Support vector regression and model reference adaptive control for optimum control of nonlinear drying process

C. KARTHIK, K. VALARMATHI, AND M. RAJALAKSHMI

ABSTRACT: In this paper, a support vector regression (SVR)-based system identification and model reference adaptive control (MRAC) strategy for stable nonlinear process input-output form is designed. In order to implement the proposed control structure, SVR-based identification methods are clearly addressed. The control of a moisture process on the paper machine illustrates the proposed design procedure and the properties of the SVR-based model identification-adaptive reference model for the nonlinear system. MRAC is widely used in linear system control areas, and neural networks (NN) are often used to extend MRAC to nonlinear areas. Some drawbacks of NN with MRAC are slow speed in learning, weak generalization ability, and a local minima tendency. To overcome this problem, SVR is used instead of NN. With the support vector regressor, a stable controller-parameter adjustment mechanism is constructed by using the model reference adaptive theory. Simulation results show that the proposed approach could reach desired performance.

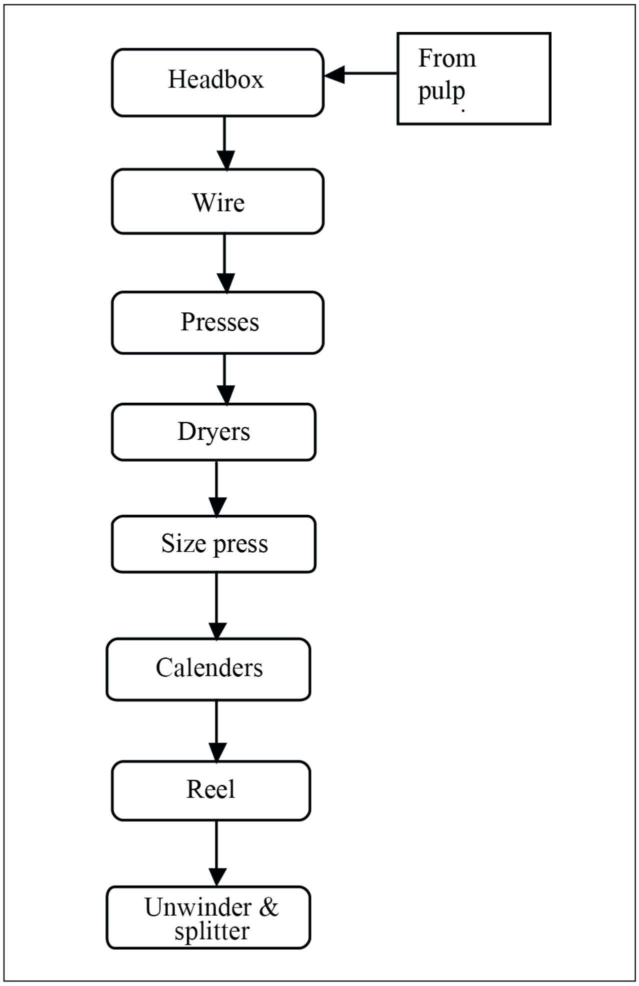
Application: Paper mills that use the SVR identified MRAC for proper reduction of moisture can achieve less power consumption, more productivity, and improved end-product quality.

Nonlinear system identification is a very important study in the control field. Nonlinear system identification has many difficulties compared to a linear system. Among the different nonlinear identification techniques, methods based on the neural network (NN) model are gradually becoming established — not only in academia but also in industrial applications. However, such methods have problems of slow convergence speed and local minima. In the last ten years, NN models and, more recently, support vector regression (SVR) have appeared as two attractive tools for modeling the system, particularly in situations where the development of phenomenological or conventional regression models becomes impractical or not easily manageable. In recent years, SVR [1-3] has provided a statistical learning theory that is preferable to NN due to its many attractive characteristics and their underlying empirical performance. The major difference between NN and SVR lie in the risk minimization principle. The NN applies the empirical risk minimization (ERM) standard to minimize error in the training data, while SVR follows structural risk minimization (SRM). In SVR, a top bound on the generalization error is minimized. Hence, it opposed to the ERM, which minimizes the prediction mistake of the training data. This equips the SVR to generalize the linear relationships. During its training stage, SVR is able to understand the relationship for making good predictions for the new input data set. Although the basis of the SVR paradigm was established in the mid-1990s, its chemical technology applications, such as obvious error detection [4,5], have appeared recently. The quality of SVR models primarily depends on a proper setting of support vector machine (SVM) parameters. The main issue is how to set these parameters for a given training and testing data set to ensure good generalization performance.

MOISTURE PROCESS

A typical machine for producing paper or paperboard is about the length of two football fields. The run time speed can be up to 2000 m/min. The machine consists of seven sections of production units. The flow of production of paper from pulping consists of three major sections: wire, press, and dryer (**Fig. 1**).

The first section of the paper machine is called the wet end. This is where the diluted pulp stock first comes into contact with the paper machine and is poured onto the machine from the headbox. A narrow aperture running across the width of the box allows the stock to flow onto the wire section with the fibers distributed evenly over the whole width of the paper machine. The machine is operated by a central control system, which monitors the paper process for various properties such as moisture content, basis weight, color, etc., and provides necessary adjustments. NN can be used effectively for the identification and control of nonlinear dynamical systems [6-11].



1. Paper manufacturing process flow.

Dryer section

The paper leaves the press section with a dry content of up to 50%–55%. Additional water is removed by vaporization in the press section. In the dryer section, the most common type of paper drying is contact drying on cylinders heated with vapor. Here, the heat energy is transferred from the outside walls of the drying cylinders to the paper surface by direct contact. The dryer section consists of a succession of drying cylinders, and as the paper web is transported over and between these cylinders, the paper alternately makes contact with the upper and the lower side. Drying takes place in different phases. In the short, first phase, only heat is transferred to the paper. There is no vaporization. This takes place during the second phase, when the wet paper starts to convey its humidity to the surrounding air as water in the paper evaporates. In the third phase, the paper surface has already been dried to the maximum extent, and heat transmission into the dry paper stimulates vaporization inside the paper.

End group

After the end of the drying process, the paper is often subjected to machine calenders. Besides machine calenders, which use steel rolls, there are also soft calenders consisting of paired rolls where one is made of steel and the other is coated with a soft, plastic material. This produces a better overall glazing effect. At the end of the machine, the paper is wound onto steel cores. Most paper machines use calenders or pope reels. To make a uniform wind at constant circumferential speed, the paper reel is pressed with opposition force.

SYSTEM IDENTIFICATION

System identification is a process that can be built from the collection of input-output data by a mathematical model of dynamic systems. The process model provides system parameters as a transfer function in terms of poles, zeros, and delay terms. This plant model is periodically identified, and the changes in its dynamic characteristics are observed. This periodic observation provides the advantages of conventional controller tuning methods, because the plant model is always at nominal operating conditions in any situation [12].

In conventional modeling methods, a model structure will select to calculate the parameters of the model. This structure selection and parameter estimation is done by optimizing an objective function using various techniques. The selection of model structure is nothing but a compromise between model accuracy and model simplicity [13-15]. Auto regressive with exogenous inputs (ARX) is one of the simplest structures for system identification. Alireza discussed that an auto regressive with moving average exogenous inputs (ARMAX) structure is a better choice for identifying the plant model than the simple ARX structure [16].

Plant identification is the first step before an attempt can be made to control the moisture content in the papermaking process. Based on the set point of moisture, the real values of moisture content were measured by an ABB distributed control system (DCS). The actual value depends on the steam pressure of the main dryer group; i.e., if there is a rise in steam pressure, the moisture content of the paper will be decreased. Thus, the model can be identified from the collected data. **Table I** shows the technical specifications of the paper machine studied here. **Figure 2** shows the method implemented to collect the process data for plant identification. Here, steam pressure data were collected from main dryers in the DCS system. The scanner provides actual information related to the manufacturing process. This information is listed in Table I.

NEURAL NETWORK IDENTIFICATION FOR MOISTURE PROCESS

In black-box identification of nonlinear dynamic systems, selection of model structures becomes a more difficult task. The multi-layer perceptron (MLP) network is most popular for learning nonlinear relationships from a set of data. Two issues primarily arise when selecting the model structure:

- Selecting the inputs to the network
- Selecting internal network architecture

Part Name	Details
DCS system AC 450 controller	Operating system: HMI, ABB Analog input: 4–20 MA Analog output: 4–20MA Software: AMPL (A Mathematical Programming Language)
Quality control scanner	Moisture: Infrared (IR) sensor Output: 4–20 MA Brand: Honeywell
Control valve	Size: 6 in., pneumatic actuated Type: air to open
I/P converter	Input: 4–20 MA Output: 0–6 Bar
Dryer	43 cylinders, 5 groups, material-milled steel
Steam pressure	3.5–4.5 bar
Steam temperature	150°C–180°C
Daily production	350 metric tons
Machine speed	700 m/min

I. Technical specifications of the paper machine studied.

In Eqs. (1) and (2), a general linear model structure and predictor form is given:

$$y(t) = g[\varphi(t,\theta),\theta] + e(t) \tag{1}$$

$$y'\left(\frac{t}{\theta}\right) = g[\varphi(t,\theta),\theta] \tag{2}$$

where $\varphi(t,\theta)$ is regression vector, 'θ' is vector containing the adjustable parameters in the NN known as weight, and g is a function realized by the NN.

