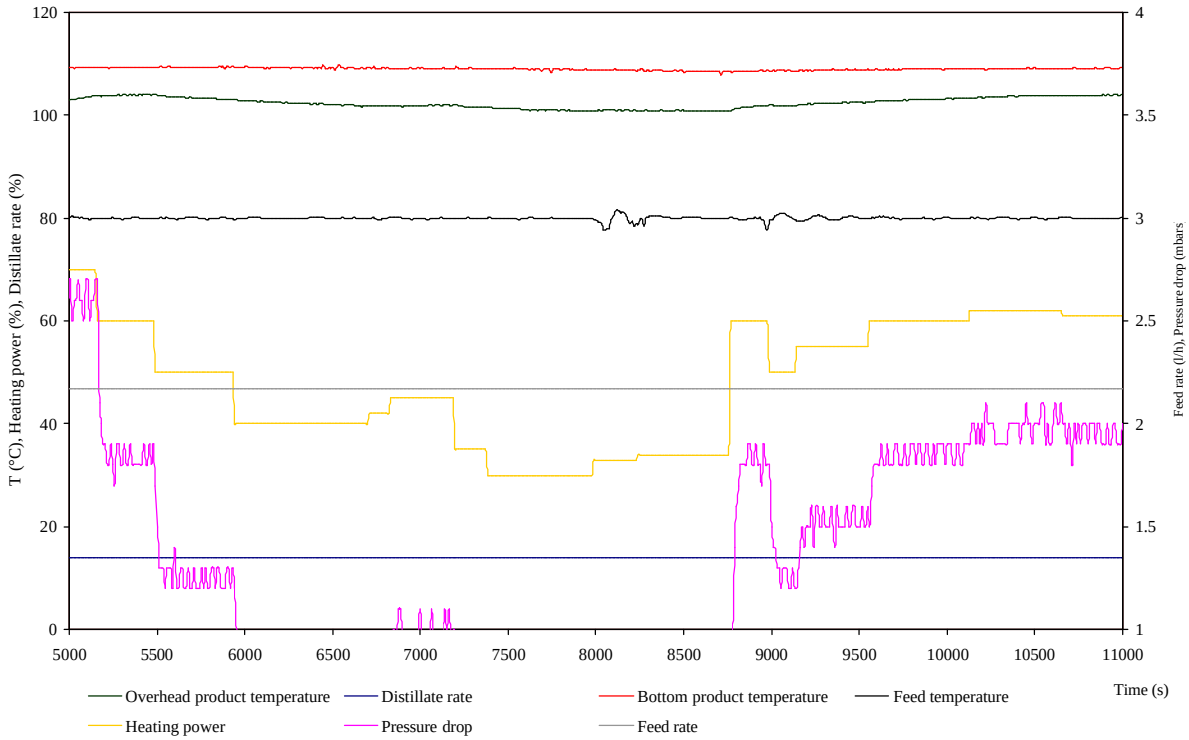

Figure 4: Relative importance of the heating power.

It is important to notice that the reflux timer has a very important influence on the distillate temperature (T_d) and quality. When the opening percentage of the reflux timer is null, all the vapors are condensed at the column overhead and routed back to the top tray. In this case, the vapor is very rich of the most volatile product, which drives the vapor temperature to decrease. Conversely, if the reflux timer opens that causes an impoverishment of the vapor of the most volatile product. The toluene being in excess compared to the methylcyclohexane, the vapor temperature increases. It is also important to point out that if R_t is excessively low, the level in the reboiler increases fairly quickly because most of the vapors are returned in the column and the feed rate (F) which remains constant sends product permanently. In the other case if R_t is too high, too

much distillate is recovered but the mixture quantity vaporized will be more important than the constant feed flowrate.

