Assisted history matching using pattern recognition technology

Alireza Shahkarami*

School of Sciences,
Saint Francis University,
Science Center 003,
P.O. Box 600, Loretto, PA, 15940, USA
Email: Ashahkarami@francis.edu
*Corresponding author

Shahab D. Mohaghegh

Benjamin M. Statler College of Engineering and Mineral Resources,
West Virginia University,
345-E Mineral Resources Bldg.,
P.O. Box 6070, Morgantown, WV, 26506, USA
Email: Shahab@wvu.edu

Yasin Hajizadeh

University of Calgary,
2500 University Dr. NW, Calgary,
Alberta, T2N 1N4, Canada
Email: yhajizad@ucalgary.ca

Abstract: This study examines the application of pattern recognition technologies to improve the time and effort required for completing successful history matching projects. The pattern recognition capabilities of artificial intelligence and data mining techniques are used to develop a surrogate reservoir model (SRM), which is then employed to perform the assisted history matching process. A well-known reservoir model, PUNQ-S3, was selected to study the potentials of the SRM in an assisted history matching process. The SRM is a prototype of a full-field reservoir simulation model that demands a low development cost and has a high implementation pace. SRMs are built based on a spatio-temporal database, which includes different types of data extracted from a few realisations of the simulation model. The SRM was coupled with the differential evolution optimisation method to construct an automated history matching workflow. The results of this study prove the SRMs' capability in assisting history matching processes.
[Received: December 3, 2015; Accepted: June 17, 2016]

Keywords: history matching; pattern recognition; surrogate reservoir models; PUNQ-S3; differential evolution.

Reference to this paper should be made as follows: Shahkarami, A., Mohaghegh, S.D. and Hajizadeh, Y. (2018) 'Assisted history matching using pattern recognition technology', *Int. J. Oil, Gas and Coal Technology*, Vol. 17, No. 4, pp.412–442.

Biographical notes: Alireza Shahkarami is an Assistant Professor of Petroleum and Natural Gas Engineering at Saint Francis University. He earned his MSc and PhD from West Virginia University. His research focuses on the implementation of artificial intelligence and data mining techniques to address conventional and unconventional problems in petroleum engineering. His research interests also include CO_2 sequestration and storage, reservoir simulation and modelling, enhanced oil recovery (EOR), and unconventional reservoirs.

Shahab D. Mohaghegh is the President and CEO of Intelligent Solutions, Inc. (ISI) and a Professor of Petroleum and Natural Gas Engineering at West Virginia University. He is pioneer in the application of artificial intelligence and data mining in the exploration and production industry. He holds BSc, MSc, and PhD degrees in Petroleum and Natural Gas Engineering.

Yasin Hajizadeh works as a Senior Data Scientist for Schlumberger in Houston and holds an Adjunct Assistant Professor appointment with the Department of Computer Science, University of Calgary. Before joining Schlumberger in 2012, he briefly experimented with being an Optimisation and Uncertainty Management Scientist for Computer Modelling Group (CMG) in Calgary. He gets excited about technology and innovation management with applications to scientific software development, machine learning, artificial intelligence, data analytics, stochastic optimisation and uncertainty quantification. He obtained his PhD in Petroleum Engineering from Heriot Watt University, Edinburgh in 2011 and a Master of Technology Management from Memorial University of Newfoundland, St. John's in 2016.

This paper is a revised and expanded version of a paper entitled 'Assisted history matching using pattern recognition technology' presented at the SPE Digital Energy Conference and Exhibition, The Woodlands, Texas, USA, 3–5 March 2015.

1 Introduction

The purpose of reservoir management is to develop strategies that maximise hydrocarbon recovery. Reservoir simulation is usually the standard decision-making tool used by industry in reservoir management workflow. A critical concern of reservoir simulation and modelling is accuracy. It is generally believed that models with higher resolutions in time and space are more accurate in terms of reservoir behaviour prediction. The new improvements in reservoir data acquisition have increased the complexity of the reservoir model, and therefore, the time required to run it. However, a compelling paradox arises: on the one hand, the model must satisfy the accuracy requirements of high resolution; on the other hand, the model 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 characterisation process. The calibrated models are then run to obtain reliable production forecasts. A simulation model tuned to match the past performance of a reservoir generally offers a higher degree of confidence to predict future reservoir behaviour. Having a trustworthy prediction of field performance directly affects on the technical and financial performance of operators.

Because of its inverse problem-solving process, history matching is tedious. Traditional history matching where reservoir parameters are adjusted manually in a trial-and-error fashion makes the operation more time-consuming and cumbersome. Assisted (automated) history matching was proposed to decrease the amount of labour required during the manual history matching. During the last two decades, efforts have been made to improve assisted history matching in a way that could be applicable in the real world (Chen and Oliver, 2010). Given the increasing rate of complexity and the high-resolution demands of reservoir models, the practicality and potential of these methods to handle highly complicated real reservoir models remain questionable. This continues to make assisted history matching a demanding research topic.

The 1960s set the stage for the earliest studies in the field of history matching (Kruger, 1961; Wahl et al., 1962; Jacquard, 1964; Jacquard and Jain, 1965; Jahns, 1966; Coats et al., 1970; Slater and Durrer, 1971). By and large, these studies sought to propose mathematical reservoir models calibrated with the aid of actual data. An important introduction in the 1990s was the use of experimental design to develop response surfaces that would replace reservoir simulation in the history matching workflow (Eide et al., 1994). Researchers aimed to move history matching from a labour-intensive user-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 the gradient-based optimisation methods, global optimisation approaches such as simulated annealing, evolutionary algorithms, and evolution strategy were proposed. Among such successful methods number: 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), the neighbourhood algorithm (Christie et al., 2002; Stephen et al., 2006; Rotondi et al., 2006; Subbey and Christie, 2003), the genetic algorithms (Castellini, 2005; Erbas and Christie, 2007), the scatter search (Sousa, 2007), the Tabu search (Yang et al., 2007), the Hamiltonian Monte Carlo (HMC) (Mohamed et al., 2009), the particle swarm optimisation (PSO) (Eberhart and Shi, 2001; Mohamed et al., 2009; Kathrada, 2009, 2010; Rwechungura et al., 2011), the ant colony optimisation (ACO) algorithm (Razavi and Jalai-Farahani, 2008; Hajizadeh et al., 2009a, 2010), the Markov chain Monte Carlo (Maucec, 2007), and the chaotic optimisation (Mantica, 2002).

The increased complexity and simulation time of reservoir models have created a bottleneck for history matching workflows. This is particularly true for history matching workflows that employ a form of population-based sampling algorithms. Depending on the number of uncertainty parameters, these algorithms require a few hundred to a few thousand simulation calls to converge to optimal regions and find history-matched solutions (Hajizadeh, 2010). Such constraints have generated well-known barriers in the application of stochastic population-based methods for real-life history matching and uncertainty quantification problems. At the same time, the limitation has incited an active area of research to reduce the simulation time of reservoir models. From the current focus of research activities, two distinct areas stand out:

1 mathematical models that improve the physics-based simulation

2 reduced order/data-driven approaches as proxies that approximate the full field simulation.

Proxy models represent an inexpensive approximation of full field simulation models, which usually entail a high computational cost. Proxy models, therefore, are frequently used in different areas of engineering (Thomas and Vernon, 1997; Storn, 1996, 1999; Liu and Lampinen, 2002). As the time and the cost required running reservoir simulation models increased, proxy models grew in popularity in the petroleum engineering field. Although they are fast and relatively easy to develop, there is still a long way to completely surpass full field reservoir simulation models in reservoir management plans, mainly due to practicality concerns. Response surface models and reduced order models constitute the most famous types of proxy models used in petroleum engineering (Goodwin, 2015). Reduced order modelling aims to transfer the high dimensional models into a meaningful representation of reduced dimensionality. In recent years, there have been some attempts in using reduced order models for history matching, uncertainty quantification, and optimisation (Cardoso, 2009; Cardoso and Durlofsky, 2010; He et al., 2011; Bazargan and Christie, 2012; Wu et al., 2013; Klie, 2013; Gildin et al., 2014).

Another approach recently going through a fast development is data-driven modelling. Data-driven modelling analyses the available data from a system using machine learning methods. Typically, data-driven modelling finds the connections between different components of a system without any explicit knowledge of the physical behaviour of these components. Statistical methods, the application of artificial neural networks (ANNs), and fuzzy logic are examples of data-driven modelling approaches. Surrogate reservoir models (SRMs), a relatively new technology in reservoir modelling and simulation, employ artificial intelligence and data mining (AI&DM) techniques and are meant either to replace or to complement existent reservoir simulation models. Examples of applications of data-driven proxy models can be found in the literature (Alimonti and Falcone, 2004; Artun et al., 2009; Graf et al., 2011; Fedutenko et al., 2012; Dzurman et al., 2013; Klie, 2015).

2 Surrogate reservoir models

'Surrogate reservoir modelling' is the terminology used to describe the new technology in reservoir modelling and simulation that employs AI&DM techniques. That they originate from the existing reservoir simulation models is important for these relatively new tools of fast track and comprehensive reservoir analysis which receive approximations of the full field three dimensional numerical reservoir models and are capable of accurately capturing the behaviour of these full field models (Mohaghegh, 2014). In this study, the SRMs are built based on ANNs. ANNs are nonlinear data-driven, fact and example based, and most importantly self-adaptive approaches. These characteristics render them an ideal modelling tool for petroleum engineering problems (Haykin, 2008; Kriesel, 2011).

The fast track modelling abilities of SRMs suit the necessity of having models with a high resolution, accuracy, and pace in the reservoir management workflow.

When the purpose of developing an SRM is to use it in a history matching study, SRM outputs describe the reservoir properties at the well location (for example, the well production). In this case, the SRM is referred to as a well-based SRM. If the outputs are at the grid level (such as pressure and saturation at grid block), the model is known as a grid-based SRM. Also, depending on the objective of the study, the training realisations required for the SRM development vary in regard to geological properties or operational conditions. For instance, a history matching study requires changing the geological characteristics, and a production optimisation analysis needs variation in operational conditions. An uncertainty assessment study might include both types of properties. In order to have a successful SRM, several important points should be considered. For instance, preparing and assembling the realisations of the 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 in the development and application of SRMs have been thoroughly discussed (Shahkarami, 2014).

