

# Robust Identification of Hydrocarbon Debutanizer Unit using Radial Basis Function Neural Networks (RBFNNs)

Masih Vafae Ayouri<sup>1\*</sup>, Mehdi Shahbazian<sup>1</sup>, Bahman Moslemi<sup>2</sup>, Mahboobeh Taheri<sup>3</sup>

<sup>1</sup>Department of Instrumentation and Automation Engineering, Petroleum University of Technology, Ahwaz, Iran

<sup>2</sup>Department of Basic Science, Petroleum University of Technology, Ahwaz, Iran

<sup>3</sup>Senior expert in R&D, Sarkhoun & Qeshm Gas treating Company, Bandar Abbas, Iran

\*Corresponding author: [masih.vafae@gmail.com](mailto:masih.vafae@gmail.com)

Received December 23, 2014; Revised January 07, 2015; Accepted January 13, 2015

**Abstract** Radial Basis Function Neural Network (RBFNN) is considered as a good applicant for the prediction problems due to its fast convergence speed and rapid capacity of learning, therefore, has been applied successfully to nonlinear system identification. The traditional RBF networks have two primary problems. The first one is that the network performance is very likely to be affected by noise and outliers. The second problem is about the determination of the parameters of hidden nodes. In this paper, a novel method for robust nonlinear system identification is constructed to overcome the problems of traditional RBFNNs. This method based on using Support Vector Regression (SVR) approach as a robust procedure for determining the initial structure of RBF Neural Network. Using Genetic Algorithm (GA) for training SVR and select the best parameters as an initialization of RBFNNs. In the training stage an Annealing Robust Learning Algorithm (ARLA) has been used for make the networks robust against noise and outliers. The next step is the implementation of the proposed method on the Hydrocarbon Debutanizer unit for prediction of n-butane (C<sub>4</sub>) content. The performance of the proposed method (ARLA-RBFNNs) has been compared with the conventional RBF Neural Network approach. The simulation results show the superiority of ARLA-RBFNNs for process identification with uncertainty.

**Keywords:** robust system identification, RBF Neural Networks, hydrocarbon debutanizer unit, support vector regression, genetic algorithm, annealing robust learning algorithm

**Cite This Article:** Masih Vafae Ayouri, Mehdi Shahbazian, Bahman Moslemi, and Mahboobeh Taheri, "Robust Identification of Hydrocarbon Debutanizer Unit using Radial Basis Function Neural Networks (RBFNNs)." *Journal of Automation and Control*, vol. 3, no. 1 (2015): 10-17. doi: 10.12691/automation-3-1-2.

## 1. Introduction

Artificial neural networks are recognized as major tools for optimization, pattern recognition and nonlinear system identification because of their learning and modeling abilities [1]. System identification is the science of building mathematical models of dynamic systems from Input–Output(I/O) pairs. It can be realized as the interface between the real world of applications and the mathematical world of model abstractions [2]. Robust identification is a method to determine the parameters of neural network when training data contaminated with noise and outliers. The intuitive definition of an outlier (Hawkins, 1980) is "an observation which deviates so much from other observations as to arouse suspicions that it is generated by a different mechanism" [3,4]. However, outliers may occur due to erroneous measurements in measurement device (sensors). When outlier exist in training data set, approximation of predicted models with neural networks is largely deteriorated [5,6].

Between the existing neural networks architectures, the Radial Basis Function Neural Networks (RBFNNs) is considered as a good candidate for approximation and

prediction due to its rapid learning capacity and simpler network structure [7]. Accordingly, the RBFNNs is a popular alternative to the Multilayer Perceptron (MLP). In contrast to MLPs, RBF networks use a localized representation of information. RBFNNs were introduced into the neural network literature by Broomhead and Lowe (1988). Radial basis function (RBF) networks typically have three layers which include an input layer, a hidden layer with a non-linear activation function and a linear output layer [3,4,7,8]. The training procedure of RBF networks is accomplished through the estimation of three kind of parameters, namely the centers and the width of basis function and finally, the synaptic weights [8]. Center selection is an important role in performance of RBFNNs. There are two kinds of methods exist in this scope. One of them is fixed selection of hidden node center (such as random selection from training data and fuzzy c-means clustering) and the other chooses this parameter systematically. SVR approach is a systematic way to define the initial structure of RBFNNs. supervised and unsupervised learning are common learning methods in RBF networks. Unsupervised learning is used to initialize the network parameters, and supervised learning is usually used for the fine tuning of the network parameters [9].

Many learning algorithms in literature have been proposed for training RBFNNs such as hybrid fuzzy clustering approach [8], orthogonal least squares (OLS) [10], gradient descent [11] and extended kalman filter (EKF) [12]. When training data contaminated with noise and outliers, traditional learning methods (least squares based) show poor performance therefore we must approach robust learning algorithm. The following robust training methods are proposed in literature for training RBFNNs: In 1995, Sanchez presented Scaled Robust Loss Function (SRLF) and conjugate gradient method for Robust learning RBFNNs [5]. Chien-Cheng Lee et al. derived new robust objective function from robust statistics (type of Hampel's M-estimator) and back propagation as learning algorithm in order to reduce the influence of outliers in training patterns [13]. Chen-Chia Chuang et al. proposed Annealing Robust Backpropagation Learning Algorithm (ARBP) that adopts the annealing concept into the robust learning algorithms to overcome the problem of training network with noise and outliers [14]. Mei-juan Su et al. (2006) illustrate that by using new SRLF and gradient descent as training method and shows strength robustness against outliers.