3.2.2 Influence of the heating power

Two pressure sensors are placed at the bottom and the top of the column, a sensor transmitter analyzes the difference of pressure in the column and sends a signal to the controller which processed it and gives the value of this difference of pressure in mbar. The heating power (Q_b) is determined according to this difference of pressure (DP). Indeed the power input of the reboiler influences the vapor flow in the column. The more vapor flows, more important frictions in the garnishing are. For this experiment, when the steady-state regime is achieved with the values given in the table 1, the heating duty input is changed by maintaining the other measurable variables constants. In fig. 4, we observe quite obviously that the variation of the heating power (Q_b) is reflected on the pressure drop (DP) across the column. We also observe the same variation of the overhead temperature (T_d) i.e. when Q_b increases, T_d increases. In fact, the increase of the vapor flow causes a faster impoverishment of the reboiler in methylcyclohexane (the most volatile product), hence the vapor temperature approaches the bubble temperature of toluene (the least volatile product). Because the opening of the reflux timer remains constant (reflux ratio is constant), this reflux timer does not correct the variation of the overhead temperature. Consequently, T_d increases. This phenomenon is also perceptible through a decrease of the heating power, the temperature and the flow of the vapor decrease in the column, therefore the temperature T_d decreases. On the other hand, it is noticeable that the variation of the heating power does not influence the preheating temperature because they are not physically connected together. It is also shown that the bottom column temperature is always constant and equal to the temperature of the least volatile component.

3.2.3. The influence of all input variables-experimental results

In conclusion, it is noted that during these experiments the most important input variable is the reflux ratio and the heating power. Table 2 collects the mean time delay determined between each input variable and the output variable. The lower time delay is, the higher the relative importance is, because if a variation of the input variable occurs the output variable would be disturbed afterwards a time delay. On another side, it is important to highlight that the time delay between the feed flow rate and the distillate temperature is 88 s. The feed flow rate has an influence on (T_d), however the response of this temperature is very low to a feed variation. Consequently, the feed flow rate has less influence on (T_d) than the other input variables. When a time delay is called infinite (represented by “8”), that means that there is no influence of the input variable upon the output variables.

Variables	R_t	Q_b	DP	Q_f	T_f	F	Q_c
T_d	22	22	33	44	66	88	8

Table 2: Mean time delay (s) between u and y.

3.3 Pattern generation

The choice of the training and test database aims at modelling the dynamic behavior of the output variables such as the overhead product temperature (T_d) according to the input variables of the processsuch as DP, R_t, Q_f, T_f and F. In order to have this database which is rich in amplitudes and in frequencies, the column behavior is modified in the temperature range (T_d = 101.5 ? T_d = 103.5 °C). The verification test subset is a set of independent data used to verify the consistency of the efficiency of the model. Figs. 5 and 6 show this database. The operation duration of the distillation column is 13 hours continuously. Fig. 6 shows the evolution of DP, R_t, Q_f and Q_b between (1650 s ? 3000 s) for more legibility.

