SPE-112124

Neural-Network based Sensitivity Analysis for Injector-Producer Relationship Identification

U. Demiryurek, F. Banaei-Kashani and C. Shahabi, University of Southern California; Frank Wilkinson, Chevron

This paper was prepared for presentation at the 2008 SPE Intelligent Energy Conference and Exhibition held in Amsterdam, The Netherlands, 25–27 February 2008.

Abstract

Determining injector-producer relationships, i.e., to quantify the inter-well connectivity between injectors and producers in a reservoir, is a complex and non-stationary problem. In this paper, we present a neural-network-based sensitivity analysis approach to address this problem. To the best of our knowledge, sensitivity analysis has never been applied for identification of the injector-producer relationships, yet we show that it is an intuitive while fundamental approach to address this problem. Sensitivity analysis is based on a theory with which the functioning of a closed system is derived by analyzing the derivatives of the output with respect to each input combination. For the injector-producer relationship identification problem, we use sensitivity analysis to determine the injector-producer relationships by varying the injection rates, i.e., the inputs to a trained neural network model of the oilfield, and analyzing the outputs, i.e., the production rates.

With our approach, we first generated a neural network to define the mapping function between each producer and its surrounding injectors based on the historical injection and production data. We employed Back-Propagation-Through-Time (BPTT) learning algorithm to train the three-layer feed-forward neural network using real data collected from 1911 to 2005. Next, we utilized the generated neural network model to apply sensitivity analysis in order to quantify the *significance* of the injectors on the corresponding producers.

We evaluated our proposed injector-producer relationship identification technique by experiments with real oilfield dataset as well as field trials. Experimental results show that our sensitivity analysis approach is not only an efficient method for identifying injector-producer relationships but also reveals significantly higher correlation accuracy as compared to the correlation typically estimated by the field engineers.

1. Introduction

Forecasting injector-producer relationships and modeling fluid flow in petroleum reservoirs is the key to recover maximum oil with reduced operation costs. Based on the analysis of the historical injection and production data from a reservoir, one can observe that the production performance is not only controlled by the injection rates itself but also maybe affected by various other factors. For example, new fractures in underground structures, fluctuation in permeability, and changes in reservoir temperature can easily impact the oil recovery. Typically, certain parameters may affect the output to a larger extent as compared with others, whereas some may have no effect on the behavior of the system. This complex behavior of oilfields renders the identification of the allocation factors between injectors and producers as an extremely hard problem.

Currently, reservoir engineers determine allocation factors between injectors and producers based on their past experience, production and injection historical data, numerical simulations. These approaches are not only time consuming but also error prone as they neglect non-linear interactions between various parameters.

In this paper, we present a novel *neural-network-based sensitivity analysis* approach to identify the injector and producer relationships in oilfields. Neural networks are mostly used when analytical model of the system either is not known or does not exist. In addition, neural networks employ historical data to learn the underlying system by developing a function that maps the input vectors to output vectors without knowing the system characteristics. In a typical neural network, the response at a particular time depends not only on the observable event but also on the past events. Moreover, neural networks [1] are adaptive systems as they arrange their weight functions on the basis of external or internal information that flows through the network. All these features of neural networks allow us to capture the complex characteristics of an oilfield and thus motivating us to develop a neural network based approach.

In our study, we first created producer-centric patterns, in which each producer was surrounded by multiple injectors, based on the real injector and producer coordinates in an active oilfield. With each pattern, we selected time varying monthly injection rates of each injector as the input vector where as oil, gas and water production as the output vector. In order to capture the mapping function between the injection rates and the production data, we first trained a three-layer feed-forward neural network using historical data. Next, we applied sensitivity analysis [2-3] to the trained neural network model to identify the injectors that have significant effect on the producers. For the injector-producer problem, we used sensitivity analysis to derive the relationships by varying the injection rates to a trained neural network model of the oilfield, and analyzing the production rates.