Model structure

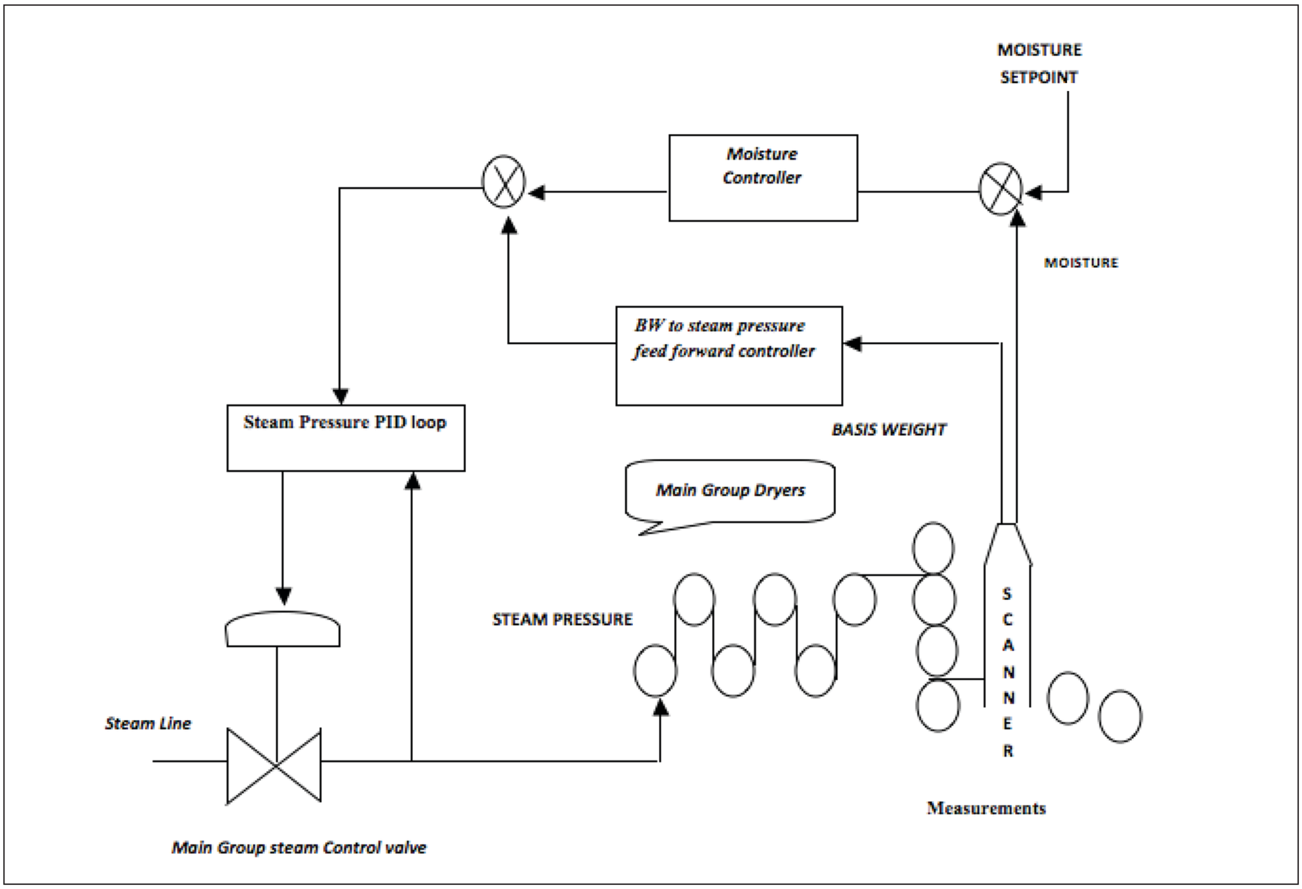
Depending on the choice of regression vector, various NN model structures are examined. The neural network ARMAX (NNARMAX) can be represented by the regression vector and predictor in Eqs. (3) and (4), respectively:

Regression Vector:

$$\varphi(t) = [y(t-1) \ldots y(t-n_a)u(t-n_k)\ldots \\ u(t-n_a-n_k+1) \in (t-1) \ldots \in (t-n_c)]^T \tag{3}$$

where:

$\in (t) =$ prediction error
$\in (t) = y(t) - \hat{y}(t|\theta)$



2. Motion-detection moisture control.

Predictor:

$$\hat{y}(t|\theta) = g(\varphi_1(t), \theta) + (c(q^{-1}) - 1)\varepsilon(t) \quad (4)$$

where:
C = polynomial in the backward shift operator

$$c(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_n q^{-n}$$

In this paper, the NNARMAX model structure was selected since it has a few advantages over other structures. Only the NN auto regressive with exogenous inputs (NNARX) model has a predictor without feedback. The remaining model types all have feedback through the choice of regressors, which in NN terminology means that the networks become recurrent in that future network inputs will depend on present and past network outputs. This might lead to instability in certain areas of the network's operating range, and it can be very difficult to determine whether or not the predictor is stable. The ARMAX structure was constructed to avoid this by using a linear MA-filter for filtering past residuals. When a particular model structure has been selected, the next choice that must be made is the number of past signals used as regressors, i.e., the model order [7]. To solve the problem of local minima, the NNARMAX model was trained several times and initialization of weights were varied at the beginning each time.

IMPLEMENTATION OF SVR FOR NONLINEAR SYSTEM IDENTIFICATION

The basic idea in SVR is to map the dataset $\{(x_1,y_1), \ldots \ldots (x_1, y_1)\} c\, Rn \times R)$ into a high dimensional feature space via non-linear mapping, wherein they are linearly correlated with the outputs [17]. The SVR formalism considers the following linear estimation function:

$$f(x,w) = (w, \emptyset(x)) + b \quad (5)$$

where $w \in R^n$ is weight vector, $b \in R$ is a constant, and $\emptyset(x)$ denotes a mapping function in the feature space.

Based on the principle of structural risk minimization, the SVR learning problem is recast as the optimization problem, which is given below:

$$\text{Min: } \frac{1}{2}\|w\|^2 + C\sum_{i=2}^{l}(\xi_i + \xi_i^*) \quad (6)$$

$$\text{Subject to: } \begin{cases} f(x_i) - y_i \le \varepsilon + \xi_i^* \\ y_i - f(x_i) \le \varepsilon + \xi_i \\ \xi_i, \xi_i^* \ge 0; i = 1, \ldots \ldots, l \end{cases}$$

where C is the regularization constant used to specify the tradeoff between the empirical risk and regularization term. Two positive slack variables, ξ_1, and ξ_1^* $i = 1,2,\ldots..,n$ can be used to measure the deviation from the boundaries of the ε -insensitive zone. That is, they represent the distance from actual values to the corresponding boundary values of ε -insensitive zone. Both C and ε are user-defined parameters. Schematic representation of the SVR using ε-insensitive loss function is illustrated in **Fig. 3**.

Applying the Lagrange function and the Karush-Kuhn-Tucker conditions to Eq. (5) yields the following dual optional form:

$$\text{Max: } -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i,x_j) + \sum_{i=1}^{l} y_i(\alpha_i^* - \alpha_i) - \varepsilon\sum_{i=2}^{l}(\alpha_i^* + \alpha_i) \quad (7)$$

$$\text{Subject to the constraints: } \begin{cases} \sum_{i=1}^{k}(\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i, \alpha_i^* \le C, i = 1, \ldots \ldots, l \end{cases}$$

where the Lagrange multipliers a_i, a_i^* satisfy the equality a_i, $a_i^* = 0$.

The Lagrange multipliers are calculated and an optimal desired weight vector of the regression hyper-plan is $w = \Sigma l_{i=1}(a_i, a_i^*)\, \emptyset(x_i)$. Hence, the general form of the SVR-based regression function can be written as:

$$f(x) = \sum_{i=1}^{l}(\alpha_i^* -)K(x, x_i) + b \quad (8)$$

where $K(x, x_i)$ is called the kernel function and the values of it equal the inner product of two vectors, x_i and x_j, and feature space $\emptyset(x_i)$ and $\emptyset(x_j)$; i.e., $K(x, x_i) = \emptyset(x_i)\emptyset(x_j)$.

The function that meets Mercer's condition can be used as the kernel function [18] and the radial basis function (RBF) is applied in this paper as kernel function, which is defined as:

$$K(x_i,x_j) = \exp\left(-\|x_i - x_j\|_2^2 \div 2\sigma^2\right) \quad (9)$$

where σ^2 is the width of the RBF.

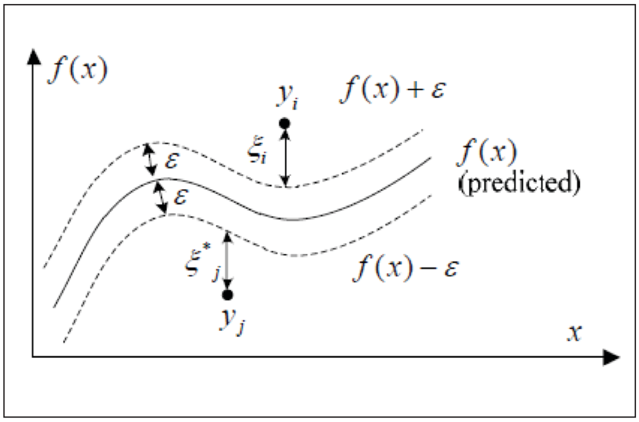
In order to design an effective model, the values of two essential model parameters of the C and ε and kernel function parameter σ^2 in SVR have to be optimally chosen in advance. The regularization parameter C determines the tradeoff cost between minimizing the training error and minimizing model complexity, which will reduce the generalization capability when it is set too small or excessive. The parameter ε defines the nonlinear mapping from the input space to some high-dimensional feature space that determines the number of support vectors, and σ^2 reflects correlation of the support vector, which also determines both the generalization capability and the prediction accuracy.

Steps in designing the SVR model

The design and development of the SVR model for any regression task involves a sequence of steps, as enumerated below:

Step 1. Data scaling/preprocessing

The input features in the train set and test set need to be scaled



3. Sketch map of the support vector regression (SVR) using ε-insensitive loss function.

properly before applying SVR [14]. This is important, as the kernel values depend on the inner products of the feature vector. Scaling prevents the domination of any feature over the other because of higher numeric values involved and also avoids numerical difficulties during calculation. Then, each attribute is linearly scaled to the range of [0, 1].

Step 2. Design of SVM model

The RBF kernel is chosen as a first choice because of its widely accepted accuracy. Further, it is capable of handling nonlinear relationships existing between the class labels and input attributes. The second reason is that the RBF kernel, unlike other kernels, has only one kernel parameter, thereby reducing the complexity of the model [19].

There are two parameters associated with the SVR model designed with RBF kernel–penalty parameter, C, and RBF kernel parameter, γ. The goal is to identify optimal (C, γ) for the classifier to accurately predict the unknown data (test data) [20]. This can be achieved by different techniques, descriptions of which follow in the next section. After designing the SVR model with the chosen kernel and optimal parameters, it is trained with the scaled input-output train set samples. Once the performance of the regression is found satisfactory in the training phase, the model is validated with test samples to determine its overall performance.

Step 3. Selection of SVM parameters

The SVR model in this work uses RBF as the kernel mapping function. Two major parameters associated with the SVR model that need to be optimized are: 1) penalty parameter (C), and 2) RBF kernel function parameter (γ). Parameter C, which represents the cost of penalty, influences the classification performance [18]. A larger C value gives a good accuracy rate in the train phase, but very poor accuracy in the test phase. Conversely, too small a C value leads to an unsatisfactory performance, making the model useless. Parameter γ, representing the kernel spread, has a strong influence on partitioning outcome in feature space. An excessive C value results in overfitting, while a disproportionately small value leads to underfitting [10]. Hence, proper selection of SVR parameters, while designing the SVR model, is highly important to achieve a good mean squared error (MSE) and high generalization ability [21]. In this paper, the grid search (GS) method is used for optimal selection of SVM parameters.

