A Hybrid Approach Based on the Combination of Adaptive Neuro-Fuzzy Inference System and Imperialist Competitive Algorithm: Oil Flow Rate of the Wells Prediction Case Study

Shahram Mollaey Berneti

Department of Electrical Engineering, Science and Research Branch, Islamic Azad University
Mazandaran, Sari, Iran

Received 15 March 2012
Accepted 15 August 2012

Abstract

In this paper, a novel hybrid approach composed of adaptive neuro-fuzzy inference system (ANFIS) and imperialist competitive algorithm (ICA) is proposed. The imperialist competitive algorithm (ICA) is used in this methodology to determine the most suitable initial membership functions of the ANFIS. The proposed model combines the global search ability of ICA with local search ability of gradient descent method. To illustrate the suitability and capability of the proposed model, this model is applied to predict oil flow rate of the wells utilizing data set of 31 wells in one of the northern Persian Gulf oil fields of Iran. The data set collected in a three month period for each well from Dec. 2002 to Nov. 2010. For the sake of performance evaluation, the results of the proposed model are compared with the conventional ANFIS model. The results show that the significant improvements are achievable using the proposed model in comparison with the results obtained by conventional ANFIS.

Keywords: ANFIS, Imperialist Competitive Algorithm, Gradient Descent, Oil Flow Rate.

1. Introduction

Oil flow rate of the wells is effective parameters on reservoir behavior simulation, field development and production allocation. Due to the complex nature of oil-water-gas three-phase flow in wells and pipes, the development, evaluation, and use of oil flow metering systems have been a major focus for the oil industry. Result of this focus is two types meter: conventional methods and multiphase flow meters (MFM). In the conventional methods, multiphase flow is split into three phases and flow of each phase is then measured by means of single-phase instrumentation at the outlets of the separator. Conventional methods are expensive to fabricate, operate and maintain; cause delays; and are not at all ideal for fast decision-making.

In contrast to conventional methods, MFM measure individual phase flow rates without first separating the phases. Due to existence of radioactive sources, manually control, expensive cost of each MFM and long distance between wells, companies cluster the wells and use MFM for these clusters in unmanned location. Consequently, data acquisition and monitoring of oil flow rate of the wells is done discretely. Therefore, it is necessary to develop a more quick and cheap ways to overcome these drawbacks. At the same time, temperature and pressure of lines which act important rule in production process from reservoir to production and separation unit is collected and stored from general existing sensor. This set of data can be utilized as input to build a model to predict oil flow rate of the wells by soft computing techniques. This model is safe for both personnel and the environment, relatively cheap, easily maintained and supervised remotely, and both installation and usage can be made very easy.

Soft computing techniques, are innovative approaches to construct a computationally intelligent system which

*sh.m.berneti@gmail.com
can be used to tackle imprecision and uncertainty involved in a complex nonlinear system. Typically, these techniques consist of several computing methodologies, including neural networks, fuzzy systems and genetic algorithm that have seen increasing interest in recent decade. One popular and robust soft computing method is neuro-fuzzy technique which is the integration of fuzzy system and neural network. A specific form of neuro-fuzzy systems is the adaptive neuro-fuzzy inference system (ANFIS), which has shown significant results in modeling nonlinear relationships. ANFIS is a multilayer neural network-based fuzzy system for modeling the complex systems where past samples are used to forecast the sample ahead. ANFIS combines the self-learning ability of NN with the linguistic expression function of fuzzy inference. It uses a hybrid learning procedure which is a combination of the gradient descent technique and least-squares method to determine the optimal distribution of membership functions and rules parameters, respectively. Using the steepest descent optimization technique in gradient descent training to minimize the error function, allowing the algorithm to easily get trapped by the local minima or flat areas of the optimization surface and the global solution may never be found. In addition, the performance of the gradient descent learning algorithm depends on the initial values of system parameters. In contrast, evolutionary algorithms are stochastic population-based search techniques that provide a more robust and efficient method for solving complex real-world problems. In the past few years, some investigations into ANFIS training using evolutionary algorithm have been published. Ho et al. used genetic learning algorithm (GA) in the ANFIS to determine the suitable membership functions. The researches by Shoorehdeli et al., Oliveira and Schirru and Pousinho et al. have proposed using particle swarm optimization (PSO) in the field of ANFIS to training the parameters in the antecedent part of a fuzzy system.

