Abstract
This study aims to examine the application of pattern recognition technologies to improve the time and efforts required for completing a successful history matching project. The pattern recognition capabilities of artificial intelligence and data mining techniques are used to develop a Surrogate Reservoir Model (SRM) and use it to perform the assisted history matching process. A well-known standard reservoir model, PUNQ-S3, was selected to examine the potentials of SRM in an assisted history matching process.

SRM is a prototype of full field reservoir simulation model that runs in a matter of seconds. SRMs are built based on a spatio-temporal database. The database includes different types of data extracted from a few realizations of the simulation model. In this study, the SRM was developed using ten geological realizations of PUNQ-S3 reservoir simulation model. The uncertain properties are distributions of porosity, horizontal, and vertical permeability. The SRM requires low development cost and has high implementation pace. The SRM was coupled with the Differential Evolution (DE) optimization method to construct an automated history matching workflow. This workflow is able to produce multiple realizations of the reservoir, which match the past performance.

The developed SRM showed a high accuracy in mimicking the behavior of reservoir simulation model. Once we select the best performing cases during history matching, we were able to also obtain reliable future forecasts for the model. The results of this study prove the capability of SRMs in assisting history matching process using population-based sampling algorithms and other computationally intensive operations in reservoir management workflows.

Introduction
The purpose of reservoir management is to develop strategies to maximize hydrocarbon recovery. Reservoir simulation is usually the standard decision making tool used by industry in this workflow. The common concern of reservoir simulation and modeling is its accuracy. It is generally believed that models with higher resolution (in both time and grid domains) are more accurate in terms of reservoir behavior prediction. The new improvements in reservoir data acquisition have increased the complexity of the reservoir model and therefore the time required to run it. Here a well-known dichotomy arises; on one hand the model must satisfy the accuracy requirements (high resolution), and on the other hand it needs to be fast enough for computationally intensive tasks such as history matching and uncertainty quantification.

History matching is an important step of any reservoir management workflow. The main objective of history matching is to improve and validate the reservoir simulation model by incorporating the observed data into the characterization process. The calibrated models are then used to obtain a reliable production forecasts. A simulation model, which has been tuned to match the past performance of a reservoir, generally offers a higher degree of confidence to predict the future. Having a trustworthy prediction of field performance has a direct impact on the technical and financial performance of operators.

History matching, by nature, is an ill-posed inverse problem. Correspondingly, classical history matching where reservoir parameters are adjusted manually in a trial-and-error fashion makes the operation more tedious and time-consuming. Assisted (automated) history matching was proposed to decrease the amount of labor required during the manual history matching. During the last two decades there have been efforts to improve assisted history matching in a way that could be applicable in the real world. Despite all the attempts, due to increasing rate of complexity and resolution in the reservoir models, there is still
hesitation about the practicality and potential of these methods to handle highly complicated real reservoir models. This makes assisted history matching still a challenging and hot research topic.

The earliest studies in the field of history matching started in 1960’s (Kruger, 1961; Wahl, et al., 1962; Jacquard, 1964; Jacquard, et al., 1965; Jahns, 1966; Coats, et al., 1970; Slater, et al., 1971). These studies mainly were based on proposing mathematical reservoir models and then calibrating these models using actual data. An important introduction in 1990s was using experimental design to develop response surfaces to replace the reservoir simulation in history matching workflow (Eide, et al., 1994). During the years after early 90’s, significant works were done to move history matching from a labor-intensive engineer-based framework to a fully or semi-fully automated approach (Tyler, et al., 1993; Palatnic, et al., 1993). In order to address the shortcomings of gradient-based optimization methods, global optimization approaches such as Simulated Annealing, Evolutionary Algorithms, and Evolution Strategy were proposed. Some of the successful methods are such as the ensemble Kalman filter (Van Leeuwen, 1999; Evensen, 2003; Haugen, et al., 2006; Aanonsen, et al., 2009; Hanea, et al., 2010; Szklarz, et al., 2011), Neighborhood Algorithm (Christie, et al., 2002; Stephen, et al., 2006; Rotondi, et al., 2006; Subbey, et al., 2003), Genetic Algorithms (Erbas, et al., 2007) (Castellini, 2005), Scatter search (Sousa, 2007), Tabu Search (Yang, et al., 2007), Hamiltonian Monte Carlo (HMC) (Mohamed, et al., 2009), Particle Swarm Optimization (PSO) (Eberhart, et al., 2001; Mohamed, et al., 2009; 2010; Rwechungura, et al., 2011; Kathrada, 2009) Ant Colony Optimization (ACO) algorithm (Razavi, et al., 2008; Hajizadeh, et al., 2009; 2010), Markov chain Monte Carlo (Maucc, 2007), and Chaotic Optimization (Manicca, 2002).

