System identification based proxy modeling of a reservoir under iWAG

B M Negash[1,*], **L D Tufa**[2] **and M Ramasamy**[2]

[1]Petroleum Engineering Department, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak Darul Ridzuan, Malaysia
[2]Chemical Engineering Department, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak Darul Ridzuan, Malaysia

[*]bmamo.negash@utp.edu.my

Abstract —The high computational cost associated with first principle based numerical reservoir models create a barrier for realizing their full potential. Currently, high fidelity proxy models, capable of mimicking the numerical reservoir models, are used to undertake some of the tasks. Nonetheless, large number of experimental simulations are required to generate modeling data. Moreover, experimental realizations are obtained by sampling static geological properties, hence, the resulting proxy model does not honor the dynamic nature of reservoirs, especially during immiscible water alternating gas injection (iWAG) process. In this study, a system identification based proxy modeling algorithm is developed for a reservoir under iWAG. A case study is used to evaluate the algorithm's performance in reproducing oil, gas and water production rates. By systematically designing the input pattern, only a single experimental run is used to generate modeling data. A blind validation and a cross plot analysis resulted in an average percentage fit of 89% and a tolerance level of up to 10%, respectively. The algorithm can have potential application for optimization of injection scenario and hence for automatic well control during iWAG.

Keywords—Numerical reservoir model; proxy model; water alternating gas injection; production prediction

1. Introduction

Different alternative methods have been studied in order to replicate the performance of full field numerical reservoir simulation models and at the same time decrease optimization time and operational cost [1]. Currently, proxy models are used to facilitate the full exploitation of the potential of reservoir simulators. These current approach of building a proxy model are based on either reduced order modeling [2], response surface based modeling [3] or surrogate reservoir modeling[4]. In addition to being expensive to build such types of proxy models, they have limitations. In reduced order modeling, the full numerical reservoir is reduced to a less complex and manageable model by upscaling the defining finite difference grid blocks and/or reducing the physics. Apart from being computationally expensive, the resulting simplified model may not truly represent the actual reservoir. In response surface based models, experimental simulation runs are undertaken at systematically selected design points to generate training and validation datasets. The design points are selected based on central composite design or Box-Behnken design of experiment techniques. Incases, where there are many

sensitive independents reservoir parameters the number of simulation runs required will be very large and hence large number of simulation runs will be required. For instance, in order to design an experiment based on the Box-Behnken sampling technique for ten sensitive and independent reservoir parameters 1025 simulation runs are required. This results in expensive computations. In addition, response surface models that are in use today does not honor the dynamic nature of reservoirs, especially when injection of a secondary fluid such as water and gas is involved.

A more recent technique that is based on spatio-temporal data on a grid block basis and artificial neural network has claimed to reduce the number of runs to 15-20 [4, 5]. Even though the time and number of run has been significantly reduced, when compared with traditional proxy modelling, making 15-20 runs still requires a significant amount of time and effort. In addition, designing and organizing of a proper spatio-temporal data is a daunting and computationally expensive task. The degree of success, particularly in practical aspects, of these approaches remains to be argued. Moreover, several runs have to be performed before a reliable model is built [6]. Development of alternative, modest and yet powerful mathematical models allow quick reservoir performance evaluation with an acceptable accuracy [7]. In this research, we developed an algorithm for potential application during the third stage of numerical reservoir modeling and simulation for secondary recovery phase. The algorithm employs system identification principles to construct proxy models. The idea of using system identification for development of a proxy model is not a new concept. However, only few studies have demonstrated the suitability of system identification for development of a proxy model of a reservoir [6]. Moreover, the studies are concentrated on injection of only a single fluid. In this study, a case study approach is used to evaluate the performance of a system identification based proxy modeling algorithm in order to develop a proxy model for a reservoir undergoing iWAG in secondary recovery mode. The case study involves a benchmark reservoir model obtained from UNISIM research group in Brazil [8, 9].