In this paper we use  $\varepsilon$ -SVR with Gaussian kernel function for systematic determination of an initial structure of RBFNNs. Another advantage of this method is robust initialization of RBFNNs against data uncertainties. When the initial structure of RBFNNs determined with  $\varepsilon$ -SVR, the next step is training the SVR parameter using Genetic Algorithm to find optimal parameters from SVR. Constructing the network using this parameter and using Annealing Robust Learning Algorithm (ARLA) in training phase. In this stage we used as M-estimator [15] as a popular method to fix the problem of parameter estimation when data contain uncertainty. The least squares based estimators residuals deviated from zero when we faced outliers in the data set therefore M-estimators are used as a robust estimation method to decrease the effect of outliers in the data set. At the point When the residuals goes outside a threshold, the M-estimator suppresses the response instead. Therefore, the M-estimator based error function is more robust to noise and outliers than the Least Mean Square (LMS) based error function. This robustness can be achieved by replacing the MSE criterion with M-estimator [6]. In this study Logistic loss function was adopted to develop a robust RBF network. Thus, the cost function is a special case of the logistic function with properly setting constants in the theory of the M-estimator. We have adopted the concepts of M-estimator and annealing to develop a feedforward network with a robust backpropagation learning algorithm. At the same time, the proposed approach has fast convergence speed and robust against outliers on the identification of the nonlinear dynamic systems contain noise and outliers.

The proposed method will be implemented on the Hydrocarbon Debutanizer unit and the capability of this method for prediction of n-butane (C4) content will be investigated. This paper is organized as follows: In Section 2 full description of the method is proposed. The implementing and its results of the proposed method on a Hydrocarbon Debutanizer unit together with a short description of the process are presented in section 3 and section 4 presents conclusions.

## 2. Methodology Description

### 2.1. Architecture of RBFNNs for Identification of Nonlinear Systems

Assume that the unknown nonlinear dynamical system is expressed by:

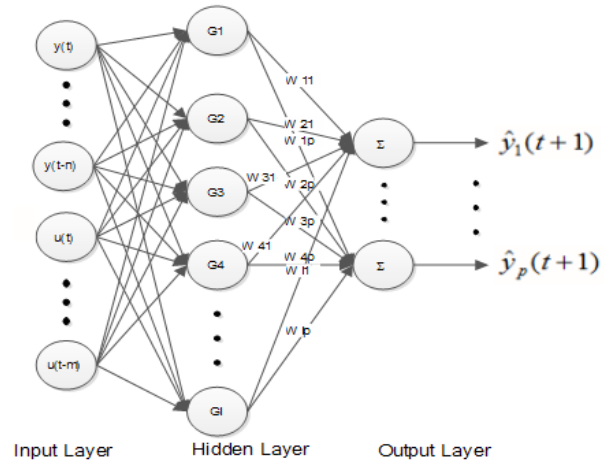
$$y(t+1) = f \left( \begin{matrix} u(t), u(t-1), \dots, u(t-m), y(t), \\ y(t-1), \dots, y(t-n) \end{matrix} \right) \quad (1)$$

Where  $m$  is input lag,  $n$  is output lag and  $f(\cdot)$  is an unknown nonlinear function needed to be identified. The goal of the identification problems is to find a suitable prediction model  $\hat{y}(t+1)$ .

$$\hat{y}(t+1) = \hat{f} \left( \begin{matrix} u(t), u(t-1), \dots, u(t-m), y(t), \\ y(t-1), \dots, y(t-n) \end{matrix} \right) \quad (2)$$

Such that  $\hat{f}$  predicts  $f(\cdot)$ .

A Radial Basis Function Neural Networks (RBFNNs) have three layers, the input layer, the hidden layer with nonlinear activation function, and the linear output layer. The architecture of RBFNNs with Gaussian basis function for identification problem is expressed in Figure 1.



**Figure 1.** Architecture of Radial Basis Function Neural Networks for Identification Problem

A RBFNNs would be shown in the form:

$$\hat{y}_o(t+1) = \sum_{j=0}^L W_{jo} G_j = \sum_{j=0}^L W_{jo} \exp\left(-\frac{X - m_j^2}{2\sigma_j^2}\right) \quad (3)$$

For  $0 \leq j \leq L$  and  $1 \leq o \leq p$ .

Where  $X = \begin{pmatrix} u(t), u(t-1), \dots, u(t-m), \\ y(t), y(t-1), \dots, y(t-n) \end{pmatrix}$  is the input vector,  $\hat{y}_o(t)$  is the  $o$ th network predicted output,  $W_{jo}$  is the synaptic weight,  $G_j$  is the Gaussian function at the  $j$ th hidden layer neuron,  $m_j$  and  $\sigma_j$  are the center and width of Gaussian function, respectively, and  $L$  is equal to the number of hidden layer neurons [16].