Increased complexity and simulation time of reservoir models has created a bottleneck for history matching workflows. This is particularly true for history matching workflows that employ a form of population-based sampling algorithms. These algorithms, depending on the number of uncertainty parameters, require few hundred to few thousand simulation calls to converge to optimal regions and find history-matched solutions. This requirement has created well-known barriers for the application of stochastic population-based methods for real-life history matching and uncertainty quantification problems. At the same time, the limitation has motivated an active area of research to reduce the simulation time of reservoir models. Two distinct areas form the current focus of research activities: 1) mathematical models to improve the physics-based simulation 2) reduced order/data-driven approaches as a proxy to full field simulation.

Proxy models as an inexpensive approximation of high computational cost full field simulation models are frequently used in different areas of engineering. By increasing the time and cost required to run the reservoir simulation models, proxy models appeared in petroleum engineering. They are fast and relatively easy to develop. However, due to the practicality concerns there is a long way to completely surpass full field reservoir simulation models in reservoir management plans. Response surface models and reduced order models are the most famous types of proxy models used in petroleum engineering. Reduced order modeling aims to transfer the high dimensional models into a meaningful representation of reduced dimensionality. They have been applied in many areas including petroleum engineering. In recent years, there have been some attempts in using reduced order models for history matching, uncertainty quantification, and optimization (Cardoso, 2009; Cardoso, et al., 2010; He, et al., 2011; Bazargan, et al., 2012; Bazargan, et al., 2013; Wu, et al., 2013) (Gildin, et al., 2014).

Another approach recently going through fast development is data-driven modeling. Data-driven modeling is based on analyzing the available data about a system using machine learning methods. This approach particularly finds the connections between different components of the system without any explicit knowledge of the physical behavior of these components. Statistical methods, application of artificial neural networks, and fuzzy logic are examples of data-driven modeling approaches. A relatively new types of data-driven models utilized in reservoir modeling and simulation are surrogate reservoir models (SRMs). SRMs are built based on artificial intelligence and data mining techniques and are meant to replace or complement the reservoir simulation models.

**Surrogate Reservoir Models**

Surrogate reservoir modeling is the terminology used to describe the new way of reservoir modeling and simulation which is based on using artificial intelligence and data mining (AI&DM) techniques. SRMs are relatively new tools for fast track and comprehensive reservoir analysis, which originate from the existing reservoir simulation models. In other words, SRMs are approximations of the full field three dimensional numerical reservoir models and are capable of accurately mimicking the behavior of these full field models (Mohaghegh, 2014). In this study, SRMs are built based on artificial neural networks. ANNs are non-linear data driven, fact and example based and most importantly a self-adaptive approach. These characteristics make them an ideal modeling tool for petroleum engineering problems (Haykin, 2008; Mohaghegh, 2011; Kriesel, 2011).