The rest of the paper is organized as follows: Section-2 briefly surveys the related work on injector-producer relationship identification problem, while Section-3 presents the construction of our neural network model and our sensitivity analysis method applied on the trained neural network. Finally, Section-4 presents experimental results performed in an active oilfield, and Section-5 concludes with directions for future work.

2. Related Work

In the past, several methods have been proposed to infer injector-producer relationship using historical injection and production rates [4-9]. Hefer e*t al.* [4] utilized Spearman rank cross correlations to estimate the relationship between injector and producer pairs where each pair consisted of an injector and a producer. They also associated these relations with geomechanics of the oilfields. Refunjol e*t al.* [5] also used Spearman rank correlation method to determine flow trends in an oilfield. In their work, they correlated time series of injection and production rates for an injector and adjacent producers while using time lags to find extreme coefficient value. Soeriawianata *et al.* [6], who also employed Spearman rank cross correlation, accounted for destructive and constructive effect of surrounding injectors to a producer. They inspired from the fact that injectors and producers operating simultaneously causes both superposition and noise. Even though, the spearman rank correlation is a very convenient non-parametric (distribution-free) statistic method, it suffers from the fact that the correlations that are calculated may give spurious negative correlations and no systematic behavior of correlations with time lag is found. Panda and Chopra [7] proposed a different approach other than the above statistics-driven methods to estimate the well interactions. Their work is similar to our approach however they do not apply the sensitive analysis over the trained neural network. They used an artificial neural network, where the multivariate data set includes the flow rate data of the well pairs and petrophysical parameters.

In most recent works, Alborteni and Lake [8] utilized Multiple Linear Regression (MLR) method to determine the interwell connectivity using resistive (transmissibility) model (RM). The linear model coefficients calculated by MLR quantitatively indicated the relation between a producer and adjacent injectors. Their method offers two solutions covering the case when the field-wide injection rate is different than the field-wide liquid production rate and the case when the two are almost equal. Yousef e*t al*. [9] extended the method proposed by Alborteni and Lake by adding capacitance (compressibility) parameter to the model. Their capacitance model (CM) approach determines two coefficients to quantify the interwell connectivity using MLR. The first coefficient is the weight that quantifies the connectivity depending on the geology and the relative position of wells. The second one is the time constant which quantifies the degree of fluid storage between wells. In both approaches, they assumed that the parameters in the models are unchanged (stationary) during the data analysis period. However, injector and producer relationship is non-stationary due to the fact that it changes constantly on the basis changing various factors such as natural geomechanical effects, bottom hole pressures, etc.

To overcome the shortcomings of the previous methods, we propose a new neural-network-based sensitivity analysis approach. Neural networks have been widely used for pattern recognition and prediction in oilfields [10-13]. Despite the fact that the neural-network-based sensitivity analysis is a popular approach to identify input and output relationships; it has never been applied to injector-producer relationship identification problem. We discuss the advantages of the neural-network-based approach in the next section.

3. Procedure

The mathematical developments of both neural network modeling and sensitivity analysis are addressed in this section.

Neural Network Modeling

Artificial neural networks are computational models of biological neural structures. A basic neural network structure consists of one input layer, one or more hidden layers and one output layer. Each layer consists of different numbers of neurons and each of these neurons connects to next layer's neuron with a weight. The neuron in each layer takes multiple inputs and produces an output on the basis of its activation function. Therefore, each input, which can be a raw data or output of other neurons, is modified by weight before fed to the next layer. Each neuron combines these weighted inputs and processes them with an activation function and a threshold value before producing its output.

Much of the interest in neural networks arises from their adaptiveness and ability to discover the underlying system by developing a function between input and output vectors on the basis of historical data. Neural networks accumulate the knowledge implicitly in connection weights between the layers. As a consequence, the knowledge can be modified by changing the weights through back-propagation. With the back-propagation rule, also referred as delta learning rule, the network first uses the input vector to produce an output, and compares this output to the desired output. In the case there is a difference, the weights are modified between the layers to further decrease the difference. This continues until the minimum desired error rate between the network produced and actual output. Fig. 1 illustrates a basic structure of a three layer neural network where x, y, z, δ represent input vector, actual output vector, desired output vector and error rate respectively; y_1 is the output of the processing element which is derived from summation of the input vector and the corresponding weights.