When utilizing an RBFNNs for the identification of nonlinear dynamical system, our objective is to determine

the values of RBFNNs parameters ( $L, W_{jo}, m_j$ ) to minimize the following objective function:

$$\sum_{i=1}^N \left\| y_i - \hat{y}_i \right\|^2 \quad (4)$$

Where  $N$  is the number of training data pairs (I/O pairs),  $y_i$  is the output of unknown nonlinear dynamical systems and  $\hat{y}_i$  is the prediction value from identified model. For determining better combination of optimal parameters, iterative methods can be applied. When constructing a RBFNNs, it is significant to define the net structure and initializes the network parameters. Most of the conventional RBFNNs approaches (Least squares based) are easily influenced by long tail noise and outliers, therefore, robust approaches are given to overcome the problem of traditional RBFNNs. These robust radial basis functions approaches mainly focus on robust learning algorithms. These algorithms assume the concept of robust estimators in the training phase. For construction the network structure we use  $\varepsilon$ -SVR approach for determining initial network structure and parameter of  $\varepsilon$ -SVR method tuned with Genetic Algorithm (GA) and using iterative methods for training RBFNNs parameters. How to use the  $\varepsilon$ -SVR-GA method to get the optimal structure of RBFNNs will be exemplified in the succeeding segment.

## 2.2. $\varepsilon$ -SVR-GA Approach for Structure Selection of RBFNNs

Suppose we are given training I/O pair data which includes  $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$  where  $X$  denotes the input space pattern and  $Y$  means output data. That is, regression function in the SVR approach is approximated by the following function as:

$$f(x, \theta) = \sum_{k=1}^N \theta_k \Phi(x_k) + b \quad (5)$$

In 1995 Vapnik [17] proposed  $\varepsilon$ -SVR as a solution for the problem is to find  $f(x, \theta)$  that minimize the subsequent risk function

$$R(\theta) = \frac{1}{N} \sum_{i=1}^N L_{\varepsilon}(y_i - f(x, \theta)) \quad (6)$$

Subject to the constraint

$$\theta^2 < C, \quad (7)$$

Where  $C$  is a constant and  $L_{\varepsilon}(\cdot)$  is the  $\varepsilon$ -insensitive loss function defined as:

$$L_{\varepsilon}(e) = \begin{cases} 0 & \text{for } |e| < \varepsilon \\ |e| - \varepsilon & \text{otherwise} \end{cases} \quad (8)$$

Where  $e = y(i) - \hat{y}(i)$  is the error between  $i$ th desired output and  $i$ th output of RBFNNs and  $\varepsilon$  is a nonnegative number.

Since an SVR approach with the insensitive loss function provides an estimated function within zone, the

initial construction of the RBFNNs can be obtained by the SVR approach that evenhandedly supplies better initialization to learning algorithm. Three kinds of parameter which can be chosen appropriately to determine the best initialization parameters of RBFNNs consist of penalty factor ( $C$ ), epsilon (coefficient of the SVR loss function) and sigma (width of kernel function). Moreover, an SVR approach with the  $\varepsilon$ -insensitive loss function can make use of a small subset of the training data, called the support vectors (SVs), to approximate the unknown functions within a tolerance band  $\varepsilon$ . The numbers of SVs are controlled by the values of tolerance band  $\varepsilon$  [17,18,19].

Shows the situation graphically. Only the points outside the shaded region contribute to the cost insofar, as the deviations are penalized in a linear fashion. It turns out that the optimization problem can be solved more easily in its dual formulation and solving this problem [19] with the Lagrange multipliers method and minimized above loss function leads to following dual optimization problem [4,16,17,19].

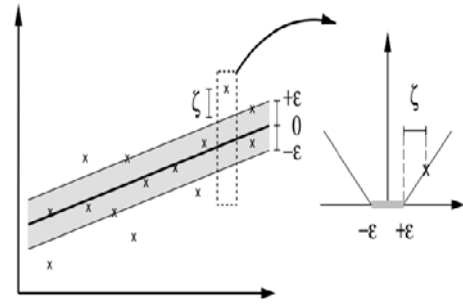


Figure 2. The soft margin loss setting corresponds for a linear SVR [19]

Minimize

$$Q(\alpha, \alpha^*) = \varepsilon \sum_{r=1}^N (\alpha_r + \alpha_r^*) - \sum_{r=1}^N y_r (\alpha_r - \alpha_r^*) + \frac{1}{2} \sum_{r,s=1}^N (\alpha_r^* - \alpha_r) (\alpha_s^* - \alpha_s) \ll \Phi(x_r) \cdot \Phi(x_s) \gg \quad (9)$$

Subject to the constraint

$$\sum_{r=1}^N \alpha_r^* = \sum_{r=1}^N \alpha_r, 0 < \alpha_r^*, \alpha_r < C \text{ for } r = 1, 2, \dots, N \quad (10)$$

Where  $Q$  is a cost function in the  $\varepsilon$ -SVR,  $\alpha, \alpha^*, \alpha_r, \alpha_r^*, \alpha_s, \alpha_s^*$  are all the nonnegative Lagrange multipliers,  $r$  and  $s$  are all indexes,  $X$  is input and  $Y$  is output. The inner product of basis function is replaced via kernel function

$$K(x_r, x_s) = \ll \Phi(x_r) \cdot \Phi(x_s) \gg \quad (11)$$

The kernel function determines the smoothness properties of solutions and should reflect a previous knowledge of the data. In the literature Gaussian kernel function often used. Hence equation (9) rewritten as

$$Q(\alpha, \alpha^*) = \varepsilon \sum_{r=1}^N (\alpha_r + \alpha_r^*) - \sum_{r=1}^N y_r (\alpha_r - \alpha_r^*) + \frac{1}{2} \sum_{r,s=1}^N (\alpha_r^* - \alpha_r) (\alpha_s^* - \alpha_s) K(x_r, x_s) \quad (12)$$

