Application of neural network to optimize oil field production

Jreou, Ghazwan N S, SPE member
Kufa University – College of engineering, Iraq – Alnajaf Al shraf P.O.Box 560,
E-mail:- Ghazwan.jreou@uokufa.edu.iq

Abstract

The development of a production strategy for oil and gas fields is an important task adopted by the directorate companies which are invested on those fields, and the main goal of their plans is to increase the ultimate field production as much as possible within economical and physical limits.

This work presents the determination of the optimum number and location of the infill production new wells under development stage of the field life for one of the four production sections of the concerning oil field in this study (Southern Iraqi Oil Field).

A feed forward neural network approach has been performed to locate the infill production wells in order to increase the total oil production from that production sector.

143 data sets of six variables (No. of injection wells, No. of production wells, Injection production ratio, Oil fraction, Cumulative of injected water and Cumulative water production ) as input layer were prepared in a bundle form in order to use directly with the selected Feed Forward Neural Network combining with the employed neural network.

These data sets were collected and arranged according to the three concentric circles – mask calculation procedure- around the prospectus new wells in order to simplify and accurate reservoir and production information for each circle calculations. One variable was gained as output layer from the constructed neural network.

Good and encouraging results were obtained that can be serve and support the development phase of the concerning field case study, in addition the results can help the directorate manager for reaching the right decision in execution manner.

Key words – (oil field production, neural network, field management, optimization).

1. Introduction

Petroleum engineers with their different disiplines try to reach the main goal in receiving field development plans, so they aim to maximize the economic recovery of oil and gas resources and ensure security of petroleum supplies in their jobs which facing them with fully recognition and difficult environment and always willing to be flexible as possible as they can in requirements to ensure that oil and gas development meets objectives.

The Oil industry in all its multi specialization with aiding of wide world reputation of big companies especially with their innovations and technical researches were jumped huge spans toward the developments and earning more knowledge in order to get more and more exploration and exploitation of natural resources in the world.

Production oil and gas fields are different in their sizes, shapes, reservoir contains and operating conditions…… etc., therefore one can search and
study the optimum conditions for putting these field into production strategy plans concerning constrains and limitations. Oil field in the world and specially whose they are classified as giant field requires many evaluations of the possible combination of the decision variable like reservoir properties, well location, production scheduling parameters and field surface facilities to obtain the economic plans when thinking of an optimization of reservoir development. By this integration of quantitative geological and geophysical analysis with appropriate flow models one can get an assessment of alternative development and completion schemes and their relative economic values.

Those optimization methods for reservoir development and production manner are published in literature and applied for searching of an objective functions with field limitations trying to maximize or minimize these functions. A mask data calculation technique was used for collecting, manipulating and preparing in the form of input data sets for using with optimization and prediction technique (FFNN).

A NN optimization technique was used to search for the optimal decision among new realization as well as it has more practical capability with these cases. Neural network [8], anon algorithmic, non-digital intensely parallel and distributive information process system that is being used more and more every day and it had been a valuable pattern recognition tools, they are capable of ending highly complex pattern with large amounts of data. Determination, production or estimation is a good line of using NN while there are an enough data sets are collected and prepared for such goal and make the NN easy to find its solution with highly precise.

Mostly, field production data are available in E&P department of the oil companies; therefore this type of information is dependent in the present work instead of other type of data that needs more verification and inspection. In addition those are actual field data. Then a suitable data (checked with simulation output) were prepared in order to use with NN depending on quality measurements and collections. Also the study concentrates on combining of mask calculation procedure with multivariable NN technique.

1.1. Objective of the study

The main objective of the current study is to present a method for inspection of the development view point by studding of increasing cumulative oil or gas production field (either sector, region gathering production point or whole of the field) taking into account all factors that are related to the objective function, by searching for and allocation of a new infill drilling wells using artificial intelligence technique (Neural network) to maximize cumulative oil production over the concerning field incorporating with mask task process for preparing and manipulating of the important production data sets as input information to the neural model.

1.2. Simple concepts of neural network.

Mohageg and other references (1, 8, 9, 11, and 12) assigned that the neural networks are well developed computational intelligence technology with many possible applications in the oil and gas industry. For this study a neural network was developed to search for the prediction production potentials for each well prior to drilling (infill drilling) for the prior election positions either for production sector or for complete field. The current study covers 36375 acres.