The major aspects of a successful neural network design are the selection of learning algorithm, input vectors, number of hidden layers and number of neurons in each hidden layer. The next section presents the methods that we used to select these features. We do not explain each feature in detail since all these are very well described in artificial intelligence literatures and it is out of our scope. Our main focus in this paper is sensitivity analysis on neural networks.

Inputs and Output Selection

Based on the real injection-production dataset from a section of an active oilfield, we first constructed a set of producer-centric patterns with which each producer was surrounded by multiple injectors. Fig. 2 shows these patterns, where red and blue markers denote injectors and producers, respectively. In this section, there are 90 different patterns whose boundaries are determined by the reservoir engineers based on the past experience and historical injection/production rates. Fig. 3 illustrates a satellite image that depicts a single pattern which includes five injectors and one producer.

To construct our neural network model of the oilfield, we selected time varying injection rates of injectors as the input vector. We selected oil, gas and water production from the corresponding producer as the output of the neural network. We also normalized the input vectors since it enhances the training process by preventing large input values from invalidating other inputs that are equally important but have smaller values. We used mean/standard deviation method to normalize the input and output vectors of our neural network.

$$X'_i = \frac{(X_i - \mu_i)}{\sigma_i}$$

where $X'_i, X_i, \mu_i, \sigma_i$ represent normalized input vector, original input vector, mean of the original input vector, and standard deviation of the input vector, respectively.

Learning Algorithm

Almost all of the previous neural network studies in petroleum engineering used the most-widely recognized back-propagation algorithm. In our study, we used Back Propagation Through Time [14] which unfolds discrete-time recurrent neural network into a multilayer feed-forward neural network where we can use classic back propagation with some modifications. In other words, recurrent network is transformed into T-multilayer feed-forward network where each layer represents one unfolding of recurrent network in time. This allows the network to remember the inputs in the past; hence, the training is more robust.

Hidden Layers

The number of neurons in the hidden layers is very important since it affects the training time and generalization property of neural networks. On the one hand, too many neurons may cause the network to memorize (over-fitting) as opposed to generalize, on the other hand a too few number of neurons would require more training time in finding optimal representation or generally result in under-fitting. We adjusted the number of neurons in hidden layer experimentally One rule of thumb in the neural network literature indicates that the number of neurons in hidden layer should be 2/3 of input neurons plus number of output neurons. [13] We determined the number of neurons in each hidden layer for each pattern mainly based on this rule of thumb, along with many trial and error experiments. For each neuron in all layers, we used the sigmoid function as the neuron's activation function.

Training

We divided the data at hand into training and test dataset based on the years. For example, for the best case scenario, we used the data from 1911 to 1985 as training set where as the data from 1985 to 2005 is used for testing the neural network model. In some cases, we utilized bagging and boosting [1] techniques to increase reliability of the networks for weak learners (data short patterns). We trained our neural network model for each producer-centric pattern consisting of a producer and surrounding multiple injectors. We first developed the network model using most of the default parameters in neural network software tool. Later, we tried different learning rules, number of iterations, activation functions and hidden layers until we find the optimal design. We employed 1) actual output/network output plot, and 2) mean square error of the entire output metrics to understand how well-trained the network.

Sensitivity Analysis

After a neural network is adequately trained, it can be used for control and/or prediction of operational parameters. However, our goal is to utilize trained neural network for identifying the important input parameters instead of prediction. For this, we use sensitivity analysis which calculates a sensitivity matrix based on the partial derivatives outputs of the network with respect to each input. The sensitivity analysis methods for neural networks have been extensively discussed in artificial intelligence literature [2, 3, 15-17]. Its effectiveness and some limitations have been shown [18]. The first approach is to add noise to each input variable, and to observe the effect upon the overall error [19]. There are some drawback in this approach such as deciding the volume of noise to add (i.e. cannot add noise to nominal variables). The second approach is to analyze the derivatives of the fan-out weights of the input units, with a high derivative indicating high sensitivity [2]. As a part of our work, we conducted sensitivity analysis on injectors to determine how each injector would affect the oil production. We utilized the sensitivity analysis method proposed by Zurada at el. [2]. This enabled us to rank the influence of the injectors for each producer in the given oilfield pattern.