2. Algorithm for proxy modelling of a reservoir under iWAG

The proposed system identification based workflow is discussed in this section. The workflow begins by designing a suitable injection pattern which is bounded by an upper and lower limit. Once a proxy model is developed it can be utilized for optimization of multiple injection scenarios [10].
The set of procedures included in the workflow are:

1. Design a suitable injection pattern and introduce it to numerical reservoir model input file
2. Run numerical reservoir simulation
3. Organize a dataset consisting of injection rate and production rates.
4. Split the dataset set into modeling and validation data set
5. Select an appropriate model structure (ARX, ARMAX, OE, BJ and their respective noise integrated versions.)
6. Select model order between 1 and 10.
7. Use the data reserved for modeling to estimate model parameters
8. Cross-validate the infinite step ahead simulated output and validation data.
9. If validation is not successful, then iterate by changing the selected model order.
10. If validation is not successful, even after exploring several model orders, then change the model structure and continue the iteration as in step 9.
11. If the validation is not successful, increase the number modeling data and repeat steps 6 to 10.
12. Repeat steps 6 to 11 until a valid proxy model is obtained.

3. System description: case study

The case study deals with developing an alternative system identification based proxy model using data collected from a reservoir model under iWAG. The reference model is a benchmark reservoir model developed by the “Research Group in Reservoir Simulation and Management (UNISIM)” [11, 12]. It was built in a high resolution geocellular model, using public data from Namorado Field, Campos Basin, Brazil. Geocellular models replicate the 3-D structure of a real reservoir, with the stratigraphic envelope, reservoir sublayers, and faults all represented in three dimensions. The original intention in developing the case was to create, test and compare new approaches related to reservoir development and management by research centers. The reservoir volume is meshed into smaller 3-D blocks, referred to as Grid blocks. The dimensions of each grid block are about 50 m (163 ft) thick

and 100m X 100m (328 X 328 ft) areally. There are a total of 93960 such grid blocks with 81,58 and 20 blocks arranged in the x, y and z directions, respectively. The water-oil and gas-oil contacts are located at a depth of 3200m and 3000m, respectively. Initial oil originally in place is calculated to be 175000 m^3 and the initial reservoir pressure was 300bar. Figure 1 presents the top view of initial fluid distribution, well locations of the UNISIM reservoir model. Figure 2 displays the top view of fluid saturation at the end of simulation period. In both Figure 1 and Figure 2 100% water is represented by blue color, 100% gas is represented by red color and 100% oil is represented by green color. The reservoir model has 17 production wells which are optimally placed and perforated at locations that favor maximum oil production. This includes high permeability, high pore volume, and high oil saturations. In addition, 11 injector wells are introduced and optimized using streamline simulation options that indicates the streamline flows. Constrains on both injection and production wells were adopted from the original model.

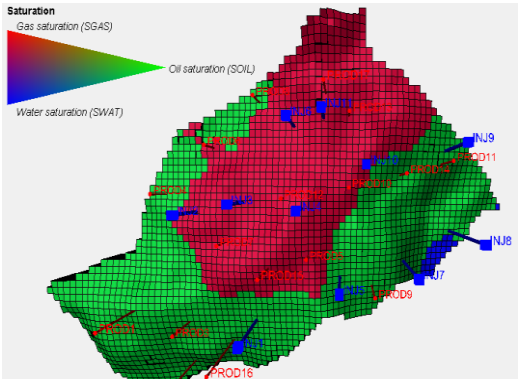

Figure 1. Initial fluid distribution of the UNISIM reservoir model

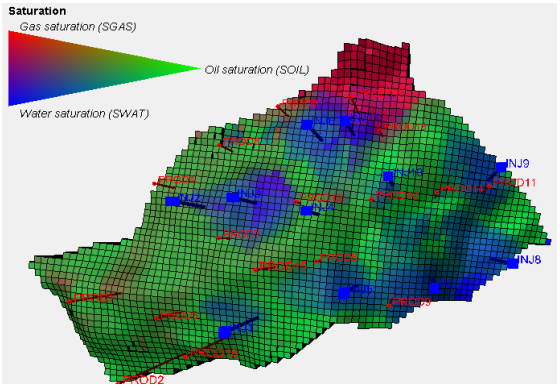

Figure 2. Fluid distribution of the UNISIM reservoir model at the end of simulation period

The reservoir model of Namorado field was modified so that all of the eleven injection wells inject water and gas alternatively while maintaining a PRBS form of injection pattern. Water and gas injection were allowed to alternate every 15 days. The resulting data were used to carryout system identification in a Two Input – Single Output (MISO) system mode where the total fluid production rate was considered as output and the total gas injection and total water injection rates were treated as two inputs to the system. Data preparation for system identification includes visual inspection and selection of an appropriate and representative dataset. Injection and production data were made available for system identification by splitting them into modeling and validation datasets. A simple model such as an ARX model is identified and its transient response is studied, then more complex model structures are investigated by using the order and time delay, obtained in the first stage, as

Table 1. Lower and upper bounds used in the design of a PRBS input during SWG.