1.2.1. The basics of artificial neural networks

Biological structure of neurons and their interconnections can be modeled in computers producing mathematical similar in functionality to their organic counter-parts. These structures are called a NN. They generally take some input, process it and give some output. Many variants and types of ANNs exist. The standard back propagation (feed forward) ANNs has been used for this work, for simple and brief description about ANNs, one can refer to the following subsection, (ref.15).
1.2.2. **Structure**

An ANN consists of nodes (neurons) with weighted connections between them. There are some nodes that receive input, some nodes that give output, and hidden nodes in between. Each node processes all its input, for example by summing it up and running the sum through a function, and propagates its result to its connected nodes until an output is given at some output node(s). Figure (1) presents the topology of a simple example ANN. For the present work, one input layer with six neurons of six variables, hidden layers and one output layer with one output neurons. Figure (1) – just explaining example of ANN structure. Input nodes receive input data set or information and pass it along to the hidden nodes through weighted connections. The received signal is processed in the hidden nodes and sent along weighted connections to output node(s) which further process the signal and produce the final output.

1.2.3. **Input and output**

Figure (1), displays an array of input nodes, in the present case 6. On the other hand, only one output node is shown. This corresponds to an input-output pattern where 6 input values produce 1 output value. The number of input and output values can vary, but this type of pattern is what ANNs typically use.

1.2.4. **Learning**

Just like animal brains, ANN’s learn from experience. By experiencing instances of a problem, an ANN receives information about how to update the weighted connections between neurons. This happens based on minimizing error in estimating output given desired output from a problem instance set with different inputs and outputs. This enables an ANN to adapt itself in a way that lets it generalize over the data and further enables it to make predictions about future (unknown) problems by acting as a function for the input. This process is referred to as back propagation.

1.2.5. **Convergence**

Convergence can be measured by observing how the error between expected target value and actual calculated value is reduced throughout training for a set of testing data other than the training data. When the error has been reduced to a point where it is no longer decreasing this generally means the ANN has converged.

1.2.6. **Data representation and interpretation**

ANNs are often built with topologies making the information at any specific node hard to interpret. The information contained in a network is often best expressed in the struck true’s total ability to produce correct output given an input. These steps are performed in the back propagation algorithm. ANNs trained on well defined input-output cases are capable of expressing a rich variety of decision surfaces without much direct regard to problem semantics. This convergence to a state that acts as a general problem solver for a given domain is an emergent feature of ANNs. It has the ability to produce a complex mathematical structure where each single node contains distributed information processing or redundancy functionality that is very hard to interpret out of the big context. ANNs are often seen as black boxes that take input and produce output, without the designer bothering too much about perfect understanding of what comes in between.

1.2.7. **Generalization**

When an ANN produce correct and acceptable within suitable output for the majority of input cases other than the ones in was trained with, it can be said to generalize well. A well trained ANN for the purposes of approximation like was used in this work should provide a smooth nonlinear mapping. This means that it should be able to interpolate to new cases that are similar but not identical to those patterns used for training. If the network is ever trained, this will result in a non-smooth mapping, and the ANN will work more like a memory with direct lookup from input to output. The structure of the ANN is therefore an important factor for correct training. A correct amount of hidden layers and nodes should be used. If more are used than are required to learn the input-output relationship, there will be more weights than necessary which can lead to overtraining of the data and bad performance when approximating unknown cases. Back propagation feed forward neural network algorithm was explained in appendix A.

2. The process of neural network technique.
The process involved gathering data used as potential input sets, including production data at known wells, then selecting the best continuous production conditions - excluding cut off periods of maintenance or work over times, or ....etc - data sets covering all variables as input to the neural model. Developing and testing various network architectures, making predictions, analyzing and applying the results in the main phases of optimal technique.

Variables considerable as potential inputs include:
1. No. of injection wells
2. No. of production wells
3. Injection production ratio. (No. of injection wells to the No. of production wells in each ring).
4. Oil fraction. (Qo/Qt).
5. Cumulative of injected water
6. Cumulative water production

The neural network was trained by (MATLAB NN toolbox), to identify the cumulative production at a set of the wells with attempt production from the formation (constrains are inherently include (BP =2660 Psia) and WC (<=40%). Cumulative production was taken as output from the chooses FFNN in barrels of oil per year over the time period of production covering by the current study.