The potential of SRMs have been previously proven, when an SRM was created for a synthetic reservoir model of a heterogeneous oilfield with 24 production wells and 30 years of production history (Shahkarami et al., 2014a). Consequently, the SRM was used as the substitute of the full field reservoir simulation model in the history matching process. By tuning only one reservoir characteristic – permeability – throughout the reservoir, the SRM performed the history matching. Overall, the study validated the potential of the SRM for a fast track and accurate reproduction of the numerical model results during history matching.

Application of SRMs as approximations of the full-field three-dimensional numerical reservoir models have been utilised in areas such as sensitivity analysis (Amini et al., 2014; Shahkarami et al., 2014a), production optimisation (Mohaghegh, 2014), and uncertainty assessment (Tayari et al., 2015).

This article furthers the investigation of SRM capabilities in order to achieve the history match of a real-life problem. For this purpose, we selected a standard test reservoir model, known as the PUNQ-S3 reservoir model in petroleum engineering literature, which 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 history matching and uncertainty quantification. The SRM was developed (trained, calibrated, and validated) using a small number of geological realisations of the PUNQ-S3. We determined that the uncertain properties in this model were the distributions of porosity and the horizontal and vertical permeabilities. In order to complete an automated history matching workflow, the newly generated SRM was coupled with a global optimisation algorithm called differential evolution (DE). The DE optimisation method is considered a novel and robust optimisation algorithm from the class of evolutionary algorithm methods (Das and Suganthan, 2009; Neri and Tirronen, 2010). The automated workflow was able to produce multiple realisations of the reservoir, which matched the reservoir past performance. The successful matches were utilised to quantify the uncertainty in the prediction of cumulative oil production.

3 DE optimisation algorithm

Storn and Price (1995) developed the DE algorithm as a stochastic population-based algorithm for continuous and real-valued numerical optimisation problems (Storn and Price, 1997; Price and Storn, 1997; Price et al., 2005). The DE belongs to the category of evolutionary algorithms, and, like other evolutionary algorithm methods such as genetic algorithms, it consists of three steps: mutation, recombination, and selection. Because of its simple mathematical structure, the DE constitutes a very effective global optimisation algorithm. The low number of control parameters renders the DE simple, fast, and easy to apply.

The DE algorithm randomly generates the first set of solutions, consisting of N vectors. After obtaining the objective function values for each of N members, the algorithm 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. Next, the scaled difference vector is added to a third randomly selected vector. After a crossover stage to increase 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, in the case of a minimisation problem. A detailed description of the DE can be found in (Shahkarami, 2014).

The DE algorithm has been recently applied to a variety of petroleum engineering case studies (Wang and Buckley, 2006; Decker and Mauldon, 2006; Jahangiri, 2007; Hajizadeh et al., 2009b, 2010; Wang and Gao, 2010; Wang et al., 2011; Mirzabozorg et al., 2013; Okano, 2013).

4 Implementing the SRM on the PUNQ-S3 problem

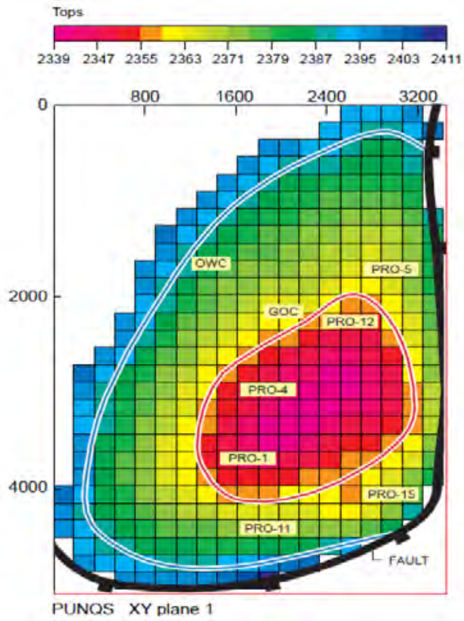
4.1 The PUNQ-S3 reservoir model

The PUNQ-S3 reservoir simulation model was built during the PUNQ project. PUNQ, which stands for the Production forecasting with Uncertainty Quantification, took shape as a study supported by the European Union and conducted by Ten European Companies, universities, and research centres (Floris et al., 2001). The reservoir model was built with the data extracted from 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 and Oliver, 2005; Gao et al., 2005; Abdollahzadeh et al., 2011).

The reservoir model consists of 19 × 28 × 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 modelled using corner-point geometry. A fault bounds the field in the east and south, and a somewhat strong aquifer borders it to the north and west. The presence of the aquifer and the resulting pressure prevents the addition of injection wells. Apart from the borders, a small gas cap in the first layer and in the centre of the dome-shaped structure impedes the drilling of wells in the first layer.

Figure 1 demonstrates the top structure of the PUNQ-S3 reservoir model. As Figure 1 indicates 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 are completed in layers 3 and 4 and PRO-15 is perforated only in layer 4.

Figure 1 The top structure of the PUNQ-S3 (see online version for colours)



Notes: 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 centre of the dome shaped structure (Floris et al., 2001).

So that the model yields similar or identical data for each of the groups involved in the PUNQ project, a 'true' case was designed. The main characteristics needed to generate the 'true' case were porosity and permeability (horizontal and vertical) distributions. The values of these properties at well sites were taken from the original field, operated by Elf Exploration and Production. Barker et al. (2001) explain the comprehensive procedure of creating the porosity and permeability distributions for the 'true' case (Barker et al., 2001; PERM, 2014). The outputs of the 'true' case were considered as actual historical data. Following are the available data provided for the PUNQ-S3 reservoir model:

- porosity and permeability values at well locations
- geological descriptions for each layer
- production history for the first eight years, for the history matching study
- cumulative production (total oil recovery) after 16.5 years, for the uncertainty quantification and production forecast study
- PVT, relative permeability and Carter-Tracy aquifer dataset all taken from the original field data
- no capillary pressure function
- gas oil contact (GOC) and water oil contact (WOC) values.

5 SRM development

The SRM is constructed based on a spatio-temporal database. Depending on the objective of the study, the database would contain different types of data resulting from different realisations of the reservoir simulation model. The main goal of the database is to teach the new model fluid flow phenomena in the reservoir. Data included in this database can be categorised as either static or dynamic. The static data refer to properties that remain constant overtime such as porosity, permeability, top depth, and thickness. In contrast, dynamic data are not necessarily fixed overtime and may include operational constraints, well production, and pressure or phase saturations at the grid blocks.

For the current project, porosity and horizontal and vertical permeabilities constitute the uncertain properties for developing the SRM and matching the field performance. These properties represent the most common uncertain reservoir characteristics employed to match the history data of PUNQ-S3 as indicated by the current literature (Abdollahzadeh et al., 2011; Li and Daoyong, 2011). Mirroring reality, these properties are measured at the well locations – through well logging and core data samples, for instance. Additionally, the provided geological descriptions of this model indicate the streaks of high porosity/permeability profiles in the reservoir (Floris et al., 2001; Barker et al., 2001). This type of information helped to generate the training realisations.

5.1 Informative simulation runs representation of reservoir uncertainties

The spatio-temporal database stockpiles information from the different realisations of the reservoir simulation model. Again, depending on the goal of study, data preparation would differ. Reservoir simulation realisations also differ from one another in the value of the variable uncertain properties. These uncertain properties represent the variables whose impact on the output of the reservoir model the study seeks to analyse.

Based on the property values provided at the well sites and on the geological descriptions, ten different realisations of the reservoir were created. In order to generate these realisations, a sampling method (Latin Hypercube) was utilised. A detailed discussion on generating these realisations has been presented (Shahkarami, 2014).

Although SRM does not need a high number of simulation runs, there are no rules to identify the exact number of realisations required to have a perfect SRM. Many criteria can increase or decrease the number of runs required for developing an SRM. The complexity of the problem, particularly the level of reservoir heterogeneity, is an important factor. In general, the run number selection is based on a rule of thumb process, in which the user's experience could be helpful. However, it is obvious that if the number of simulation runs is too small, the SRM may not be able to catch the uncertainty and the variation in the parameters. In this situation, the surrogate reservoir model might even show good results for the training samples; here is where the validation step plays an important role. Although the SRM might have a good performance on the training samples, it will fail to create the same quality for the validation set. Therefore, the validation examples will expose the lack of required information in the training samples. Alternatively, if the number of simulation runs is too large, there is no need to develop an SRM since the solution is close to the original problem, which entails a high number of simulation runs. This is a problem that occurs frequently for the case studies involving geo-statistical-based proxy models. In these cases, the cost of developing the proxy model is too high and is not justified. In this study, we selected ten realisations based on the experience of previous research studies (Sampaio et al., 2009; Shahkarami et al., 2014a, 2014b; Amini et al., 2014). The robustness of the validation results, discussed in the results section, indicates that the decision to select ten runs was right.

5.2 *Reservoir delineation and tier system*

Data summarisation is an essential task during SRM development. One method of data summarisation entails delineating the reservoir into different segments and calculating an average of data over each reservoir segment. In order to divide the reservoir into the abovementioned segments, we used feedback from the available geological descriptions and we employed a modified version of the Voronoi diagrams (Erwig, 2000; Gomez et al., 2009). Figure 2 depicts the designed drainage areas created by means of the modified Voronoi theory. Consequently, every drainage area is divided into four tiers. The first tier represents the well block, which has a significant impact on the well behaviour. The second tier includes the first row of grid blocks around the well block. The third tier is composed of the next row of grid blocks around the second tier. Finally, the remaining 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 sketches an example of the designed tier system in this study.