Fast track modeling abilities of SRMs suits the necessity of having models with high resolution, accuracy, and pace in the reservoir management workflow. Since the advent of SRMs in 2006 (Mohaghegh, et al.) many successful examples of their applications have been published (Mohaghegh, et al., 2009; Jalali, et al., 2009; Amini, et al., 2012; Amini, et al., 2014; Shahkarami, et al., 2014a). Mohaghegh et al. (2012a; 2012b) have discussed the results of several projects involving surrogate reservoir models for the fast track analysis of numerical simulation models. Other publications regarding the SRMs can be found in variety of reference materials (Mohaghegh, 2009; Mohaghegh, 2011; Mohaghegh, 2014; Shahkarami, et al., 2014a; Amini, et al., 2014).
When the purpose of developing of SRM is to use in a history matching study, SRM outputs are at the well location (for example well production). In this case the SRM is referred as well-based SRM. If the outputs are at the grid level (such as pressure and saturation at grid block) they are known as grid-based SRM. Also depending on the objective of the study, the training realizations (required for SRM development) are varied in geological properties or operational conditions. For instance, a history matching study requires changing the geological characteristics and a production optimization analysis needs variation in operational conditions. An uncertainty assessment study might include both types of these properties. In order to have a successful surrogate reservoir model, there are important points that should be considered. Preparation and assembling the realizations of reservoir simulation in a way that suits the features of AI&DM techniques are critical. The skill and knowledge of the user in reservoir engineering as well as the basics of AI&DM techniques play an important role for this purpose. The details of development and application of SRMs have been thoroughly discussed here (Shahkarami, 2014).

In our previous study (Shahkarami, et al., 2014b), an SRM was created for a synthetic reservoir model of a heterogeneous oilfield with 24 production wells and 30 years of production history. Consequently, the SRM was used as the substitute of reservoir simulation model in history matching process. The history matching was performed using the SRM and by tuning only one reservoir characteristic (permeability) throughout reservoir. The result proved the potential of SRM for fast track and accurate reproduction of the numerical model results during history matching.

This article is the further investigation of SRM capabilities in order to achieve the history match of a real-life problem. Thus, a standard test reservoir model, known as PUNQ-S3 reservoir model in the petroleum engineering literature, was selected. The PUN-Q-S3 reservoir model represents a small size industrial reservoir engineering model (Floris, et al., 2001). This model has been formulated to test the ability of various methods in the history matching and uncertainty quantification. The surrogate reservoir model was developed (trained, calibrated, and validated) using a small number of geological realizations of PUNQ-S3 reservoir model. The uncertain properties in this model are distributions of porosity, horizontal, and vertical permeability. In order to complete an automated history matching workflow, the developed SRM was coupled with a global optimization algorithm called Differential Evolution (DE). DE optimization method is considered as a novel and robust optimization algorithm from the class of evolutionary algorithm methods. This automated workflow is able to produce multiple realizations of the reservoir which match the past performance. The successful matches were utilized to quantify the uncertainty in the prediction of cumulative oil production.

**Differential Evolution Optimization Algorithm**

The Differential Evolution (DE) algorithm was developed by Storn and Price (1995) as a stochastic population based algorithm for continuous and real-valued numerical optimization problems (Storn, et al., 1997; Price, et al., 1997; Price, et al., 2005). DE belongs to the category of evolutionary algorithms and, like other EA methods (such as genetic algorithms), it has mutation, recombination, and selection steps. Because of simple mathematical structure, DE is considered as a very effective global optimization algorithm. The low number of control parameters, makes DE simple, fast, and easy to apply. The details about DE can be found somewhere else (Shahkarami, 2014).

The DE algorithm randomly generates the first set of solutions, consisting of \( N_p \) vectors. After obtaining the objective function values for each of \( N_p \) members, the algorithms randomly combines two vectors among the current population and calculates the difference vector between these two members. The difference vector is then multiplied by a real number called the scaling factor \( F \in [0, 2] \) that controls the perturbation of this vector. The scaled difference vector is then added to a third randomly selected vector. After a crossover stage to increase the population diversity, objective function values are evaluated for each member of the population. Each trial vector is now compared against the population vector of the same index and wins the competition if it has a lower objective function value.

DE recently has been applied to a variat of petroleum engineering case studies (Wang, et al., 2006) (Decker, et al., 2006; Jahangiri, 2007; Hajizadeh, et al., 2009; 2010; Wang, et al., 2010; Wang, et al., 2011; Mirzabozorg, et al., 2013) (Okano, 2013).