Step 4. Grid search

GS is the most common and simple method used to determine the appropriate values for parameters C and γ. The GS method adopts a v-fold cross validation technique [13]. In a v-fold cross validation, the whole training set is divided into v subsets of equal size. Sequentially, one subset is tested using the SVM regression trained on the remaining (v-1) subsets. Thus, each instance of the training set is predicted once and the cross-validation accuracy is the percentage of data samples that are correctly classified [21]. In this work, GS using fivefold cross validation is used. All pairs of (C, γ) were tried and the one with highest cross-validation accuracy was selected [15]. Exponentially growing sequences of C and γ is a practical method to identify optimal parameters. The sequences of $C = \{2^{-5},\ldots,2^{15}\}$ and $\gamma = \{2^{-15},\ldots,2^{5}\}$ are used in this SVR experiment.

MODEL REFERENCE ADAPTIVE CONTROL

The model reference adaptive scheme is based on the observation that the difference between the output of the plant and the output of the reference model (subsequently called plant-model error) is a measure of the difference between the real and the desired performance [22]. This information is used by the adaptation mechanism to directly adjust the parameters of the controller in real time in order to asymptotically force the plant model error to zero. This scheme used for control applications may be called a model reference adaptive system (MRAS). Nonlinear characteristics of a system validate the use of adaptive controllers such as model reference adaptive control (MRAC) and SVR [17]. The model reference adaptive system gives the desired response to the reference signal [23]. Thus, the response of the command model acts as reference for the process to the normal feedback loop. The adaptive mechanism used for adjusting the parameters in an MRAS can be obtained in two ways: by using a gradient method or by applying Lyapunov stability theory [24]. The mechanism uses the error (e), the nonlinear process output (y), and the input signal (u_c) to adjust the parameters of the control system. These parameters are adjusted so as to minimize the error. As this paper discusses, the parameters are varied so that the cost function J (θ) is reduced.

Comparison of plant and model output provides error information that can be used for finding the cost function in Eq. (12). This cost function is to optimize the controller output:

$$\text{Cost function: } J(\theta) = \frac{1}{2}e^2(\theta) \tag{10}$$

where e is $y_{plant\ output} - y_{model\ output}$

MIT rule is $\frac{d\theta}{dx} = -\gamma \frac{\delta J}{\delta\theta} = -\gamma e \frac{\delta e}{\delta\theta}$ (11)

where $\frac{d\theta 1}{dt}$ is $-\gamma\left(a_{mu_c}\middle|s^3 + a_m\right)e$

and $\frac{d\theta 2}{dt}$ is $\cdot\gamma\left(a_{my_{plant}}\middle|s^3 + a_m\right)e$

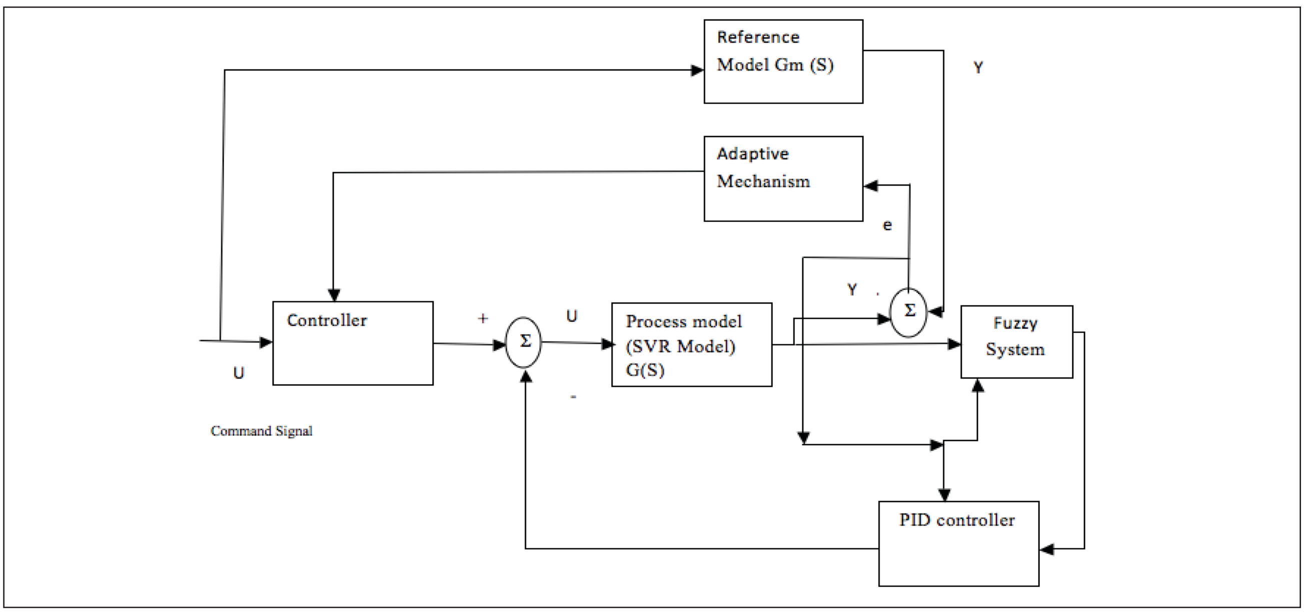
Also, the controller parameters 01 and 02 are updated from $\frac{d\theta 1}{dt}$ and $\frac{d\theta 2}{dt}$.

Fuzzy model reference adaptive control

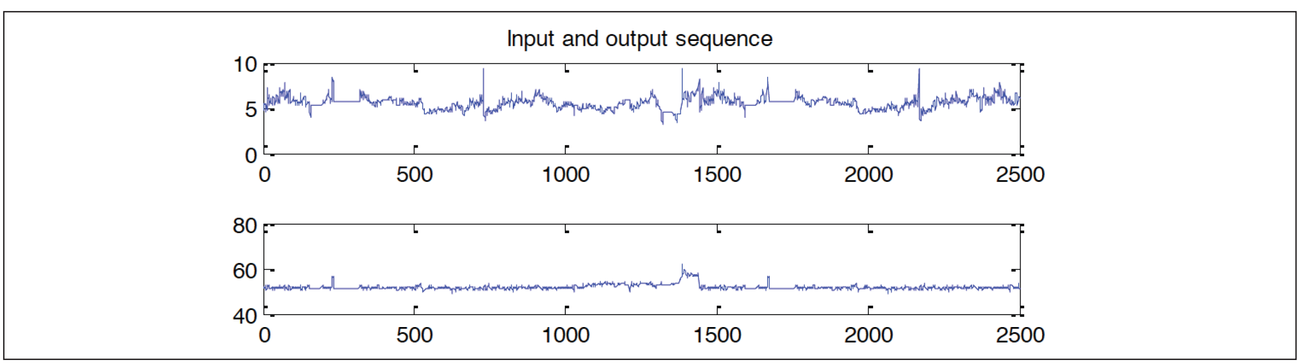
Fuzzy systems based on MRAC have been reported by a number of researchers. The major components are the reference model, fuzzy logic controller (FLC), and an adaptation mechanism. The reference model incorporates the desired performance characteristics of the system response. The FLC is implemented in a reference adaptive control system based on the gradient descent algorithm method [25]. The adaptation mechanism is involved to adjust the characteristic of the FLC in response to the error signal *e*(*t*) ,which is the output difference of the reference model and plant. An inverse plant model is designed as an adaptive mechanism for the correction value of $\Delta U(t)$ applied to the $U(t)$, which is the output signal. The adaptive mechanism updates the information in order to apply that correction via the FLC [19,20]. The updating algorithm tries to modify the FLC characteristic such that its output is altered by the value ΔU (*t*). The adaptation is affected by adjusting the value of the output fuzzy sets [9]. In the proposed method, the error that can be measured between the moisture and the output of a reference model is applied to an FLC, which can force the system to act as the model by adjusting the knowledge base of the fuzzy controller or by adding a correction signal to the fuzzy controller output [22]. The structure of fuzzy modified model reference adaptive control (FMRAC) is monozygotic to that of the FLC, as illustrated in **Fig. 4**.

RESULTS AND DISCUSSION

The identification process was performed using a conventional and support vector regressor. A total of 2000 data samples were taken that describe the behavior of the moisture



4. Fuzzy-modified model reference adaptive control (FMMRAC) structure.



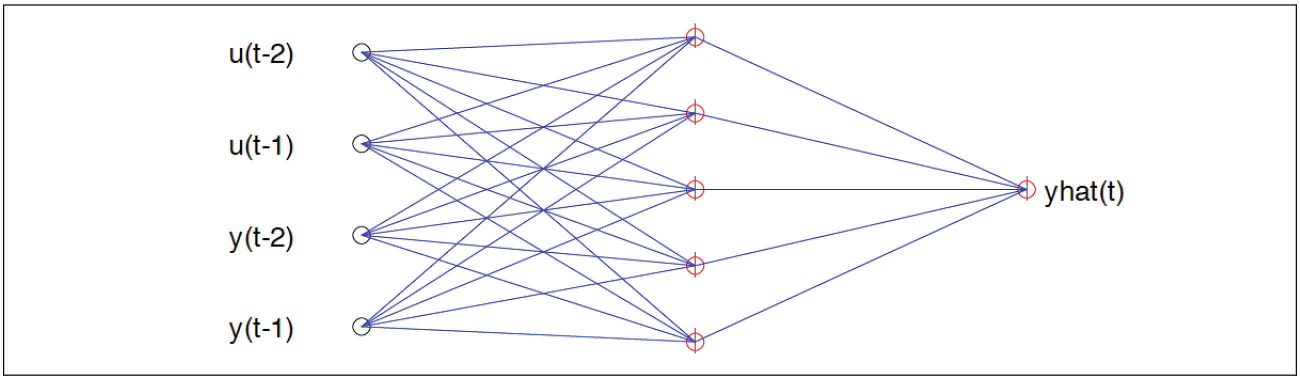
5. Input and output response.

process, and 1260 data samples were from a single shift. Another 2500 data samples were collected from two shifts a day. The data set was then divided into 70% for training and 30% for testing. For training data of input-output pairs, output was "moisture of a paper reel" and input was "pressure variation in the cylinder." **Figure 5** shows the input and the output response of the moisture data from the selected training and testing data set.