Recently, a new evolutionary algorithm called imperialist competitive algorithm (ICA) has been proposed by Atashpaz-Gargari and Lucas. Unlike PSO and GA that mimic the natural behavior, this algorithm is inspired by socio-political evolution of human. It simulates the process of powerful imperialistic countries dominate the colonies with imperialistic competition. Although this algorithm occurred lately, it has been successfully utilized in several engineering applications, such as control, data clustering, industrial engineering, flowshop scheduling, supply chain, heat transfer. In this paper, a hybrid algorithm combining the imperialist competitive algorithm with the gradient descent method is proposed for tuning the membership functions required to achieve a lower error. The proposed method can make use strong global searching ability of the ICA algorithm and strong local searching ability of the gradient descent method, simultaneity. The introduced model is then applied to predict oil flow rate of the wells.

This paper is organized as follows: In Section 2, the ANFIS structure and learning algorithm are reviewed. The ICA method is explained in Section 3. The simulation results and conclusions are given in Sections 4 and 5, respectively.

2. Adaptive Neuro-Fuzzy Inference System (ANFIS)

The adaptive neuro-fuzzy inference system (ANFIS), first introduced by Jang, is a multilayer feed-forward network which incorporates fuzzy reasoning and neural network learning algorithm to map relation between the input and output data. In the ANFIS, fuzzy inference system (FIS) is used to model non-linear systems and the neural network is used to tune input and output membership function parameters. The generation of fuzzy rules from numerical data or expert knowledge and adaptively construct a rule base makes this system very powerful in modeling numerous processes.

2.1. Architecture of ANFIS

The architecture of the ANFIS that is used for present work is based on the first-order Takagi–Sugeno (T-S) model. For a first-order T-S model, a typical rule set with two fuzzy if-then rules can be defined as:

Rule 1: if \( x \) is \( A_1 \) and \( y \) is \( B_1 \) then \( z_1 = p_1 x + q_1 y + r_1 \)

Rule 2: if \( x \) is \( A_2 \) and \( y \) is \( B_2 \) then \( z_2 = p_2 x + q_2 y + r_2 \)

where \( p, q \) and \( r \) are linear output parameters which are adjusted during the training stage. These two rules are illustrated in Fig. 1a.

Fig. 1b shows a possible architecture of ANFIS consists of five layers to implement these two rules. The functions of each layer are described as follows.

First layer:
Each node in this layer is an adaptive node and makes the membership grade for the input variables. Output of the node \( i \) is defined as follows:

\[
O^1_i = \mu_{A_i}(x), \quad i = 1, 2
\]

\[
O^2_i = \mu_{B_{i-2}}(y), \quad i = 3, 4
\]

In this layer, a hybrid algorithm combining the imperialist competitive algorithm with the gradient descent method is proposed for tuning the membership...
where \( x, y \) are the inputs to node \( i \), and \( \mu_A(x), \mu_B(x) \) are the membership function of the linguistic variables \( A_i \) and \( B_i \), respectively.

