# Nonlinear Identification of Ph Process using Support Vector Machine

M.Rajalakshmi[*1], S.Kalyani[2, Ph.D], S.Jeyadevi[3,], Ph.D C.Karthik[4]
Assistant Professor[*1,4], Professor[2,3],
Kamaraj College of engineering Technology[1,2,3],Electronics & Instrumentation Department,Kalasalingam University[4], Electronics & Instrumentation Department, Tamilnadu.

## ABSTRACT

This paper discusses the application of support vector machine in the area of identification of nonlinear dynamical systems. The aim of this paper is to identify suitable model structure for nonlinear dynamic system. In this paper, Adaptive Neuro Fuzzy Inference Systems (ANFIS) and Support Vector Regression (SVR) models are applied for identification of highly nonlinear dynamic process. The results obtained by ANFIS and SVR are compared. The simulation results show that SVR is very effective to identify the nonlinear system.

**Keywords**

pH process, SVR, ANFIS, Dynamic System

## 1. INTRODUCTION

Nonlinear system identification is a very important study in control field. Nonlinear system identification has many difficulties compared to linear system. Among the different nonlinear identification techniques, method based on *ANN* models is gradually becoming established not only in the academia but also in industrial applications. But this method has problems of slow convergence speed and local minima. In the last ten years, the artificial neural systems (ANNs) and more lately, the support vector regression (SVR) have appeared as two attractive tools for non-linear modeling, particularly in situations where the development of phenomenological or conventional regression models becomes impractical or not easily manageable. In latest years, the support vector regression (SVR) [15-17] which is a statistical learning theory becomes more attractiveness over ANN due to its many attractive characteristics and a undertaking empirical performance. The major difference between conventional ANNs and SVM lies in the risk minimization principle. The conventional ANNs apply the empirical risk minimization (ERM) standard to minimize the error on the training data, while SVM follows Structural Risk Minimization (SRM). In SRM, a top bound on the generalization error is minimized. Hence it opposed to the ERM, which minimizes the prediction mistake on the training data. This equips the SVR to generalize the I/O relationships. Thus it can able to understand the relationship during its training stage for making good predictions for new input data set. Although the base of the SVR paradigm was laid down in the mid 1990s, its chemical technology applications such as obvious error detection [18,19] have appeared recently. Although, numerous SVM regression submission investigations are presented by "expert" users having good understanding of SVM methodology. The quality of SVM models mainly depends on a proper setting of SVM parameters. The main issue is how to set these parameters for a given training and testing data set to ensure good generalization performance.

In this paper, we propose a new method for solving the nonlinear identification problems. The method utilizes support vector regression and ANFIS. The rest of the paper is organized as follows. Section 2 and 3 deals with the nonlinear process (pH process) and basic system identification procedure. Section 4 gives a brief introduction to SVR. The design and role of the parameters in the performance of SVR is discussed in section 5. Section 6 deals with neuro-fuzzy model description. Experimental results with discussions are given in section 7. Finally section 8 gives concluding remarks.

## 2. PH PROCESS

The pH process is very important in many industrial applications. It is important to appreciate the diverse nature of these pH control applications because a small amount of acid or base near the set point shows drastic change in pH gain and its characteristics [14]. As shown in Fig. 1, acetic acid is fed to the reactor with a constant flow rate and sodium hydroxide is introduced to the reactor. The mathematical model of the pH neutralization process proposed by McAvoy et al. and reproduced below is used in this work to simulate the pH process. The pH variables used in this study are volume of the continuous stirred tank reactor (V), base flow rate of pH process ($F_b$), acid flow rate ($F_a$) and concentration of acid (Ca) and concentration of base (Cb) [24, 25].