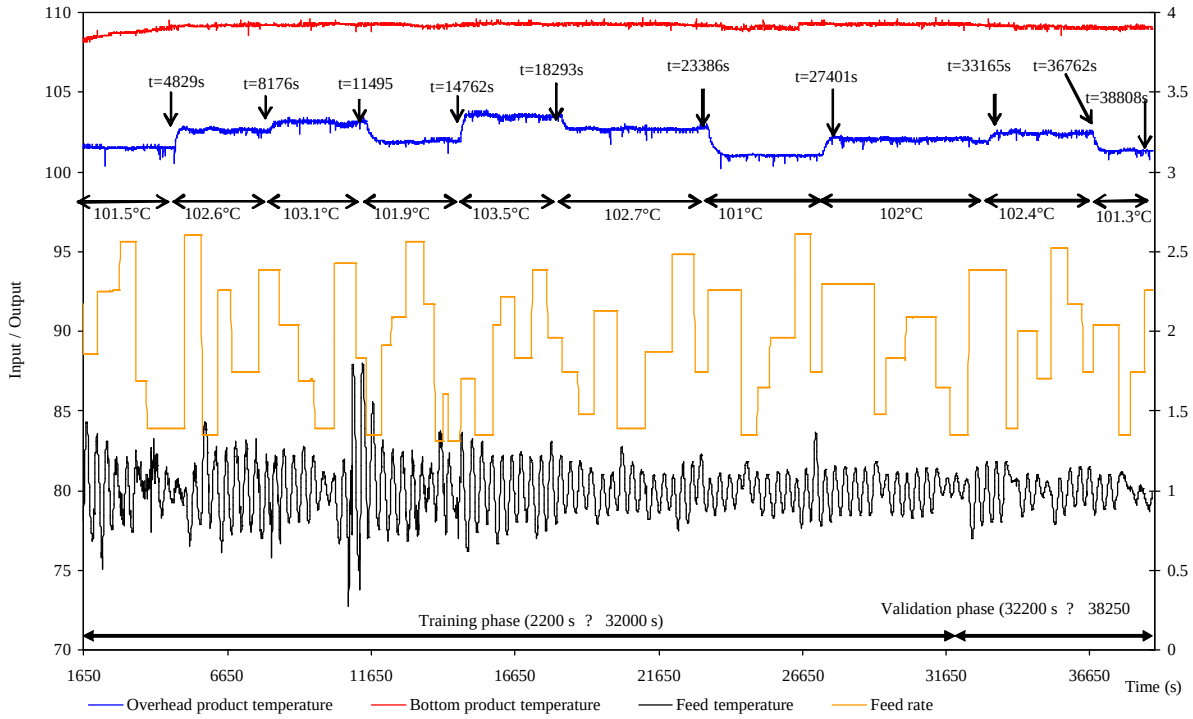

Figure 5: Training and test database.

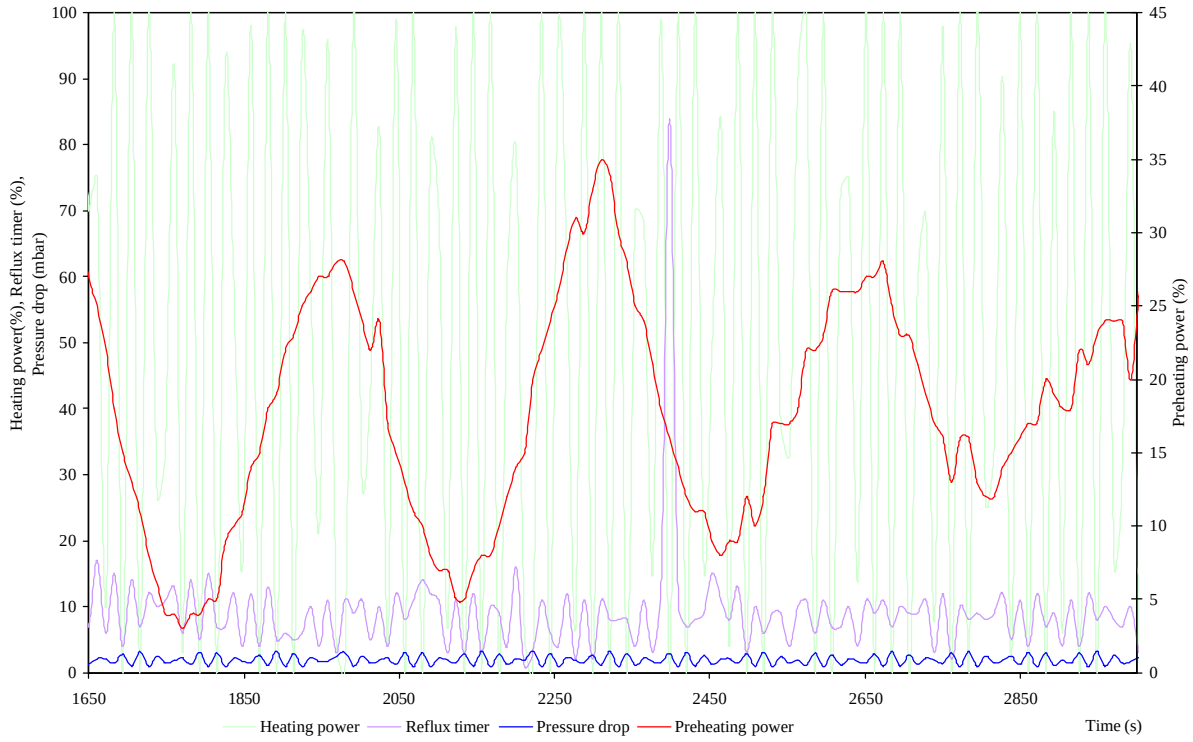

Figure 6: Training and test database.

3.4 Training scheme and net development

The validation error will normally decrease during the initial phase of training, as does the training set error.

However, when the network begins to overfit the data, the error on the validation set will typically begin to rise, an early stopping procedure to stop the learning process was used for improving generalization, and the weights at the minimum of validation error are returned. Also, the accurate number of hidden neurons cannot be achieved from a universal formula. It is determined by the user and can vary from zero to any finite number. Networks with too many parameters tend to memorize the input patterns, while those with too few hidden parameters may not be able to simulate a complex system at all.