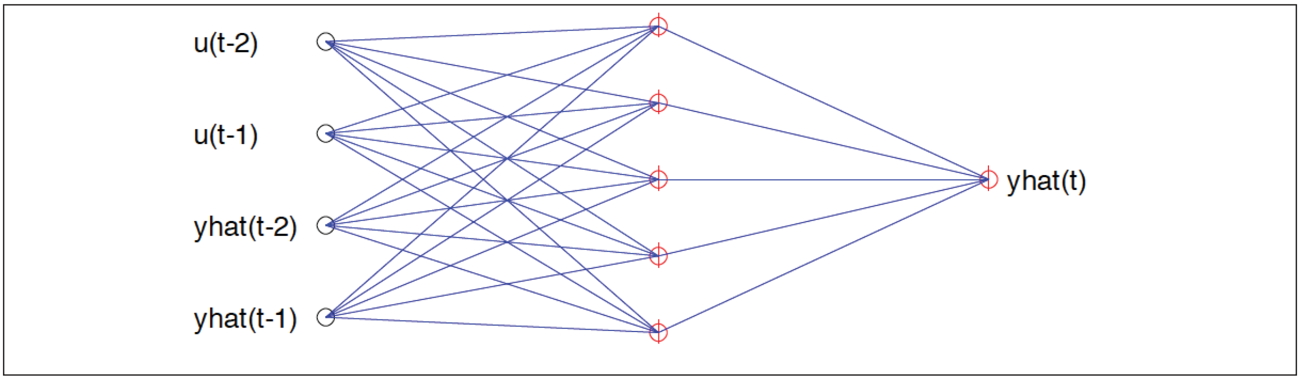
Here, the NN regression models were applied to the divided set of data. The model structure selection consists of two subproblems: choosing a regressor structure and choosing network architecture. For all regressor network architecture, five hidden "tanh" units and one linear output unit were chosen.

From **Figs. 6-8**, the NN structure for different nonlinear identification-based NNs is presented. NN-based nonlinear structure is generally defined in terms of NNARX structure, and the structure is used to identify the nonlinear system. Based on the identification, the correlation function can be plotted as in **Figs. 9-11**, where NNARX, NNARMAX, and neural network output error (NNOE) structure were analyzed for the empirical process.

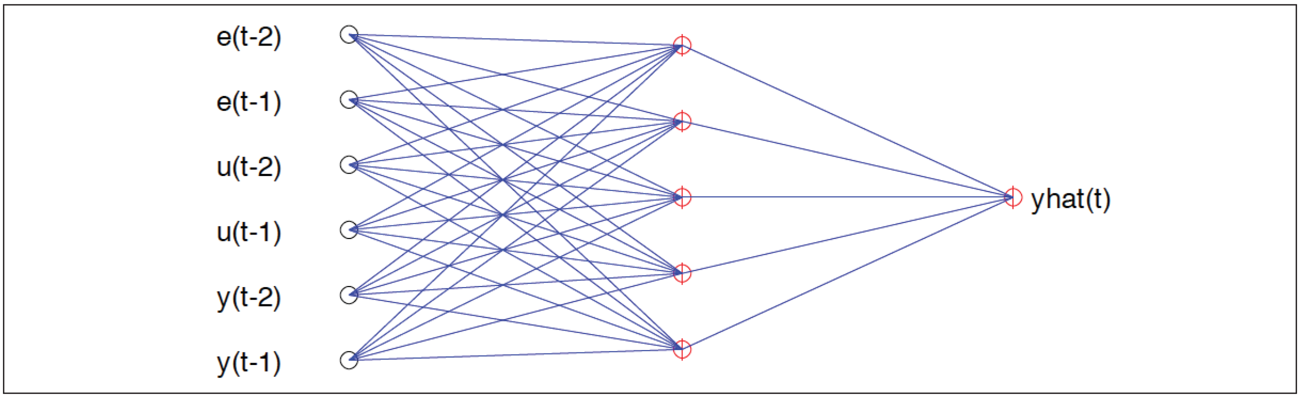
Figures 9-11 depict the correlation plot for the prediction error by the different models for the nonlinear system. The comparison provides detailed performance analysis of the different NN mod-



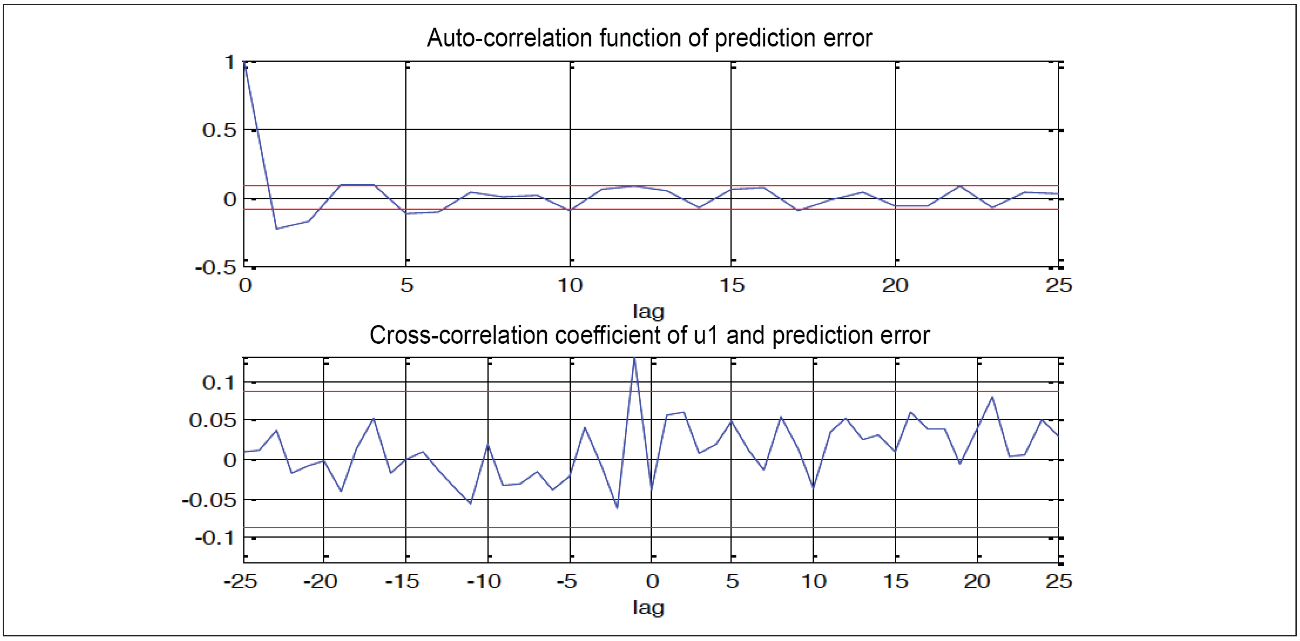
6. Network structure for neural network auto regressive with exogenous inputs (NNARX) model.



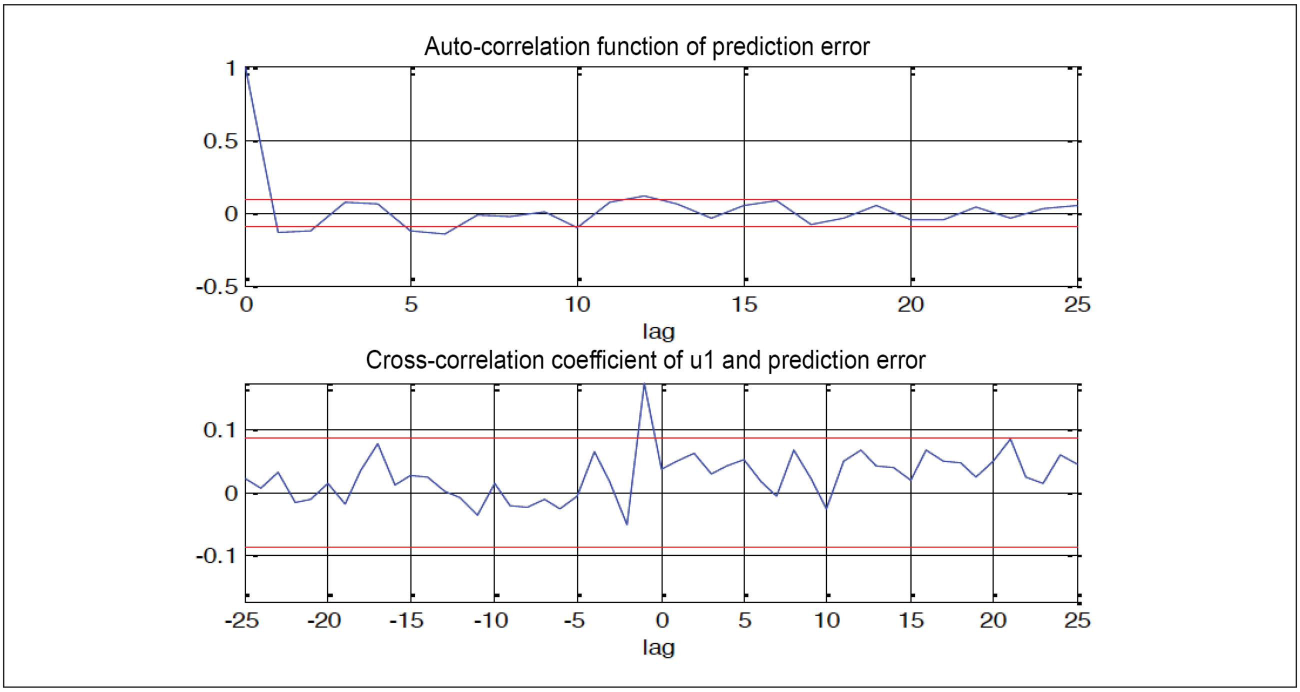
7. Network structure for neural network output error (NNOE) model.