$$V\frac{dxa}{dt} = F_a C_a - (F_a + F_b)X_a \qquad (1)$$

$$V\frac{dxb}{dt} = F_b C_b - (F_a + F_b)X_b \qquad (2)$$

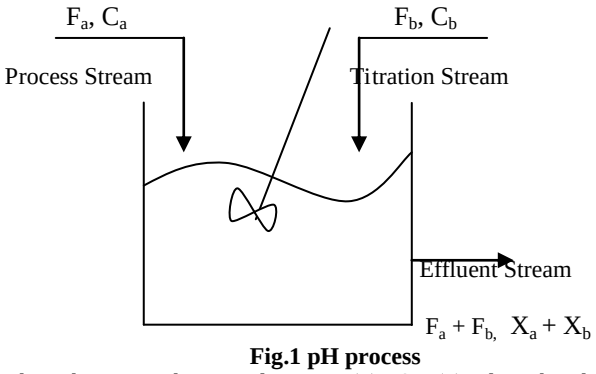



**Fig.1 pH process**

The above mathematical Eqn. (1) & (2) describe how the concentration of the acidic and basic components, $X_a$ and $X_b$ change dynamically with time subject to the input streams, $F_a$ and $F_b$. The

relationship can be described by a nonlinear equation known as the titration or characteristics curve. The titration curve varies in two ways discussed as two cases [2].

Case 1: A weak influent is neutralized by a strong reagent

Case 2: A strong influent is neutralized by a strong reagent.

The word strong and weak is used to characterize the degree of ionic dissociation in an aqueous solution. Strong reagents completely dissociate into their hydrogen or hydroxyl ions whereas weak reagents are only partially ionized [24, 25]. Consider an acetic acid (weak acid) denoted by HAC being neutralized by a strong base NaOH (sodium hydroxide) in water. The reaction between HAC and NaOH:

$$H_2O \rightleftharpoons H^+ + OH^+ \quad (3)$$

$$HAC \rightleftharpoons H^+ + AC^- \quad (4)$$

$$NaOH \rightleftharpoons Na^+ + OH^- \quad (5)$$

Invoking the electro neutrality condition, the sum of the ionic charges in the solution must be zero as shown in Eqn. (6).

$$[N_a{}^+] + [H^+] = [AC^-] + [OH^-] \quad (6)$$

The [X] denotes the concentration of the X ion. The equilibrium relations also hold for water and acetic acid.

$$K_a = \frac{[AC^-][H^+]}{[HAC]} \quad (7)$$

Where, $X_a = [HAC] + [AC^-]$

$X_b = [Na^+]$

Combining Eqn. (6) and (7) gives:

$$[H^+]^3 + [H^+]^2\{K_a + X_b\} + [H^+]\{K_a(X_b - X_a) - K_w K_a = 0 \quad (8)$$

Let $pH = -log_{10}[H^+]$ and $pK_a = -log_{10}K_a$ .The titration curve is represented mathematically is shown in Eqn. (9),

$$X_b + 10^{-pH} - 10^{-pH-14} - \frac{X_a}{1 + 10^{PK_a - pH}} = 0 \quad (9)$$

Where, $K_a$ and $K_b$ is the dissociation constant of acetic acid at 5℃ ($K_a$ -1.778x$10^{-5}$& $K_w = 10^{-14}$ ). Eqn. (9), clearly indicates that a strict static nonlinear relation exists between the states $X_a$ and $X_b$ and the output pH variable.

## 3. BASIC SYSTEM IDENTIFICATION PROCEDURE

System identification (SI) is a methodology for building mathematical models of dynamic systems from experimental data, i.e., using measurements of system input and output (I/O) signals.

The process of SI requires the following steps:

- Measurement of the IO signals of the system in time or frequency domain.
- Selection of a candidate model structure.
- Choice and application of a method to estimate the value for the adjustable parameters in the candidate model structure.
- Validation and evaluation of the estimated model to see if the model is right for the application needs, preferably with a different set of data.

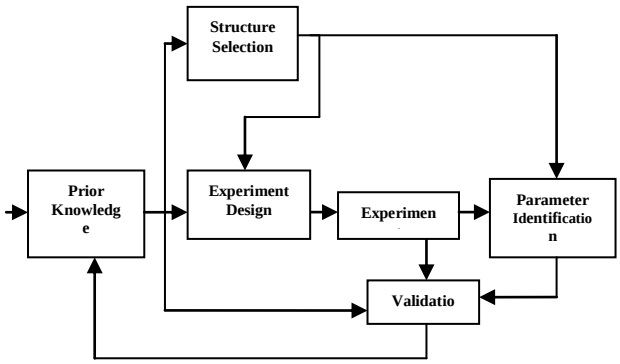



**Fig.2 Flow chart of the system identification process**

Empirical system modeling generally involves the following aspects [3].