**Implementation of SRM on PUNQ-S3 Problem**

**PUNQ-S3 Reservoir Model**

The PUNQ-S3 reservoir simulation model was built during the PUNQ project. PUNQ, which stands for Production forecasting with Uncertainty Quantification, was a mutual study supported by European Union and conducted by 10 European companies, universities, and research centers (Floris, et al., 2001). The reservoir model is based on a real field operated by Elf Exploration and Production (Floris, et al., 2001; Barker, et al., 2001) and is used widely as a standard synthetic test case to investigate the capability of different methods of history matching and uncertainty quantification (Floris, et al., 2001; Barker, et al., 2001; Gu, et al., 2005; Gao, et al., 2005; Abdollahzadeh, et al., 2011).

The reservoir model consists of \( 19 \times 28 \times 5 \) grid blocks (180 m by 180 m), of which a total of 1,761 grid blocks are active. The geometry of the field has been modeled using corner-point geometry. The field is bounded to the east and south by a fault. There is also a fairly strong aquifer located in the north and west of reservoir. Due to presence of aquifer and the resulting pressure support, no injection wells were considered. In addition, there is a small gas cap in the first layer and in the center of
the dome shaped structure. In order to avoid the gas production from the gas cap no well has been perforated in the first layer.

Figure 1 demonstrates the top structure of the PUNQ-S3 reservoir model. As it is shown in this figure, there are six production wells drilled in the reservoir. Layers one and two are left without perforation. The other layers are completed for different wells: wells PRO-1, PRO-4 and PRO-12 are perforated in layers 4 and 5. The wells PRO-5 and PRO-11 have been completed in layers 3 and 4 and well PRO-15 has been perforated only in layer 4.

![Figure 1: The top structure of PUNQ-S3 model. The field is bounded to the east and south by a fault, and links to the north and west to a fairly strong aquifer. In addition there is a small gas cap in the center of the dome shaped structure (Floris, et al., 2001).](image)

In order to provide similar and identical data for everyone, a “true” case has been designed. The main characteristics to generate the true case are porosity and permeability (horizontal and vertical) distributions. The values of these properties at well sites are taken from the original model. The comprehensive procedure of creating the porosity and permeability distributions for true case can be found somewhere else (Barker, et al., 2001; 2014). Then the outputs of the true case are considered as the actual historical data. Following are the available data provided for PUNQ-S3 reservoir model:

- Porosity and permeability values at well locations
- Geological descriptions for each layer
- Production history for the first 8 years (for history matching study)
- Cumulative production (total oil recovery) after 16.5 years (for uncertainty quantification and production forecast study)
- PVT, relative permeability and Carter-Tracy aquifer dataset all taken from original field data
- There is no capillary pressure function
- Gas oil contact (GOC) and water oil contact (WOC) values

**SRM Development**

The artificial intelligence based reservoir models including SRMs are constructed based on a spatio-temporal database. Depending on the objective of study this database contains different types of data. The source of data is from different realizations of the reservoir simulation model. The data are extracted from these realizations to create the database. The main goal of this database is to teach the model the process of fluid flow phenomena in the reservoir. From one point of view the data in this database can be categorized as static and dynamic data. The static data refer to the properties, which are constant overtime such as porosity, permeability, top depth, and thickness. Also dynamic data are those ones, which are not necessarily fixed overtime like operational constraints, the production at wells, pressure, and phase saturations at the grid blocks, etcetera.

In this study, the uncertain properties for developing the SRM and matching the field performance include: porosity, vertical, and horizontal permeabilities. These properties are the most common uncertain reservoir characteristics used in the literature to match history data of PUNQ-S3 reservoir model. Similar to what happens in reality, it is assumed that these properties are measured at the well locations (well logging and core data samples). Also the geological descriptions of this model indicate the streaks of high porosity/permeability profiles in the reservoir (Floris, et al., 2001; Barker, et al., 2001). These types of information were used to generate training realizations.
Informative Simulation Runs Representation of Reservoir Uncertainties

The source of information in spatio-temporal database is the different realizations of the reservoir simulation model. Again, depending on the goal of study, the preparation of these realizations would be different. These realizations differ from each other in the variable uncertain properties. These uncertain properties are the variables, which study would like to investigate their impacts on the output of reservoir model.