The following is the definition of sensitivity analysis of our three layer feed-forward neural network model where $\vec{x} = [x_1, x_2 ..., x_n]$, $\vec{h} = [h_1, h_2 ..., h_m]$, $\vec{y} = [y_1, y_2 ..., y_p]$ represent the input, hidden and output layers respectively. Given this information we can define the sensitivity of trained output with respect to an individual input as:

$$S_{pn} = \frac{\partial y_p}{\partial x_n} \tag{1}$$

Using the standard notation of an error back propagation through time, the derivative of sensitivity can be expressed in terms of network weights as show below:

$$\frac{\partial y_p}{\partial x_n} = y'_p \sum_{m=1}^{m} w_{pm} \frac{\partial h_n}{\partial x_n} = y'_p \sum_{m=1}^{m} w_{pm} h'_m v_{mn} \tag{2}$$

where v_{mn} denotes the weight value between input and hidden layer and w_{pm} denotes the weight value between the hidden and the output layer. Here y'_p, h'_m represents the derivative of output and hidden layer activation function, respectively.

Equation (2) defines the sensitivity of the production with respect to one injector at time t. Since our training set includes dataset in time series for each producer and injector in the given pattern, we define the pattern sensitivity matrix in order to include all entries in the training set.

$$S^p = Y'WH'V \tag{3}$$

where W and V are the output and the hidden layer weight matrices, respectively, and:

$$\begin{aligned} Y' &= diag(y'_1,....y'_p) \\ H' &= diag(h'_1,....h'_p) \end{aligned} \tag{4}$$

Equation (3) defines the sensitivity matrix for a specific training data vector at time t. However, each training pair produces a different sensitivity matrix. In order to apply sensitivity analysis, the sensitivity matrix must be evaluated over the entire training dataset. For this, we used the following metric:

$$S_{pn,avg} = \sqrt{\frac{\sum_{d=1}^{D} [S_{pn}]^2}{D}} \tag{5}$$

Using the metric in Equation (5) we can now define the significance Φ of any input parameter over the output using

$$\Phi i = \max_{p=1..P} \{ \sqrt{\frac{\sum_{d=1}^{D} [S_{pn}]^2}{D}} \} \tag{6}$$

Finally, we can now sort the significance of each input over output as shown by the example below (where Φ_1 represent least significance):

$$\Phi_3 > \Phi_2 = \Phi_4 > \Phi_1$$

The basic idea behind the neural-network-based sensitivity analysis is that each input channel to the network is offset slightly and the corresponding change in the output(s) is reported. The input channels that produce low sensitivity values can be considered insignificant which in turn represent the injector and producer relationship in the pattern. We can summarize the algorithm as follows:

Step 1: Train the neural network on the original training set
Step 2: Calculate all input-output sensitivities for each training pair (Eq. 3)
Step 3: Calculate the sensitivity matrix over entire training dataset (Eq. 5)
Step 4: Calculate the significance of each input over output(s) (Eq. 6)

4. Experimental Results

We performed numerous experiments in order to evaluate our proposed technique by using an active oil reservoir dataset as well as field trials. We observed that neural network based sensitivity analysis correctly ranked the injectors efficiencies around a producer.

Oilfield Dataset Experiments

In the first experiment, we selected a pattern with five injectors and one producer active since 1984. Fig. 4 illustrates the producer-centric pattern where “I” and “P” represents surrounding injectors and the producer, respectively. Consequently, we generated an input vector of injection rates from time series data as $\vec{I} = [I_1, I_2 ..., I_5]$ and an output vector of production rates for oil, gas and water as $\vec{P} = [P_{oil}, P_{gas}, P_{water}]$

For the purpose of this test, we divided the available dataset into training (75%) and test (25%) subsets. We trained a 5-8-3 (input, hidden, output layers respectively) neural network with learning rate =0.1 and momentum= 0.4. Both the inputs and outputs are normalized to range [0, 1]. We trained the network for an average output unit error of 0.005 using sigmoid as the activation function for both hidden layers and output layers. Later, we quantified the sensitivity of each injector to the producer on the basis of algorithm described in Section-3.