Therefore, the solution of the SVR method [15] is in the form of the following linear expansion of kernel function:

$$f(x, \alpha, \alpha^*) = \sum_{k=1}^N (\alpha_k^* - \alpha_k) K(x, x_k) + b \quad (13)$$

Note that only some of  $(\alpha_k^* - \alpha_k)$ s are not zeros and the corresponding vectors  $x_k$  are termed support vectors (SVs). That is,  $\#SVs \leq N$ , SVs is number of SVs. In this paper, we use the Gaussian kernel function is used in kernel of  $\varepsilon$ -SVR and then relation (13) can be rewritten as

$$f(x, \lambda) = \sum_{k=1}^{\#SVs} \lambda_k \exp\left(-\frac{x-x_k^2}{2\sigma_k^2}\right) + b \quad (14)$$

Where  $\lambda_k = (\alpha_k^* - \alpha_k) \neq 0$  and  $x_k$  are SVs and  $b$  is bias of the network.

To design an effective SVR model, the values of SVR parameters have to be chosen carefully. For this purpose, we use Genetic Algorithm as an advanced tool to find the best solution. The concept of GA was developed by Holland and his coworkers in the 1960s and 1970s [20]. Genetic algorithms (GAs) are well appropriate for searching global optimal values in complex search space (multi-modal, multi-objective, non-linear, discontinuous, and highly constrained space), combined with the fact that they work with raw objectives only when compared with conventional techniques [21,22].

The procedure of GA for finding the best solution can be summarized below:

1. Choose a randomly generated population.
2. Calculate the fitness of each chromosome in the population.
3. Create the offspring by the genetic operators: selection, crossover and mutation.
4. Check the termination condition. If the new population does not satisfy the termination condition, repeat steps 2 up to 4 for the generated offspring as a new starting population

The SVR parameters are as follows:

- Regularization parameter  $C$ , which determines the tradeoff cost between minimizing the training error and minimizing the complexity of the model.
- Spread parameter ( $\sigma^2$ ) of the kernel function which defines the width of Gaussian kernel function.
- Epsilon parameter ( $\varepsilon$ ) of the loss function which determine the number of SVs. A small  $\varepsilon$  value allows more points to be outside the  $\varepsilon$ -tube and results in more SVs and a large  $\varepsilon$  value results in less SVs and probably in a smoother regression function.

### 2.3. Annealing Robust Learning Algorithm (ARLA) for Training RBFNNs

An ARLA is proposed as a learning algorithm for training RBFNNs. An important feature of ARLA that adopts the annealing concept in the cost function of the robust backpropagation learning algorithm is suggested in [4]. Therefore, the ARLA can overcome the existing problems in the traditional backpropagation learning algorithm when data contaminated with outliers. The

intuitive definition of an outlier (Hawkins, 1980) is “an observation which deviates so much from other observations as to arouse suspicions that it is generated by a different mechanism” [3,4]. However, outliers may occur due to erroneous measurements in measurement device(sensors). Outliers in training data set can cause substantial deterioration of the approximation realized by a neural network architecture [5,6]. Statistical techniques are often used which are sensitive to such outliers, and negative results may have been affected by them, and the most robust and resistant methods have been developed since 1960 and less sensitive to outliers. Robustness is the key issue for system identification.

A cost function for ARLA is defined as

$$E_{ARLA}(t) = \frac{1}{N} \sum_{i=1}^N \rho[e_i(t); \beta(t)] \quad (15)$$

Where

$$e_i(t) = y_i - f(x_i) = y_i - \sum_{j=1}^L W_j \exp\left\{-\frac{\|x - m_j\|^2}{2\sigma_j^2}\right\} \quad (16)$$

$t$  is number of epochs,  $e_i(t)$  is the error between the  $i$ th desired output and the  $i$ th output of the trained network at epoch  $t$ ,  $\beta(t)$  is a deterministic annealing schedule acting like the cut-off points (scale estimator) and  $\rho(\cdot)$  is Logistic loss function and defined as

$$\rho(e_i, \beta) = \frac{\beta}{2} \ln\left(1 + \frac{e_i^2}{\beta}\right) \quad (17)$$

Our objective function defined as

$$J(h) = \frac{1}{N} \sum_{i=1}^N \rho[e_i(t); \beta(t)] \quad (18)$$

In a Logistic loss function using annealing schedule (Scale Estimator) as a threshold for the rejection of outliers. Usually, the scale estimator can be chosen in two ways. One is to obtain the scale estimator based on some robust statistic theories such as the median errors and the median of absolute deviation (MAD) and the other way of defining the scale estimator is to count out the pre-specified percentage of points [14]. This loss function with scale estimator degrade the effect of noise and outliers in dataset.