1. Structure identification or selection: Model structure selection or identification is closely connected to the control objective, i.e. the model must be appropriate for the specific controller design chosen. It must also reproduce adequately the physical phenomena under consideration, e.g. a Lur'e model is capable of capturing any steady-state behavior, but a Hammerstein model cannot exhibit output multiplicities. Finally, the selected model structure must have an adequate parameterization.
2. Experiment design- Once a model structure has been selected, the extent to which it provides a reasonable approximation to observed system responses can depend very strongly on the input sequences used to elicit these responses. Thus the input sequences must be chosen to obtain the maximum information about the system behavior.
3. Parameter estimation- Given a model structure and input and output sequences, the goal is to find the model parameters that best correspond to the system behavior. An adequate structure selection is important, because it will determine the possible techniques for parameter estimation, as well as the number of parameters to be estimated.
4. Validation of the final model.

## 4. SUPPORT VECTOR RESESSION AND PARAMETERS ANALYSIS

The basic idea in SVR is to map the dataset $\{(x_1, y_1), \ldots \ldots (x_1, y_1)\} c\, R^n \times R$ into a high dimensional feature space via non-linear mapping, wherein they are correlated linearly with the outputs [12]. The SVR formalism considers the following linear estimation function:

$$f(x, w) = \big(w, \emptyset(x)\big) + b \quad (10)$$

Where,

$w \in R^n$ is weight vector,

$b \in R$ is a constant,

$\emptyset(x)$ 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:

$$\text{Min:} \frac{1}{2}\|w\|^2 + C\sum_{i=2}^{l}(\xi_i + \xi_i^*) \quad (11)$$

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 trade-off between the empirical risk and regularization term. Two positive slack variables, $\xi_I$ *and* $\xi_I^*$ , $i$ =1,2,…..,$n$ 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. and a schematic representation of the SVR using ε-insensitive loss function is illustrated in Fig.3.

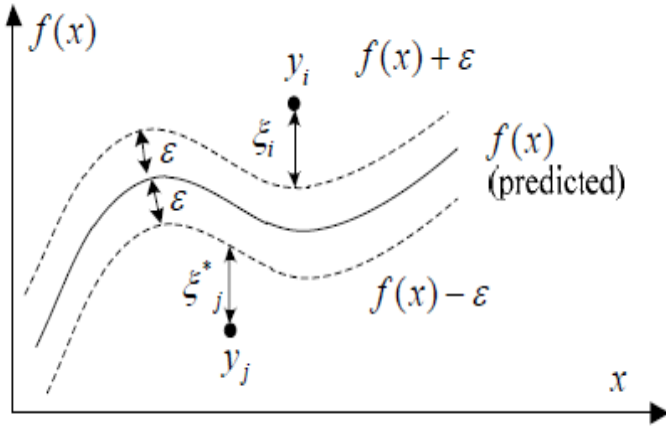



**Fig.3 Sketch map of the SVR using $\varepsilon$-insensitive loss function**

**By applying Lagrange function and Karush-Kuhn-Tucker conditions to the Eqn. (10), yields the following dual optional form**:

Max: $-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^*-\alpha_i)(\alpha_j^*-\alpha_j)K(x_i,x_j)+\sum_{i=1}^{l}y_i(\alpha_i^*-\alpha_i)-\varepsilon i=2l(\alpha i*+\alpha i)$ (12)

Subject to the constraints:

$$\begin{cases}\sum_{i=1}^{k}(\alpha_i-\alpha_i^*)=0\\ 0\le\alpha_i,\alpha_i^*\le C, i=1,\dots\dots,l\end{cases}$$

Where, the Lagrange multipliers $\alpha_i,\alpha_i^*$ satisfy the equality $\alpha_i\alpha_i^*=0$.