Fig. 5 is a sketch of the sensitivity profile which illustrates the evaluation of the significance of each injector to the producer. In this case, injectors I2, I4 and I1 are the most sensitive injector to the producer where as injector I5 has very low relevance to the production. Therefore, we can conclude that small changes on I2, I4 and I1 results in relatively large changes on the producers.

In order to verify the results of the first experiment, we preformed two additional tests. In the second experiment, we trained a 4-8-3 neural network by removing the least effective injector (I5). We observed that the output (production rate) of the trained neural network is minimally affected. As shown in Fig. 5, neural network trained on pruned training set converges to similar production rate. As a third experiment, we trained the same neural network this time removing most effective (I2) injector. Fig. 6 shows that removing most effective injector from the system causes significant drop (28%) in the production rate. We can therefore conclude that our algorithm identified the least and most sensitive injectors correctly.

We conducted similar experiments to evaluate our approach for multiple injectors and producers cases. We inferred that our algorithm is extendable to the case of multiple injectors and producers since it works as efficient as multiple injector and single producer case.

Field Trials

We also evaluated our proposed technique in a real field trial. For this evaluation, we used various maintenance operations (i.e., injector shutdown) data of the oilfield subsection. In the first field trial experiment, we found the least sensitive injectors, using sensitivity analysis as described in Section-3, for 14 patterns in which different injectors were shutdown. We expected that shutdown of the least effective injector would cause minimum impact on the oil recovery of that particular pattern. As another field trial experiment, we identified the most effective injectors for 18 different patterns. This time, we expected the production would significantly drop since the most effective injectors were shutdown in each pattern. Based on the analysis of production data for the selected patterns, we observed that we successfully identified the least and most effective injectors with the success rate of 86% and 83% respectively as illustrated in Fig. 8. Table 1 represents the number of correct and false classifications in determining least and most effective injectors for selected patterns.

5. Conclusion and Future Work

The main objective of our study was to use sensitivity analysis on a trained neural network to determine the injector and producer relationships. In the first phase, we constructed producer-centric patterns based on the real coordinates of injectors and producers from an active oilfield. In the second phase, we trained a neural network for a given pattern in the oilfield until the training error decreases to an acceptable level. For training, we used time varied injection rates as the input vector and oil, gas and water production as the output vector. Finally, in the third phase, we applied sensitivity analysis on the trained network as described in Section-4 to determine the injectors that have a significant influence on the producers. We observed that our technique yields better results on the old patterns due to abundance of training dataset. In addition, we evaluated our proposed technique by using an active oil reservoir dataset. The experimental results on the real dataset show that our sensitivity analysis approach is not only an efficient method for identifying injector-producer relationship but also reveals significantly higher correlation accuracy as compared to the correlation estimated by the field engineers.

This analysis may assist field engineers in making field management decisions to optimize fluid flow, improve ultimate recovery, minimize injection costs, and predict water breakthrough more accurately.

In the future, we like to extend our work in two ways: 1) we like to add other available time varied datasets (i.e.: bottom hole pressure, permeability, etc.) from injectors and producers to train our neural network model more efficiently, and 2) we plan to develop algorithms using the fuzzy curves approach to capture the nonlinear injector-producer relationships and compare it with neural-network-based sensitivity analysis. Fuzzy curves use fuzzy membership functions to establish the relationship between the input variables and output variable.