In this study the initial model had few parameters and the hidden neurons are gradually added during learning until the optimal result is achieved in the test subset. To establish a suitable NARMAX model order for the studied distillation column, neural networks of increasing model order can be trained and their performance on the training data compared using the loss function (or mean squared error), *LF*. This function is expressed by the following equation:

$$LF = \frac{1}{N}\sum_{i=1}^{N} e^2(t) \tag{13}$$

where $e(t)=y(t)-\hat{y}(t)$ represents the prediction error and N is the data length. In order to select the optimal number of hidden neurons, tests were performed by varying the number of neurons between 5 and 25 for the temperature (T_d) modelling.

3.4.1 Modelling of the temperature T_d

In this case, the equation 1 becomes according to the table 2.

$$T_d(t) = f(T_d(t-1), R_t(t-2), \Delta P(t-3), Q_f(t-4), T_f(t-6), F(t-8), T_d(t-1) - \hat{T}_d(t-1)) \tag{14}$$

where T_d and $\hat{T}_d$ represent the actual and predicted distillate temperatures. The evolution of the loss function is given in fig. 7 for the training and the test database. This figure shows that the neurons number in the hidden layer in the test phase is 11 neurons. The same result is found for the training phase.

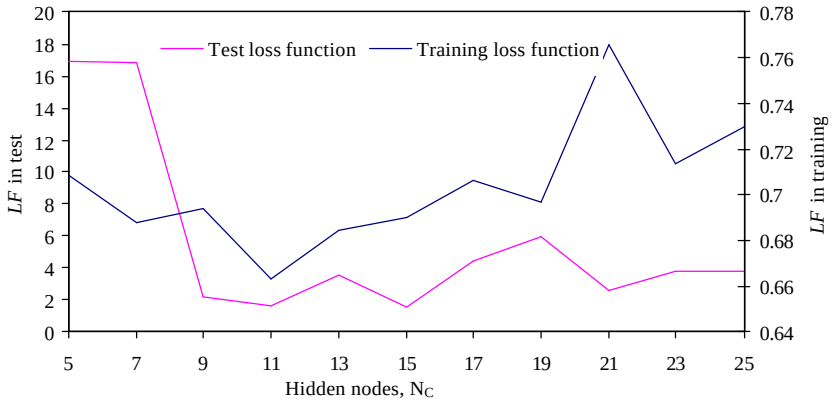

Figure 7: Loss functions.

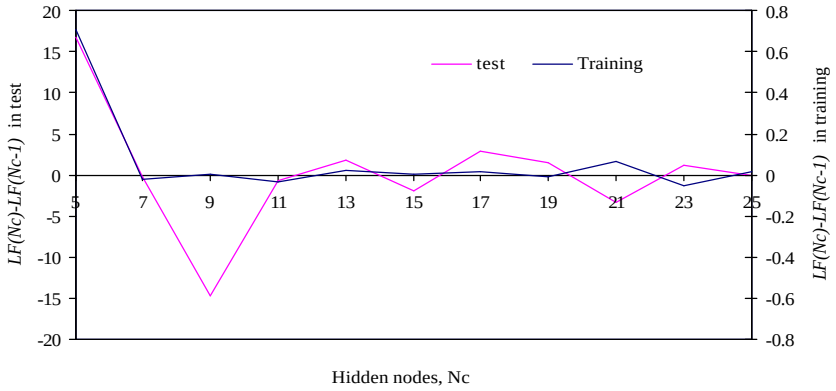

Figure 8: Trends in loss functions.

The decision can be assisted by examining the trend in the *LF* (fig. 8). This figure indicates that a number of hidden neurons *Nc* = 9 causes a large reduction in the *LF* compared to others *Nc* (13, 15). But this choice is not being retained because this number is too high. The model which has a structure (6-11-1) exhibits the acceptable *LF*. But this model may not be the best choice, because there is a trade-off between the model complexity (i.e. size) and accuracy. A small decrease in the *LF* may be rejected if it is at the expense of enlarging the model size. Thus, the decision procedure for selecting a parsimonious model using the *LF* consists in deciding for each increase in model order whether any reductions in the *LF* are worth the expense of a larger model. The difficult trade-off between model accuracy and complexity can be clarified by using model parsimony indices from linear estimation theory [26], such as Aikeke's Information Criterion (AIC), Rissanen's Minimum Description Length (MDL) and Bayesian Information Criterion (BIC). The validation phase thus makes it possible to distinguish the model describing correctly the dynamic behavior of the process. These statistical criteria are defined as follows:

$$AIC = ln\left(\frac{N}{2}LF\right) + \frac{2n_w}{N} \tag{15}$$

$$MDL = ln\left(\frac{N}{2}LF\right) + \frac{2n_w ln(N)}{N} \tag{16}$$

$$BIC = ln\left(\frac{N}{2}LF\right) + \frac{n_w ln(N)}{N} \tag{17}$$

where n_w is the number of model parameters (weights in a neural network).
Hence, the AIC, MDL and BIC are weighted functions of the *LF* for the validation data set, which penalize for reductions in the prediction errors at the expense of increasing model complexity (i.e. model order and number of parameters). Strict application of these statistical criteria means that the model structure with the minimum AIC, MDL or BIC is selected as a parsimonious structure. However, in practice, engineering judgment may need to be exercised. Fig. 9 shows the evolution of AIC, MDL and BIC criteria according the *LF* in the validation database.

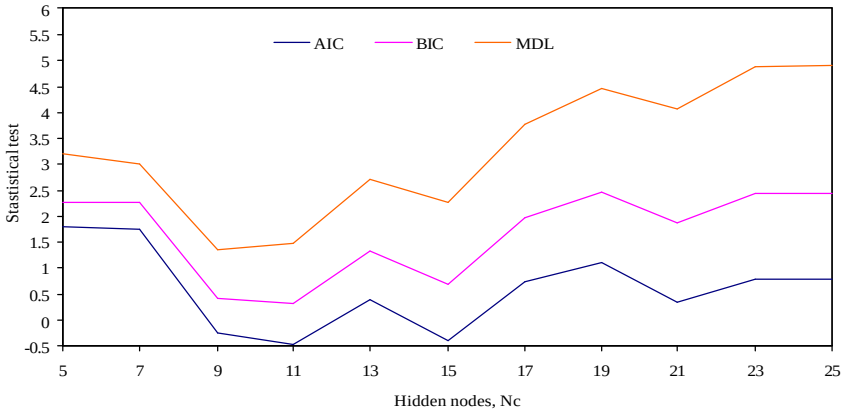

Figure 9: Evolution of the statistical criteria.

A strict application of the indices would select the model 6-11-1 because it exhibits the lowest of three indices for all the model structures compared.

3.4.2 Residual analysis

Once the training and the test of the NARMAX model have been completed, it should be ready to simulate the system dynamics. Model validation tests should be performed to validate the identified model. Billings et al. [27] proposed some correlations based model validity tests. In order to validate the identified model, it is necessary to evaluate the properties of the errors that affect the prediction of the outputs of the model, which can be defined as the differences between experimental and simulated time series. In general, the characteristics of the error are considered satisfactory when the error behaves as white noise, i.e. it has a zero mean and is not correlated [2, 27]. In fact, if both these conditions are satisfied, it means that the identified model has captured the deterministic part of the system dynamics, which is therefore accurately modeled. To this aim, it is necessary to verify that the auto-correlation function of the normalized error $e(t)$, namely $fee(t)$, assumes the values 1 for t=0 and 0 elsewhere; in other words, it is required that the function behaves as an impulse. This auto-correlation is defined as follows [27, 28]:

$$fee(t)=E(e(t-t)e(t)]=d(t) \quad \forall t, \tag{18}$$

where e is the model residual. $E(X)$ is the expected value of X, t is the lag.
This condition is, of course, ideal and in practice it is sufficient to verify that $fee(t)$, remains in a confidence band usually fixed at the 95%, which means that $fee(t)$ must remain inside the range $\pm\frac{1.96}{\sqrt{N}}$, with N the number of testing data on which $fee(t)$ is calculated. Billings et al. [27] also proposed tests for looking into the cross-correlation among model residuals and inputs. This cross-correlation is defined by the following equation:

$$fue(t)=E(u(t-t)e(t)]=0 \quad \forall t, \tag{19}$$

To implement these tests (18, 19), u and e are normalized to give zero mean sequences of unit variance. The sampled cross-validation function between two such data sequences $u(t)$ and $e(t)$ is then calculated as:

$$fue(t)=\frac{\sum_{t=1}^{N-t}u(t)e(t+t)}{\left[\sum_{t=1}^{N}u^2(t)\sum_{t=1}^{N}e^2(t)\right]^{1/2}} \tag{20}$$

If the equations (18, 19) are satisfied then the model residuals are a random sequence and are not predictable

from inputs and, hence, the model will be considered as adequate. These correlations based tests are used here to validate the neural network model. The results are presented in fig. 10. In these plots, the dash dot lines are the 95% confidence bands.