Figure 2 The drainage areas assigned to the wells in different layers based on a modified version of the Voronoi diagram (see online version for colours)

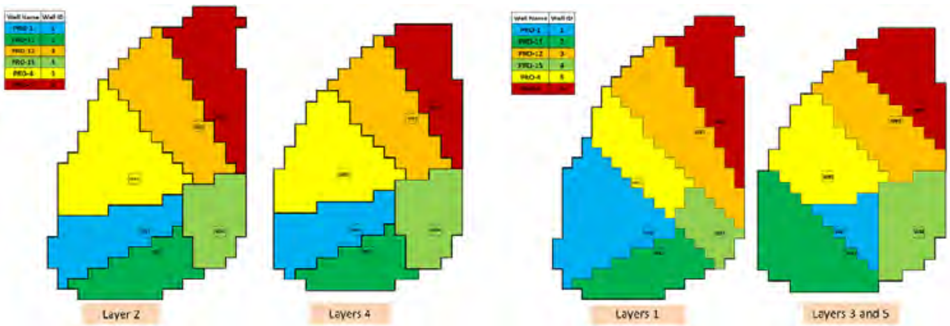
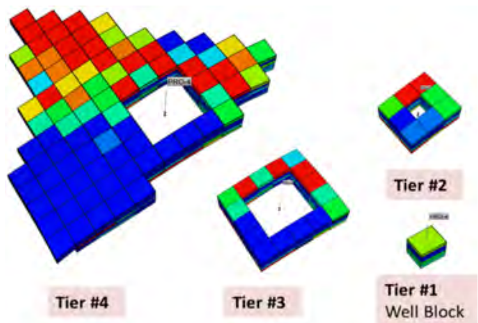


Figure 3 The designed tier system for the PUNQ-S3 reservoir model (see online version for colours)



5.3 SRM inputs and outputs

Due to the complexity of the PUNQ-S3, a higher number of input parameters had to be considered in this case study compared to the previous examples (Shahkarami et al., 2014). The oil production rate constituted the main constraint during the reservoir's eight years of production. Thus, the outputs of the PUNQ-S3 that need to be matched were: the well bottom-hole pressure, the gas production rate, and the water production rate. For each one of these outputs, one ANN was created (yielding three ANNs in total). The outputs are considered at each time step.

This particular SRM is well-based. That means that the SRM predicts the outputs at the well location for different time steps. Therefore, the spatio-temporal database includes different records. To be exact, the number of records is equal to the number of wells multiplied by the number of time steps. Each record has inputs and output(s). The inputs should always be provided. At the training phase, the outputs should also be known. However, during the predictive phase, we provide the inputs and the SRM (ANNs) will predict the outputs.

The PUNQ-S3 model comprises five layers and a total of six wells. As elucidated above, each drainage area was divided into four tiers and three uncertain parameters were considered: the porosity and the horizontal and vertical permeabilities. Therefore, we identified 360 uncertain or adjustable parameters (5 *layers* × 4 *tiers* × 6 *wells* × 3 *properties*), which could be tuned to match the history data. In order to build the SRM, we needed to include at least 60 parameters for each well. These do not contain the other types of data such as thickness and top for each tier. Overall, we created a database with more than 120 inputs. Figure 4 summarises the types of inputs and outputs in the spatio-temporal database for the PUNQ-S3. Selecting the right inputs raises considerable challenges, and many artificial-based models fail at this step (Zubarev, 2009; Mohaghegh et al., 2012).

Figure 4 Inputs and outputs available in the spatio-temporal database for the PUNQ-S3 SRM (see online version for colours)

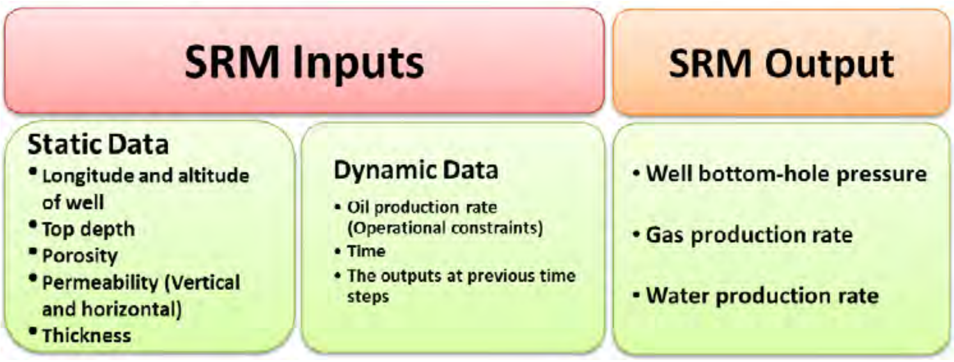


Table 1 to Table 3 show the selected inputs for the three ANNs. Selecting the inputs of the SRM entails certain pre-processing steps (Shahkarami, 2014). The inputs generally include static and dynamic data. All three ANNs share time and oil production rate as their common inputs. In addition, we included the output of each ANN at the time steps behind, as an input for the next time step. For example, to predict the well bottom-hole pressure at time t, the well bottom-hole pressures at one $(t - 1)$ and two $(t - 2)$ time steps behind are employed as ANN inputs.

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

Static inputs		*Dynamic inputs*	*Output*
Latitude (X)	@ Well block	Well bottom-hole pressure @ $(t - 1)$ and $(t - 2)$	
Longitude (Y)			
Horizontal permeability	@5 layers and 4 tiers	Time	Well bottom-hole pressure
Vertical permeability	@5 layers and 4 tiers		
Thickness	@5 layers and 4 tiers	Oil production rate (well constraints)	
Top depth	@5 layers and 4 tiers		

Table 2 Selected inputs for the gas production rate network

Static inputs		*Dynamic inputs*	*Output*
Latitude (X)	@ Well block	Gas production rate @ $(t - 1)$ and $(t - 2)$	
Longitude (Y)			
Horizontal permeability	@5 layers and 4 tiers	Time	Gas production rate
Vertical permeability	@5 layers and 4 tiers		
Thickness	@5 layers and 4 tiers	Oil production rate (well constraints)	
Top depth	@5 layers and 4 tiers		

Table 3 Selected inputs for the water production rate network

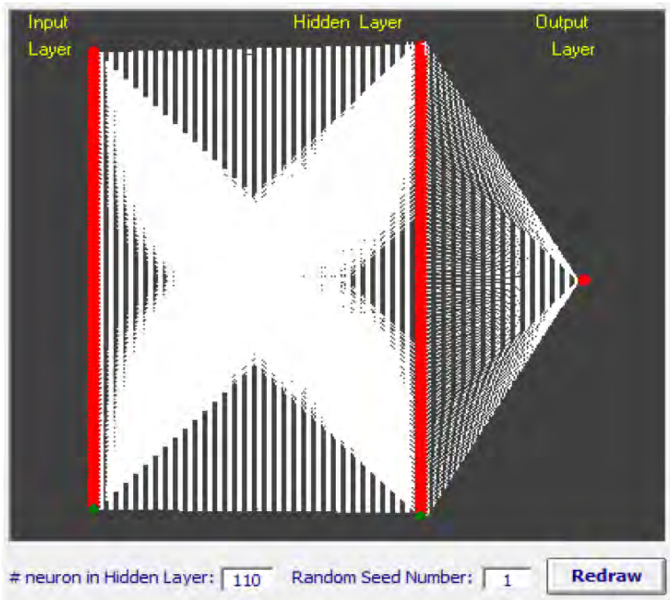
Static inputs		*Dynamic inputs*	*Output*
Latitude (X)	@ Well block	Water production rate @ $(t - 1)$ and $(t - 2)$	
Longitude (Y)			
Horizontal permeability	@5 layers and 4 tiers	Time	Water production rate
Vertical permeability	@5 layers and 4 tiers		
Thickness	@5 layers and 4 tiers	Oil production rate (well constraints)	
Top depth	@5 layers and 4 tiers		

5.4 Training, calibrating, and validating the ANNs

The SRM training includes three different steps: training (learning), calibration, and validation. Consequently, the spatio-temporal database is divided into three sets: the training or learning set, the calibration set, and the validation or verification set. The training set constitutes the part of the database used directly to train the ANNs, which are adapted to this set to match the provided outputs. The calibration set is not used to adjust the outputs but to assure that any increase in accuracy over the training dataset will lead to an increase in accuracy over a set of data not used for training the ANNs. The calibration dataset helps determine when the training stops, and also prevents over-fitting the ANNs. We do not want the ANNs just to memorise the behaviour of the training set. A well-trained ANN will perform with accuracy beyond the learning set. Lastly, the verification set represents the part of the database that verifies the predictability of the trained ANNs, and, subsequently, this dataset is not used to train the ANNs.

In this study, 80% of the database was used as a training set, 10% for calibration and the last 10% for the validation set. Corresponding to the total number of outputs we created three networks. All three networks contain one hidden layer (Figure 5). The SRM integrates three neural networks after the training process is completed. 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 (see online version for colours)



As a further validation step, 'Blind verification' tests the robustness of SRM. The term 'blind' indicates a set of realisation(s) that has not been used during the training process. These blind testing sets are complete realisations of the reservoir, whereas the verification set used in the training process is a randomly selected portion of the spatio-temporal database.

6 Automated history matching

During the final stage, we combined the developed and validated SRM with the DE optimisation algorithm to construct an automated history matching workflow.

The following equations are among the most common objective functions used for the history match process (CMG, 2013). The objective function computes the relative difference between the SRM results and the measured production data. Equation (1) calculates the relative differences at the well level. The subscripts i and t represent well and time, respectively. Nt(i, j) is the total number of measured data points for each well i and property j. $Y^{s}_{i,t}$ are the predicted production by SRM and $Y^{m}_{i,t}$ are the measured production data. ΔY_i^m is the scale calculated by subtracting the maximum and minimum of measured production data for well i. N(i) is also the total number of properties required to be matched (for example oil production, gas production, water cut). Although for a real case, measurement error should be considered in the calculation, we assume this kind of error does not exist in this study.

In practice, it is common to consider that the quality and importance of measured data may be different for some specific properties, time intervals, and wells. Therefore, some weighting factors ($tw_{i,j}$) are present in these equations. For instance, let us assume the measured data for a specific well over a particular time period have been recorded at a higher resolution using a better quality recording device. The user might want to value this part of data more than the rest. This would be possible simply by increasing the weighting factors for the specific well and time periods. In this study, we assume all the wells, time steps, and properties have the same impacts and the corresponding weighting factors are equal to one.