8. Network structure for neural network auto regressive with moving average exogenous inputs (NNARMAX) model.



9. Auto-correlation and cross-correlation of prediction error in NNARX model.



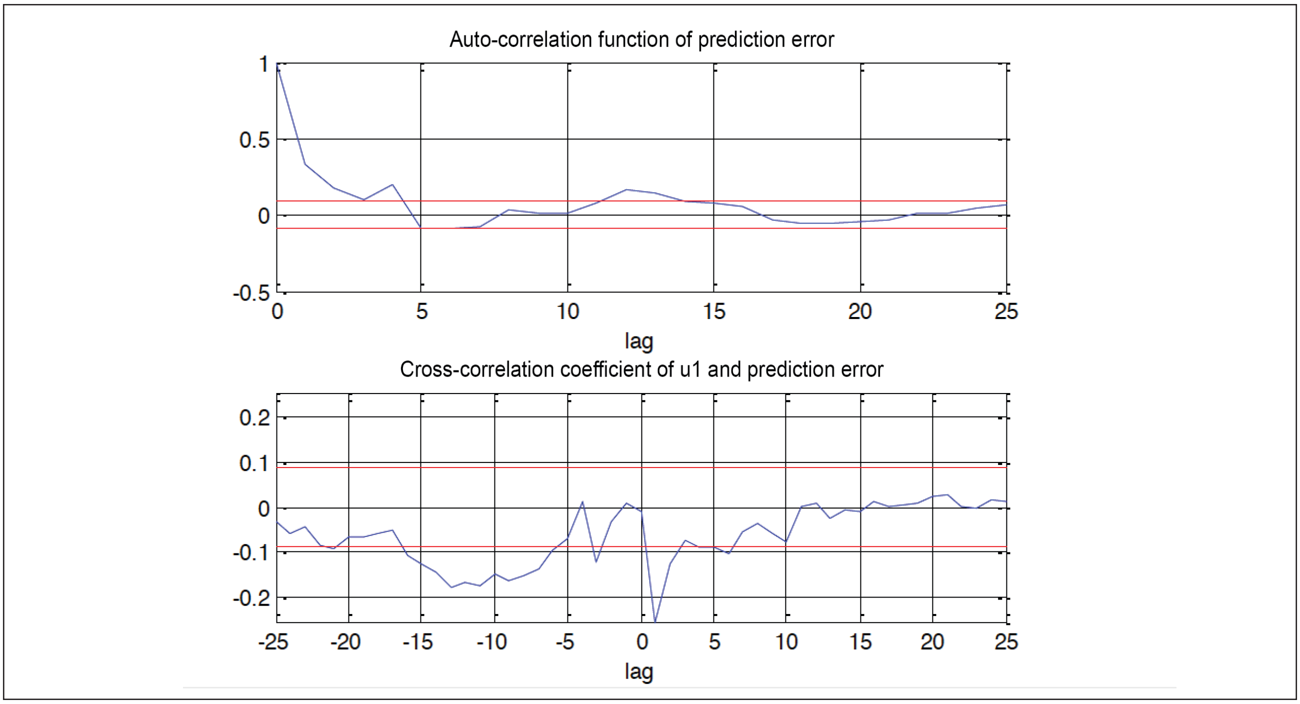
10. Auto-correlation and cross-correlation of prediction error in NNARMAX model

els. The cross-validation procedure is the best procedure to validate models [24]. During the procedure, the estimated NNARX, NNARMAX, and NNOE models are applied to a data set that differs from the one used for estimation. Based on this analysis, a fitness of 93.24% is obtained from the NNARMAX model. Model validation can also be done by applying residual analysis and correlation analysis to the experimental data. The residual analysis gives the auto correlation of the output residuals and the cross correlation between the input to the process [26].

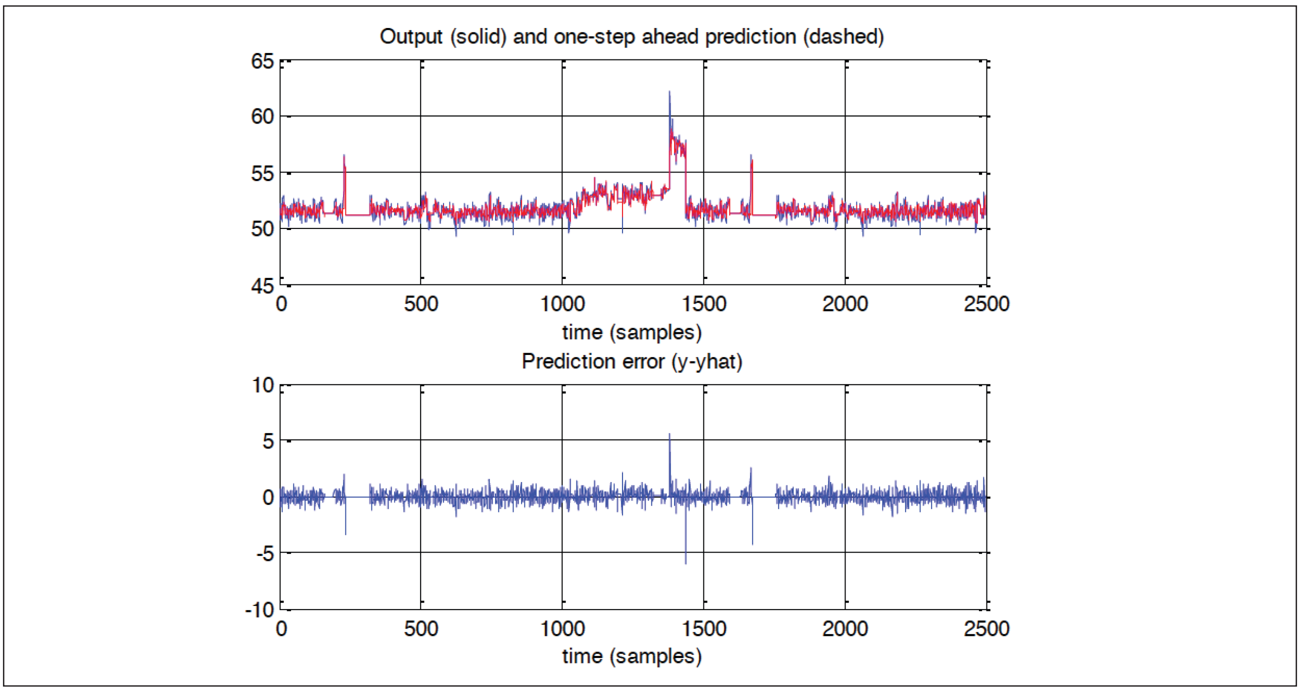
The validation responses obtained from the various NR structures are plotted in **Figs. 12-14**. In this analysis, the performance of an NNARMAX model provides a minimized MSE value.

Implementation of SVR identification of moisture process

The LIBSVM software developed by Chang and Lin has been used for the design and testing of the SVM model [27]. Pro-



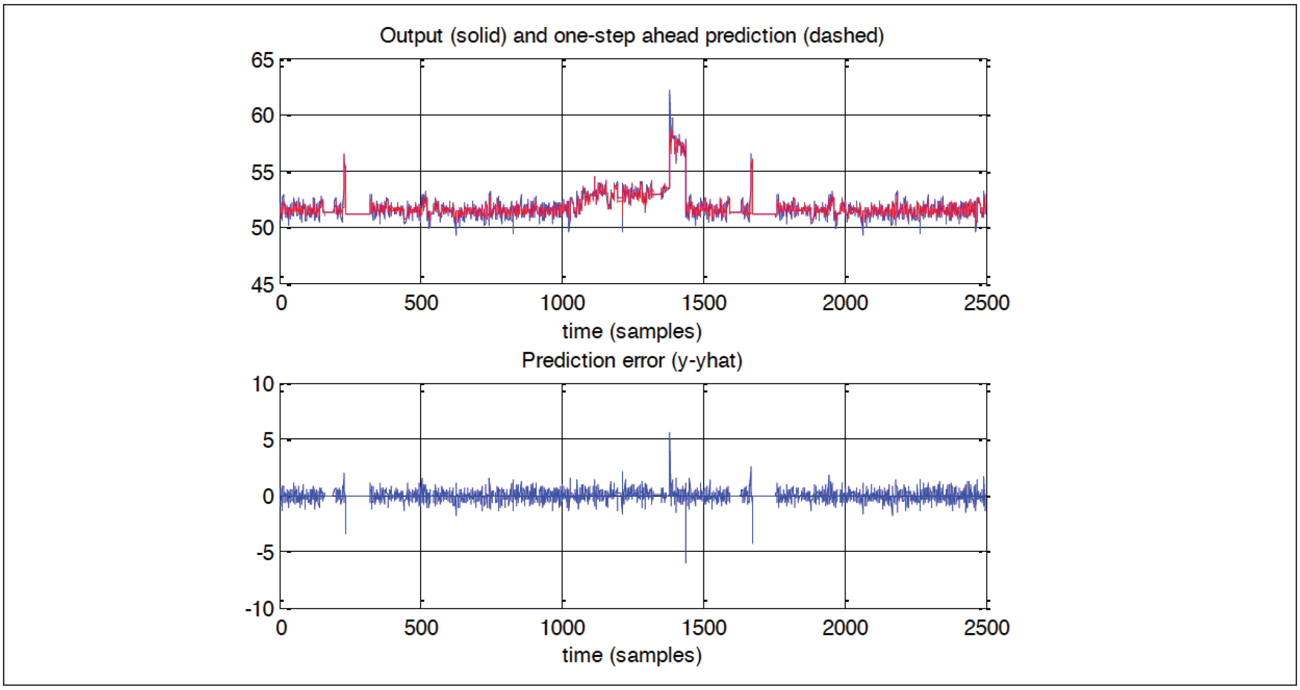
11. Auto-correlation and cross-correlation of prediction error in NNOE model.