Well Name	Gas injection Sm^3/day	Water injection Sm^3/day	Water injection Sm^3/day	Water injection Sm^3/day
	Lower bound		Upper bound	
Well 1	40000	1500	140000	4500
Well 2	45000	1750	145000	4750
Well 3	50000	2000	150000	5250
Well 4	55000	2250	155000	5550
Well 5	60000	2500	160000	5750
Well 6	65000	2750	165000	6000
Well 7	40000	3000	155000	6250
Well 8	45000	3250	160000	6500
Well 9	50000	3500	165000	6750
Well 10	55000	3750	165000	7000
Well 11	60000	4000	155000	7250

initial approximation. The use of simpler model allows to obtain insight into the order and delay of more complex model structures that are capable of identifying the system more accurately. In this study, the inputs, gas injection rates and water injection rates were designed as a PRBS switching between two bounds as shown in Table 1, and at the same time alternating every 15 days.

Injection of water and gas alternatively allows piston like displacement as displayed in Figure 3. This has the potential to effectively sweep the oil with only a small amount of residual oil.

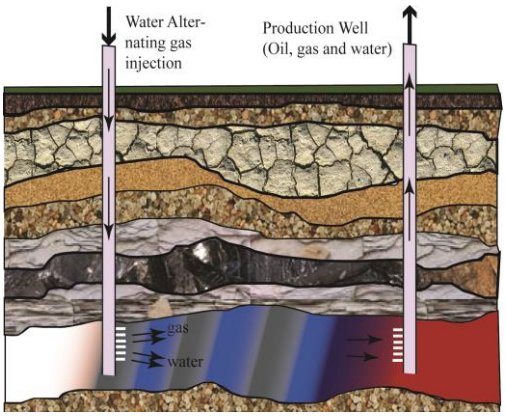

Figure 3. Physics of a reservoir flow during iWAG.

4. Results and Discussion

This section presents the results and discussions of modeling total oil, gas and water production rates, as a function of total gas injection and total water injection rates. The models obtained in this study can be useful for optimization of water and gas injection rates during iWAG.

4.1. Modeling oil production

This system has two inputs, namely total gas injection rate and water injection rate, and one output, viz. oil production rate. Figure 4 presents the modeling and validation data used in modeling oil production as a function of gas and water injection rates. A total of 2666 data at a frequency of 1/day were obtained from the reservoir simulator. The first 1500 data points were used for modeling and the next 10 data were removed to fulfill the requirements of predictive models. The remaining 1156 data points were used for cross validation.

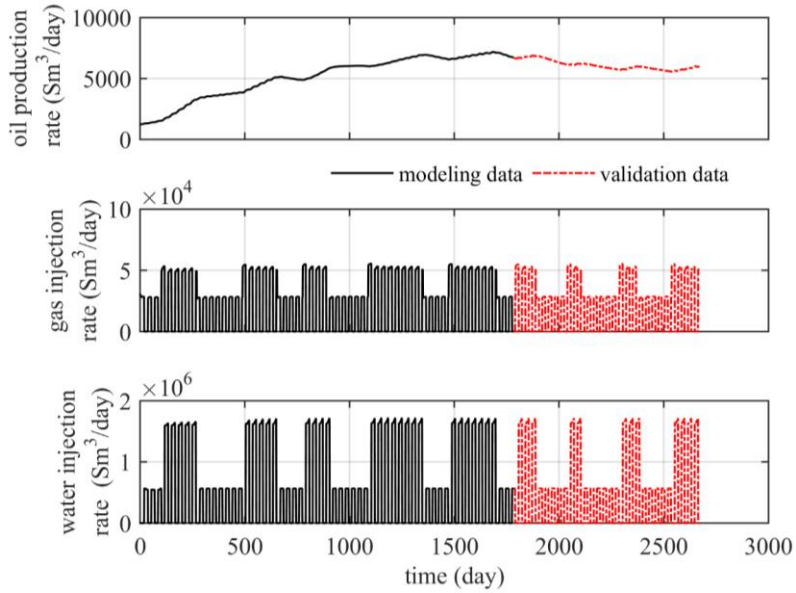

Figure 4. input and output data for modeling and validation of total oil production during iWAG injection.