- Individual well objective function

$$OF_i = \frac{1}{\sum_{j=1}^{N(i)} tw_{i,j}} \sum_{j=1}^{N(i)} \frac{\sqrt{\frac{\sum_{t=1}^{Nt(i,j)} \left(Y_{i,j,t}^s - Y_{i,j,t}^m\right)^2}{Nt(i,\,j)}}}{\Delta Y_{i,j}^m} .100\%.tw_{i,j} \tag{1}$$

It is also common to define a global objective function in order to have calculations in the field level. Equation (2) describes the global objective function using the well level objective function, which we defined in equation (1). Here OF_{global} denotes the global objective function, OF_i represents the objective function for well i, and N_w stands for the total number of wells. w_i is the defined weight for well i. Again, in this study, we consider all the measured data points being equally important and the weighting factors being equal to one.

- Global (field) objective function

$$OF_{global} = \frac{1}{\sum_{i=1}^{N_w} w_i} \sum_{i=1}^{N_w} w_i OF_i \tag{2}$$

7 Results

The development and application of the SRM in the history matching of PUNQ-S3 yielded interesting results. We randomly selected two wells (Wells PRO-1 and PRO-4) out of six wells to represent the results.

7.1 The SRM training

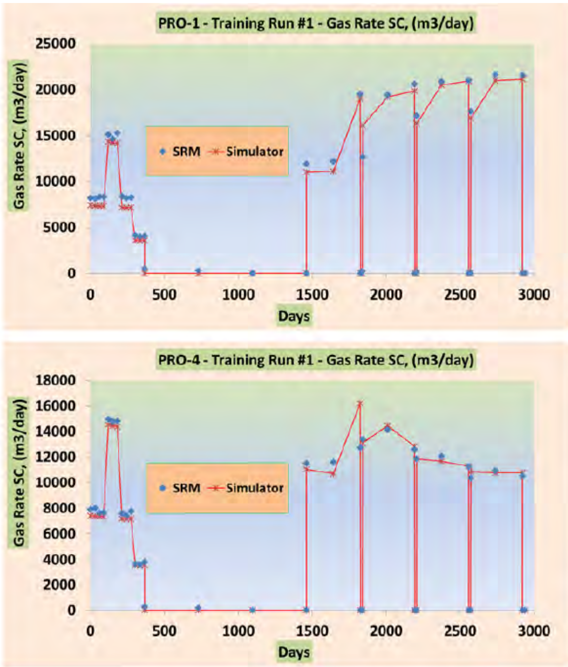
The first set of results belongs to a training realisation used to train the SRM.

Figure 6 Comparison of the well bottom-hole pressure profile generated by the SRM (indicated by blue markers) with the similar results from a numerical simulator (CMG-IMEXTM) for a training realisation of the PUNQ-S3, wells PRO-1 and PRO-4 (see online version for colours)

Figure 6 depicts a comparison between the well bottom-hole pressure results generated by the SRM and similar results collected from the simulator CMG-IMEXTM (CMG, 2013). Figure 6 portrays two profiles corresponding to the bottom-hole pressure profiles of PRO-1 and PRO-4. The blue markers indicate the SRM results over a period of eight years (almost 3,000 days) of the reservoir life compared to the simulator outputs showed by the red line. The oil rate is the main constraint during this period. Consequently, we matched the bottom-hole pressure, the gas production rate, and the water production rate. As it is obvious visually, the training is carried out very well for the bottom-hole pressure. The SRM is able to capture the fluctuation in the bottom-hole pressure of the training data.

Figure 7 depicts the results of the gas production rate for a training realisation. It is a comparison between the results from the SRM and from the simulator for the wells PRO-1 and PRO-4 in the PUNQ-S3. The quality of the training results is high. The SRM has detected the shut-in periods (zero gas rate) and production peaks very well.

Figure 7 Comparison of the gas production rates generated by the SRM (depicted by the blue markers) with the similar results from the numerical simulator for a training realisation of the PUNQ-S3 (see online version for colours)



Across the eight years of history data, only one well, the PRO-11, has water breakthrough. Figure 8 compares the results of the water production rate generated by the SRM and analogous results obtained from the simulator. Although no water production can be observed during the first five years of history data, we begin to observe some gradually thereafter. Noticeably, the SRM is able to properly capture the zero water production rate.

Figure 8 Comparison of the water production rate generated by the SRM (depicted by the blue markers) with the similar results from the numerical simulator for a training realisation of the PUNQ-S3 (see online version for colours)

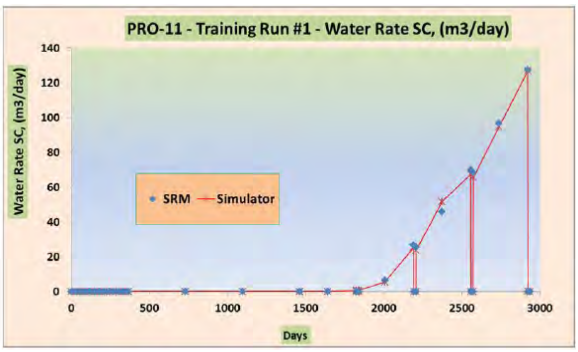
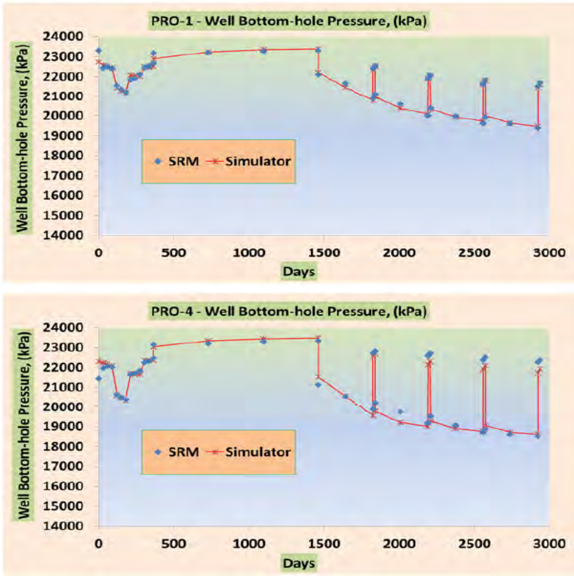
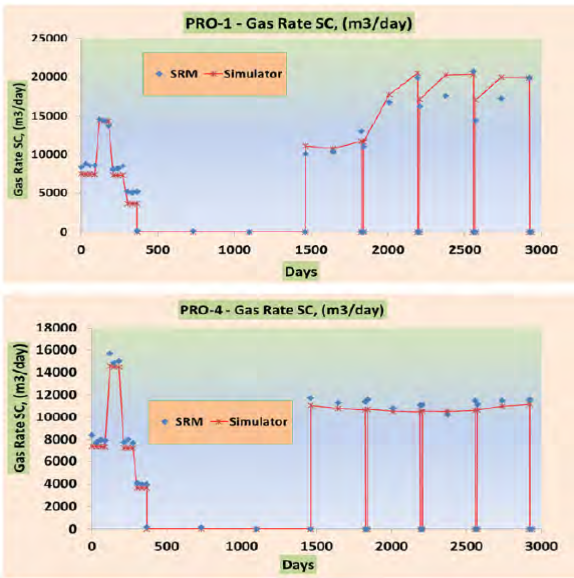


Figure 9 Validating the SRM using a blind run (see online version for colours)



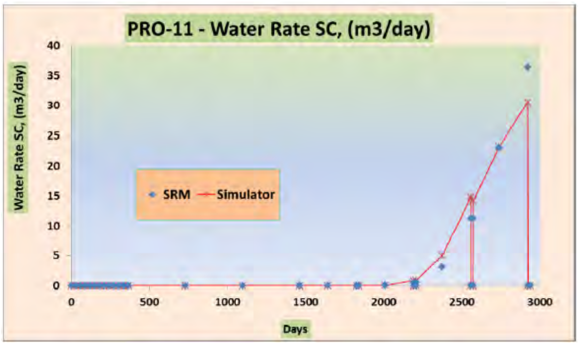
Note: Comparison of the well bottom-hole pressure from the SRM with the results from the numerical simulation model for the PUNQ-S3.

Figure 10 Validating the SRM using a blind run (see online version for colours)



Note: Comparison of the gas production rate profile from the SRM with the results from the numerical simulation model for the PUNQ-S3.

Figure 11 Validating the SRM using a blind run (see online version for colours)



Note: Comparison of the water production rate profile, in well PRO-11, from the SRM with the results from the numerical simulation model for the PUNQ-S3.

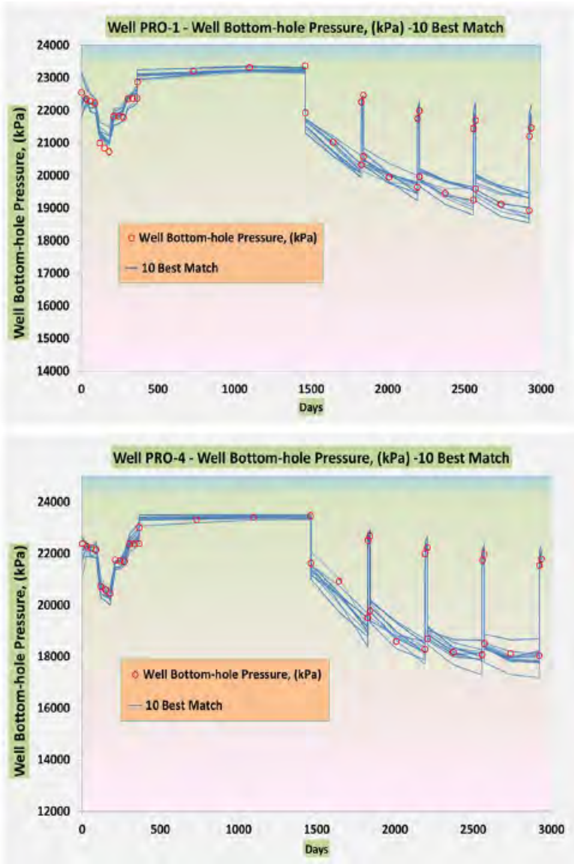
7.2 *The blind verification run*

In order to insure that the SRM has a good performance on the set of data not utilised during the training process, we implemented the trained SRM on a completely blind realisation. Figure 9 depicts the validation results for the well bottom-hole pressure profiles of wells PRO-1 and PRO-4 in the blind run. Like in the training results, the blue markers represent the SRM results while compared with the simulator outcome, denoted by the red line. Figure 10 and Figure 11 illustrate the validation results for the gas production and the water production rates, respectively. Although the blind verification realisation was not used during the training stage, the SRM shows a good accuracy over these data.