2.1. Mask calculation technique.

A data mask calculation technique is easy way to handle many wells in each step better than manipulation of each well variable. In addition this technique is good to use with NN, because the latest has its ability to handle huge data information with one time of calculation runs. This mask is square in shape and equal in cells for columns and rows directions, also the mask has three or more circles depending on the size of the region under study. Once the user elect the new infill wells location on grid map, then can use the mask to put its first circle on the election well location and gathering all type of production data for the wells falling within each other circles, therefore collection, verification, calculations and manipulation all of the data types will be reduced. The procedure can be written by any programming language (visual FORTRAN for example), and incorporated with NN model files directly. The used mask in this study consists of three circles.

3. Results and discussions.

Results were evaluated by inspecting the predicted map of cumulative production and performing statistical testing including a correlation between the predicted and actual Cumulative production which gives an idea about the correlation coefficient of predicted results (R = 0.95). Then the results were used as an expected system designed to reduce exploitation risk.

In another side, the obtained results of the current study are varying according to the each step of the implemented optimization technique – NN model. The technique approach can be achieved in steps agreement with Zurada, J.1992, presenting in the section of results and discussions (paragraph 3.3).

3.1. Problem analysis

Total oil production (cumulative oil production) or can be reflected to the net present value over a certain time period is the objective function for the current study where seeking for the field development by optimization technique. This study concerning with the cumulative oil production.

3.2. Data sets collection and preparation

The first step of constructed neural network is to prepare data sets that are professional to use in solving the problem, while the type of the data sets are not always available like offsets. Production or reservoir characterization data. Since the production data is readily available for all the wells in the field, data sets contain (143 sets of six input variables – listed before) which are selected as input for the neural model in this study with one variable as output node. The specific production and other related data are gathered for every production well in the Concerning field region, that is shown in figure (2), while the figure (3) shows the gird map of the studied region.

A data mask technique is used for preparing and prospecting drilling location of a new infill Wells. The data mask is shown in figure (4) which is5*5 grids covering 5511.356 acres, each Square grid (3330) feet *(2883.78) feet equals (9602987.4) square feet – Dimensions are taken from Hamadallah, S.M. 1999. The ranges of input and output data sets are listed in the table (1) which is collected for every well within the boundaries of the data masks.
Since the number of input data sets (6 nodes) to the NN is fixed, two neural network models were attempted, in order to make them more recognized effectively and accurate in prediction by using of the employed neural network software (MATLAB NN toolbox). Therefore the numbers of data in each set have to be constant for every well location according to the used neural network architecture.

Three rings are implemented labeled X1, X2 and X3 forming a bull – eye target with the drilling location at the center of the ring X1 in order to overcome the difference of wells number for each mask acres area, as mentioned by Boomer, R.J. 1995, therefore an average data for each ring is implementing and normalized, in order to be ready used in input phase of the NN model. The output of the NN is cumulative oil production for each year covering the time period Spanned from 1980 to 2005, all the data are collected from 40 production wells and 20 Injection wells; The NN has been trained with 100 random best selections, while 43 data sets are used for testing and verifying the accuracy of the constructed NN model and preventing any overtraining case if happen.

Table (1), Input and output data sets ranges for each variable.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Injection wells</td>
<td>1 - 5</td>
</tr>
<tr>
<td>No. of Production wells</td>
<td>1 - 10</td>
</tr>
<tr>
<td>(Ratio = Inj. / Prod.)</td>
<td>0.1 - 21</td>
</tr>
<tr>
<td>Oil Fraction - Fo</td>
<td>0.833667 – 1.0</td>
</tr>
<tr>
<td>Cum. Of water injection (STBW)</td>
<td>61685 – 26994661</td>
</tr>
<tr>
<td>Cum. Of Water Production (STBW)</td>
<td>521.2129 – 3896482</td>
</tr>
<tr>
<td>Output</td>
<td>13264 - 43284200</td>
</tr>
</tbody>
</table>

3.3. Neural network procedure steps

3.3.1. Learning

Potential input variables were tested a giants the output to determine the best variables for Training phase of the technique, a series of neural network were developed to test different Architectures, these architectures were compared based on its behavior during its goal reaching and given tolerance accuracy on based of the input data sets size.