Based on the property values provided at well sites and also the geological descriptions, ten different realizations of the reservoir were created. In order to create these realization a sampling method (Latin Hypercube) was utilized. The details of generating these realizations are provided here (Shahkarami, 2014).

Reservoir Delineation and Tier System

Data summarization is an essential task during SRM development. One way of data summarization is to delineate the reservoir to different segments and make an average of data over the segments of reservoir. A common way of delineation process is based on a famous theory known as Voronoi graph theory (Erwig, 2000; Gomez, et al., 2009). In this study we used the feedback from the available geological descriptions and modified some of Voronoi diagrams. Figure 2 depicts the designed drainage areas. Consequently each one of these drainage areas are divided into four tiers. The first tier is the well block, which has a significant impact on the well behavior. The second tier includes the first row of grid blocks around the well block. The third tier is the next row of grid blocks around the second tier. Finally the rest of grid blocks in the drainage area are summed up in the fourth tier. The average value of reservoir characteristics was calculated at each tier and assigned to the corresponding well and tiers in the database. Figure 3 is a scheme of the designed tier system in this study.

SRM Inputs and Output

Due to the complexity of PUNQ-S3 reservoir model we face a higher number of input parameters in this case study compared to the previous examples (Shahkarami, et al., 2014b). The oil production rate is the main constraint during the eight years of production for this reservoir. Thus the outputs of this model, which need to be matched, are well bottom-hole pressure, gas production rate, and water production rate. For each one of these outputs one ANN was created (totally three ANNs). PUNQ-S3 model has five layers and a total of six wells. Each drainage area was divided into four tiers and also we have three uncertain parameters (porosity, vertical, and horizontal permeability). Consequently there are 360 uncertain (adjustable) parameters ($5 \text{ layers} \times 4 \text{ tiers} \times 6 \text{ wells} \times 3 \text{ properties}$), which can be tuned in order to match the history data. In order
to build the SRM, just considering a single well we need to at least include 60 parameters for each well. These do not contain the other types of data such as thickness and top for each tier. Generally we have a database with more than 120 inputs. Figure 4 summarizes the types of input and output in the spatio-temporal database for PUNQ-S3 reservoir model. Making sure to select the right inputs is not an easy task to do. Many of artificial based models fail in this step (Zubarev, 2009; Mohaghegh, et al., 2012a).

![SRM Inputs](image1)

**Table 1:** Selected inputs for the well bottom-hole pressure network

<table>
<thead>
<tr>
<th>Static Inputs</th>
<th>Dynamic Inputs</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude (X)</td>
<td>Well bottom-hole Pressure @ (t-1) and (t-2)</td>
<td>Well bottom-hole Pressure</td>
</tr>
<tr>
<td>Longitude (Y)</td>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Horizontal Permeability</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
<tr>
<td>Vertical Permeability</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
<tr>
<td>Thickness</td>
<td>@ 5 Layers and 4 Tiers</td>
<td>Oil production rate (operational constraints)</td>
</tr>
<tr>
<td>Top</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2:** Selected inputs for the gas production rate network

<table>
<thead>
<tr>
<th>Static Inputs</th>
<th>Dynamic Inputs</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude (X)</td>
<td>Gas production rate @ (t-1) and (t-2)</td>
<td>Gas production rate</td>
</tr>
<tr>
<td>Longitude (Y)</td>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Horizontal Permeability</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
<tr>
<td>Vertical Permeability</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
<tr>
<td>Thickness</td>
<td>@ 5 Layers and 4 Tiers</td>
<td>Oil production rate (operational constraints)</td>
</tr>
<tr>
<td>Top</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3:** Selected inputs for the water production rate network

<table>
<thead>
<tr>
<th>Static Inputs</th>
<th>Dynamic Inputs</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude (X)</td>
<td>Water production rate @ (t-1) and (t-2)</td>
<td>Water production rate</td>
</tr>
<tr>
<td>Longitude (Y)</td>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Horizontal Permeability</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
<tr>
<td>Vertical Permeability</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
<tr>
<td>Thickness</td>
<td>@ 5 Layers and 4 Tiers</td>
<td>Oil production rate (operational constraints)</td>
</tr>
<tr>
<td>Top</td>
<td>@ 5 Layers and 4 Tiers</td>
<td></td>
</tr>
</tbody>
</table>