7.3 *History matching*

Once the validation stage completed, we used the SRM in an automated history matching workflow. A notable benefit of the automated history matching workflow is the ability to offer multiple realisations that match the field data. In this study, we selected the top ten best matches. Figure 12 displays the results of history matching for the well bottom-hole pressure for wells PRO-1 and PRO-4. Each diagram compares the results of the top ten matches (marked by the blue lines) with the actual data (demarcated by the red circles). The top ten matches are the result of ten different realisations of the reservoir characteristics. The visible range in Figure 12 depicts the uncertainty of reservoir properties.

Figure 12 History matching results of the well bottom-hole pressure for wells PRO-1 and PRO-4 (see online version for colours)

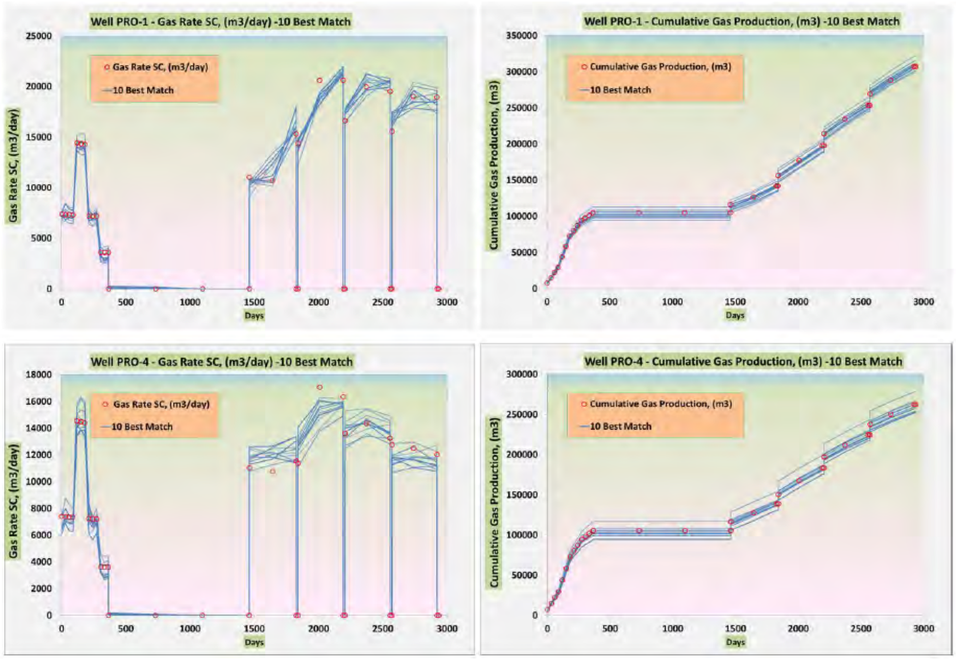


Note: Comparison of ten best matches (depicted by the blue lines) coming from the SRM with the actual data (indicated by the red circles).

Figure 13 demonstrates the comparison between the ten best matches (delineated again by the blue lines) and the actual data (represented through the red circles), for the gas production of wells PRO-1 and PRO-4. The left side graphs indicate the gas production rate, while the right side diagrams denote the cumulative gas production.

Figure 14 compares the results of the top ten matches with the actual data for the water production. In Figure 14, the diagram on the left illustrates the water rate production, and the diagram on the right side depicts the cumulative water production, both for well PRO-11.

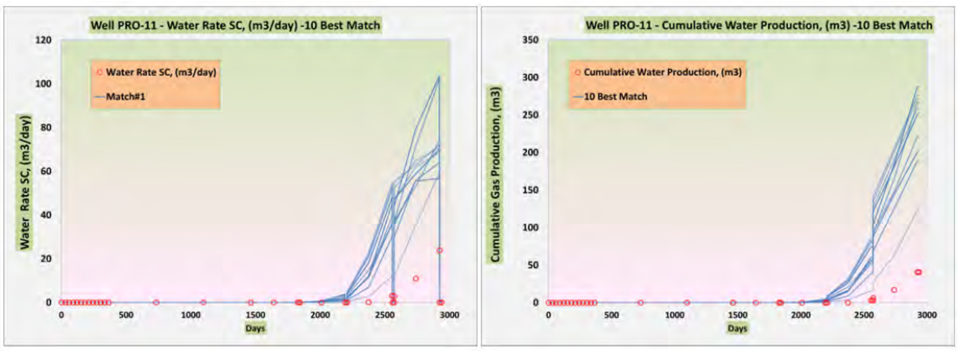
Figure 13 History matching results of the gas production, (a) the production rates (b) the cumulative production (see online version for colours)



(a) (b)

Notes: Comparison of ten best matches (marked by the blue lines) coming from the SRM with the actual data (marked by the red circles). The data belong to the wells PRO-1 and PRO-4.

Figure 14 History matching results of the water production, (a) the production rates (b) the cumulative production (see online version for colours)



(a) (b)

Notes: Comparison of ten best matches (marked by the blue lines) coming from the SRM with the actual data (marked by the red circles). In this study, we have just one well (PRO-11) with water breakthrough during eight years of production history.

Figures 15 and 16 display the distributions of porosity and horizontal permeability for the ten best history matched realisations. The 'true' maps of porosity and permeability are also included in Figures 15 and 16.

Figure 15 Ten best matched porosity distributions compared with the 'true' porosity distributions of PUNQ-S3 reservoir model (see online version for colours)

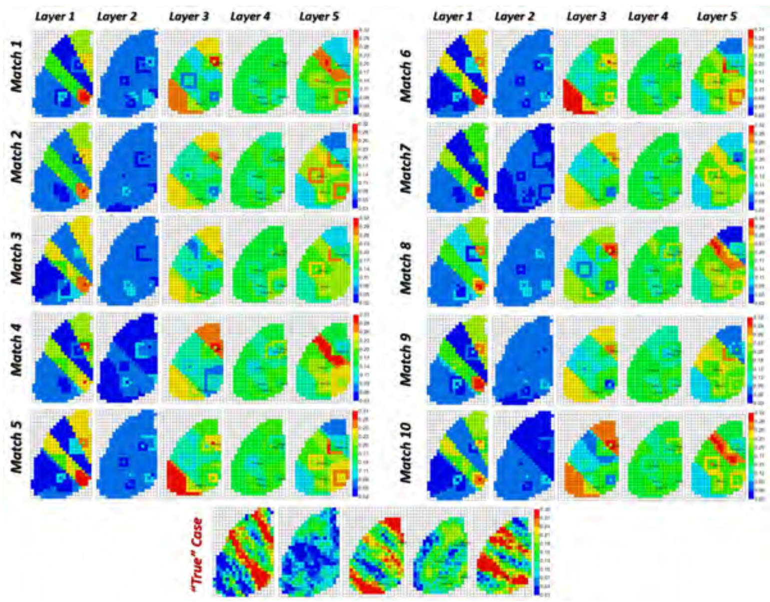
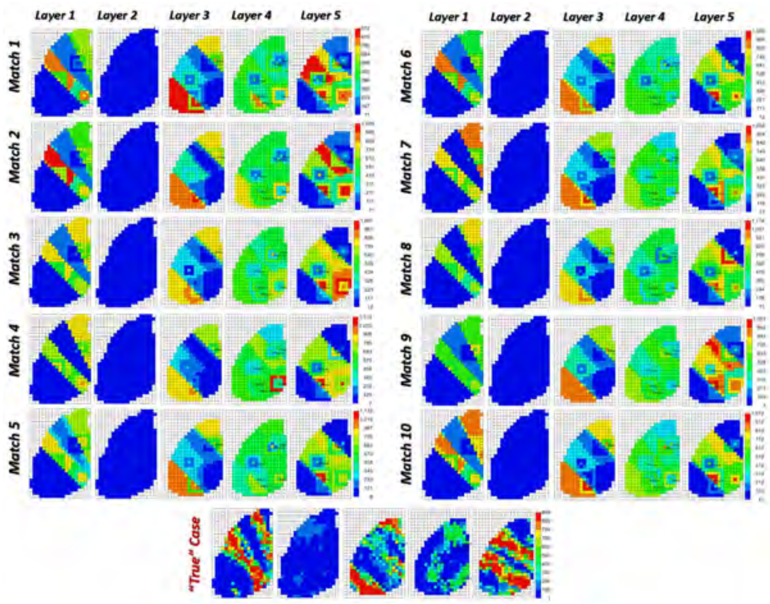


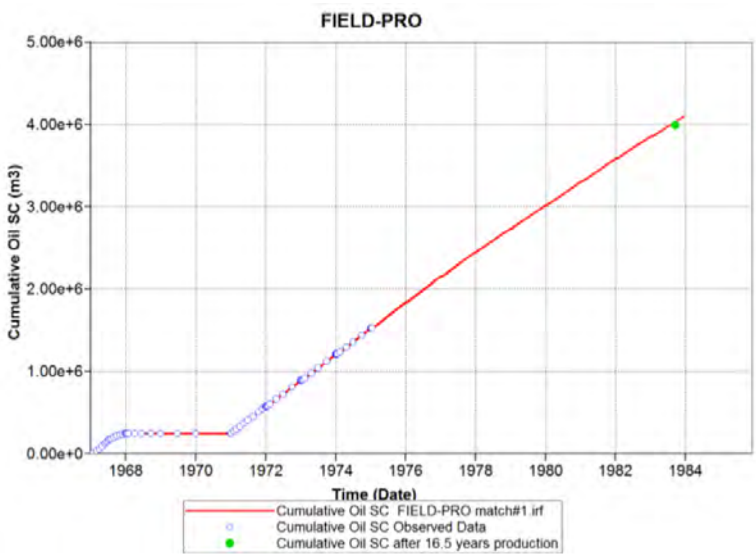
Figure 16 Ten best matched horizontal permeability distributions compared with the 'true' horizontal permeability distributions of PUNQ-S3 reservoir model (see online version for colours)



7.3 Importing the matched reservoir characteristics into the simulator

The developed SRM was aimed to substitute an industrial reservoir simulator CMG-IMEXTM (CMG, 2013) during the history matching process. Thus, we designed an automated SRM-based history matching workflow. This workflow can provide multiple realisations of the reservoir, which match the actual data. We chose ten best matches. These ten realisations were imported into the simulator so that we can observe the performance of the simulator with inputs coming from the SRM. Figure 17 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). We used eight years of field data for history matching purposes. In addition to the eight years of history data, PUNQ project has published the field cumulative oil production after 16.5 years. Many studies have used this data for future forecast comparisons (Floris et al., 2001; Barker et al., 2001). The green point in Figure 17 represents the provided value for the field cumulative oil production after 16.5 years.