![Diagram of ANFIS system: (a) first-order T-S model; (b) corresponding architecture.](image)

Second layer:
There is no parameter adjusted in this layer, so each node is a fixed node labeled \( \Pi \). Every node in this layer calculates the firing strength of a rule via multiplying the incoming signals from the previous layer. The output \( O_i^2 \) of the node \( i \) can be calculated as:

\[
O_i^2 = w_i = \mu_A(x) \times \mu_B(y)
\]

Third layer:
The nodes in this layer are fixed node labeled \( N \). Each node calculates the ratio of the \( i \)th rule’s firing strength to the sum of all rules’ strengths. Consequently, \( O_i^3 \) is normalized firing strength defined as follows:

\[
O_i^3 = \bar{w}_i = w_i / (w_1 + w_2)
\]

Fourth layer:
Every node in this layer is an adaptive node, and calculates the contribution of \( i \)th rule towards the overall output which is defined as:

\[
O_i^4 = \bar{w}_i \times f_i = w_i \times (p_i x + q_i y + r_i)
\]

where \( \pi_i \) is the output of layer 3, and \( \{ p_i, q_i, r_i \} \) are the parameter set which are adjusted during the training stage.

Fifth layer:
The single node of this layer is a fixed node labeled \( \Sigma \), which computes the overall output by summing all the incoming signals.

\[
O_i^5 = \sum_{i=1}^{3} \pi_i f_i
\]

It can be observed that layer 1 and 4 are adaptive layers. Note that the number of nodes in layer 1 and the dimension of layer 4 determine the number of fuzzy sets and the number of fuzzy rules, respectively.

2.2. Learning algorithm of ANFIS
ANFIS uses the hybrid learning algorithm, which is a combination of gradient descent technique and the least-squares method to determine the optimum value for the non-linear parameters of the membership functions and the linear parameters on the fuzzy rules, respectively. The hybrid learning algorithm is composed of two passes22,25. In the forward pass, the node outputs go forward until layer 4 and the linear parameters are identified by least-squares method, while the premise parameters are kept constant in the current cycle through the training set. In the backward pass, the error signals propagate backward and the non-linear parameters are updated by gradient descent method, while the linear parameters are held constant.

The gradient descent method has a several drawbacks. One of these drawbacks is that the performance of the gradient descent method depends on the initial values of system parameters. This shortcoming can be removed by global searching ability of the evolutionary algorithms such as imperialist competitive algorithm.

3. Imperialist Competitive Algorithm (ICA)
The imperialist competitive algorithm (ICA) is a new evolutionary algorithm to search for the best solution by simulating the human's socio-political evolution. This evolutionary algorithm has proven its superior capabilities in convergence rate and global optima achievement over conventional evolutionary algorithms such as genetic algorithm, and simulated annealing26-33.

Fig. 2 shows the flowchart of the ICA. Like other evolutionary algorithms the ICA starts with initial populations called countries. There are two types of countries: colony and imperialist (in optimization terminology, countries with the least cost) which together form empires. Colonies of the initial population are divided among imperialists based on their power which are inversely proportional to their cost.
3.1. Generating initial empire

Supposing the dimension for a searching space is $N$ and algorithm starts with the $N_{\text{pop}}$ initial country. The country express as $1 \times N$ array that defined as bellow:

$$
country = [p_1, p_2, p_3, \ldots, p_{N_{\text{pop}}}]$$

(7)

The cost of a country is found by evaluating the cost function $f$:

$$
Cost = f(country) = f([p_1, p_2, p_3, \ldots, p_{N_{\text{pop}}}])
$$

(8)

The $N_{\text{imp}}$ of the most powerful countries (countries with minimum cost) selected to form the empires. The remaining countries ($N_{\text{col}}$) are colonies that each belong to an empire. The initial number of colonies of an empire is in convenience with its power. The colonies divide among empires in proportion the normalized cost of an imperialist defined as:

$$
C_n = c_n - \max_i \{c_i\}
$$

(9)

where $c_n$ is the cost of $n$th imperialist and $C_n$ is its normalized cost.

The normalized power of each imperialist is defined by:

$$
P_n = \frac{C_n}{\sum_{i=1}^{N_{\text{imp}}} C_i}
$$

(10)

Then the initial number of colonies of an imperialist will be:

$$
N.C_n = \text{round}\{P_n\cdot N_{\text{col}}\}
$$

(11)

where $N.C_n$ is the initial number of colonies of $n$th empire and $N_{\text{col}}$ is the number of all colonies. To divide the colonies among imperialists, $N.C_n$ of the colonies is chosen randomly and assigned them to each imperialist. The colonies together with the imperialist form the $n$th empire.