Based on Gradient-Descent as a type of learning algorithm the parameters of RBFNNs ( $W_{jo}, m_j, \sigma_j$ ) updated as

$$\Delta W_{jo} = -\eta \frac{\partial J}{\partial W_{jo}} = -\eta \sum_{i=1}^N \varphi(e_i, \beta) \frac{\partial e_i}{\partial W_{jo}} \quad (19)$$

$$\Delta m_j = -\eta \frac{\partial J}{\partial m_j} = -\eta \sum_{i=1}^N \varphi(e_i, \beta) \frac{\partial e_i}{\partial m_j} \quad (20)$$

$$\Delta \sigma_j = -\eta \frac{\partial J}{\partial \sigma_j} = -\eta \sum_{i=1}^N \varphi(e_i, \beta) \frac{\partial e_i}{\partial \sigma_j} \quad (21)$$

$$\varphi(e_i; \beta) = \frac{\partial \rho(e_i, \beta)}{\partial e_i} = \frac{e_i}{1 + (e_i)^2 / \beta(h)} \quad (22)$$



Where  $\eta$  is a learning rate,  $\varphi(\cdot)$  is usually called the influence function. When outliers exist in dataset, they have a major impact on the predicted results. In the ARLA, the properties of the annealing schedule  $\beta(h)$  have [12]:

- $\beta_{initial}, \beta(h)$  has large values for the first epochs;
- $\beta(h) \rightarrow 0^+$  for  $h \rightarrow \infty$ ;
- $\beta(h) = \frac{k}{h}$  for any h epoch, where k is a constant.

The procedure of robust learning algorithm for RBFNNs described as follows:

### 2.3.1. Robust learning Procedure

**Step 1:** Initializing the structure of RBFNNs using  $\varepsilon - SVR - GA$  as shown in section 2.2.

**Step 2:** Compute the predicted output of network and its error for training data.

**Step 3:** Find the values of annealing schedule  $\beta(h) = \frac{k}{h}$  for each epoch, where  $k = 2 * \max\{|e_i|_{initial}\}$ .

**Step 4:** Update the network parameters such as (synaptic weights ( $W_{jo}$ ), width of Gaussian function ( $\sigma_j$ ) and the centers ( $m_j$ )).

**Step 5:** Compute the robust cost function explain in (18).

**Step 6:** If the stopping criteria are not satisfied, then go to Step 2; otherwise terminate the training stage.

## 3. Case Study, Implementation and Results

### 3.1. Hydrocarbon Debutanizer Process

Debutanizer column is an essential component of the gas recovery unit in the petroleum refineries. It is used to recover butane from the light end product containing C2–C8 hydrocarbons [23]. Nevertheless, the debutanizer column demonstrates high dimensional coupling with severe nonlinearity and cooperative set of operational constraints. A debutanizer is a type of fractional distillation column used to recover light gases (C1–C4) and Liquefied Petroleum Gas (LPG) from the overhead distillate before producing light naphtha during the refining process. Distillation is the process of heating a liquid to vapor and condensing the vapors back to liquid in order to separate or purify the liquid. The main equipment of this process consist of a distillation column, reboiler and condenser. The debutanizer condenser condenses the overhead vapor and the debutanizer overhead pressure control valves controls the overhead system. The reflux from the top of the debutanizer consists of the collected condensed hydrocarbon (light hydrocarbon). There are three manipulated variables for the distillation column which are the feed flow rate, reflux flow rate and reboiler duty. The feed flow rate controls the feed to the column, the debutanizer reboiler control valve controls the reboiler temperature while the debutanizer bottom level controller controls the bottom product (Heavier Hydrocarbon) level. The debutanizer reflux control valve controls the ratio of the liquid and distillate

flow rate at the top of the column. This column is a challenging process because a highly nonlinear multivariable process, involves a great deal of interactions between the variables, has lag in many variable of the control system (Dynamic System), all of which makes it difficult system to be modeled by linear techniques [24]. Therefore we use nonlinear models for the prediction of properties and identification of this process. At this time, the composition of the debutanizer products is measured through tedious and time consuming laboratory measurements. Consequently, prediction of product quality is an important issue for solving these problems. To maintain the product quality at a desired level, it is necessary to predict the top and bottom compositions of the debutanizer column quickly with a high degree of precision[25].Therefore robustness issues are important for contaminated data. The debutanizer column considered in this study is a fifteen-stage multi-component distillation column fed by two input streams consisting of a mixture of light hydrocarbons [26]. The two input feed streams to debutanizer are compositions of the light hydrocarbons containing i-butane, n-butane, i-butene, i-pentane, n-pentane, n-hexane, n-heptane and n-octane. The Process Flow Diagram (PFD) of this process presented in the Figure 3.

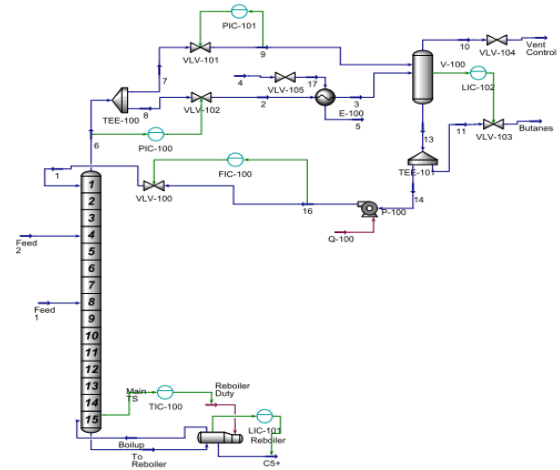


Figure 3. PFD of Debutanizer column

### 3.2. Results and Discussion

To demonstrate the validity of the proposed ARLA-RBFNNs identification method in practical modeling applications, the proposed method has been used for identification of Debutanizer unit.