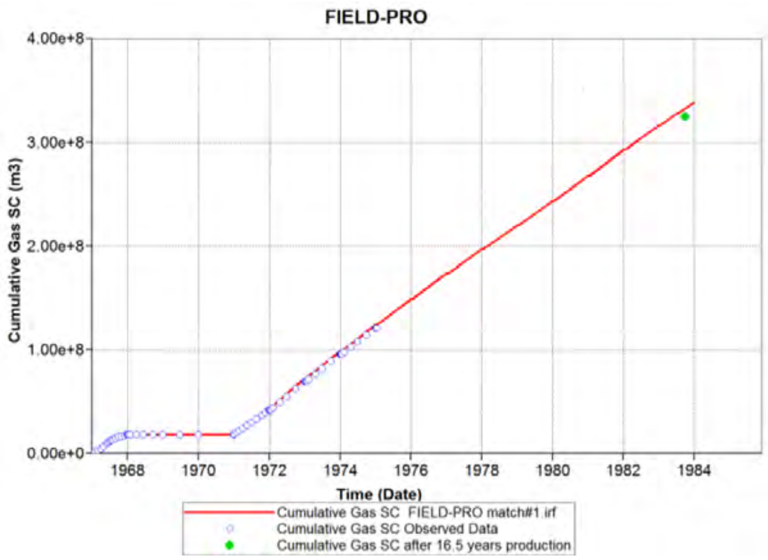
Figure 17 Comparison between the SRM-based history matching results (match #1) and actual data for cumulative oil production (see online version for colours)



Notes: The red line represents the matched realisation and the blue circles indicate the actual field data (eight years of production history). The green point also displays the cumulative production for a true case after 16.5 years.

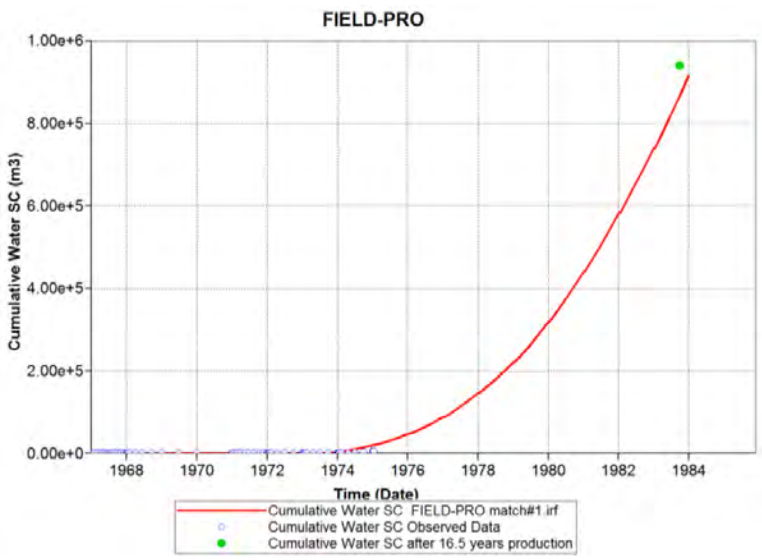
Similar to the oil production, Figure 18 shows the comparison between the simulator results and actual field data for cumulative gas production.

Figure 18 Comparison between the SRM-based history matching results (match #1) and actual data for cumulative gas production (see online version for colours)



Notes: The red line represents the matched realisation and the blue circles are the actual field data (eight years of production history). The green point also displays the cumulative production for a true case after 16.5 years.

Figure 19 Comparison between the SRM-based history matching results (match #1) and actual data for cumulative water production (see online version for colours)



Notes: The red line represents the matched realisation and the blue circles indicate the actual field data (eight years of production history). The green dot also displays the cumulative production for a true case after 16.5 years.

The water breakthrough occurs during the seventh year at one well only (PRO-11). Therefore, except for the well PRO-11 that produces water, the other wells have an overall negligible water production. Figure 19 compares the cumulative water production between the simulator and the actual data.

8 Discussion and concluding remarks

In this study an SRM was developed for the PUNQ-S3 reservoir simulation model. The PUNQ-S3 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 render PUNQ-S3 a unique case for the study of the SRMs potential for history matching. The variable properties employed to create the SRM are the porosity and the permeability (horizontal and vertical) distributions. In order to train the SRM, ten realisations of PUNQ-S3 simulation model were generated. An extra realisation (the 11th case) was used to verify the trained SRM.

One important feature of an effective history matching workflow is its automation ability. Therefore, the developed SRM was coupled with the DE optimisation algorithm. The objective functions were created to calculate the misfit values between the actual data and the measured results (SRM). The goal of the history matching process was to match eight years of history data available for three different properties. These properties include the well bottom-hole pressure, the gas production rate, and the water production rate. This workflow was able to report multiple realisations of the reservoir that matched the actual data. 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.

8.1 SRM training quality

Figure 6, Figure 7, and Figure 8 show the results of SRM during the training process for the well bottom-hole pressure, the gas production rate, and the water production rate, respectively. These graphs portray the comparison between the SRM results with the simulator outputs for two representative wells, PRO-1 and PRO-4. The superb match between the results of the SRM and those from the simulator proves that the SRM has been well-trained. The ability of the SRM to capture the zero values of gas and water production rates appears clearly in Figure 7 and Figure 8.

8.2 SRM further validation quality

Figure 9, Figure 10, and Figure 11 depict the performance of the SRM on a completely unseen realisation of the reservoir during the training process. This step, referred to as ‘blind realisation’, reveals the robustness of the SRM. The quality of the match for the blind case, as seen in these graphs, is not as good as the training realisations given by Figure 6, Figure 7, and Figure 8. This is an expected response of the SRM to a set of data not used in the SRM training. For example, in Figure 10 illustrating the validation results for the gas production rate, the SRM gives a slightly overestimated prediction for the well PRO-4. One important point about data-driven approaches like the SRM is that the quality of the prediction performance is very good as long as the inputs are at the same range as they were in the training sets.

8.3 History matching quality

The trained and verified SRM was used to perform history matching. Figure 12, Figure 13, and Figure 14 designate the results of ten best history matched realisations. The results represent the outputs of the SRM. These graphs capture the comparison between the ten best matches and the actual data for the well bottom-hole pressure, the gas production rate, the cumulative gas production, the water production rate, and the cumulative water production. The matches for the well bottom-hole pressure and the gas production rate are satisfactory. However, when it comes to the water production rate (Figure 14), the matches are not as good as the well bottom-hole pressure and the gas production rate. In fact, the SRM overestimates the water production. This observation reveals one important characteristic of the SRM and other similar data-driven approaches. These methods are developed based on data, and the inability to provide sufficient information during the training stage will cause problems during the prediction step. Among the three properties that the SRM was developed to predict, the water rate production has the minimum amount of information in the training realisations. If we go back to the PUNQ-S3 reservoir model, out of six wells there is only one well (well PRO-11) that produces water. In addition, water production begins at the end of the production profile. In other words, out of 37 data points, there are only three non-zero water production points, which makes just 8% of the data. As it is clear from Figure 14 and Figure 19, the matches capture the zero values of water production very well. For the non-zero values of water production, although the SRM knows that they are not zero, the results indicate an overestimation. One way to address this issue is to provide more non-zero examples of water production rates during the training step.

Figures 15 and 16 compare the ten best matched distributions of porosity and horizontal permeability with the ‘true’ reservoir model. One common problem of history matching approaches is that the matched property distributions not always ensure the geologic consistency. In this study, the implemented reservoir delineation and tier system, Figures 2 and 3, constrained our history matching workflow in order to maintain the geologic consistency. For the PUNQ-S3 reservoir, a part of provided data is the geological descriptions for each layer (Floris et al., 2001). This information guided us in designing the reservoir delineation and tier system. However for a green reservoir with little geologic information available, this step of SRM development could be a challenge.

Figure 17 displays the results of the best achieved match imported into the simulator. Figure 17 compares the results of this realisation with the actual data for the field cumulative oil production. The results indicate 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 future production. Figure 18 and Figure 19 represent the same comparison for the field cumulative gas and water production. For gas production, we note the same overestimating behaviour; in contrast, the water production has been slightly underestimated.

In the term of coupling the SRM with an optimisation algorithm, we selected the DE based on the results of previous studies (Wang and Buckley, 2006; Decker and Mauldon, 2006; Jahangiri, 2007; Hajizadeh et al., 2009b, 2010; Wang and Gao, 2010; Wang et al., 2011; Mirzabozorg et al., 2013; Okano, 2013). We concentrated on studying the performance of the SRM as a data-driven technique to assist history matching. However, we found out that the DE is a powerful algorithm in identifying optimum values. The run-time of the SRM takes fractions of a second, and therefore its run time and computational cost are not problematic. Nevertheless, the DE was able to converge to the optimum solutions within the couple hundreds runs.

The general results in this study demonstrate the robustness of the SRM in history matching for the PUNQ-S3 problem. Numerous studies have used the PUNQ-S3 reservoir model to test methods of history matching, and many of these studies investigate different optimisation methods for automated history matching. Generally, these optimisation algorithms have been coupled with a commercial simulator. The reported numbers of simulation runs for history matching in the PUNQ-S3 reservoir model are in the order of thousands (Hajizadeh et al., 2010, 2009a, 2009b; Abdollahzadeh et al., 2011). In our study, the required simulation runs to create and validate the SRM (11 runs) are very few. 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 a higher computational cost. In such a case, using a numerical reservoir simulator for history matching would pose a major computational issue. The application of the SRM for history matching purposes would be a great asset in the reservoir management workflow.