Training, Calibrating and Validating the ANNs

The training process of an SRM includes three different steps: training (learning), calibration, and validation procedures. Based on that, the spatio-temporal database is divided into three sets: the training or learning set, calibration set, and validation or verification set. The training set is part of the data shown to the ANNs during the training process. The ANNs are adapted to this set to match the provided outputs (reservoir simulation results). On the other hand, the calibration set is not used to adjust the outputs. This set is utilized to assure that any increase in accuracy over the training data set will lead to an increase in accuracy over a data set that ANNs has not seen them during the training part. This set of data is helpful in determining when the training should be stopped. Finally, the verification set is a part of the database used to verify the predictability of the trained
ANNs, and, subsequently, this data set is not used to train the ANNs.

In this study, 80% of the database was used as training set. The share of calibration and validation sets is also 10% for each one. For each output we have one network; therefore, a total three networks are set. All three networks contain one hidden layer (Figure 5). Consequently the SRM is created by integrating three neural networks after the training process is completed. It is worth mentioning that the elapsed time to perform the training process (learning, calibration, and verification) is negligible, particularly when it is compared to the reservoir simulation run-time.

**Figure 5: ANN structure used for training the SRM**

A further validation step is applied to test the robustness of the SRM. This step is referred to as “Blind Verification”. The term “blind” indicates a set of realization(s) that has not been used during the training process. These blind testing sets are complete realizations of the reservoir, whereas the verification set used in the training process is a randomly selected portion of the spatio-temporal database.

Finally the developed and validated SRM was combined with DE optimization method to build up the automated workflow for history matching. The objective functions estimate the misfits between SRM estimation and actual data. In order to begin the automated workflow, the ranges of properties are given (by user) to the optimization algorithm. These ranges are from the geological description provided for PUNQ-S3 model. In this workflow the range of properties, objective functions, and DE parameters (such as stoppage criteria) are the only items that user provides. The automated workflow was set to be stopped automatically after 3000 runs of SRM (stoppage condition). However, the process can be stopped at any time or reaching a specific value of misfit.

**Results**

In this section we present some results for development and application of SRM for PUNQ-S3 problem. The first set of results belongs to a training realization used to train the SRM.

Figure 6 is the comparison of well bottom-hole pressure results generated by SRM and similar results from simulator. The figure portrays six profiles corresponding to six wells. Blue markers are SRM results over eight years (almost 3000 days) of reservoir life compared to simulator (CMG) outputs (red line). As it was mentioned earlier, the oil rate is the main constraint during this period. Therefore, we have well bottom-hole pressure, gas rate, and water rate date to match.
Figure 6: Comparison of well bottom-hole pressure profile generated by the SRM (blue markers) with the similar results from a numerical simulator (CMG) for a training realization of PUNQ-S3 reservoir model.

Figure 7 depicts the results of gas rate production for a training realization. It is a comparison between the results of SRM and simulator for six wells in PUNQ-S3 reservoir model. For eight years of history data, it is just one well (well PRO-11) that has water breakthrough. Figure 8 compares the results of water rate production generated by SRM and same results from simulator. For the first five years there is no water production, and then slowly we observe some water production. However, SRM is able to properly capture the zero water production.
Figure 7: Comparison of gas production rates generated by the SRM (blue markers) with the similar results from a numerical simulator (CMG) for a training realization of PUNQ-S3 reservoir model

Figure 8: Comparison of water production rate generated by the SRM (blue markers) with the similar results from a numerical simulator (CMG) for a training realization of PUNQ-S3 reservoir model

Blind Verification Run

The training results show good performance of SRM during the training process. It sounds the SRM has captured the behavior of wells through the database. Nevertheless, SRM should show the same quality over the realizations that have not been used during the training process. For this purpose we implemented the trained SRM on a completely blind realization. Figure 9 is the validation results for well bottom-hole pressure. Similar to the training results, blue markers represent SRM results while compared with simulator outcome (red line).
Figure 9: Validating the SRM using a blind run. The comparison of well bottom-hole pressure from the SRM with the results of numerical simulation model for PUNQ-S3 problem.