In the identification problems, the training input-output data are obtained by feeding a signal  $x(t)$  to the MISO system and measuring corresponding outputs  $y(t+I)$ . An objective of the debutanizer is minimizing the n-butane (C4) content in the debutanizer bottom flow, this concentration is chosen as the output variable which must be estimated by the proposed algorithm.

After gathering 1000 data samples from the debutanizer, 60 percent of given data selected randomly picked up as the training data (600 sample) for modeling and the remaining samples were used for testing purpose. some of the original data set are replaced by artificially generated Gaussian noise, Cauchy noise and outliers for evaluating the robustness of the proposed algorithm against

uncertainty. The following Figures Demonstrates the data set contaminated with artificially Gaussian noise and Cauchy noise with outliers.

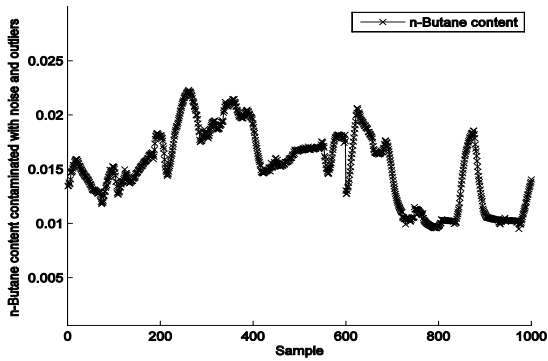


Figure 4. Contaminated data with Gaussian noise

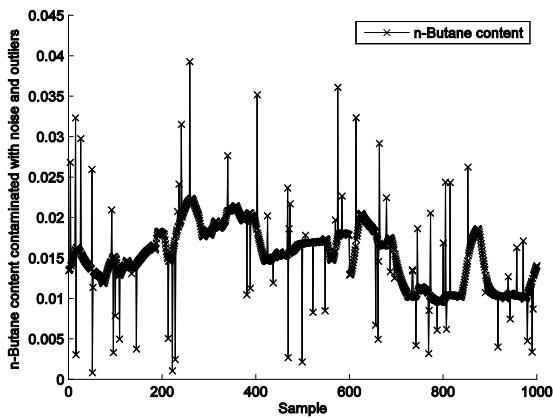


Figure 5. Contaminated data with Gaussian noise and 7% outlier (STD=5)

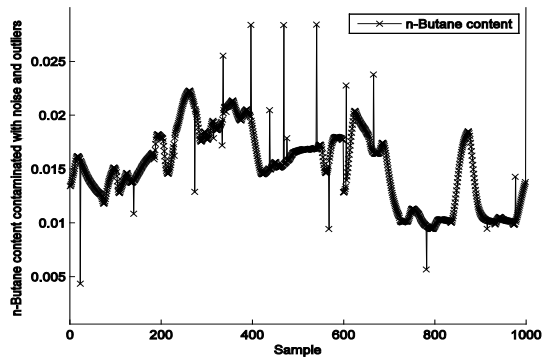


Figure 6. Contaminated data with Cauchy noise

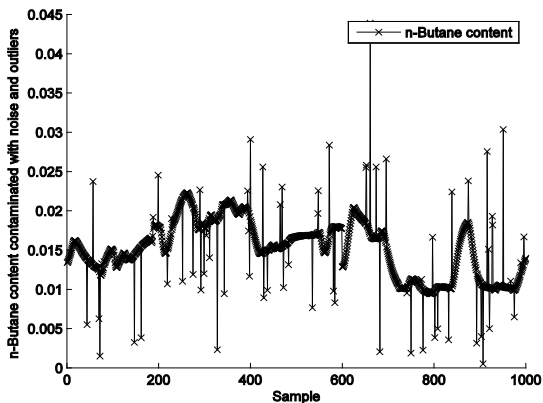


Figure 7. Contaminated data with Cauchy noise and 7% outlier (STD=5)

Figure 4, shows contaminated data with Gaussian noise, Figure 5, illustrates contaminated data with Gaussian noise and 7% outlier when standard deviation (STD) is equal to 5, Figure 6, represents Contaminated data with Cauchy and Figure 7, shows the Contaminated data with Cauchy noise and 7% outlier when STD is equal to 5.

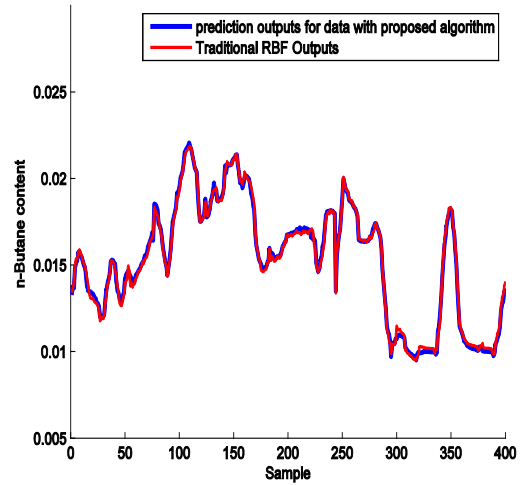


Figure 8. Output prediction with the proposed algorithm versus traditional RBF for Gaussian noise