SRMs require a low development cost during the development and implementation stages. An SRM usually requires a few realisations of the reservoir simulation models. The main limitation of SRMs is that they are example based and case-subjective. SRMs are developed based on the training examples and their accuracy is tied to the quality of the training examples. It also means that there is no general SRM that serves all problems and scenarios. Depending on the problems, SRMs are trained and validated using the examples of the problem. For instance, an SRM developed for a history matching a reservoir will not be suited for the history matching another reservoir. In addition, the SRM will not be useful for another study like reservoir production optimisation. Nevertheless having enough information, it is convenient to build the right SRM for the right job.

Acknowledgements

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 the development of the SRM, respectively.

Nomenclature

F	Scaling factor in the DE algorithm
i	Subscript counter for well
j	Subscript counter for property
GOC	Gas oil contact
N	Number of members for a solution vector in the DE algorithm
N(i)	Total number of properties
Nt(i, j)	Total number of measured data points for each well i and property j
Nw	Total number of wells
OF_{global}	Global objective function
OF_i	Objective function for well i
t	Time step counter
$tw_{i,j}$	Weighting factor for well i and property j
w_i	Weighting factor for well i
WOC	Water oil contact
$Y^S_{(i,t)}$	Predicted production by SRM
$Y^m_{(i,t)}$	Measured production data
ΔY_i^m	Difference between maximum and minimum of measured production data for well i

References

Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C. and Vallès, B. (2009) 'The ensemble Kalman filter in reservoir engineering – a review', *SPE Journal*, Vol. 14, No. 3, pp.393–412.

Abdollahzadeh, A., Reynolds, A., Christie, M.A., Corne, D., Davies, B.J. and Williams, G. (2011) 'Bayesian optimization algorithm applied to uncertainty quantification', *SPE EUROPEC/EAGE Annual Conference and Exhibition*, Society of Petroleum Engineers, Vienna, Austria.

Alimonti, C. and Falcone, G. (2004) 'Integration of multiphase flowmetering, neural networks, and fuzzy logic in field performance monitoring', *SPE Production & Facilities*, Vol. 19, No. 1, pp.25–32.

Amini, S., Mohaghegh, S.D., Gaskari, R. and Bromhal, G. (2014) 'Pattern recognition and data-driven analytics for fast and accurate replication of complex numerical reservoir models at the grid block level', *SPE Intelligent Energy Conference and Exhibition*, Society of Petroleum Engineers, Utrecht, The Netherlands.

Artun, E., Ertekin, T., Watson, R.W. and Al-Wadhahi, M.A. (2009) 'Development of universal proxy models for screening and optimization of cyclic pressure pulsing in naturally fractured reservoirs', *International Petroleum Technology Conference*, Doha, Qatar.

Barker, J., Cuypers, M. and Holden, L. (2001) 'Quantifying uncertainty in production forecasts: another look at the PUNQ-S3 problem', *SPE Journal*, Vol. 6, No. 4, pp.433–441.

Bazargan, H. and Christie, M. (2012) 'Efficient polynomial chaos proxy-based history matching and uncertainty quantification for complex geological structures', *SPE Kuwait International Petroleum Conference and Exhibition*, Society of Petroleum Engineers, Kuwait City, Kuwait.

Cardoso, M.A. (2009) 'Reduced-order models for reservoir simulation', *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers, New Orleans, Louisiana.

Cardoso, M.A. and Durlofsky, L.J. (2010) 'Use of reduced-order modeling procedures for production optimization', *SPE Journal*, Vol. 15, No. 2, pp.426–435.

Castellini, A. (2005) 'Quantifying uncertainty in production forecast for fields with significant history: a West African case study', *International Petroleum Technology Conference*, Doha, Qatar.

Chen, Y. and Oliver, D. S. (2010) 'Ensemble-based closed-loop optimization applied to Brugge field', *SPE Reservoir Evaluation & Engineering*, Vol. 13, No. 1, pp.56–71.

Christie, M., MacBeth, C. and Subbey, S. (2002) 'Multiple history-matched models for teal south', *The Leading Edge*, Vol. 21, No. 3, pp.286–289.

CMG (2013) *Computer Modelling Group Manual*.

Coats, K., Dempsey, J. and Henderson, J. (1970) 'A new technique for determining reservoir description from field performance data', *Society of Petroleum Engineers Journal*, Vol. 10, No. 1, pp.66–74.

Das, S. and Suganthan, P. (2009) 'Differential evolution: a survey of the state-of-the-art', *IEEE Transactions on Evolutionary Computation*, Vol. 15, No. 1, pp.4–31.

Decker, J. and Mauldon, M. (2006) 'Determining size and shape of fractures from trace data using a differential evolution algorithm', *The 41st US Symposium on Rock Mechanics (USRMS): '50 Years of Rock Mechanics – Landmarks and Future Challenges*, American Rock Mechanics Association, Golden, Colorado.

Dzurman, P.J., Leung, J.Y.W., Zanon, S.D.J. and Amirian, E. (2013) 'Data-driven modeling approach for recovery performance prediction in SAGD operations', *SPE Heavy Oil Conference Canada*, Society of Petroleum Engineers, Calgary, Alberta, Canada.

Eberhart, R. and Shi, Y. (2001) 'Particle swarm optimization: developments, applications and resources', *Proceedings of the 2001 Congress on Evolutionary Computation*, Vol. 1, pp.81–86, Seoul, South Korea.

Eide, A., Holden, L., Reiso, E. and Aanonsen, S. (1994) 'Automatic history matching by use of response surfaces and experimental design', *4th European Conference on the Mathematics of Oil Recovery*, Roros, Norway.

Erbas, D. and Christie, M. (2007) 'Effect of sampling strategies on prediction uncertainty estimation', *SPE Reservoir Simulation Symposium*, Houston, Texas.

Erwig, M. (2000) *The Graph Voronoi Diagram with Applications*, John Wiley & Sons, Inc., SL.

Evensen, G. (2003) 'The ensemble Kalman filter: theoretical formulation and practical implementation', *Ocean Dynamics*.

Fedutenko, E., Yang, C., Card, C. and Nghiem, L.X. (2012) 'Forecasting SAGD process under geological uncertainties using data-driven proxy model', *SPE Heavy Oil Conference Canada*, Society of Petroleum Engineers, Calgary, Alberta, Canada.

Floris, F., Bush, M., Cuypers, M., Roggero, F. and Syversveen, A. (2001) 'Methods for quantifying the uncertainty of production forecasts', *Petroleum Geoscience*, Vol. 7, No. 10, pp.87–96.

Gao, G., Zafari, M. and Reynolds, C. (2005) 'Quantifying uncertainty for the PUNQ-S3 problem in a Bayesian setting with RML and EnKF', *Reservoir Simulation Symposium*, Society of Petroleum Engineers, Houston, Texas.

Gildin, E., Ibrahim, A. and Ghasemi, M. (2014) 'Reduced order modeling in reservoir simulation using the bilinear approximation techniques', *SPE Latin America and Caribbean Petroleum Engineering Conference*, Society of Petroleum Engineers, Maracaibo, Venezuela.

Gomez, Y., Khazaeni, Y., Mohaghegh, S.D. and Gaskari, R. (2009) 'Top down intelligent reservoir modeling', *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers, New Orleans, Louisiana.

Goodwin, N. (2015) 'Bridging the gap between deterministic and probabilistic uncertainty quantification using advanced proxy based methods', *SPE Reservoir Simulation Symposium*, Houston, Texas, USA.

Graf, T., Zangl, G., May, R. S., Hartlieb, M., Randle, J.M. and Al-Kinani, A. (2011) 'Candidate selection using stochastic reasoning driven by surrogate reservoir models', *SPE Reservoir Evaluation & Engineering*, Vol. 14, No. 4, pp.413–422.

Gu, Y. and Oliver, D. (2005) 'History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter', *SPE Journal*, Vol. 10, No. 2, pp.217–224.

Hajizadeh, Y. (2010) 'Ants can do history matching', *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers, Florence, Italy.

Hajizadeh, Y., Christie, M. and Demyanov, V. (2009a) 'Ant colony optimization for history matching', *SPE EUROPEC/EAGE Annual Conference and Exhibition*, Society of Petroleum Engineers, Amsterdam, The Netherlands.

Hajizadeh, Y., Christie, M. and Demyanov, V. (2009b) 'Application of differential evolution as a new method for automatic history matching', *2009 Kuwait International Petroleum Conference and Exhibition*, Society of Petroleum Engineers, Kuwait City, Kuwait.

Hajizadeh, Y., Christie, M. and Demyanov, V. (2010) 'History matching with differential evolution approach; a look at new search strategies', *SPE EUROPEC/EAGE Annual Conference and Exhibition*, Society of Petroleum Engineers, Barcelona, Spain.

Hanea, R.G., Przybysz-Jarnut, J.K., Krymskaya, M., Heemink, A. and Jansen, F.D. (2010) 'The choice of the 'best' assisted history matching algorithm for subsurface characterization in reservoir engineering', *SPE EUROPEC/EAGE Annual Conference and Exhibition*, Society of Petroleum Engineers, Barcelona, Spain.

Haugen, V.E., Natvik, L-J., Evensen, G., Berg, A.M., Flornes, K.M. and Naevdal, G. (2006) 'History matching using the ensemble Kalman filter on a north sea field case', *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers, San Antonio, Texas, USA.

Haykin, S. (2008) *Neural Networks and Learning Machines*, Prentice Hall, New York City, NY, USA.

He, J., Sarma, P. and Durlofsky, L.J. (2011) 'Use of reduced-order models for improved data assimilation within an EnKF context', *SPE Reservoir Simulation Symposium*, Society of Petroleum Engineers, The Woodlands, Texas, USA.

Jacquard, P. (1964) 'Theory of the interpretation of pressure measurements', *Revue IFP*, March, Vol. 19, No. 3, pp.297–338.

Jacquard, P. and Jain, C. (1965) 'Permeability distribution from field pressure data', *SPE*, Vol. 4, No. 4, pp.281–294.