The identification of a characteristic LTI model had started by identifying a simple ARX model. The model order and time delay of the identified ARX model was then used as initial approximation for further investigation into more complex model structures. More complex Polynomial LTI model structures (ARMAX, OE, and BJ) with and without noise integration had been thoroughly searched for a model that gives minimum error or best fit during validation. This was done by systematically changing the time delays and orders of the defining polynomials in both the deterministic and stochastic components of each model structures until a model is successfully identified. As a result, a discrete-time Box-Jenkins Polynomial with Noise Integration model that exhibits a one-time delay and fourth order in all the coefficient polynomials is found to be an optimal estimator in terms of the percentage fit obtained. The model was named BJI (4-4-4-4-1). The percentage fit to modeling and validation data were found to be 99.94% and 89.96%, respectively. The mean squared error (MSE) and Final prediction error (FPE) were 1.037and 1.085. The BJI (4-4-4-4-1) model is presented below:

$$y(t) = \left[\frac{B(q)}{F(q)}\right]u(t) + \left[\frac{C(q)}{D(q)(1-q^{-1})}\right]e(t) \quad (1)$$

where:

$$B1(q) = [3.813 \quad 2.276 \quad -0.1771 \quad -0.508] * e^{-5}$$
$$B2(q) = [-1.734 \quad 6.035 \quad 5.963 \quad -0.641] * e^{-7}$$
$$C(q) = [-1.232 \quad -0.5181 \quad 0.9541 \quad -0.1607]$$
$$D(q) = [-2.37 \quad 1.403 \quad 0.3641 \quad -0.3927]$$
$$F1(q) = [-1.29 \quad 0.2356 \quad -0.4558 \quad 0.5118]$$
$$F2(q) = [-0.8214 \quad -0.819 \quad 0.1862 \quad 0.4544]$$

A comparison plot of the identified BJI (4-4-4-4-1) model's infinite step simulated output and validation data is shown in Figure 5. It can be observed that the model has successfully represented oil production rate as a function of gas and water injection rates.

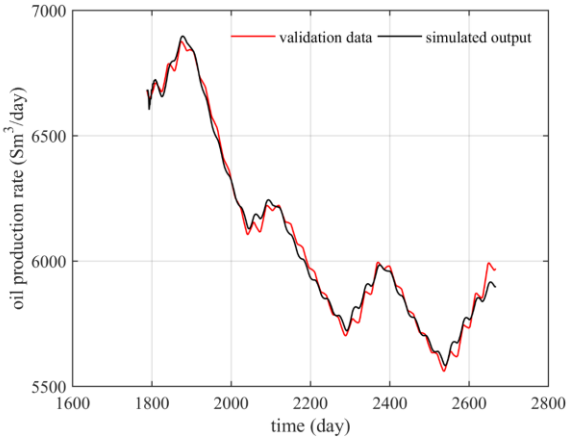

Figure 5. Validation of the BJI (4-4-4-4-1) mode.

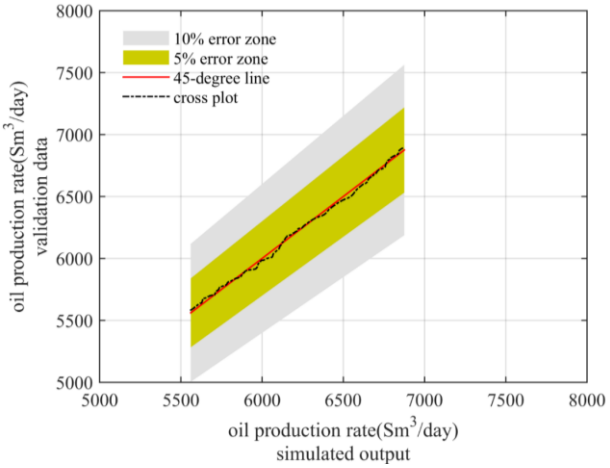

Figure 6. Cross plot for BJI (4-4-4-4-1).