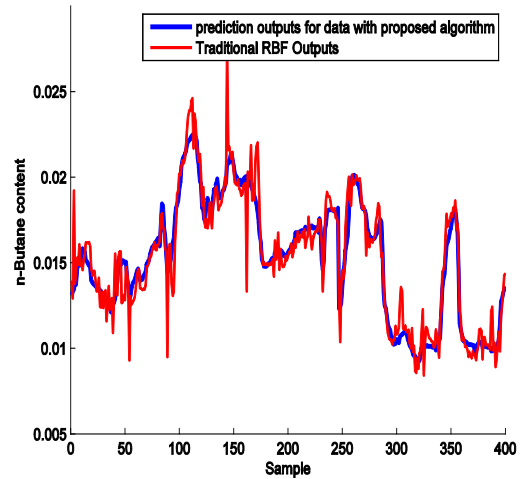


Figure 9. Output prediction with the proposed algorithm versus traditional RBF for Gaussian noise and 7% outliers (STD=5)

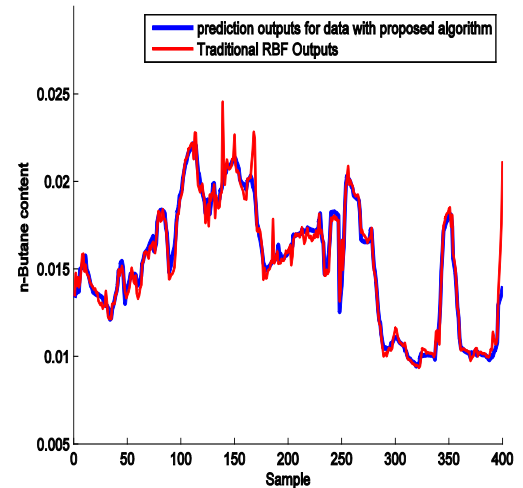
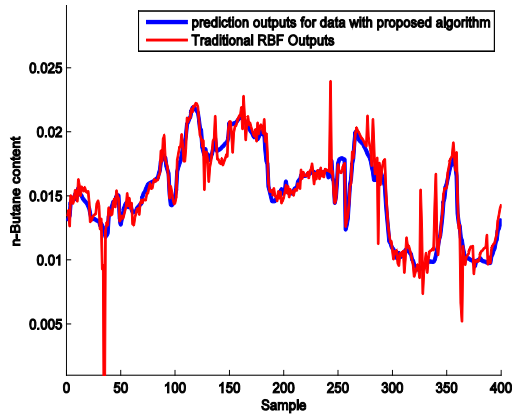


Figure 10. Output prediction with the proposed algorithm versus traditional RBF for Cauchy noise



**Figure 11.** Output prediction with the proposed algorithm versus traditional RBF for Cauchy noise and 7% outliers (STD=5)

Inputs applied to  $\varepsilon$ -SVR which parameters optimized with GA and initial parameters for RBF structure selection are determined. After this step, using robust learning algorithm to train the proposed network and then determined a prediction outputs. Simulation results show the effectiveness of the proposed algorithm versus traditional RBF. Traditional RBF consists of **MATLAB** Toolbox for creating RBFNNs with **newrb** function. The width of basis functions and number of neurons in this function replace with calculated value from GA-SVR (initialization method).

**Table 1.** Comparison results of the prediction error for the proposed method and traditional RBF in various case studies

	Proposed method		Traditional RBFNNs	
	RMSE	CC	RMSE	CC
Output data contaminated with Gaussian noise	0.00022869	0.9977	<b>0.00012979</b>	<b>0.9992</b>
Output data contaminated with Gaussian noise and 7% outliers with STD=5	<b>0.00025760</b>	<b>0.9972</b>	0.0011	0.9530
Output data contaminated with Cauchy noise	<b>0.00025001</b>	<b>0.9977</b>	0.00068912	0.9757
Output data contaminated with Cauchy noise and 7% outliers with STD=5	<b>0.00028294</b>	<b>0.9978</b>	0.0014	0.9191

The results from the demonstrate that when the data set is contaminated with Gaussian noise the proposed method reveals a good performance. Although the conventional RBF approach (least square estimators) has slightly better performance; this is because the conventional approach is an optimal estimator (Maximum Likelihood Estimator) for Gaussian noise. Since, the data set contain non-Gaussian (Cauchy) noise and outliers, performance of the conventional least squares estimator is getting worse and robust training methods have been used for achieving better performance.

Evidently, the proposed method (ARLA-RBFNNs) has superior performance and is more robust against outliers than traditional RBF.

## 4. Conclusion

In this paper, the radial basis function networks with the support vector regression and the robust learning algorithm is developed for the system identification of nonlinear plant with outliers. To design an effective SVR model, suitable values of parameters in SVR have been chosen using GA. The SVR approach has been used to determine the number of hidden nodes, the initial parameters of the kernel, and the initial weights of the proposed neural networks. Using the annealing robust learning algorithm to adjust the parameters of the model,

The Figure 8 - Figure 11 depicts the comparison results between proposed algorithm and traditional RBF.

The simulation results show that the proposed approach significantly improves the robustness against outliers versus traditional methods and follows real value very smoother than conventional approach.

The Root Mean Square Error (RMSE) and Correlation Coefficient (CC) of test data are used to measure the performance of the learned networks. The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (23)$$

Where  $y_i$  is the desired output and  $\hat{y}_i$  is the output of the proposed method. And the value of the Correlation Coefficient(CC) is given by:

$$CC = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{Y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{Y}})^2 (y_i - \bar{y})^2}} \quad (24)$$

Where  $\bar{y}$  is the mean desired output and  $\bar{\hat{Y}}$  is the mean predicted output of the proposed method. **Table 1** shows the comparison result of the prediction performance for proposed network and traditional RBFNNs.