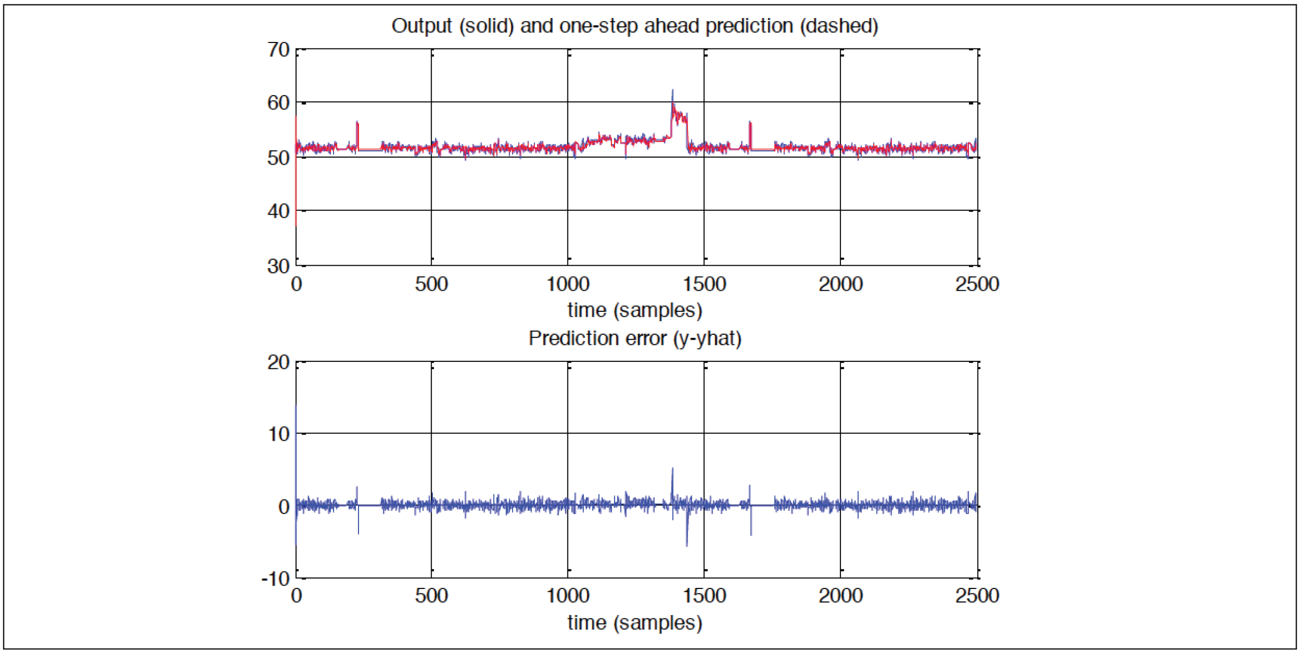
11. Auto-correlation and cross-correlation of prediction error in NNOE model.

grams for parameter selection of SVM by GS techniques are written in the MATLAB environment. The SVR method can be implemented because it appeared the best tool for modeling the nonlinear system [1,3]. In this paper, the SVM model uses a radial basis function (RBF)-type kernel function. The performance of any SVM regression model will strongly depend on the selection of SVM parameters. The penalty parameter (C) and RBF kernel parameter (c) are associated with SVM [28-33]. The selection of parameters is performed by the GS technique, and the results are shown in **Table II**. The selection of parameters was done with the fivefold cross validation procedure.

In **Fig. 15**, SVR model response uses the GS parameter selection method. The best values of SVR parameters obtained for a maximum cross validation error of 0.00068879 as seen from Table II are $2\gamma = 5$ and $2C = 10000$. The collect-

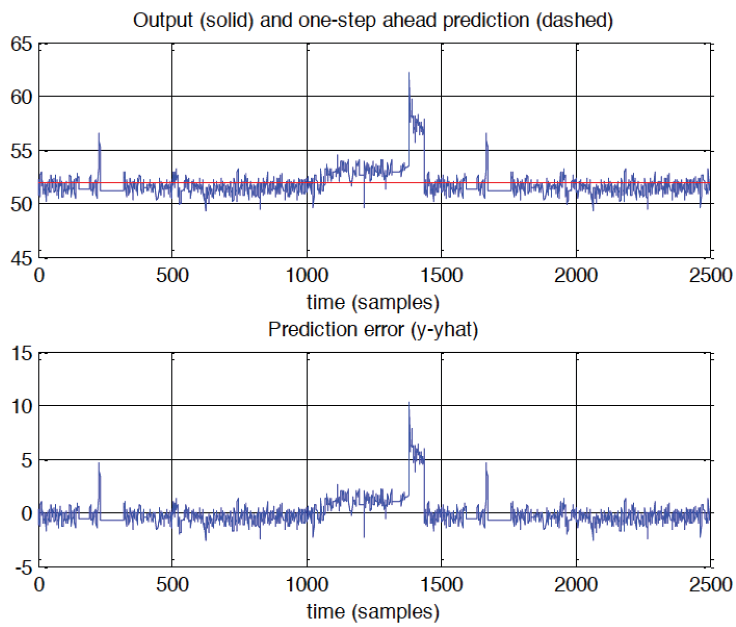


12. Validation response of NNARX model.



13. Validation response of NNARMAX model.

ed data from the DCS were plotted as shown in **Fig. 5**. From the collection, 54% of data is given for input testing and the remaining are given for input training. Input cylinder pressure and output moisture data are given for training and testing of SVR methods. In Fig. 15, the validation response of an SVR is plotted, and the response shows its capability to better reduce MSE than previous methods. A performance comparison of various identification structures in reducing MSE is given in **Table III**. The obtained model was identified as suitable and given to the adaptive controller for optimized control of the process. Based on the previous comparison made with an NN structure nonlinear identification, the fitness value and the correlation MSE of the NNARMAX model gives better performance. Ultimately, to improve the value of fitness, an advanced NN strategy was implemented — SVR. Here, the RBF-based kernel function provides better performance than the value of previous comparative methods listed in Table II, which shows the detailed analysis of SVR within



14. Validation response of NNOE model.

Γ	C=0.1	C=1	C=10	C=100	C=1000	C=10000
2^{-3}	0.683743	0.683589	0.0680289	0.0679112	0.672649	0.0682098
2^{-2}	0.683728	0.0683454	0.0678852	0.678734	0.0680817	0.0701059
2^{-1}	0.0683573	0.0683362	0.06781	0.067812	0.0678762	0.0699764
2^{0}	0.0683367	0.0683029	0.0678166	0.0677027	0.0675849	0.00712111
2^{1}	0.0683047	0.068134	0.0677606	0.0677062	0.0676804	0.0672988
2^{2}	0.0681362	0.0679116	0.0676907	0.0675335	0.0675084	0.0708155
2^{3}	0.067917	0.0678562	0.0677108	0.0678762	0.0680621	0.0687694
2^{4}	0.0678477	0.0677686	0.0679467	0.0682526	0.0680976	0.0709433
2^{5}	0.0677513	0.0677195	0.0682725	0.0068338	0.0068310	**0.00068879**

II. Cross validation error.

a prescribed range of γ and C values. The best values of SVR parameters obtained for a maximum cross validation mean squared error of 0.000689, as seen from Table II, are $2\gamma = 5$ and $2C = 10000$; i.e., nearly 214.

In Table III, various identification structure methods were compared with SVR performance. Here, the performance index was chosen as MSE, and the least value of MSE is obtained from SVR.

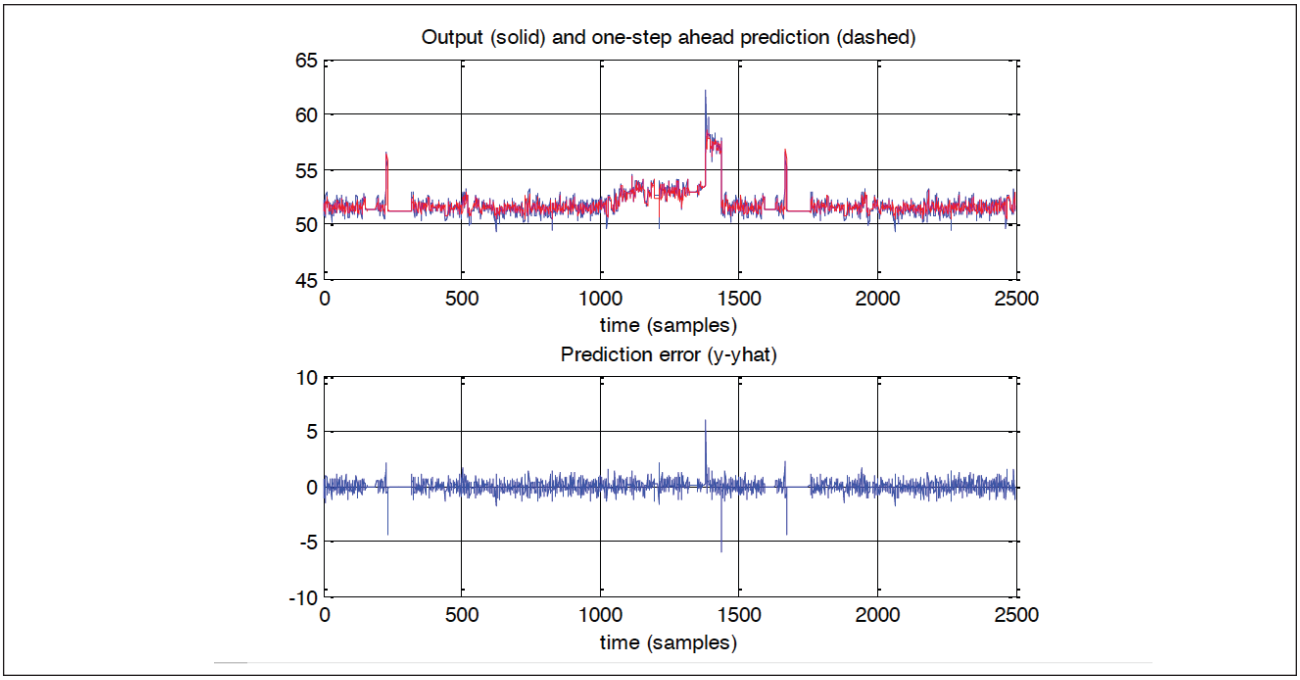
Controller implementation

Based on the previously made comparison, the best performance was obtained from the SVR model, so it was also applied as a model to the MRAC. The combination of SVR iden-