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

$f(x)=\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)K(x,x_i)+b$ (13)

Where, $K(x,x_i)$ is called the kernel function and the values of it equals the inner product of two vectors, $x_i$ and$x_j$ , in the feature space $\emptyset(x_i)$ and$\emptyset(x_j)$; that is, $K(x,x_i)=\emptyset(x_i)\emptyset(x_j)$ .

Any function that meets Mercer's condition can be used as the kernel function [13]. The radial basis function (RBF) is applied in this paper as kernel function, which is defined as:

$K(x_i,x_j)=\exp(-\|x_i-x_j\|_2^2\div 2\sigma^2)$ (14)

Where, $\sigma^2$ is the width of the RBF.

In order to design an effective model, the values of the two essential model parameters ($C$ ,ε ) and one kernel function parameter ( $\sigma^2$ ) in SVR have to be chosen optimally 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 was set too small or excessive. The parameter ε defines the non-linear mapping from the input space to some high-dimensional feature space which determines the number of support vectors and $\sigma^2$ reflects correlation of the support vector, which also determines both the generalization capability and the prediction accuracy.

## 5. DESIGN AND DEVELOPMENT OF SUPPORT VECTOR REGRESSION

5.1. Steps in design of SVR

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

5.1.1. Data scaling/preprocessing

The input features in train set and test set needs to be scaled properly before applying SVR [23]. 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. We recommend each attribute to be linearly scaled to the range of [0, 1].

5.1.2. Design of SVM model

5.1.2.1. Choice of Kernel.

The Radial Basis Function (RBF) kernel is chosen as a first choice because of its wide known accuracy. Further, it is capable of handling non-linear relation existing between the class labels and input attributes. The second reason is that RBF kernel, unlike other kernels, has only one kernel parameter, thereby reducing the complexity of the model.

4.1.2.2. Adjusting the Kernel Parameters.

There are two parameters associated with 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) [23]. This can be achieved by different techniques, description of which follows in the next section.

5.1.3. Training and testing the SVM model

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 training phase, the model is validated with test samples to access its overall performance.

5.2. Selection of SVM parameters

SVR model, in our work, uses RBF as the kernel mapping function. Two major parameters associated with SVR model that needs to be optimized are (i) Penalty parameter (C) and (ii) RBF kernel function parameter (γ). Parameter C representing the cost of penalty influences the classification performance [23]. A larger C value gives a good accuracy rate in train phase, but very poor accuracy in test phase. On the contrary, 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 over-fitting, while a disproportionately small value leads to under-fitting [20]. Hence, proper selection of SVR parameters, while designing the SVR model, is highly important to achieve good mean squared error (MSE) and high generalization ability [23]. In this paper, for optimal selection of SVM parameters, grid search (GS) method is used.

5.2.1. Grid Search (GS)

Grid search is the most common and simple method used to determine the appropriate values of parameter C and γ. Grid search

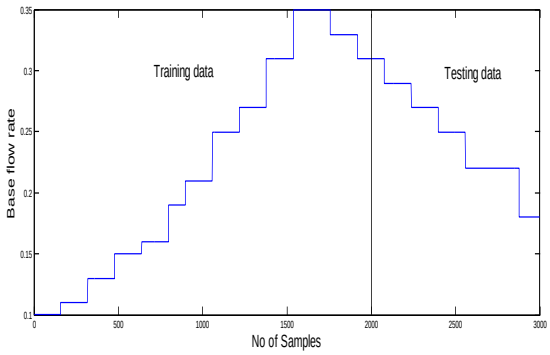
method adopts v-fold cross validation technique. In a v-fold cross validation, we divide the whole training set into v subsets of equal size. Sequentially one subset is tested using the SVM regression trained on the remaining (v-1) subsets. Thus, each instance of the train set is predicted once and the cross-validation accuracy is the percentage of data samples that are correctly classified [21]. In this work, Grid Search using 5-fold cross validation is used. All pairs of (C, γ) were tried and the one with highest cross-validation accuracy was selected. We also realized that trying exponentially growing sequences of C and γ is a practical method to identify optimal parameters. We have used the sequence of C = $\{2^{-5},\ldots..,2^{15}\}$ and γ = $\{2^{-15},\ldots,2^{5}\}$ in the SVR experiment.