3.2 Moving colonies of an empire toward the imperialist

The imperialist countries try to improve their colonies and make them a part of themselves. This fact is modelled by moving all colonies toward their relevant imperialist. Fig. 3 shows this movement. In this figure, $x$ and $\theta$ are random variable with uniform distribution and $d$ is the distance between the colony and the imperialist.
A Hybrid Approach Based on the Combination of ANFIS and ICA

During the movement, a colony might reach to a position with lower cost than the imperialist. In this case, the imperialist and the colony change their positions. Then the colonies will be assimilated by the imperialist in its new position.

3.3. The total power of an empire

The total power of an empire is defined by the power of imperialist country plus a percentage of mean power of its colonies.

\[ T.C_n = \text{Cost(imperialist)} + \xi \text{mean(Colony of empire)} \]  

(12)

where \( T.C_n \) is the total cost of the \( n \)th empire, and \( \xi \) is a positive number which is considered to be less than 1. A small value for \( \xi \) implies that the total power of an empire to be determined by just the imperialist and increasing it will increase the role of the colonies in determining the total power of an empire.

3.4 Imperialistic competition

Due to the nature of imperialist that like to achieve more colonies, all empires attempt to take the possession of colonies of other empires and control them. During the competition, the power of weaker and powerful empires are decreased and increased, respectively. This competition is modelled by just selecting some of the weakest colonies of the weakest empires and making a competition between all empires to possess the selecting colonies. To start the competition, first, the possession probability of each empire which is proportionate to the total power of the empire, is found. The normalized total cost is obtained by:

\[ N.T.C_n = T.C_n - \max_i(T.C_i) \]  

(13)

where \( T.C_n \) and \( N.T.C_n \) are the total cost and the normalized total cost of \( n \)th empire, respectively. Then, the possession probability of each empire is calculated by:

\[ P^i_n = \frac{N.T.C_n}{\sum_{i=1}^{N_{\text{emp}}} N.T.C_i} \]  

(14)

To divide the selected colonies among empires, vector \( P \) is formed as:

\[ P = [P_1, P_2, P_3, \ldots, P_{N_{\text{emp}}} ] \]  

(15)

Then the vector \( R \) with the same size as \( P \) whose elements are uniformly distributed random numbers is created:

\[ R = [r_1, r_2, r_3, \ldots, r_{N_{\text{emp}}} ] \]  

(16)

Then vector \( D \) is formed by subtracting \( R \) from \( P \):

\[ D = P - R = [D_1, D_2, D_3, \ldots, D_{N_{\text{emp}}} ] \]  

(17)

Referring to vector \( D \), the selected colonies are given to an empire whose relevant index in \( D \) is maximized. In this competition when an empire loses all of its colonies, will be collapsed. The ultimate result of the imperialistic competition is that one imperialist will remain.

3.5 Convergence

At the end, only one empire will be exists and all the other countries will be its colonies. In this stage imperialist and colonies have the same position and power, which means the algorithm converges to the best solution.

4. Simulation Results

In the present work, ANFIS model combined with ICA (ANFIS-ICA) was applied to predict oil flow rate of the wells by using the data set of 31 wells in one of the northern Persian Gulf oil fields of Iran. Data set are collected in three month periods for each well from...
Dec. 2002 to Nov. 2010. To prevent the saturation problem and consequently the low rate training of the models, all data are normalized into the interval [-1, 1]. Before training a fuzzy inference system, the data set was divided into two different sets, namely, training and testing sets. The training set was used for the training process, during which the membership functions were tuned and the fuzzy rules were generated. The testing set was used to guarantee the model generalization and to avoid over-fitting the model to the training data set. Train and test sets must be different and were selected randomly from original data set. Among 600 total data sets, 390 data sets were selected for training and 210 data sets were selected for testing.