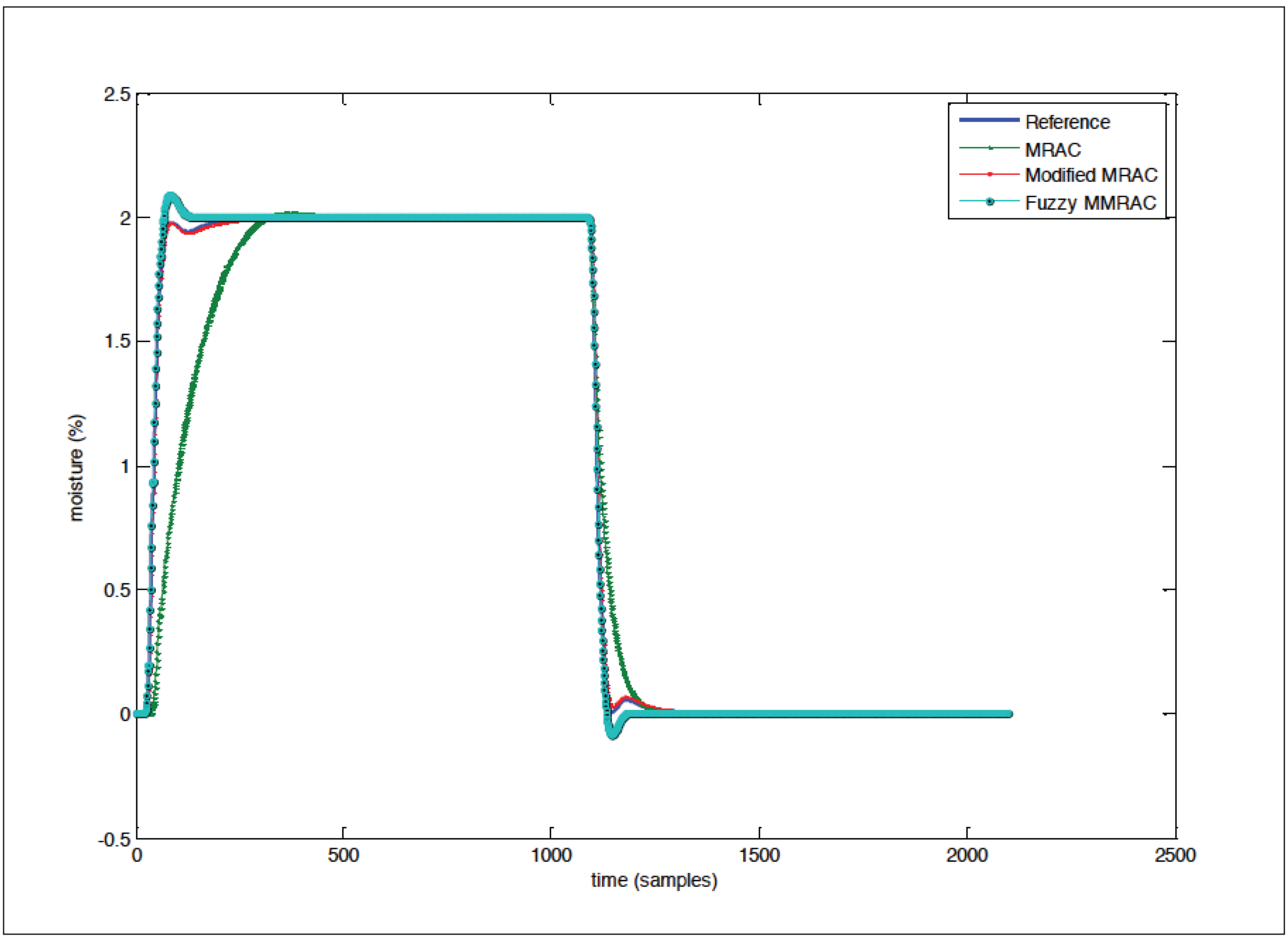
Method	MSE
NNARX	0.003232
NNARMAX	0.002176
NNOE	0.03154
SVR	0.000689

MSE = mean square error; NNARX = neural network auto regressive with exogenous inputs; NNARMAX = neural network auto regressive with moving average exogenous inputs; NNOE=neural network output error; SVR = support vector regression.

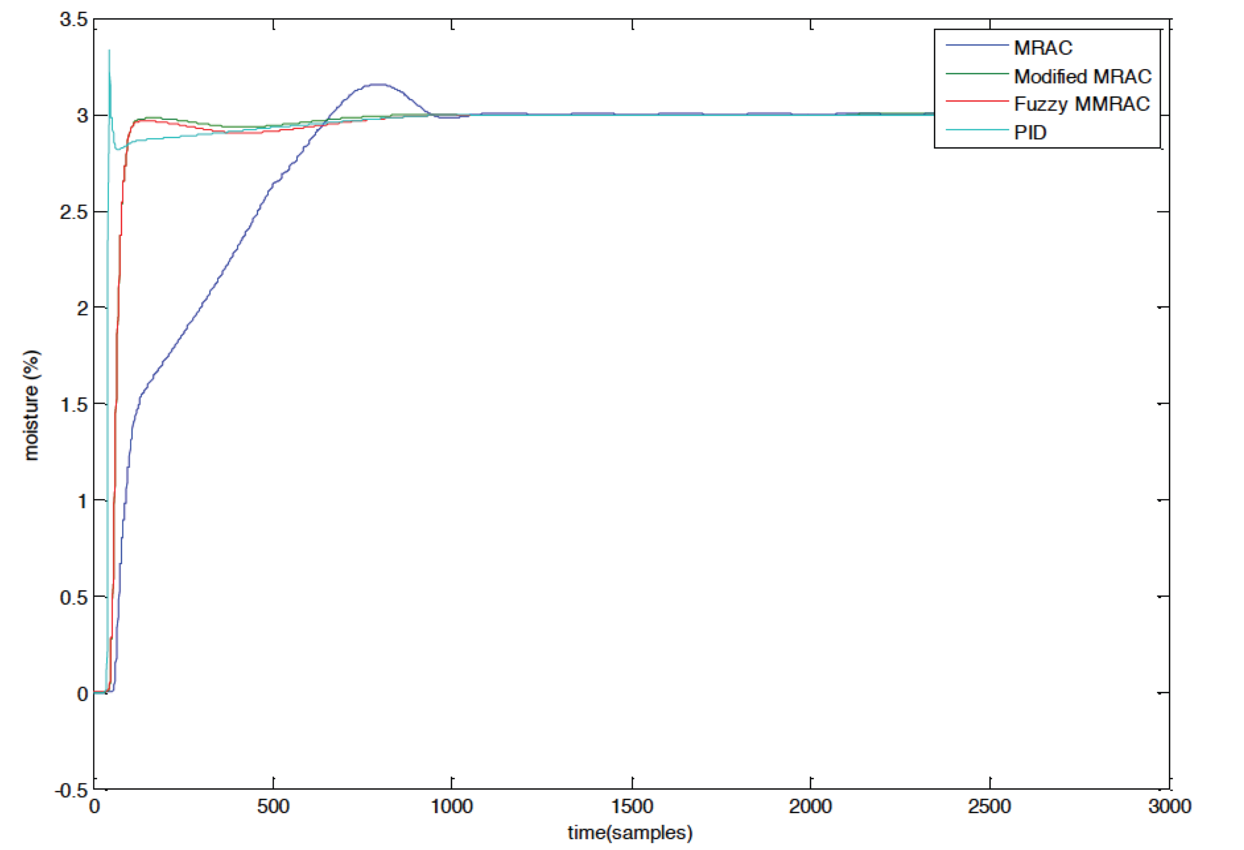
III. Performance comparison of various identification structures.



15. Validation response of SVR model.



16. Servo response analysis of different controllers.



17. Process input for step input of 3 to the command signal.

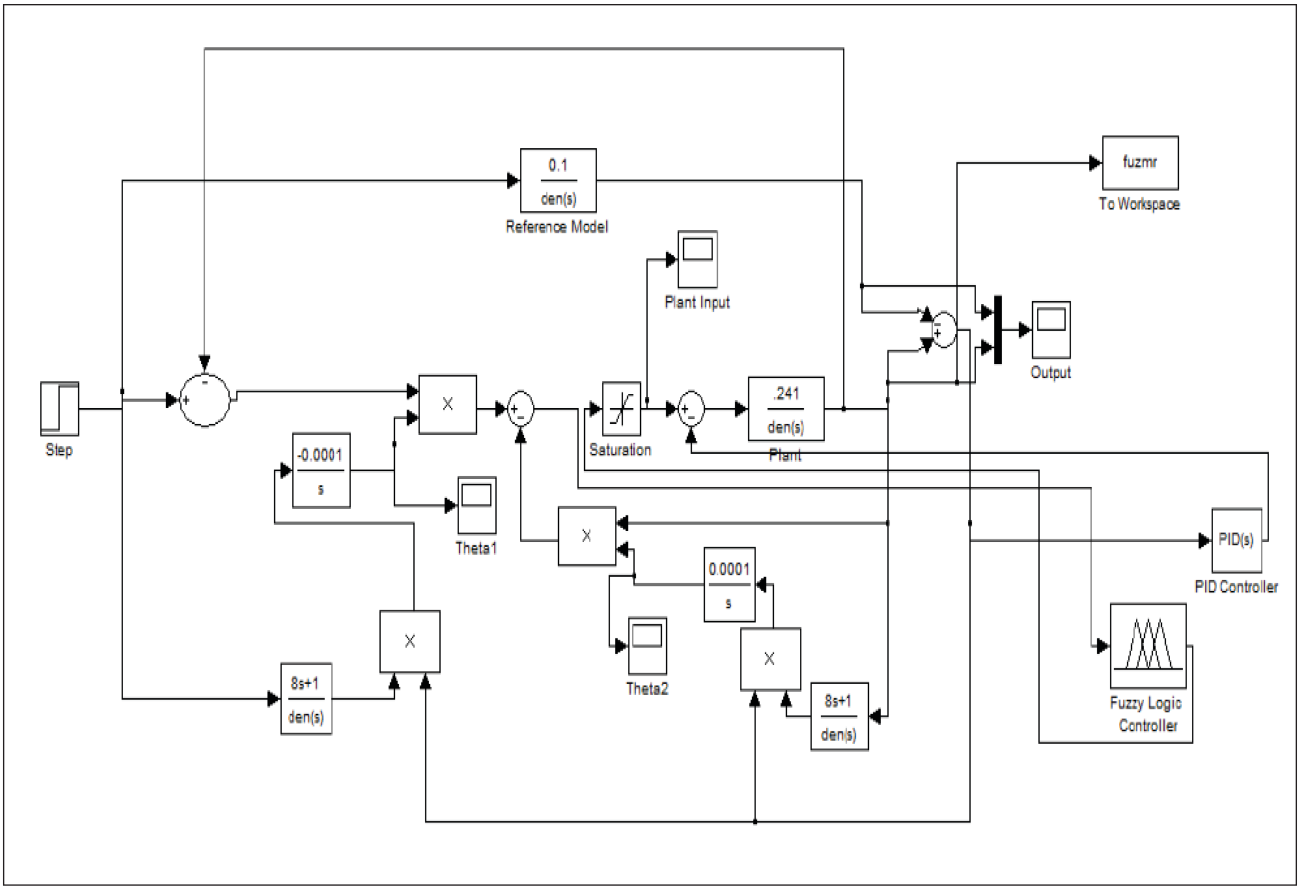
tification and MRAC procedure provides an SVR-based MRAC (SVR-MRAC) strategy. This may provide an efficient control strategy for the nonlinear process, which is plotted in **Fig. 16**. Figure 16 demonstrates the performance of FMMRAC with different control methods. The proposed SVR-MRAC method exactly follows the reference signal. Here, a varying input signal is given to the system for analyzing the proposed method with servo response, proving the effectiveness of the proposed controller strategy. Input to the MRAC controller is varied and the response was obtained as shown in Fig. 16. The servo response was observed by varying the moisture value in the range of 4.3%–18%, and the dynamic performance of various controllers was identified and compared in **Table IV** and shown pictorially in Fig. 16. Based on the comparison, the FMMRAC provides good performance with a sufficient performance index and exactly follows the reference signal. So, the combination of SVR-modeled FMMRAC was identified as the optimized controller for the nonlinear process.