## 6. NEURO-FUZZY MODEL

Neuro-Fuzzy system comes from a standard framework referred to as adaptive networks that unifies each neural networks and fuzzy models. The fuzzy model below the framework of adaptive networks is named as ANFIS (Adaptive Network based Fuzzy Inference System), that possess some advantages over neural networks. ANFIS was planned to predict the longer term actions of the method in an endeavor to formalize a systematic approach using the input-output data set [8]. ANFIS needs the antecedent membership functions and fuzzy rules to be outlined before the training. Whenever the selected iteration range is reached or the training error goal is reached the checking data technique is stopped, which is further utilized for ANFIS training method [24]. To stop over fitting of the testing data set the checking data set uses the validation data with suitable format. Over fitting may be detected once the distinction between the output from checking FIS and also the output data starts increasing whereas the training error continues to be decreasing until reaching a minimum price [6].

## 7. RESULTS AND DISCUSSION

The identification process was performed using ANFIS and SVR.A total of 3000 data samples are taken which describes the behavior of the pH process [2]. The data set is then divided into 70% for training and 30% for testing data set. Given training data of input - output pairs where output is "pH of the solution" and input is "Base flow rate". Figure 4 shows the pH of the solution from the selected training and testing data set. The LIBSVM software developed by Chang and Lin has been used for the design and testing of SVM model [22]. Programs for parameter selection of SVM by GS techniques are written in Matlab environment.



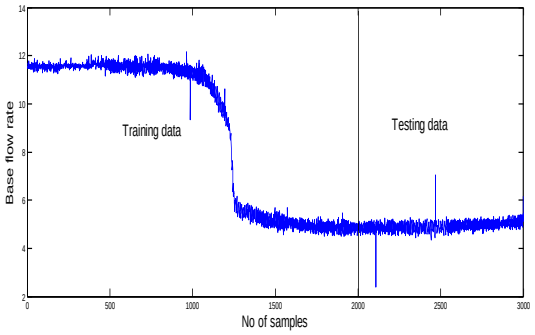




Fig.4 Input and Output Response

In this paper, the SVM model uses Radial Basis Function (RBF) type kernel function. The performance of any SVM regression model will depend strongly on the selection of SVM parameters. There are two parameters associated with SVM:– (i) C, Penalty parameter and (ii) c, RBF Kernel parameter. The selection of parameters is performed by grid search (GS) techniques as discussed in Section 5 and the results are shown in Table 5.The selection of parameters was done by 5 fold cross validation procedure.

**Table 1. Cross Validation Error.**

| γ | C=0.1 | C=1 | C=10 | C=100 | C=1000 | **C=10000** |
|---|---|---|---|---|---|---|
| $2^{-3}$ | 11.87 | 2.25 | 1.59 | 1.44 | 1.17 | 1.108 |
| $2^{-2}$ | 9.87 | 1.80 | 1.53 | 1.22 | 1.12 | 1.09 |
| $2^{-1}$ | 6.68 | 1.60 | 1.37 | 1.104 | 1.125 | 0.977 |
| $2^{0}$ | 3.33 | 1.49 | 1.20 | 1.108 | 1.018 | 0.8755 |
| $2^{1}$ | 1.76 | 1.27 | 1.07 | 1.021 | 0.841 | 0.72 |
| $2^{2}$ | 1.45 | 1.126 | 0.99 | 0.819 | 0.671 | 0.544 |
| $2^{3}$ | 1.15 | 0.94 | 0.796 | 0.619 | 0.526 | 0.52 |
| $2^{4}$ | 0.92 | 0.76 | 0.579 | 0.517 | 0.52 | 0.499 |
| **$2^{5}$** | 0.72 | 0.55 | 0.502 | 0.50 | 0.46 | **0.442** |

Fig.5 shows the SVR model response using Grid Search parameter selection method. The best values of SVR parameters obtained for a maximum cross validation error of 0.442 as seen from Table 1, are $2^{\gamma}$ = 5 and $2^{C}$ = 10,000 (i.e) nearly $2^{14}$.