The successful results indicated that the proposed ARLA-RBFNNs method can be used as a reliable technique for the system identification from data contamination with outliers. This algorithm has been implemented on Hydrocarbon Debutanizer unit for prediction of normal butane (C4) concentration. The simulation results show that the proposed approach overcome the problems of identification with outliers and have more accurate and smoother results than the conventional approaches.

## Acknowledgement

The authors would like to thank the technical and financial supports provided by the National Iranian Gas Company-Sarkhoon & Qeshm Gas treating Company (NIGC-SQGC).

## References

- [1] G. Apostolikas and S. Tzafestas, "On-line RBFNN based identification of rapidly time-varying nonlinear systems with optimal structure-adaptation," *Mathematics and Computers in Simulation*, vol. 63, pp. 1-13, 2003.
- [2] L. Ljung, "Perspectives on system identification," *Annual Reviews in Control*, vol. 34, pp. 1-12, 2010.
- [3] C.-N. Ko, "Identification of nonlinear systems with outliers using wavelet neural networks based on annealing dynamical learning

- algorithm," *Engineering Applications of Artificial Intelligence*, vol. 25, pp. 533-543, 2012.
- [4] C.-C. Chuang, J.-T. Jeng, and P.-T. Lin, "Annealing robust radial basis function networks for function approximation with outliers," *Neurocomputing*, vol. 56, pp. 123-139, 2004.
- [5] A. Sánchez and V. David, "Robustization of a learning method for RBF networks," *Neurocomputing*, vol. 9, pp. 85-94, 1995.
- [6] C.-C. Lee, Y.-C. Chiang, C.-Y. Shih, and C.-L. Tsai, "Noisy time series prediction using M-estimator based robust radial basis function neural networks with growing and pruning techniques," *Expert Systems with Applications*, vol. 36, pp. 4717-4724, 2009.
- [7] D. S. Broomhead and D. Lowe, "Radial basis functions, multi-variable functional interpolation and adaptive networks," DTIC Document 1988.
- [8] A. D. Niros and G. E. Tsekouras, "A novel training algorithm for RBF neural network using a hybrid fuzzy clustering approach," *Fuzzy Sets and Systems*, vol. 193, pp. 62-84, 2012.
- [9] K.-L. Du and M. N. Swamy, *Neural networks in a softcomputing framework*: Springer, 2006.
- [10] S. Chen, C. F. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *Neural Networks, IEEE Transactions on*, vol. 2, pp. 302-309, 1991.
- [11] N. B. Karayiannis, "Gradient descent learning of radial basis neural networks," in *Neural Networks, 1997., International Conference on*, 1997, pp. 1815-1820.
- [12] D. Simon, "Training radial basis neural networks with the extended Kalman filter," *Neurocomputing*, vol. 48, pp. 455-475, 2002.
- [13] C.-C. Lee, P.-C. Chung, J.-R. Tsai, and C.-I. Chang, "Robust radial basis function neural networks," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 29, pp. 674-685, 1999.
- [14] C.-C. Chuang, S.-F. Su, and C.-C. Hsiao, "The annealing robust backpropagation (ARBP) learning algorithm," *Neural Networks, IEEE Transactions on*, vol. 11, pp. 1067-1077, 2000.
- [15] P. J. Huber, *Robust statistics*: Springer, 2011.
- [16] Y.-Y. Fu, C.-J. Wu, C.-N. Ko, and J.-T. Jeng, "Radial basis function networks with hybrid learning for system identification with outliers," *Applied Soft Computing*, vol. 11, pp. 3083-3092, 2011.
- [17] V. N. Vapnik and V. Vapnik, *Statistical learning theory* vol. 2: Wiley New York, 1998.
- [18] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, pp. 199-222, 2004.
- [19] J. H. Holland, *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*: U Michigan Press, 1975.
- [20] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," *Reliability Engineering & System Safety*, vol. 91, pp. 992-1007, 2006.
- [21] B. Üstün, W. Melsens, M. Oudenhuijzen, and L. Buydens, "Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization," *Analytica Chimica Acta*, vol. 544, pp. 292-305, 2005.
- [22] A. K. Jana, A. N. Samanta, and S. Ganguly, "Nonlinear state estimation and control of a refinery debutanizer column," *Computers & Chemical Engineering*, vol. 33, pp. 1484-1490, 2009.
- [23] N. Mohamed Ramli, M. Hussain, B. Mohamed Jan, and B. Abdullah, "Composition Prediction of a Debutanizer Column using Equation Based Artificial Neural Network Model," *Neurocomputing*, vol. 131, pp. 59-76, 2014.
- [24] M. Behnasr and H. Jazayeri-Rad, "Robust data-driven soft sensor based on iteratively weighted least squares support vector regression optimized by the cuckoo optimization algorithm," *Journal of Natural Gas Science and Engineering*, vol. 22, pp. 35-41, 2015.
- [25] S. Ferrer-Nadal, I. Yélamos-Ruiz, M. Graells, and L. Puigjaner, "On-line fault diagnosis support for real time evolution applied to multi-component distillation," *Computer Aided Chemical Engineering*, vol. 20, pp. 961-966, 2005.