Nomenclature

$\vec{x}$ = *input layer vector*
$\vec{h}$ = *hidden layer vector*
$\vec{y}$ = *output layer vector*
δ = *error rate*
X' = *normalized input vector*
μ = *mean of the original input vector*
σ = *standard deviation*
v_{mn} = *weight value between input and hidden layer*
w_{pn} = *weight value between hidden and output layer*
y'_p = *derivative of the output value*
h'_m = *hidden layer activation function*
Φ = *significance*
S = *sensitivity matrix*
W = *output layer weight matrix*
V = *hidden layer weight matrix*
I = *surrounding injector*
P = *producer*

Acknowledgements

This research was funded in part by the Center of Excellence for Research and Academic Training on Interactive Smart Oilfield Technologies (CiSoft); CiSoft is a joint University of Southern California-Chevron initiative and in part by NSF grant IIS-0238560 (PECASE), IIS-0324955 (ITR). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

We also thank to Dr. Vega Sankur and Dr. Iraj Ershaghi for their support and valuable suggestions.

References

1. Mitchell, T.:"Machine Learning" McGraw-Hill, NY (1997).
2. Engelbrecht, AP., Cloete, I., Zurada, JM.: "Determining The Significance of Input Parameters Using Sensitivity Analysis", IWANN (May 1994) 382.
3. Zurada, JM, Malinowski, A.: "Sensitivity Analysis for Minimization of Input Data Dimension for Feedforward Neural Network" IEEE ISCAS (Jun. 1994) 447.
4. Heffer, K.J., Fox, R.J., and McGill, C.A.: "Novel Techniques ShowLinks between Reservoir Flow Directionality, Earth Stress, Fault Structure and Geomechanical Changes in Mature Waterfloods," paper SPE 30711 presented at the SPE ATCE,Dallas, TX, October 1995.
5. Refunjol, B.T., and Lake, L.W.: "Reservoir Characterization Based on Tracer Response and Rank Analysis of Production and Injection Rates," presented at the 4th International Reservoir Characterization Technical Conference, Houston, TX, March 2-4, 1997.
6. Soeriawanta, T and Kelkar,M.:"Reservoir Management Using Production Data" paper SPE 52224 presented at the 1999 SPE Mid-Continent Operation Symposium, Oklahama City, Oklahama, 28-31 March
7. Panda M., Chopra A.: "An Integrated Approach to Estimate Well Interactions", paper SPE 39563 presented at the 1998 India Oil and Gas Conference, New Delhi, Feb 17-19
8. Alejandro Albertoni and Larry W. Lake, "Inferring Interwell Connectivity only from Well-Rate Fluctuations in Waterfloods," SPE 83381, 2003.
9. Yousef, P. Gentil, J. L. Jensen and L. W. Lake, "A Capacity Model to Infer Interwell Connectivity From Production and Injection Rate Fluctuations," SPE 95322, presented at the SPE ATCE 2005, Dallas, Oct., 2005.
10. Sugai, K., Nishikiori, N.: "An Integrated Approach to Reservoir Performance Monitoring and Analysis", paper SPE 100995 presented at the 2006 Asia Pacific Oil & Gas Conference, Adeliade, Sep. 11-13.

11. Lechner, J., Zangl, G.: "Treating Uncertainties in Reservoir Performance Prediction with Neural Networks" paper SPE 94357 presented at the 2005 Europec/EAGE Conference, Madrid, Sep. 13-16.
12. A, JK. : "Neural Networks: A New Tool for the Petroleum Industry?", paper SPE 27561 presented at the 1994 European Petroleum Conference, Aberdeen, March 15-17
13. Al-Fattah S.: "Predicting Natural Gas Production Using Artificial Neural Networks", paper SPE 68593 presented at the 2001 Hydrocarbon Economics and Evaluation Symposium, Dallas, April 2-3.
14. Werbos, P. J.:"Backpropagation Through Time: What It Does and How to Do It" *IEEE*, (Aug. 1978) 1550.
15. Engelbrecht, P.:"A New Pruning Heuristic Based on Variance Analysis of Sensitivity Information", IEEE Neural Networks (Nov 2001) 1386.
16. Engelbrecht, P.:"Selective Learning for Multilayer Feedforward Neural Networks", Fundamental Informatica (Feb. 2001). 295
17. Zurada, JM., Malinowski, A., Usui S.:"Perturbation Method For Deleting Redundant Inputs of Perceptron Networks", Neurocomputing (Feb. 1997) 177.
18. Engelbrecht, P., Cloete, I.:"Incremental Learning Using Sensitivity Analysis", IJCNN, (Jun 1999) 1350.
19. Montano, J., Palmer, A.:"Numeric Sensitivity Analysis Applied to Feedforward Neural Networks", Neural Computing, (Nov 2003) 119.