The constructed NN model has been trained with 100 data sets which were prepared in previous step, many tries were done in order to reach the best behavior of the trained network and reach its goal with minimum error percentage, figure (5) shows the best trained behavior of the learning NN and reflect the degree of coincidental step with the following learning parameters:-
- Net. Train parameters. Show 50
- Net. Train parameters. lr = 0.69
- Net. Train parameters. mc = 0.9
- Net. Train parameters. Epochs = 400
- Net. Train parameters. goal = 0.001
- Input layer with 6 variables
- Middle layer = hidden layers = one layer (50 nodes)
- Output layer with one variable

3.3.2. Validation (Testing phase).

For more confidence and dependency in using with the constructed and learned NN model and to use it with the next step – prediction and simulation, the network was inspected against data sets that were chosen and separated for this phase of the technique process covering all the ranges of the input variables. Good results were obtained according to the statistical measurements. The final structure of the used NN is shown in figure (6).Table (2) contains the data sets for testing and validation, where figure (7) reflect the precision fact of the NN model.

Table (2), Validation data sets ranges for each variable.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Injection wells</td>
<td>1 - 5</td>
</tr>
<tr>
<td>No. of Production wells</td>
<td>3 - 10</td>
</tr>
</tbody>
</table>
3.3.3. Simulation

Once the NN trained and inspected by testing and validation phase, the final network was used to predict cumulative production over the entire region, with interesting time period. This predicted production was further refined by excluding production at location when no Production wells were present. This is the last phase of the present optimization technique which is forming the Prediction step or the objective goal of the present study that was searching for the optimal location of the prospecting drilling infill production wells with verifying maximum cumulative oil production per year, figure (8) shows the locating of these new wells over the concerning oil field region in the present study, as shown from that figure the priority of the execution plane of the infill wells drilling within the concerning oil field region follows the new suggested locations symbolic by letters O, P, and M. The predicted locations of the infill production wells were arranged and tabulated according to their additional cumulative oil production, these results were pointed in table (3).

Table (3) Incremental percentage of cumulative production from the new infill production wells.

<table>
<thead>
<tr>
<th>WELL No.</th>
<th>Cumulative Production</th>
<th>Incremental percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-13000</td>
<td>170943951</td>
<td>10.521</td>
</tr>
<tr>
<td>O-15000</td>
<td>464833388</td>
<td>28.609</td>
</tr>
<tr>
<td>P-16000</td>
<td>457576152</td>
<td>28.162</td>
</tr>
</tbody>
</table>

3.3.4. Comparison results

A comparison was made for the obtained results of the present work with those taken from reference (14), whenever for the same concerning oil field region using of Genetic Algorithm (GA) artificial intelligence technique and net present value criteria (NPV), figure (9) shows the optimum location of the new infill wells by the reference (14), while figure (7) shows the results from using of NN with and mask task technique depending on cumulative oil production criterion. It appears that there is an agreement between these search works.

Al modafer (ref.14), in his study allocate three infill wells along the longitudinal axis of the anticline structure of the field region under study which they are symbolic triangle shape. Especially the crest of the structure still deserves connate water saturation – SWC, which is good prediction. Our new infill wells allocated along the crest of the structure also, with low water saturation, so this is good chance for long production time of free water oil (WC ≤ 40%) from those new wells.


1. Artificial Neural Network is one of the most important ways to predict a lot of reservoir parameters or functions, as well as search for new sites of infill wells, which is an important uses of this technique in conjunction with simulators because of their relationship with development operations and production of oil and gas fields.

2. Both Artificial Neural Network and Data Mask technique gave a good and easy way for the preparation and data management in every phase of prediction and optimization technique.

3. The constructed Artificial Neural Network model (FFNN), had good training and validating Results according to the data sets used in both of training and validation steps, and obtained good behavior after many trials and reflecting high confidence to the prediction or simulation stage.