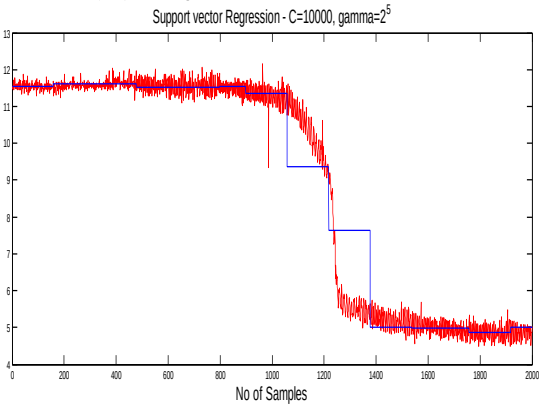



**Fig.5. SVR model Response**

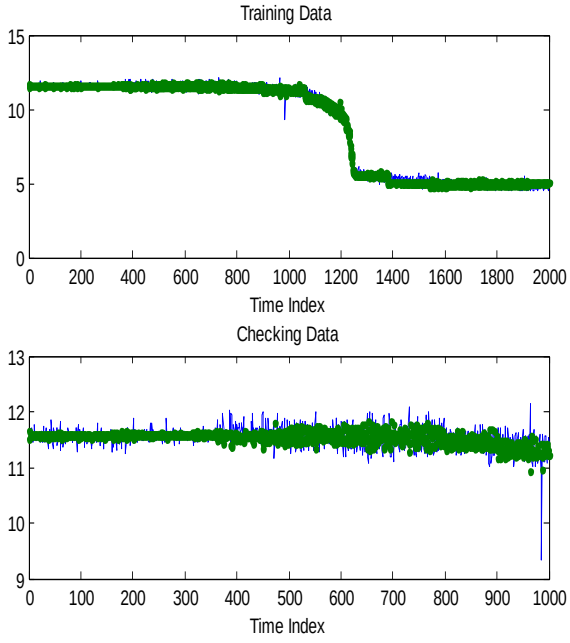
To use ANFIS for system identification, the first step is proper selection of input. That is, to determine which variables should be the input arguments to an ANFIS model. For simplicity, we suppose that there are 10 input candidates (y(k-1), y(k-2), y(k-3), y(k-4), u(k-1), u(k-2), u(k-3), u(k-4), u(k-5), u(k-6)), and the output to be predicted is y(k). A heuristic approach to input selection is called sequential

forward search (seqsrch), in which each input is selected sequentially to optimize the total squared error. This can be done by the function seqsrch. For input selection; another more computationally intensive approach is to perform an exhaustive search (exhsrch) on all possible combinations of the input candidates. The function that performs exhaustive search is exhsrch, which selects 3 inputs from 10 candidates. However, exhsrch usually involves a significant amount of computation if all combinations are tried. For instance, if 3 is selected out of 10, the total number of ANFIS models is M (10, 3) = 120. Fortunately, for dynamical system identification, we know that the inputs should not be from either of the following two sets of input candidates exclusively:

$$Y = \{y(k-1), y(k-2), y(k-3), y(k-4)\} \quad (15)$$

$$U = \{u(k-1), u(k-2), u(k-3), u(k-4), u(k-5, uk-6\} \quad (16)$$

By taking two inputs from Y and one from U, inputs to ANFIS is suitably formed. The total number of ANFIS models is then M (4, 2)*6=36, which is much less. An ANFIS prediction on both training data sets is shown in Fig 6 and the comparison between SVR and ANFIS shown in Table 2.




MSE=0.6798

**Fig.6 Training and Checking of ANFIS model response**

**Table 2. Comparison of MSE**

| Method | Testing + Validation | MSE |
|---|---|---|
| ANFIS | 2000+1000 | 0.442 |
| SVR | 2000+1000 | 0.6798 |

## 5. CONCLUSION

In this paper to overcome the difficulty of neutralized pH value in pH process, the application of ANFIS and SVR models are applied. From the results, it can be concluded that the SVR is the best performed model compared to ANFIS model from the collection of input and output data as shown in Table 2. In SVR there is no local minima and it scales relatively well to high dimensional data. It has trade-off between regression complexity and error can be controlled explicitly via C and γ. Hence over fitting is avoided for any fixed C and γ. Hence it's proved that support vector regression is suitable for highly nonlinear process. Further work will focus on the design, analysis and implementation of suitable controller for the proposed model.