Jahangiri, H.R. (2007) 'Production optimization using smart well technology with differential evolution algorithm', *Graduate Student Symposium*, University of Southern California.

Jahns, H. (1966) 'A rapid method for obtaining a two-dimensional reservoir description from well pressure response data', *SPE*, Vol. 6, No. 4, pp.315–327.

Kathrada, M. (2009) *Uncertainty Evaluation of Reservoir Simulation Models using Particle Swarms and Hierarchical Clustering*, PhD thesis, Heriot Watt University, Edinburgh.

Klie, H. (2015) 'Physics-based and data-driven surrogates for production forecasting', *SPE Reservoir Simulation Symposium*, Society of Petroleum Engineers, Houston, Texas, USA.

Kriesel, D. (2011) *A Brief Introduction to Neural Networks* [online] http://dkriesel.com (accessed 15 November 2015).

Kruger, W.D. (1961) 'Determining areal permeability distribution by calculations', *J. Pet. Tech.*, Vol. 13, No. 7, p.691.

Li, H. and Daoyong, T.Y. (2011) 'Estimation of multiple petrophysical parameters for the PUNQ-S3 model using ensemble-based history matching', *SPE EUROPEC/EAGE Annual Conference and Exhibition*, Society of Petroleum Engineers, Vienna, Austria.

Liu, J. and Lampinen, J. (2002) 'A fuzzy adaptive differential evolution algorithm', *2002 IEEE Region 10 Conference on Computers, Communications, Control and Power*, Vol. 9, No. 6, pp.606–611.

Mantica, S.C. (2002) 'Combining global and local optimization techniques for automatic history matching production and seismic data', *SPE Journal*, Vol. 7, No. 2, pp.123–130.

Maucec, M.D. (2007) 'Streamline-based history matching and uncertainty, Markov-chain Monte Carlo study of an offshore oil field', *SPE Annual Technical Conference and Exhibition*, Anaheim, California.

Mirzabozorg, A., Nghiem, L., Chen, Z. and Yang, C. (2013) 'Differential evolution for assisted history matching process: SAGD case study', *SPE Heavy Oil Conference Canada, Calgary*, Society of Petroleum Engineers, Alberta, Canada.

Mohaghegh, S.D. (2014) 'Converting detail reservoir simulation models into effective reservoir management tools using SRMs: case study – three green fields in Saudi Arabia', *Int. J. Oil, Gas and Coal Technology*, Vol. 7, No. 2, pp.115–131.

Mohaghegh, S.D., Liu, J.S., Gaskari, R., Maysami, M. and Olukoko, O.A. (2012) 'Application of surrogate reservoir models (SRM) to an onshore green field in Saudi Arabia: case study', *North Africa Technical Conference and Exhibition*, Society of Petroleum Engineers, Cairo, Egypt.

Mohamed, L., Christie, M.A. and Demyanov, V. (2009) 'Comparison of stochastic sampling algorithms for uncertainty quantification', *Reservoir Simulation Symposium*, Society of Petroleum Engineers, Woodlands, Texas.

Neri, F. and Tirronen, T. (2010) 'Recent advances in differential evolution: a survey and experimental analysis', *Artif. Intell. Rev.*, Vol. 33, Nos. 1–2, pp.61–106.

Okano, H. (2013) 'Reservoir model history-matching and uncertainty quantification in reservoir performance forecast using Bayesian framework', *SPE Reservoir Characterisation and Simulation Conference and Exhibition*, Society of Petroleum Engineers, Abu Dhabi, UAE.

Palatnic, B., Zakirov, L., Haugen, S. and van Roosmalen, J. (1993) 'New approaches to multiple history matching', *Seventh European Symposium on Improved Oil Recovery*, Moscow, Russia.

Petroleum Engineering & Rock Mechanics Group (PERM) (2014) http://www3.imperial.ac.uk/earthscienceandengineering/research/perm/punq-s3model/truthcase/procedure (accessed March 2014).

Price, K. and Storn, S. (1997) 'Differential evolution', *Dr. Dobbs J.*, Vol. 22, No. 4, pp.18–24.

Price, K., Storn, R. and Lampinen, J. (2005) *Differential Evolution: A Practical Approach to Global Optimization*, Springer, Berlin, Germany.

Razavi, F. and Jalai-Farahani, F. (2008) 'Ant colony optimization: a leading algorithm in future optimization of petroleum engineering processes', *9th International Conference on Artificial Intelligence and Soft Computing*, Zakopane, Poland.

Rotondi, M., Nicotra, G., Godi, A., Contento, F.M., Blunt, M.J. and Christie, M. (2006) 'Hydrocarbon production forecast and uncertainty quantification: a field application', *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers, San Antonio, Texas, USA.

Rwechungura, R.W., Dadashpour, M. and Kleppe, J. (2011) 'Application of particle swarm optimization for parameter estimation integrating production and time lapse seismic data', *Offshore Europe*, Society of Petroleum Engineers, Aberdeen, UK.

Sampaio, T.P., Ferreira Filho, V.J.M. and Neto, A.D.S. (2009) 'An application of feed forward neural network as nonlinear proxies for use during the history matching phase', *Latin American and Caribbean Petroleum Engineering Conference*, Society of Petroleum Engineers, Cartagena de Indias, Colombia.

Shahkarami, A. (2014) *Assisted History Matching Using Pattern Recognition Technology*, PhD dissertation, West Virginia University.

Shahkarami, A., Mohaghegh, S.D., Gholami, V. and Haghighat, S.A. (2014a) 'Artificial intelligence (AI) assisted history matching', *SPE Western North American and Rocky Mountain Joint Meeting*, Society of Petroleum Engineers, Denver, Colorado.

Shahkarami, A., Mohaghegh, S.D., Gholami, V., Haghighat, A. and Moreno, D. (2014b) 'Modeling pressure and saturation distribution in a CO_2 storage project using a surrogate reservoir model (SRM)', *Greenhouse Gas Sci. Technol.*, Vol. 4, No. 3, pp.289–315.

Slater, G. and Durrer, E. (1971) 'Adjustment of reservoir simulation models to match field performance', *Society of Petroleum Engineers Journal*, Vol. 11, No. 3, pp.295–305.

Sousa, S. (2007) 'Scatter search metaheuristic applied to the history matching problem', *SPE Annual Technical Conference and Exhibition*, Anaheim, California, USA.

Stephen, K.D., Soldo, J., Macbeth, C. and Christie, M.A. (2006) 'Multiple model seismic and production history matching: a case study', *SPE Journal*, Vol. 11, No. 4, pp.418–430.

Storn, R. (1996) 'Differential evolution design of an IIR-filter with requirements of magnitude and group delay', *IEEE Conference on Evolutionary Computation*, Nagoya, Japan, pp.268–273.

Storn, R. (1999) 'System design by constraint adaptation and differential evolution', *IEEE Transactions on Evolutionary Computation*, Vol. 3, No. 1, pp.22–34.

Storn, R. and Price, K. (1995) *Differential Evolution – A Simple and Efficient Adaptive Scheme for Global Optimization Over Continuous Spaces*, Technical Report for International Computer Science Institute, Berkeley.

Storn, R. and Price, K. (1997) 'Differential evolution – a simple and efficient heuristic for global optimization over', *J. Glob. Optim.*, Vol. 11, No. 4, pp.341–359.

Subbey, S. and Christie, M. (2003) 'A strategy for rapid quantification of uncertainty in reservoir performance prediction', *SPE Reservoir Simulation Symposium*, Houston, Texas.

Szklarz, S., Hanea, R.G. and Peters, E. (2011) 'A case study of the history matching of a sector of the Norne field using the ensemble Kalman filter', *SPE EUROPEC/EAGE Annual Conference and Exhibition*, Society of Petroleum Engineers, Vienna, Austria.

Tayari, F., Blumsack, S., Dilmore, R. and Mohaghegh, S.D. (2015) 'Techno-economic assessment of industrial CO2 storage in depleted shale gas reservoirs', *Journal of Unconventional Oil and Gas Resources*, Vol. 11, pp.82–94.

Thomas, P. and Vernon, D. (1997) 'Image registration by differential evolution', *Irish Machine Vision and Image Processing Conference*, Magee College, University of Ulster, Northern Ireland.

Tyler, K., Svanes, T. and Omdal, S. (1993) 'Faster history matching and uncertainty in predicted production profiles with stochastic modeling', *68th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers*, Society of Petroleum Engineers, Houston.

Van Leeuwen, P.J. (1999) 'Comment on 'data assimilation using an ensemble Kalman filter technique', *Monthly Weather Review*, Vol. 127, No. 6, pp.1374–1377.

Wahl, W., Mullins, L., Barham, R. and Bartlett, W. (1962) 'Matching the performance of Saudi Arabian oil fields with an electrical model', *Journal of Petroleum Technology*, Vol. 14, No. 11, pp.1275–1282.

Wang, C. and Gao, J. (2010) 'Enhancing resolution of seismic data based on waveform inversion with cooperative coevolutionary differential evolution algorithm', *SEG Denver 2010 Annual Meeting*, Denver, CO.

Wang, C., Gao, J., Yang, H., Zhao, W. and Zhu, Z. (2011) 'Waveform inversion of cross-well data with cooperative coevolutionary differential evolution algorithm', *SEG San Antonio 2011 Annual Meeting*, San Antonio, Texas.

Wang, J. and Buckley, J.S. (2006) 'Automatic history matching using differential evolution algorithm', *International Symposium of the Society of Core Analysts*, Trondheim, Norway.

Wu, X-H., Bi, L., Yang, Y. and Wang, X. (2013) 'Multiphase upscaling using approximation techniques', *SPE Reservoir Simulation Symposium*, Society of Petroleum Engineers, The Woodlands, Texas, USA.

Yang, C., Ngheim, L. and Card, C. (2007) 'Reservoir model uncertainty quantification through computer-assisted history matching', *SPE Annual Technical Conference and Exhibition*, Anaheim, California, USA.

Zubarev, D. (2009) *Pros and Cons of Applying Proxy-Models as a Substitute for Full Reservoir Simulations*, Society of Petroleum Engineers, New Orleans, Louisiana.