4. Clear and quite improvement appeared with additional cumulative oil production for the
production field sector under the current study, especially with the wells M, O and P, among those elected locations of infill wells. This became a good decision maker to which of the well can we implement for the future production plan and development field strategy view point.

5. ANN used and proved to be an effective tool within prediction and optimization process, especially when dealing with developmental ideas of the oil and gas production, from the standpoint of oil reservoir management and field directorate for future strategic planning.

**Notation**

BP = Bubble point pressure (psia).
STBW = Stock tank water barrel
STWO = Stock tank oil barrel
SWC = connate water saturation
WC = Water cut, percentage.
Lr = learning rate
Mc = momentum coefficient (additional training parameter).
Qo = oil production (STB/D)
Qt = total production (STB/D).

**Appendix – A**\(^{16}\)

The back propagation algorithm is used to compute the necessary corrections. The algorithm can be decomposed in the following four steps:

i) Feed-forward computation.

ii) Back propagation to the output layer.

iii) Back propagation to the hidden layer.

iv) Weight updates.

The algorithm is stopped when the value of the error function has become sufficiently small.

**First step: feed-forward computation**

The vector \( o \) is presented to the network. The vectors \( o^{(1)} \) and \( o^{(2)} \) are computed and stored. The evaluated derivatives of the activation functions are also stored at each unit.

**Second step: back propagation to the output layer**

We are looking for the first set of partial derivatives \( \frac{\partial E}{\partial w_{ij}} \), we can collect by simple inspection all the multiplicative terms which define the back propagated error \( \delta_j^{(2)} \)

\[
\delta_j^{(2)} = a_j^{(2)}(1-a_j^{(2)})o_j^{(2)} - t_j \]  
\[
\tag{A1}
\]

and the partial derivative we are looking for is

\[
\frac{\partial E}{\partial w_{ij}} = \left[ a_j^{(2)}(1-a_j^{(2)})o_j^{(2)} - t_j \right]a_i^{(1)} \]  
\[
\tag{A2}
\]

Remember that for this last step we consider the weight \( w_{ij}^{(2)} \) to be a variable and its input \( a_i^{(1)} \) a constant.

\[
O_i \frac{\partial E}{\partial w_{ij}} = \delta_j^{(2)} \]  
\[
\tag{A3}
\]

At the input side of the edge with weight \( w_{ij}^{(2)} \) we have \( a_i^{(1)} \) and at the output side the back propagated error \( \delta_j^{(2)} \).

**Third step: back propagation to the hidden layer**

Now we want to compute the partial derivatives \( \frac{\partial E}{\partial w_{ij}} \). Each unit \( j \) in the hidden layer is connected to each unit \( q \) in the output layer with an edge of weight \( w_{jq}^{(2)} \), for \( q = 1, \ldots, m \). The back propagated error up to unit \( j \) in the hidden layer must be computed taking into account all possible backward paths. The back propagated error is then

\[
\delta_j^{(1)} = a_j^{(1)} (1 - a_j^{(1)}) \sum_{q=1}^{m} w_{jq}^{(2)} \delta_q^{(2)} \]  
\[
\tag{A3}
\]

Therefore the partial derivative we are looking for

\[
\frac{\partial E}{\partial w_{ij}} = \delta_j^{(1)}a_j \]  
\[
\tag{A4}
\]

The back propagated error can be computed in the same way for any number of hidden layers and the expression for the partial derivatives of \( E \) keeps the same analytic form.

**Fourth step: weight updates.**

After computing all partial derivatives the network weights are updated in the negative gradient direction. Learning constant defines the
step length of the correction. The corrections for the weights are given by

\[
\Delta w_{ij}^{(2)} = -\gamma o_i^{(1)} \delta_j^{(2)} \quad \text{for } i = 1, \ldots, k+1, j = 1, \ldots, m
\]

\[
\Delta w_{ij}^{(1)} = \gamma o_i^{(1)} \delta_j^{(1)} \quad \text{for } i = 1, \ldots, n+1, j = 1, \ldots, k
\]

(A5)  
(A6)

It is very important to make the corrections to the weights only after the back propagated error has been computed for all units in the network. Otherwise the corrections become intertwined with the back propagation of the error and the computed corrections do not correspond any more to the negative gradient direction. Some authors fall in this trap [16]. Note also that some books define the back propagated error as the negative traversing value in the network. In that case the update equations for the network weights do not have a negative sign (which is absorbed by the deltas), but this is a matter of pure convention.

