ARTIFICIAL NEURAL NETWORK BASED IDENTIFICATION OF THE GAS VOLUME FRACTION IN AN ELECTRICAL SUBMERSIBLE PUMP

Carlos C. Rodriguez, Alberto L. Serpa, Jorge L. Biazussi*, Natache A. Sassim*
Faculty of Mechanical Engineering, *Petroleum Studies Center (Cepetro)
University of Campinas, Campinas, Brazil
carlos_cortes_47@hotmail.com, serpa@fem.unicamp.br, biazussi@unicamp.br, natache@cepetro.unicamp.br

ABSTRACT
The electrical submersible pump (ESP) under multiphase flow is very common in the oil industry. These pumps present frequent premature failures when the gas flow is high. In addition to this, a further increase of the gas may fill most of the pump impeller, making the flow rate to decrease down to zero, known as gas locking. Due to lack of information and mathematical models that can be used in real time for this type of pumps, experimental studies are usual in this area. This paper applies artificial neural network (ANN) modeling for the volume gas fraction identification in ESP. The algorithm uses experimental data collected directly from the system for different gas fractions, such as pressure, flow rate, mechanical torque, elevation, etc. This model uses a back propagation learning algorithm and multi-layer perceptron neural network, where different structures are analyzed to find the optimal number of hidden layers. Results show that the system is able to identify the volume gas fraction in the pump with a very good accuracy.

KEY WORDS

1 Introduction
The artificial lift methods are widely used in the oil industry to increase or start the production on wells whose energy reservoir is not sufficient to raise the fluid naturally to the surface. The electrical submersible pump is a very common method. These pumps are dynamic devices that use kinetic energy to increase liquid pressure. They work successfully when handling incompressible fluids with low viscosities, but are severely impacted by free gas or highly compressible fluids.

Due to the oil reservoir depletion, the pressure decreases down to bubble point. Because of this, the produced fluid is a multiphase gas-liquid mixture, and the changes in the fluid properties can be a serious problem for the performance of the ESP. The gas presence depends on the relative amount of gas and liquid, and the performance varies from a slight deterioration to a complete blockage, known as gas locking [1]. Before gas locking occurs, another phenomenon, known as surging take place. This phenomenon causes instabilities in the pressure-flow behavior and is characterized by a strong decrease in pumping capacity [2]. Therefore, the operating conditions and volume of gas knowledge in the ESP are essential for a good performance of the system.

Several investigators have studied, both experimentally and theoretically, the behavior of centrifugal pumps used in oil wells and other applications. Lea and Bearden [2] classified the two-phase flow performance of an ESP into four categories: non-gas interference, gas interference, intermittent gas locking and gas locking. A first empirical correlation for centrifugal pumps operating with a gas-liquid mixture was developed in 1986 [3]. The performance of the ESP under multiphase flow are presented in [4] through an empirical correlation which provides a critical gas volume fraction $\lambda$ for a stable operation. In this correlation, $\lambda$ is a function only of the inlet pressure. Another model in function of the normalized flow rate and other operational conditions, such as liquid flow and angular velocity, is presented in [5]. In 2011, Monte Verde [6] determined the characteristic curves for one prototype of ESP operating with an air-water mixture with a gas volume fraction in the range of 0-10% at different rotational speed, intake pressure, and liquid flow rate.

In this context, majority of empirical correlations were developed to find the surging point and a few models allow to find the gas volume fraction for noncritical conditions. In addition to this, in the real operation of the pump, not enough data are available for applying this model. On the other hand, some input and output signals are collected in real time, providing an alternative to generate models to identify the gas volume fraction in the ESP.

Nowadays, ANNs are being applied to a lot of real world and industrial problems, from functional prediction and system modeling, where physical processes are not well understood or are highly complex, similar to two-phases flow in ESP. The principal advantage of ANN is the ability to learn and approximate relationships between input and output decoupled from the size and complexity of the problem [7]. An artificial neural network is a model consisting of many processing neurons in layered...
structures, and the learning process consists of relating given outputs to given inputs by adjusting the weights of neuronal connections [8].

Different researches about the behavior of the ESP based on ANN have been developed as alternative methods to detect the gas flow, although the principal interest is the cavitation, Nasiri et al. [9] show that the behavior of the centrifugal pumps can be approximated through ANN. Marques et al. [10] study the prediction of volume fraction in three phase flow and different regimes using artificial neural network, getting good results for each regime.

The method presented here employs some selected data collected directly from the system, such as intake pressure, output pressure, liquid flow rate, gas flow rate, mechanical torque and rotation speed as inputs of a multi-layered neural network, and predicts the gas volume fraction in three different ESP.