The structure of the proposed ANFIS model is shown in Fig. 4. Temperature and pressure of lines were taken as input parameters and the corresponding oil flow rate of the wells was used as output. In the used ANFIS model, the number of fuzzy rules is 25, the number of linear parameters is 75, the number of nonlinear parameters is 30, the total number of parameters is 105 and the number of nodes is 75. Each input parameters has five membership functions. Triangular type function was used as membership function.

ICA was applied to find optimal initial parameters of the membership function and the mean square error used as a cost function in this algorithm. The aim in the proposed algorithm is minimizing this cost function. The following assumptions were made here: the number of countries, the imperialists and number of generation

---

Fig. 5. Triangular membership functions for ANFIS-ICA model, (a) initial, (b) final.
were considered 40, 4 and 100, respectively; parameter \( \xi \) was set to 0.1.

In order to evaluate the performance of the ANFIS-ICA model, ANFIS model with the same architecture of ANFIS-ICA model was built by default conditions in MATLAB commercial software. Figs. 5 and 6 show the initial and final membership functions of the two input parameters in the ANFIS-ICA model and ANFIS model, respectively. There is obviously a considerable change in the shape of membership functions due to use of ICA algorithm. Figs. 7 and 8 show the comparison between predicted and measured oil flow rate values at training and validation phases for both ANFIS-ICA and ANFIS model, respectively.

To assess the models performance, mean square error (MSE) and efficiency coefficient \( R^2 \) were used. These statistical parameters show an average behavior of error in the model performance and are overall statistics that do not show the error distribution over results. Table 1 gives the value of these parameters for
the two different models. As seen in this table, it was obtained that prediction performance of the ANFIS-ICA model is better than the ANFIS model, where all the values of MSE are smaller, and all correlation coefficients are also closer to unity.

In Figs. 9 and 10 the measured oil flow rate values versus predicted values by ANFIS-ICA and ANFIS models to verification data points are shown, respectively. These figures and results from table 1 reveal that the combination of global search ability of ICA with local search ability of gradient descent makes the ANFIS achieve the capability of avoiding being trapped in local optima.

5. Conclusions

This paper proposed a new hybrid approach based on the combination of adaptive neuro-fuzzy inference system and imperialist competitive algorithm. The ICA was applied to determine the most suitable initial membership functions of the two input parameters in the ANFIS. Due to the combination of the global search ability of the imperialist competitive algorithm and local search ability of the gradient descent method, the proposed model can avoid the local minimum problem in the gradient descent method. The effectiveness of the proposed method was shown by applying it to predict the oil flow rate of the wells using data set of 31 wells in one of the northern Persian Gulf oil fields of Iran. Based on open literature, the application of this model to oil flow rate prediction is both novel and effective. The performance of the ICA based ANFIS was compared with conventional ANFIS in terms of statistical parameters such as mean square error and efficiency coefficient. The comparison results show that by employing the ICA with ANFIS, ANFIS can obtain
the more optimum distributed membership functions to describe the mapping relation in the input and output process parameters.

![Fig. 9. $R^2$ (ANFIS-ICA model)](image9)

![Fig. 10. $R^2$ (ANFIS model)](image10)

| TABEL 1. Comparison between the performances of ANFIS-ICA model and ANFIS model |
|----------------------------------|----------------------------------|
|                                  | **Training**                     | **Validation**                  |
| **MSE**                          | $R^2$                            | **MSE**                         | $R^2$ |
| ANFIS-ICA                        | 0.0064                           | 0.9916                          | 0.0077 | 0.985 |
| ANFIS                            | 0.0379                           | 0.956                           | 0.0256 | 0.9544 |

Co-published by Atlantis Press and Taylor & Francis
Copyright: the authors

206
References