Figure 6 presents a cross plot obtained by plotting simulated output vs. validation data on a linear and square plot. It can be seen that the 45-degree line closely matches the cross plot indicating that the model does, neither underestimated nor overestimated during validation.

4.2. Modeling gas production

The previous session outlined and discussed modeling and validation of total oil production rate as a function of total water and total gas injection. This section applies the same procedure to model total gas production as a function of total gas and total water injection rates. Figure 7 shows the input and output data collected at a sampling frequency of 1/day from year 2005 to end of simulation year. A total of 2666 data at a frequency of 1/day were obtained from the reservoir simulator. A simple visual inspection, shows that the first 500 data set is not representative of the dataset observed at a later time, hence, it was not considered during the modeling stage. The remaining data were used for identification. Data from 500 to 1500 were used for modeling and the next 10 data were removed to

fulfill the requirements of predictive models. The remaining 1156 data points were used for cross validation.

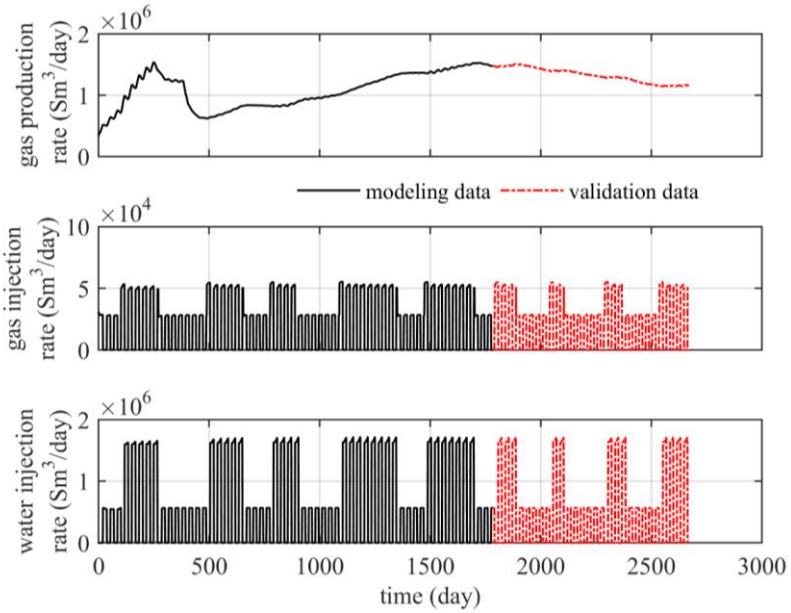

Figure 7. Input and output data for modeling gas production during iWAG injection.

At first an ARX and ARIX model structures, a simple model structures, were searched for by systematically changing their model order and time delay. The best estimating ARX or ARIX model's order and time delay were then utilized as initial approximation for further investigation into more complex model structures. of Consequently, a Box-Jenkins model of order seven in all the numerators and denominators and a one-time delay was found to produce best estimation in terms of calculated percentage fit. The model was named BJ (7-7-7-7-1) and it was cross-validated using data that has not been employed during the modeling stage. Figure 8 presents a comparison between validation data and simulated output of the identified model. The fit to estimation data was 99.86% and the fit to validation data was 91.06%. Final prediction error (FPE) and mean square error (MSE) were 6076 and 9892, respectively. By removing the first 500 points, the percentage fit has improved when compared to a similar identification process with all the data points included. The best match which captured the gas production rate dynamics during identification and while using the whole set of data was 60.71%. This case serves as an example of the importance of inspecting of input and output data prior to splitting the data into modeling and validation.