2 Methodology

2.1 Experimental Data

The experimental data used in this work was collected in the ESP bench test (Fig. 1) in the LABPETRO, a laboratory from the “Petroleum Studies Center” (Cepetro-UNICAMP). The data base used here are available in [11].

![ESP Bench test diagram](image)

The tests were conducted with air-water for three different intake pressures, 100, 300, 500 kPa, and three rotation velocities, 2400, 3000 and 3500 rpm for the P23 and P47 pumps, and 1800, 2400 and 3000 rpm for the P100 pump. For each pump, 36 performance curves were generated.

Since the objective is to identify the gas volume fraction, it is necessary to determine the volume flow rate of gas. This was calculated considering the air as an ideal gas in the intake conditions by the Equation (1):

\[ V_g = \frac{nRT}{P} \]

where \( V_g \), \( P \) and \( T \) are the volume, the pressure and the temperature of gas in the intake conditions respectively, \( n \) is the amount of substance in moles and \( R \) is the universal gas constant. Tests are performed with a constant homogeneous gas void fraction (\( \lambda \))

\[ \lambda = \frac{q_g}{q} \]

where, \( q_g \) is the volumetric gas flow rate and \( q \) is the total volumetric flow of mixture, both measured under the conditions of pressure and temperature of the ESP suction.

2.2 Artificial Neural Network

In this investigation, building a model to identify the gas volume fraction in ESP is the main target. For the model, a multilayer perceptron (MLP) based feed forward ANN has been used with back propagation training. In the Fig. 2, the process as a generalized model is depicted.

![Generalized ANN model diagram](image)

MLP is composed of simple perceptron in a hierarchical structure forming a feed forward topology with one or
more hidden layers between input and output. The neurons are expanded by a threshold factor and the sigmoid function, given by Equation (2), is employed as the activation function, where \( x \) is the neuron input.

\[
f(x) = \frac{1}{1 + e^{-x}} \quad (3)
\]

ANN a set of inputs or inputs-outputs pairs and appropriate training mechanism to adjust the weights (w) and biases (b) of neuron interconnections. For this model, the supervised learning is used, which means that the inputs and outputs patterns are provided and the learning process is done by comparison between networks output and the correct expected output to determine the error.

The numbers of hidden layer and number of neurons in each hidden layer is usually a trial and error process. Despite this, there are several criteria to define the starting point. Two hidden layers were selected here for the model because they can represent a function with complex shape and can represent an arbitrary decision boundary to arbitrary accuracy with rational activation function [12]. The initial number of neurons in every hidden layer was 2/3 of the size of the input layer plus the size of the output layer according to [12].

Back-propagation (BPN) learning algorithms are the most commonly used to train an ANN and this has been adopted in this model. The objective is to minimize the instantaneous value of mean square error (MSE). The MSE is calculated and the BPN uses it to adjust the value of the weights on the neural connections in the hidden layers. This process is repeated until the MSE has a low value. The MSE is given by:

\[
MSE(W) = \frac{1}{2n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i(u, W))^2 \quad (4)
\]

where \( \hat{Y} \) is the outputs calculated by the ANN, \( Y \) is the real output value of the system, \( n \) denotes the number of elements for \( \hat{Y} \) and \( Y \), \( u \) is the input vector and \( W \) is the ANN matrix of weights.

2.3 ANN Training

The ANN was trained using the MATLAB® 8.3 neural network toolbox with back propagation and Levenberg-Marquardt algorithm. For each pump, 75% of the data were selected randomly and used during the training phase and the remaining 25% were used to test and validate the trained network. For training, a MSE of \( 10^{-6} \), a minimum gradient of \( 10^{-7} \) and maximum iteration number (epoch) of 200 for the P23 and P47 pumps, and 250 for the P100 pump, were used. The training process would stop if one of these conditions is achieved. The initial weights and biases of the ANN were selected randomly.

The number of hidden layer for each ANN was 2 and the number of neurons in the two hidden layer was 32. The input layer has nodes representing the data measured experimentally and the target value of the output can have any value representing the percentage of gas in the ESP.

<table>
<thead>
<tr>
<th>Table 2. ANN Inputs-Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ANN Inputs</strong></td>
</tr>
<tr>
<td>( P_{in} ) [kPa]</td>
</tr>
<tr>
<td>( P_{out} ) [kPa]</td>
</tr>
<tr>
<td>( Q_{Total} ) [m³/h]</td>
</tr>
<tr>
<td>( P_{mech} ) [W]</td>
</tr>
<tr>
<td>( \omega ) [rad/s]</td>
</tr>
</tbody>
</table>

The inputs and target of the ANN are shown in Table 2. Since the system is deterministic and the stability is more complex than for linear systems, the NNFIR (Neural Network Finite Impulse Response) nonlinear model was used for the inputs values [13]. In this model, the output \( y(k) \) correspond to a weighted sum of past delayed values of the input [14]:

\[
y(k) = \sum_{n=0}^{T} (W(n)x(k - n)) \quad (5)
\]

where \( x \) is the values of the input. Figure 3 shows the NNFIR model, where \( q^{-1} \) represent the unit delay operator, i.e., \( x(k - 1) = q^{-1}x(k) \).