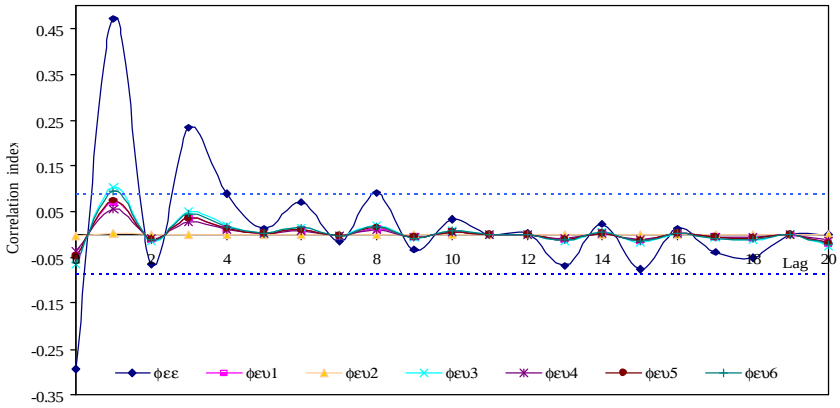

Figure 10: Correlation test results for a 6-11-1.

The evolution of the correlation of the NARMAX model is inside the 95% confidence bands (fig. 10). Few points are outside this confidence interval. In addition, the NARMAX cross-correlation is low. This explains the independence of the residual signal from the input one. Therefore, this model is considered a reliable one for describing the temperature (T_d) profile. Fig. 11 shows the difference between the experimental overhead product temperature and those simulated by the neural model.

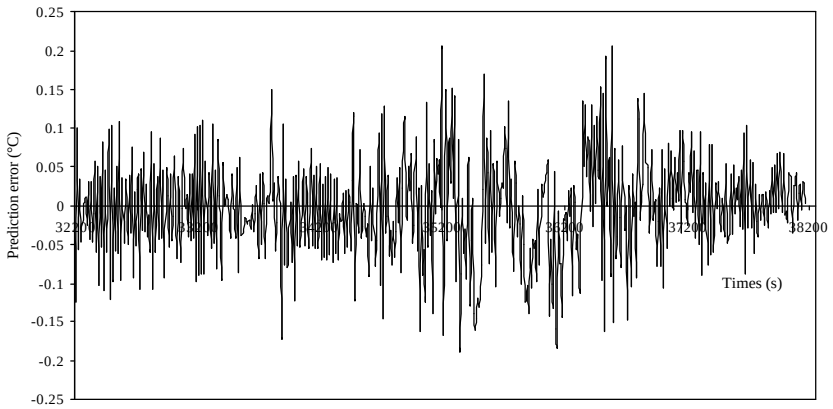

Figure 11: Prediction error of the temperature (T_d).

After analyzing this figure, it emerges that the NARMAX model 6-11-1 for modelling the temperatures (T_d) ensures satisfactory performances as it is indeed able to correctly identify the dynamics of the distillation column. Therefore, the found neural model is considered a reliable one for describing the dynamic behavior of this column. This validation phase is used with the neural weights found in the training phase. There is a good agreement between the learned neural model and the experiment in the validation phase. This result is important because it shows the ability of the neural network with only one hidden layer to interpolate any nonlinear function [21]. In conclusion the main advantage of the neural approach consists in the natural ability of neural networks in modelling nonlinear dynamics in a fast and simple way and in the possibility to address the process to be modeled as an input-output black-box, with little or no mathematical information on the system.

4 Conclusion

This work aims to identify process dynamics by means of a NARMAX model. The identification of the

system dynamics by means of input-output experimental measurements provides a useful solution for the formulation of a reliable model. In this case, the results showed that the model is able to give satisfactory descriptions of the experimental data. The developed neural model is used in a recursive scheme in order to test their ability to perform the behavior prediction. Finally, the identified neural model will be useful for the detection and the isolation (FDI) of faults which can occur through the process dynamics.