Figure 10 and Figure 11 are the validation results for gas production and water production rates, respectively.
Once the robustness of SRM was assured, it is ready to be used in history matching process. One advantage of the automated history matching workflow is to offer multiple realizations that match the field data. In this study we selected top ten best matches. Figure 12 displays the results of history matching for well bottom-hole pressure. There are six profiles of bottom-hole pressure for six wells present in PUNQ-S3. Each diagram compares the results of top ten matches (blue lines) with the actual data (red circles).
Figure 12: History matching results of well bottom-hole pressure. The comparison of 10 best matches (blue lines) coming from SRM with the actual data (red circles)

Figure 13 and Figure 14 demonstrate the comparison between the ten best matches (blue lines) and the actual data (red circles) for gas production. On the left side we have the gas production rate and the right side diagrams show the cumulative gas production.
Figure 13: History matching results of gas production (rates are on the left side and cumulative production are on the right side of the figure). Comparison of 10 best matches (blue lines) coming from the SRM with the actual data (red circles). The data belong to the wells PRO-1, PRO-11, PRO-12 and PRO-15.
Figure 14: History matching results for gas production (rates are on the left side and cumulative production are on the right side of the figure). Comparison of 10 best matches (blue lines) coming from the SRM with the actual data (red circles). The data belong to wells PRO-4 and PRO-5.

Finally, Figure 15 compares the results of top ten matches (blue lines) with the actual data (red circles) for water production. In this figure, the diagram on the left is water rate production and the right side depicts the cumulative gas production both for well PRO-11.

Figure 15: History matching results of water production (rates are on the left side and cumulative production are on the right side of the figure). Comparison of 10 best matches (blue lines) coming from the SRM with the actual data (red circles). In this study we have just one well (PRO-11) with water breakthrough during eight years of production history.

Importing the Matched Characteristics into the Simulator

In this study, the developed SRM substituted an industrial reservoir simulator (CMG-IMEX™ (CMG, 2013)) during history matching process. Therefore, we designed an automated SRM-based history matching workflow. This workflow is able to provide multiple realizations of reservoir, which match the actual data. We chose to have ten best matches. These ten realizations were imported into the simulator in order to observe the performance of simulator with the inputs coming from the SRM. Figure 16 demonstrates the field cumulative oil production results of the simulator after importing the matched properties (match # 1) from the SRM into the simulator. This graph compares the simulator results (red line) with the actual field cumulative oil production (blue circles). As it was mentioned earlier there are eight years of field data available for history...
matching purposes. In addition to eight years of history data, we have field cumulative oil production after 16.5 years. This value can be used for future forecast comparison. It should note this information was not used during the history matching process. The green dot in the Figure 16 represents this value.

![Figure 16](image)

**Figure 16**: Comparison between the SRM based history matching results (match #1) and actual data for cumulative oil production. The red line represents the matched realization and the blue circles are the actual field data (eight years of production history). The green dot also displays the cumulative production for the true case after 16.5 years.

Similar to the oil production, Figure 17 shows the comparison between the simulator results and actual field data for the cumulative gas production.

![Figure 17](image)

**Figure 17**: Comparison between the SRM based history matching results (match #1) and actual data for cumulative gas production. The red line represents the matched realization and the blue circles are the actual field data (eight years of production history). The green dot also displays the cumulative production for the true case after 16.5 years.

As it was mentioned earlier, the water breakthrough happens at the seventh year for just one well (PRO-11). Therefore, except for the well PRO-11 that produces water, the other wells have a very little water production. Figure 18 compares the cumulative water production between the simulator and actual data.
Figure 18: Comparison between the SRM based history matching results (match #1) and actual data for cumulative water production. The red line represents the matched realization and the blue circles are the actual field data (eight years of production history). The green dot also displays the cumulative production for the true case after 16.5 years.