The lag space for the NNFIR model was a choice by trial and error, taking an enough large lag space and reducing it gradually until finding the optimal number of delays [13]. The optimal lag space for each ANN was 7, and the input layer for each network has 40 nodes.

3 Results and Discussions

Performance evaluations of different models were compared based on MSE. After training and validation, the performance of the ANNs was evaluated through the linear regression by comparing the measured data and the ANN output data.
In order to validate the ANN performance to identify the gas volume fraction in electrical submersibles pumps through experimental data, different networks were trained and validated for three different databases corresponding to P23, P47 and P100 model pumps.

The specific network architecture and lag space were chosen by trial and error method, trying to find the best response of the system and minimizing the error (MSE) with the output generated by the ANN. The previously mentioned structure was chosen because after some tests it generated good results.

Figures 4, 5 e 6 show the comparison of volume gas fraction measured and volume gas fraction approximated by the ANNs for P23, P100, and P47 pumps respectively, where time samples correspond to the dataset measured experimentally for each pump. The amount of data measured for each pump was different. The P100 pump dataset has the larger number of samples, 1443, followed by the P23 pump with 724, and P47 pump dataset with 552 samples.

We can notice that the predictions of the ANN are good, predicting the gas volume fraction in the pump with accuracy, even with abrupt changes in the data and different time samples. However, the P100 pump data set present a better response, may be due to the more number of experimental measured data, and consequently the greater availability of data for training an validate the ANN.

The network identified the gas volume fraction between 0% and in some cases more than 20%. For the higher gas fractions, it is possible that the pump was at the surging point. It is possible to say that the ANNs can be a good option to identify this region, allowing to change the operation status of the pump before reaching the gas locking point.

The linear regression calculated for the data set generated by the output of ANN is an indicator of the good performance of the network, i.e., the accuracy of prediction. Figures 7, 8 and 9 show the linear regression for each ANN output and the ideal tendency for every dataset, knowing as identity function, i.e., the slope is one and vertical-axis intercept is zero.

A good network performance evaluated through linear regression means that the slope should be as close as possible to one and the vertical-axis intercept should be zero. Table 3 contains the values for each data regression. In all three cases the trend generated by the data set obtained by networks are good, meaning that the performance of the ANNs to predict the volume gas fraction is also good.
We can see in Figures 7, 8, 9 and in the Table 3 that the network corresponding to the P100 pump has the best performance. On the other hand, the network corresponding to the P47 pump has the most distant parameters from the identity linear regression function, and this may be related directly to the amount of available data for the training and validation of each ANN.

Table 3. Parameters of linear equations given by regression

<table>
<thead>
<tr>
<th>Pump</th>
<th>Slope</th>
<th>Vertical axis intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>P23</td>
<td>0.9929</td>
<td>0.0390</td>
</tr>
<tr>
<td>P47</td>
<td>0.9928</td>
<td>0.0527</td>
</tr>
<tr>
<td>P100</td>
<td>0.9985</td>
<td>-0.0071</td>
</tr>
</tbody>
</table>

4 Conclusions

An identification of the gas volume fraction in an electrical submersible pump approach based on Artificial Neural Network with inlet pressure, outlet pressure, gas-liquid flow, mechanical power and angular velocity as input parameters has been studied. The network was trained using back propagation algorithm with supervised learning and achieved with good accuracy the percentage of gas in each pump. From the results, it is observed that the performance of the networks is directly affected by the amount of data available for training and validation, more data would be necessary for a better ANN model. Despite this, the three datasets used in this study have the sufficient elements to generate good approximations of the gas volume fraction in ESPs. Several models and architectures for the networks were studied. The network configuration and the NNFIR model presented good results to approximate the gas volume fraction. Larger lag spaces were tested for this model, but the improvement in the approach does not justify the greater computation time.

Acknowledgments

The authors gratefully acknowledge the financial support of Queiroz Galvão E&P S.A. (QGEP) and Fund for Support to Teaching, Research and Extension - FAEPEX - UNICAMP,
grant number 2145/15. The authors also wish to thank to CEPETRO for its support in conducting this research work.

References