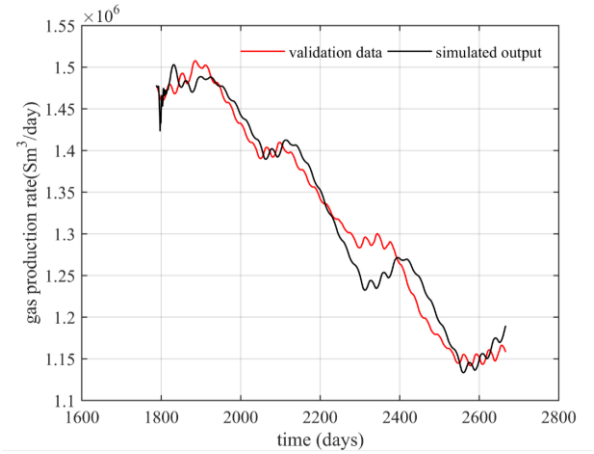

Figure 8. Validation of the BJ (7-7-7-7-1) model

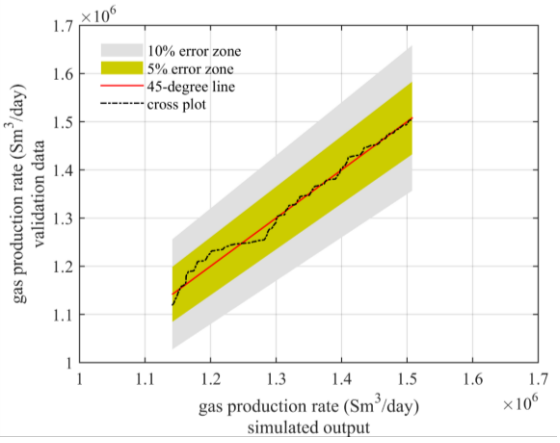

Figure 9. Cross plot analysis for model BJ (7-7-7-7-1).

The discrete time BJ (7-7-7-7-1) model is presented below:

$$y(t) = \left[\frac{B(q)}{F(q)}\right]u(t) + \left[\frac{C(q)}{D(q)}\right]e(t) \tag{2}$$

where

$B1(q) = [0.005366 \quad 0.01955 \quad -0.01959 \quad 0.0143 \quad -0.007743 \quad 0.00717 \quad -0.004015]$

$B2(q) = [-40.53 \quad 70.24 \quad -29.34 \quad 32.58 \quad 66.9 \quad 32.84 \quad 1.513] * e^{-5}$

$C(q) = [0.02011 \quad 0.03306 \quad 0.0005555 \quad -0.01909 \quad -0.01029 \quad 0.0004809 \quad 0.0006449]$

$D(q) = [-2.176 \quad 1.547 \quad -0.8752 \quad 0.8192 \quad -0.4003 \quad 0.2673 \quad -0.1819]$

$F1(q) = [-2.421 \quad 2.024 \quad -0.6362 \quad -0.135 \quad 0.2282 \quad -0.0415 \quad -0.01869]$

$F1(q) = [-2.01 \quad 1.199 \quad -0.9998 \quad 1.845 \quad -1.234 \quad 0.06368 \quad 0.1357]$

Figure 9 presents a cross plot analysis of an infinite step ahead prediction and validation data. The plot was made by first concatenating the two data and sorting them in ascending order. The plot is almost in line with the 45-degree line indicating that the identified mode has neither underestimated nor overestimated during validation.

4.3. Modeling water production

In addition to modeling oil and gas production, modeling water production is also an essential task. This section demonstrates the application of system identification in building a proxy model for capturing the relationship between total water production rate, total water injection rate and total gas injection rate during water alternating gas flooding. Data for identification were collected by proper excitation of the Namorado field's numerical reservoir model. The same simulation run that produced the gas and oil production rates in the previous two sessions also provided water production rate. Figure 10 presents the input and output data used for modeling and cross validation of a polynomial model. A total of 2666 data at a frequency of 1/day were obtained from the reservoir simulator. However, a simple visual inspection of the data revealed that the reservoir's response has only begun after 500 days of injection and production. This signifies that the first 500 data is not useful for identification. Hence, data points from 500 to 1500 were used for modeling and the next 10 data were removed to fulfill the requirements of predictive models. The remaining 1156 data points were used for cross validation.