Symbols definition

\( O^{(1)} \) = computed vector for input  
\( O^{(2)} \) = stored vector  
\( \delta_j^{(2)} \) = back propagated error signal  
\( E \) = error function  
\( \frac{\partial E}{\partial w_{ij}} \) = partial derivatives in respect to the weight  
\( \gamma \) = learning constant defines the step length of the correction  
\( t \) = target vector

References


[12]. Yan Pan and, Roland N. Horn. 1998. Improved methods for multivariate optimization
of field development scheduling and well placement design, paper SPE 49055.


Figure (1) - Basic structure for multilayer perceptron network
Figure (2) – Southern oil field map

Figure (3) – Grid map of the concerning area under study.
Figure (4) – Schematic diagram of data mask technique

Figure (5) - Schematic diagram of training NN behavior
Figure (6) – Neural Network Model Structure

Figure (7) - Comparison between measured and calculated Cum. Production (STB)
Figure (8) – A new suggested locations of infill drilling wells.

Figure (9) – A new suggested locations of infill drilling wells (after reference 14).
تطبيق الشبكات العصبية في أمثلية انتاج حقل نفطي

غوان نوري سعد، عضو جمعية مهندسي البترول، جامعة الكوفة، كلية الهندسة، العراق
الرقابة الأشرف ص.ب. 560
 البريد الإلكتروني: Ghazwanns@eng.kuiraq.com

الخلاصة:

إن تطوير الخطط الاستراتيجية للحول النفطي والغازية تعتبر من المهام الرئيسية للشركات المستمرة
للكن التحول، حيث أن الهدف الرئيسي من هذه الخطط هو زيادة الإنتاج المتراكم بأكبر كمية ممكنة ويشمل أقتصادي في
بمتطلبات الحدود الموضوعية لتلك الخطط.

العمل البحثي المقدم هو لإيجاد أفضل عدد من مواقع الأبار الإنتاجية البيئية كأبار جديدة ولمرحلة تطويرية في الإنتاج الحفلي
وقطع واحد عبر الحياة الإنتاجية للحقل المتكون من أربع قطاعات (حفظ الرميلة الجنوبي العراقي).

استخدم أسلوب الشبكات العصبية لتحديد مواقع تلك الأبار الإنتاجية البيئية لغرض زيادة كمية الإنتاج النفطي من تلك القطاع.

البيانات الإنتاجية تم الحصول عليها وهي بثبيط 143 سبت من البيانات ولستة متغيرات (عدد أبار الحقن، عدد أبار الإنتاج، نسبة
أبار الحقن للإنتاج، نسبة النفط المتبقي، كمية الماء المحقون والتراكم وكمية ماء المنتجة المتراكمة) حيث أعدت على هيئة خريطة
متكاملة لتسهيل وتطبيق الشبكة المستخدمة وهي (FFNN) كأداة مستخدمة من خلال أخير البرامج المستخدمة في هذا
المجال.

البيانات المستخدمة تم أعدادها من خلال استخدام تقنية التقاط الماسك وذلك بتمثيل المنطقة المستخدمة للحساب على شكل ثلاثة
حقائق متحدة المركز حول الموقع المفترض للبن الانتاجي كأسلوب آسرع في حصر نوع البيانات المطلوبة ونوعيتها دون حلقة
التشويش على عمل الشبكة العصبية المختارة ونفس الوقت هو أسرع ودقيق للبيانات الإنتاجية والمكمنية لكل حلقة
حساب خاص الموضوع.

النتائج المستخلصة جيدة ومشجعة لدعم هذه الفكرة في تطبيق التطور الحفلي على القطاع الإنتاجي من الحقل النفطي المختار قيد
الدراسة، وهي بمثابة خطة موضوعية أمام إدارة المكمنين الحفلي لغرض الوصول إلى طبيعة القرار الصعب والمطلوب في هذا
المضارع كي ينتقل إلى مرحلة التطبيق.

الكلمات المفتاحية (الانتاجية، الشبكات العصبية، الإدارة الحفلي، الأمثلية).