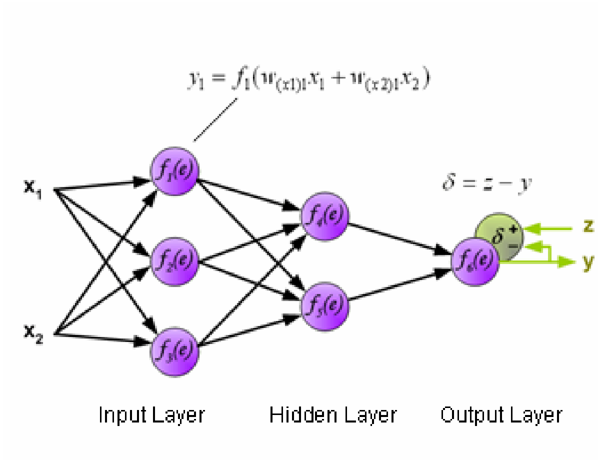

Figure 1: An example of the three-layer neural network

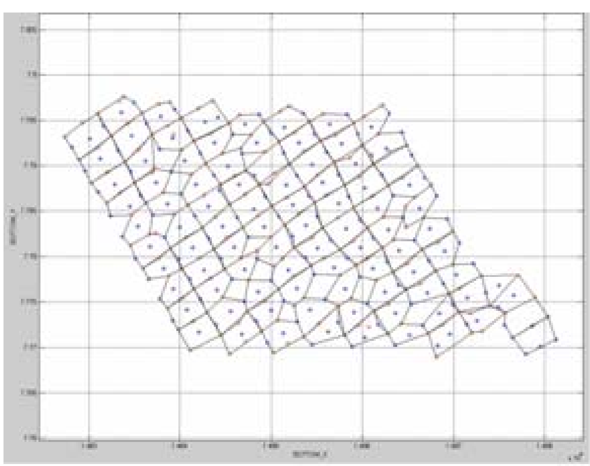

Figure 2: A producer-centric pattern layout in an oilfield

Figure 3: A satellite image of a producer-centric pattern in an oilfield

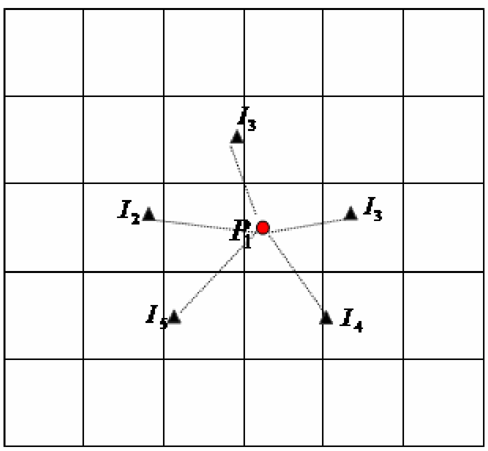

Figure 4: A producer-centric pattern with five injectors

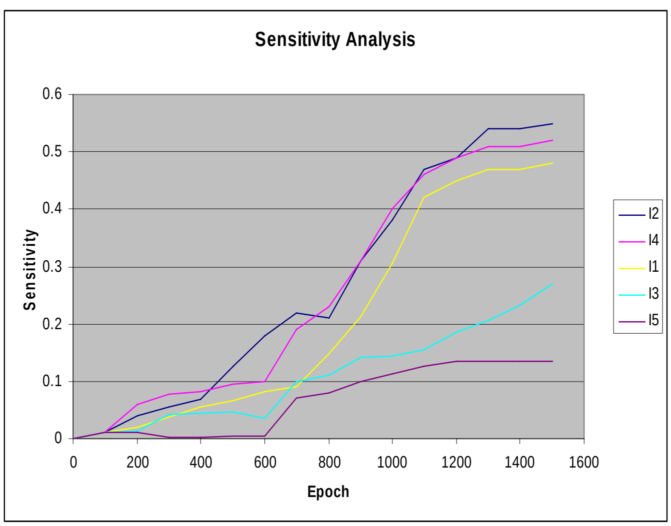

Figure 5: Sensitivity analysis output for a pattern with five injectors

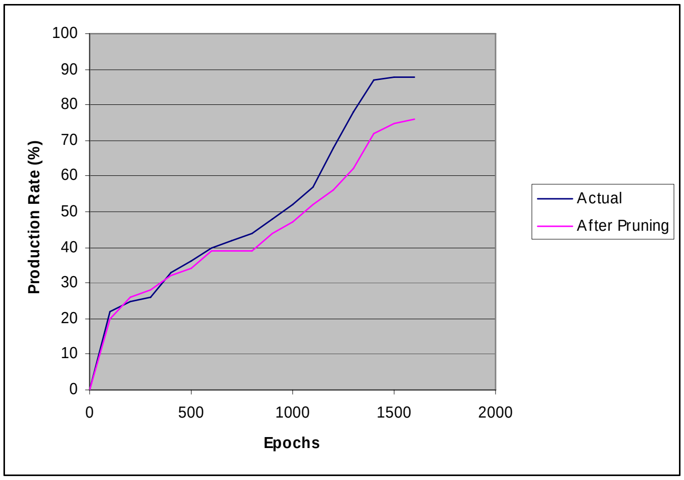