Discussion and Concluding Remarks

In this article a surrogate reservoir model was developed for the PUNQ-S3 reservoir simulation model. The PUNQ-S3 model is widely accepted as a standard reservoir simulation model to test the ability of different methods on history matching and uncertainty quantification. The characteristics of this model make PUNQ-S3 model a unique case to study the potential of SRM for history matching. The variable properties to create the SRM are porosity and permeability (horizontal and vertical) distributions. In order to train the SRM, ten realizations of PUNQ-S3 simulation model were generated. An extra realization (11\textsuperscript{th} case) was used to verify the trained SRM.

One important feature of an effective history matching workflow is the automation ability. Therefore, the developed SRM was coupled with the DE optimization algorithm. The objective functions were created to calculate the misfit values between the actual data and measured results (SRM). The goal of history matching was to match eight years of history data available for three different properties. These properties include the well bottom-hole pressure, gas production rate, and water production rate. The stoppage condition for the automated history matching workflow was 3,000 times calling the SRM. Also this workflow was able to report multiple realizations of the reservoir that match the actual data. Thus, ten best matches were selected to be used for the future forecast. The future forecast consists of predicting the field cumulative oil production after 16.5 years production. Beside the eight years of history data, the PUNQ project provides the field cumulative oil production after 16.5 years for the purpose of future production comparison.

Figure 6, Figure 7, and Figure 8 show the results of SRM during the training process for the well bottom-hole pressure, gas production rate, and water production rate, respectively. These graphs portray the comparison between the SRM results with the outputs of simulator. The superb match between the results of SRM and simulator proves that SRM has been well-trained. The ability of SRM to capture the zero values of gas and water production rates is clear in the Figure 7 and Figure 8.

Figure 9, Figure 10, and Figure 11 depict the performance of SRM on a completely unseen realization of reservoir during the training process. This step, referred to as blind realization, shows the robustness of the SRM. The quality of the match for the blind case, as it is seen in these graphs, is not as good as the training realizations (Figure 6, Figure 7, and Figure 8). This is a normal and expected response of SRM to the set of data which are not used in the SRM training.

The trained and verified SRM was used to perform the history matching. Figure 12, Figure 13, Figure 14, and Figure 15 are the results of ten best history matched realizations. The results are the outputs of SRM. These graphs are the comparison between the ten best matches with the actual data for the well bottom-hole pressure, gas production rate, cumulative gas production, water production rate, and cumulative water production. The matches for the well bottom-hole pressure and gas production rate are excellent. For the water production rate (Figure 15), the matches are acceptable. The SRM matched the zero values of water production very well; however for the non-zero values, the quality of matches decrease (although they are still acceptable).

Figure 16 displays the results of the best achieved match imported into the simulator. This figure compares the results of this realization with the actual data for the field cumulative oil production. The results show a good match for the eight years of
available history. Also, this graph predicts the field cumulative oil production for the next 8.5 years. At the end of this time period, the prediction performance has been compared with the reported value. Although the match shows an excellent quality, the prediction is slightly overestimating the future production. Figure 17 and Figure 18 are the same comparison for the field cumulative gas and water production. For the gas production, we see same overestimating behavior; however water production has been slightly underestimated.

The general results in this study demonstrate the robustness of SRM in history matching and future prediction for the PUNQ-S3 problem. Numerous studies have used the PUNQ-S3 reservoir model to test the methods on history matching. Many of these studies are the investigation of different optimization methods for automated history matching. Generally these optimization algorithms have been coupled with a commercial simulator. The reported numbers of simulation runs for history matching the PUNQ-S3 reservoir model are in the order of thousands realizations (Hajizadeh, et al., 2010; 2009; Abdollahzadeh, et al., 2011). In this study, the required simulation runs to create and validate the SRM (eleven runs) are extremely less. Although, the run-time is not an issue for the PUNQ-S3 reservoir simulation model; in reality, a typical reservoir simulation model is more time-consuming to run and requires higher computational cost. In such a case using a numerical reservoir simulator for history matching would be a huge computational issue. In conclusion, the application of SRM for history matching purposes would be a great asset in the reservoir management workflow.

Acknowledgement

The authors wish to extend their appreciation to Computer Modeling Group (CMG) and Intelligent Solution Inc. (ISI) for providing the software applications for reservoir simulation and for development of the SRM, respectively.

Reference