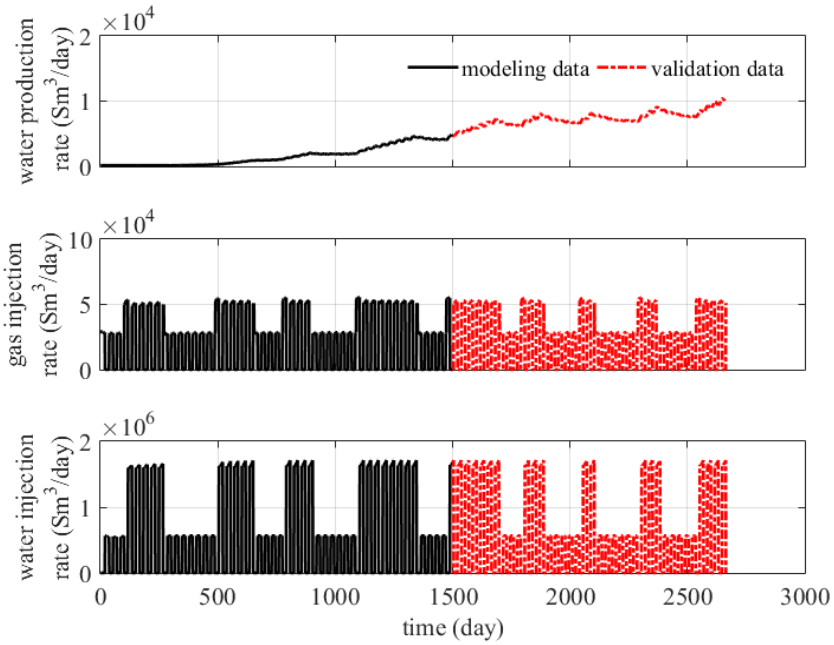

Figure 10. input and output data for modeling and validation of total water production during gas injection.

Identification of a suitable model has beggun by first exploring a simpler polynomial model structure, ARX, and its noise integrated version, ARIX. We found that an ARX model whose order is eight in both the numerators and denominators and a zero-time delay was the best estimating model in

terms of the percentage fit exihibited. This model was named ARX (8-8-0). Using ARX (8-8-0) as initial estimate, more complex model structures were searched. We found a noise integrated Box-Jenkin model that has a model order of four in all the numerator and denominator polynomilas and a zero-time delay to best fit the validation data. This model was named BJI (4-4-4-4-0) and was validated using the validation data. The fit to estimation and independent validation data were found to be 99.86% and 60.52%. Moreover, the final prediction error and mean square error were calculated to give 2.727 and 425.2, respectively. Figure 11 presents a graphical comparison between validation data and simulated output. It is observed that the identified model has closely predicted total water production during validation. The structure and polynomial coefficients of BJI (4-4-4-4-0) model are shown below:

$$y(t) = \left[\frac{B(q)}{F(q)}\right]u(t) + \left[\frac{C(q)}{D(q)(1-q^{-1})}\right]e(t) \tag{3}$$

where:

$B1(q) = [0.0001924 \quad 0.0002775 \quad 0.0001607 \quad 3.495e^{-5}]$ $D(q) = [-1.782 \quad -0.1408 \quad 1.731 \quad -0.7927]$

$B2(q) = [1.961 \quad -1.076 \quad 3.861 \quad -1.877] * e^{-6}$ $F1(q) = [-0.4819 \quad -0.9691 \quad 0.1753 \quad 0.3016]$

$C(q) = [-0.2918 \quad -1.438 \quad 0.3392 \quad 0.5495]$ $F2(q) = [-2.015 \quad 1.437 \quad -0.5185 \quad 0.09636]$

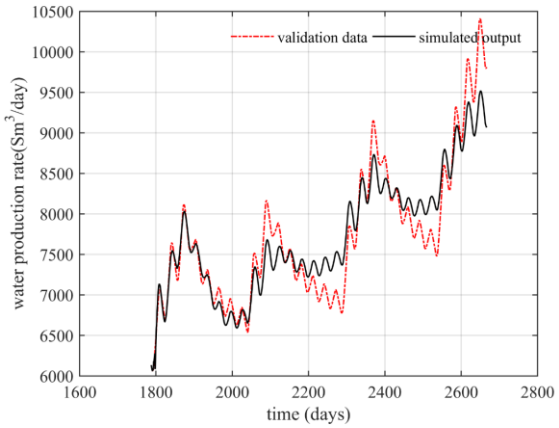

Figure 11. Validation of model BJ (7-7-7-7-0).

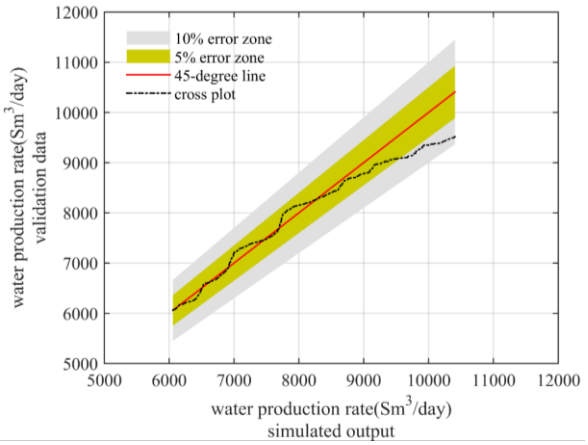

Figure 12. Cross plot analysis of model BJI (4-4-4-4-0).