References

[1] Chetouani, Y., Fault detection in a chemical reactor by using the standardized innovation, *Process Safety and Environmental Protection*, 84:27-32, 2006.
[2] Cammarata, L., Fichera, A., and Pagano, A., Neural prediction of combustion instability, *Applied Energy*, 72:513-528, 2002.
[3] Chetouani, Y., Modelling and prediction of the dynamic behaviour in a reactor-exchanger using NARMAX neural structure, *Chemical Engineering Communications Journal*, 194:691-705, 2007.
[4] Chetouani, Y., Use of Cumulative Sum (CUSUM) test for detecting abrupt changes in the proccess dynamics, *International Journal of Reliability, Quality and Safety Engineering*, 14:65-80, 2007.
[5] Engell, S. and Fernholz, G., Control of a reactive separation process, *Chemical Engineering and Processing*, 42:201-210, 2003.
[6] Nanayakkara, V.K., Ikegami, Y., and Uehara, H., Evolutionary design of dynamic neural networks for evaporator control, *International Journal of Refrigeration*, 25:813-826, 2002.
[7] Sharma, R., Singh, K., Singhal, D., and Ghosh, R., Neural network applications for detecting process faults in packed towers, *Chemical Engineering and Processing*, 43:841-847, 2004.
[8] Himmelblau, D.M., Applications of artificial neural networks in chemical engineering, *Korean J. Chem. Eng.*, 17:373–392, 2000.
[9] Mu, J., Rees, D., and Liu, G.P., Advanced controller design for aircraft gas turbine engines, *Control Engineering Practice*, 13:1001-1015, 2005.
[10] Fung, E.H.K., Wong, Y.K., Ho, H.F., and Mignolet M.P., Modelling and prediction of machining errors using ARMAX and NARMAX structures, *Applied Mathematical Modelling*, 27:611-627, 2003.
[11] Wu, S.M. and Ni, J., New approaches to achieve better machine performance, *Proc. of the USA–Japan Symposium on Flexible Automation*, 1063–1068, 1988.
[12] Fung, E.H.K. and Chung, A.P.L., Using ARMA models to forecast workpiece roundness in a turning operation, *Applied Mathematical Modelling*, 23:567–585, 1998.
[13] Teja, S.R. and Jayasingh, T., Characterisation of ground surface profiles––a comparison of AR, MA, ARMA modelling methods, *International Journal of Machine Tools and Manufacture*, 33:103–109, 1993.
[14] Chen, S. and Billings, S.A., Representation of nonlinear systems—The NARMAX model, *International Journal of Control*, 49:1013–1032, 1989.
[15] Thomson, M., Schooling S.P., and Soufian, M., The practical application of a nonlinear identification methodology, *Control Engineering Practice*, 4:295–306, 1996.
[16] Lamedica, R., Prudenzi, A., Sforna, M., Caciotta, M., and Cencellli, V.O., A neural network based technique for short-term forecasting of anomalous load periods, *IEEE Transactions on Power Systems*, 11:1749–1756, 1996.
[17] McMenamin, J.S. and Monforte, F.A., Short-term energy forecasting with neural networks, *Energy Journal*, 19:43–61, 1998.
[18] Narendra, K.S. and Parthasarathy, K., Identification and control of dynamical systems using neural networks, *IEEE Trans. on Neural Networks*, 1:4–21, 1990.
[19] Ljung, L., *System Identification, Theory for the User*, Prentice Hall, Englewood Cliffs, New Jersey, 1999.
[20] Zaknich, A., *Neural networks for intelligent signal processing*, Singapore, World Scientific, 2003.
[21] Cybenko, G., Approximation by superpositions of a sigmoidal function, *Math. Control Signals Systems*, 4:303–312, 1989.
[22] Rumelhart, D.E., Hinton, G.E., and Williams, R.J., Learning representations by back-propagating errors, *Nature*, 323:533–536, 1986.
[23] Hertz, J., Krogh, A., and Palmer, R.G., *Introduction to the theory of neural computation*, Addison-Wesley, Redwood City, CA, 1991.
[24] Nguyen, D. and Widrow, B., Improving the learning speed of 2-layer neural networks by choosing initial

values of the adaptive weights, *International Joint Conference of neural networks*, 3: 21-26, 1990.
[25] Garson, G.D., Interpreting neural-network connections weights, *AI Expert*, 6:47-51, 1991.
[26] Ljung, L., *System identification toolbox user's guides*, The Math Works, Natick, 2000.
[27] Billings, S.A. and Voon, W.S.F., Correlation based model validity tests for nonlinear models, *International journal of Control*, 44:235–244, 1986.
[28] Zhang, J. and Morris, J., Process modelling and fault diagnosis using fuzzy neural networks, *Fuzzy Sets and Systems*, 79:127-140, 1996.