## 6. REFERENCES

[1] K. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," IEEE Transactions on Neural Networks, vol. 1, pp. 4–27, 1990.

[2] K. Valarmathi, D. Devaraj and T. K. Radhakrishnan. "Intelligent techniques for system Identification and controller tuning in ph process". Brazilian Journal of Chemical Engineering Vol. 26, No. 01, pp. 99 - 111, January - March, 2009

[3] Lennart Ljung, System Identification - Theory for the User, 2nd Ed, PTR Prentice Hall, Upper Saddle River, N. J. (1999).

[4] L.R. Medsker, L.C. Jain, Recurrent neural networks: design and applications, Boca Raton, FL: CRC Press, 2000.

[5] Alireza Karbasi, "Comparison of NNARX, ANN and ARIMA Techniques to Poultry Retail Price Forecasting", Thesis, University of Zabol, 2009.

[6] Heba Al-Hiary, Malik Braik, Alaa Sheta, and Aladdin Ayesh, "Identification of A Chemical Process Reactor Using Soft Computing Techniques ", IEEE, 2008.

[7] M. Norgaard, O. Ravn, Poulsen, and L. K. Hansen, "Neural Networks for Modelling and Control of Dynamic Systems". Springer, London, 2000.

[8] L.Ricker,http://depts.washington.edu/control/larry/te/download.html,Computers and Chemical Engineering, 1995.

[9] A. K. Jain, J. Mao, and K. K. Mohiuddin, "Artificial neural networks:A tutorial," IEEE Computer Special Issue on Neural Computing, pp. 31–44,1996.

[10] H.L. Qin and X.B. Li. A chaotic search method for global optimization on tent map. Electric Machines and Control, 2004, 8(1):67-70.

[11] Juan Angel Resendiz-Trejo,Wen Yu,Xiaoou Li " Support Vector Machine for Nonlinear system on-Line identification",International Conference on Electrical and Electronics engineering(ICEEE2006), Mexico.

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

[13] V.N. Vapnik. The Nature of Statistical Learning Theory Springer, New York, 2000.

[14] S R Navghare, Dr. G L Bodhe, Shruti," Design of Adaptive pH Controller using ANFIS", International Journal of Computer Applications (0975 – 8887) Volume 33– No.6, November 2011.

[15] V. Vapnik, The Nature of Statistical Learning Theory, springer Verilag, New York,1995.

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

[17] V. Vapnik, S. Golowich, A. Smola, Adv. Neural Inform. Proces. Syst. 9(1996) 281.

[18] M. Agarwal, A. M. Jade, V. K. Jayaraman, B. D. Kulkarni, Chem. Eng.Prog. 98 (2003) 100.

[19] L. B. Jack, A. K. Nandi, Mech. Sys. Sig. Proc. 16 (2002) 372.

[20] Lin S, Ying K, Chen S, Lee Z. Particle swarm optimization for parameter determination and feature selection of support vector machines. Exp Syst Appl 2008;35(4):1817-24.

[21] Min J, Lee Y. Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Exp Syst Appl 2005;28(4):603–14.

[22] Chang C, Lin C. LIBSVM: a library for support vector machines; 2001<http://www.csie.ntu.edu.tw/cjlin/libsvm>.

[23] S.Kalyani , K.S.Swarup " Pattern analysis and classification for security evaluation in power networks", Electrical power and energy systems 44 (2013) 547–560 Elsevier, 2012.

[24] M.Rajalakshmi, C.Karthik "Nonlinear identification of pH process using NNARX model", CiiT International Journal of Artificial Intelligent Systems and Machine Learning, Print: ISSN 0974 – 9667 & Online: ISSN 0974 – 9543, July 2012.

[25] M.Rajalakshmi, C.Karthik **"**Non Linear Structure Identification Of pH Process**"** in IEEE - ICAESM 2012, EGS Pillay Engineering college, Nagapattinam, cited by Scopus, ISBN 978-1-4673-0213-5,Mar 30-31, page 45.