Crossplot analysis was also conducted to find out whether the model has overestimated, underestimated or well estimated the output. This is realized by first sorting and plotting simulated output vs. validation data and comparing the plot with a straight 45-dgree line. Figure 12 shows that the model does not underestimate or overestimate the production of water at the early stage of the production.

5. Conclusion

The large computational time required to make simulation runs inhibit the full exploitation of the potential of numerical reservoir models. In this study, an algorithm that infuses system identification via LTI model structures and numerical reservoir models is proposed and evaluated. The resulting outcome is a proxy models to replace the role of numerical reservoir simulators. As opposed to current methods of proxy modeling, system identification based proxy models require only a single experimental run with and a pseudo random binary sequence form of injection pattern. The application is demonstrated by using a benchmark case study obtained from UNISIM, a research group established in 1996 and focused on studies related to numerical simulation and management of petroleum reservoirs. The case study involves a reservoir model with 18087 active grid blocks with 11 injectors and 19 producers which are optimally placed. Proxy models are developed to represent oil, gas and water production rates as a function of alternating water and gas injection rate. A percentage fitness test followed by whiteness test and cross plot analysis were used to measure the performance of the developed models. As a result, an average percentage fit of 89% is obtained indicating that the

algorithm is reasonably efficient. The crossplot analysis also reveal that the developed proxy models are accurate and have a tolerance level of up to 10%. With this performance the algorithm can be useful for well setting and production optimization purposes.

References

[1] Zubarev D I 2009 Pros and cons of applying proxy-models as a substitute for full reservoir simulations *SPE Annual Technical Conference and Exhibition* (4-7 October 2009, New Orleans, Louisiana)
[2] Zangl G, Graf T and Al-Kinani A 2006 Proxy modeling in production optimization *SPE Europec/EAGE Annual Conference and Exhibition* (12-15 June 2006, Vienna, Austria)
[3] Negash B M, Vasant P M and Jufar S R 2018 Application of artificial neural networks for calibration of a reservoir model *Intelligent Decision Technologies* 12 (1) pp 67-79
[4] Alenezi F and Mohaghegh S 2017 Developing a Smart Proxy for the SACROC Water-Flooding Numerical Reservoir Simulation Model *SPE Western Regional Meeting* (23-27 April 2017 Bakersfield, California)
[5] Lutidze G and Younis R 2018 Proxy reservoir simulation model for IOR operations *Saint Petersburg 2018*
[6] Negash B M, Tufa L D, Ramasamy M and Awang M B 2017 System Identification Based Proxy Model of a Reservoir under Water Injection *Modelling and Simulation in Engineering* **2017** pp 1-10
[7] He Q 2016 *Investigating Continuously Updated History Matching Using Smart Proxy* (*Surrogate Reservoir Model*) (West Virginia University)
[8] Ljung L 1998 System identification *Signal analysis and prediction* (Springer) pp 163-173
[9] Nelles O 2013 *Nonlinear system identification: from classical approaches to neural networks and fuzzy models* (Springer Science & Business Media)
[10] Negash B M, Vel A and Elraies K A 2017 Artificial Neural Network and Inverse Solution Method for Assisted History Matching of a Reservoir Model *International Journal of Applied Engineering Research* **12** (11) pp 2952-62
[11] Gaspar A T, Santos A, Maschio C, Avansi G, Hohendorff Filho J and Schiozer D 2013 Study Case for Reservoir Exploitation Strategy Selection based on UNISIM-I Field ed: Campinas-SP, Brazil, http://www.unisim.cepetro.unicamp.br/unisim-i
[12] Gaspar A T, Avansi G D, Maschio C, Santos A A and Schiozer D J 2016 UNISIM-IM: Benchmark Case Proposal for Oil Reservoir Management Decision-Making *SPE Trinidad and Tobago Section Energy Resources Conference* (13-15 June, Port of Spain, Trinidad and Tobago)