Tuning Method	Dynamic Performance Specifications			Performance Index
	T_r	T_s	M_p, %	MSE
MRAC	512	670	0.0	4.2926
MMRAC	11	16.32	1	0.3342
FMMRAC	11	15.221	0. 0	0.04406

T_r = rise time; T_s = settling time; M_p = peak overshoot; MSE = mean square error; MRAC = model reference adaptive control; MMRAC = modified MRAC; FMMRAC = fuzzy-modified MRAC.

IV. Comparison of performance indices of different controllers.

Figure 17 shows the different responses for reference moisture value. To control the moisture content, MMRAC and FMRAC show the better performance compared to other methods. The efficient collection of data provides a suitable model with the least MSE values from a support vector concept. This can be effectively controlled by a fuzzy modified adaptive system. **Figure 18** shows the simulation model of proposed fuzzy modified MRAC. Proportional-integral-derivative (PID)-tuned MRAC can be embedded with fuzzy knowledge to provide proper control action. This performance is clearly described in Table IV. The PID



18. Simulation model of fuzzy modified MRAC (FMMRAC).

controller-tuned system has been removed, which is shown in Figure 18, so it will act as a simple modified MRAC system. The step input is given to the proposed SVR-MRAC system and the response was identified as shown in Fig. 17, where performances were compared.

CONCLUSION

The nonlinear moisture process was identified as an over-damped second order process using standard NN-based nonlinear identification procedures such as NNARX, NNARMAX, and NNOE. Then, the model accuracy was increased by using an SVR-based system identification procedure. The validation results show that the model obtained is capable of explaining the main dynamic characteristics of the nonlinear moisture process. The proposed modification of the MRAC scheme has provided improved performance in transient, as well as steady state, analysis when compared to MRAC and PID controllers. The application of fuzzy-based modified MRAC has given better results in transient and steady-state performance than those of the earlier methods of MRAC, modified MRAC, and PID controllers. The simulation results described the superiority of the fuzzy-based modified MRAC, which is proposed over the other three methods studied. The proposed modified MRAC and fuzzy-based modified MRAC give improved transient and steady-state performance in the control of nonlinear processes. Future work aims to identify the moisture process by using a multiple-model approach and to design the proposed fuzzy-based modified MRAC for each local model for each fuzzy set. **TJ**

LITERATURE CITED

1. Vapnik, V., *The Nature of Statistical Learning Theory*, Springer Verilag, New York, 1995.
2. Vapnik, V., *Statistical Learning Theory*, John Wiley, New York, 1998.
3. Vapnik, V., Golowich, S., and Smola, A., *Adv. Neural Inform. Proces. Syst.* 9: 281(1996).
4. Agarwal, M., Jade, A.M., Jayaraman, V.K., et al., *Chem. Eng. Prog.* 98(1): 57(2003).
5. Jack, L.B. and Nandi, A.K., *Mech. Sys. Sig. Proc.* 16(2002): 373(2002).
6. Narendra, K. and Parthasarathy, K., "Identification and control of dynamical systems using neural networks," *IEEE Trans. Neural Networks* 1(1): 4(1990) .
7. Norgaard, M., Ravn, O., Poulsen, N.K., et al., *Neural Networks for Modelling and Control of Dynamic Systems*, Springer, London, 2000.
8. Medsker, L.R. and Jain, L.C., *Recurrent Neural Networks: Design and Applications*, CRC Press, Boca Raton, FL, USA, 2000.

9. Vieira, J., Dias, F.M., Mota, A., "Artificial neural networks and neuro-fuzzy systems for modelling and controlling real systems: a comparative study," *Engineering Applications of Artificial Intelligence* 17(3):265(2004).

10. Jain, A.K., Mao, J., and Mohiuddin, K.K., "Artificial neural networks: A tutorial," *IEEE Computer Special Issue on Neural Computing*, 1996, p. 31–44.

11. Qin, H.L. and Li, X.B., "A chaotic search method for global optimization on tent map," *Electric Machines and Control* 8(1): 67(2004).

12. Valarmathi, K., Devaraj, D., and Radhakrishnan, T.K., *Braz. J. Chem. Eng.* 26(1): 99(2009).

13. Karthik, C. and Rajalakshmi, M., "Nonlinear identification of pH process using NNARX model," *CiiT International Journal of Artificial Intelligent Systems and Machine Learning, (0974 – 9543)* 4(8): (2012).

14. Karthik, C., and Rajalakshmi, M., "Non Linear Structure Identification of pH Process," *IEEE - ICAESM 2012, Int. Conf. Adv. Eng., Sci. Management, (978-1-4673-0213-5)*, 2012, p. 45.

15. Åström, K.J. and Hagglund, T., *PID Controllers: Theory, Design, and Tuning*, Instrument Society of America, Research Triangle Park, NC, USA, 1994.

16. Alireza, Karbasi, "Comparison of NNARX, ANN and ARIMA Techniques to Poultry Retail Price Forecasting," Master's thesis, University of Zabol, Zabol, Iran, 2009.

17. Min, J. and Lee, Y., *Exp. Syst. Appl.* 28(4): 603(2005).

18. Vapnik, V.N., *The Nature of Statistical Learning Theory*, Springer, New York, 2000.

19. Rajalakshmi, M., Kalyani, S., Jeyadevi, S., et al.,"Nonlinear Identification of pH Process using Support Vector Machine," *IJCA Proc., Int. Conf. Innovations in Intelligent Instrumentation, Optimization and Electrical Sciences (ICIIIOES)* (2): 36(2013).

20. Gretton, A., Doucet, A., Herbrich, R., et al., "Support vector regression for black box system identification," *IEEE Workshop Stat. Signal Process., 11th*, 2001.

21. Cristianini, N. and Shawe-Taylor, J., *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Publishing House of Electronics Industry, Beijing, 2004.

22. Aström, K. and Wittenmark, B., *Adaptive Control*, Addison-Wesley, Boston, 1989.

23. Reay, D.S. and Dunniganu, M.W., *IEE Proc. - Control Theory and Applications* 144(6): 605(1997).

ABOUT THE AUTHORS

Generally, process industries — like paper mills — have a number of nonlinear process parameters that must be controlled to obtain better performance. Of these nonlinear parameters in the paper industry, basis weight and moisture are complicated to control. Improving the quality of the controller strategy for these nonlinear processes leads to energy reduction, increased production rate, and economic profit. This is the motivation behind the authors to work in this area of research.

The application of fuzzy-based modified model reference adaptive control (MRAC) has given better results in transient and steady-state performance than that of the earlier methods of MRAC, modified MRAC, and PID control proposed by the researchers. The proposed work is to identify the dryer model using support vector regression (SVR) to develop a suitable controller. Finally, the obtained SVR model is optimized with the fuzzy modified MRAC (FMMRAC) controller.

The most difficult part in the research was the model identification of the dryer process. With the help of the SVR concept, this difficulty was addressed.

The most interesting part in the research was the FMMRAC concept. The concept of the fuzzy inference system is key to the reference model where it is compared with the actual SVR model. The FMMRAC ensures the stability of the control system.

The FMMRAC provides improved transient and steady-state performance in the control of nonlinear processes. This controller concept may be implemented in an industrial scanner for optimum controller performance. Development of modern scanners with the concept of FMMRAC will ensure the quality of production with less consumption of energy and more economic profit.


Karthik


Valarmathi


Rajalakshmi

The next step in our research is to identify the moisture process by using a multiple-model approach and to design the proposed FMMRAC for each local model of each fuzzy set. The consolidated work will extend further to identification of basis weight, ash content, color, and online monitoring with the development of a modern scanner.

Karthik is assistant professor, Kalasalingam University, Tamilnadu, India. Valarmathi is professor, PSR Engineering College, Tamilnadu, India. Rajalakshmi is assistant professor, Kamaraj College of Engineering & Technology,Tamilnadu, India. Email Karthik at karthikmtech86@gmail.com.

24. Navghare, S.R., Bodhe, G.L., and Shruti,"Design of Adaptive pH Controller using ANFIS", *Int. J. Computer Applications (0975–8887)* 33(6): 41(2011).

25. Spooner, J.T., Ordonez, R., and Passino, K.M., "Stable direct adaptive control of a class of discrete time nonlinear systems," *Proc. IFAC World Congr., 13th*, International Federation of Automatic Control, Laxenburg, Austria, 1996, p. 343.

26. Kamalasadan, Sukumar and Ghandakly, Adel A., *IEEE Trans. Instrum. Meas*. 56(5): 1786(2007).

27. Chang, C., and Lin, C. *LIBSVM — A Library for Support Vector Machines*. Available [Online] https://www.csie.ntu.edu.tw/~cjlin/libsvm/ <15Feb2016>.

28. Kalyani, S. and Swarup, K.S., *Int. J. Electrical Power & Energy Systems* 44(1): 547(2013).

29. Silva Neto, J.L. and Huy, H. Le, "An improvement fuzzy learning algorithm for motion control applications," *Proc. IEE, Institute of Electrical and Electronics Engineers*, New York, 1998, Vol. 1, p. 1.

30. Lin, S., Ying, K., Chen, S., et al., *Exp. Syst. Appl.* 35(4): 1817(2008).

31. Ljung, Lennart, *System Identification - Theory for the User*, 2nd edn., PTR Prentice Hall, Upper Saddle River, NJ, USA, 1999.

32. Al-Hiary, Heba, Braik, Malik, Sheta, Alaa, et al., "Identification of a chemical process reactor using soft computing techniques," *IEEE Int. Conf. Fuzzy Syst*., 2008, p. 845.

33. Resendiz-Trejo, Juan Angel, Yu, Wen, and Li, Xiaoou," Support vector machine for nonlinear system on-line identification," *Int. Conf. Electr. Electron. Eng., 3rd*, Institute of Electrical and Electronics Engineers, New York, 2006.