Figure 6: Production rate after pruning least effective injector

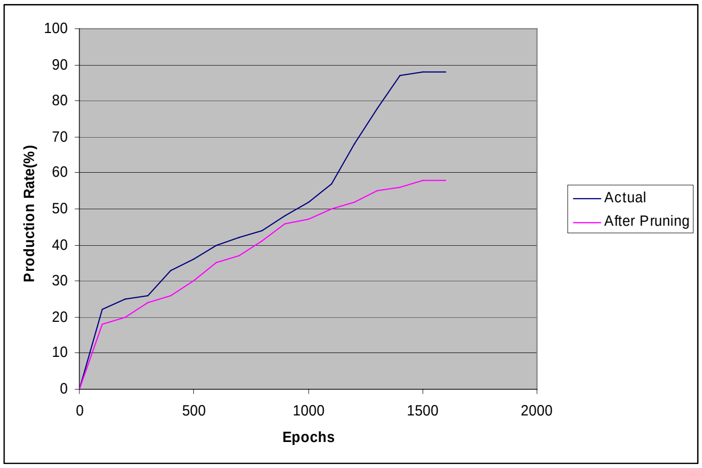

Figure 7: Production rate after pruning most effective injector

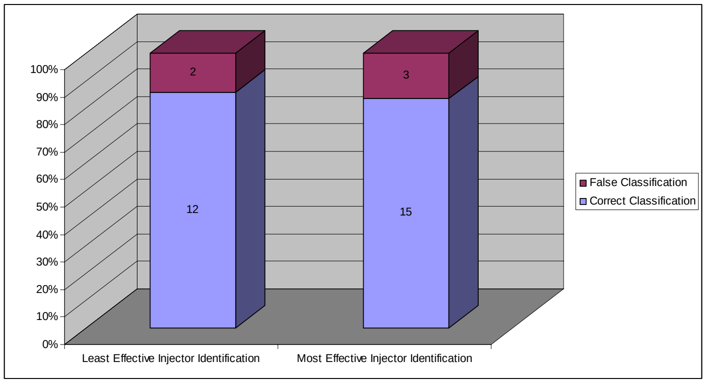

Figure 8: Correct and false classification for least/most effective injector identification

Table 1: Correct and false classification summary

	Least Effective Injector Identification	Most Effective Injector Identification
Correct Classification #	12	15
False Classification #	2	3
Total	**14**	**18**