

Research Article

System Identification Based Proxy Model of a Reservoir under Water Injection

Berihun M. Negash, Lemma D. Tufa, M. Ramasamy, and Mariyamni Bt. Awang

Universiti Teknologi Petronas, 32610 Seri Iskandar, Perak Darul Ridzuan, Malaysia

Correspondence should be addressed to Berihun M. Negash; bmamo.negash@utp.edu.my

Received 30 August 2016; Revised 6 December 2016; Accepted 26 January 2017; Published 23 February 2017

Academic Editor: Feng Xiao

Copyright © 2017 Berihun M. Negash et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Simulation of numerical reservoir models with thousands and millions of grid blocks may consume a significant amount of time and effort, even when high performance processors are used. In cases where the simulation runs are required for sensitivity analysis, dynamic control, and optimization, the act needs to be repeated several times by continuously changing parameters. This makes it even more time-consuming. Currently, proxy models that are based on response surface are being used to lessen the time required for running simulations during sensitivity analysis and optimization. Proxy models are lighter mathematical models that run faster and perform in place of heavier models that require large computations. Nevertheless, to acquire data for modeling and validation and develop the proxy model itself, hundreds of simulation runs are required. In this paper, a system identification based proxy model that requires only a single simulation run and a properly designed excitation signal was proposed and evaluated using a benchmark case study. The results show that, with proper design of excitation signal and proper selection of model structure, system identification based proxy models are found to be practical and efficient alternatives for mimicking the performance of numerical reservoir models. The resulting proxy models have potential applications for dynamic well control and optimization.

1. Introduction

Reliable prediction is a natural requirement to the design of an optimal management strategy whereby important decisions on future investment, facility development, and expansion will be made [1]. Typically, in the petroleum industry, forecasting oil, gas, and water production rates is undertaken on a regular basis. Some of the factors that determine precision and accuracy of these forecasts include quality of geological and geophysical (G&G) data, quality and quantity of observed production rate and pressure data, and modeling procedures employed.

Decline curve analysis (DCA) and numerical reservoir simulation are classical methods that are frequently employed to forecast reservoir performance [2–6]. Each of these methods has strengths, limitations, and restrictions [7]. In practice, it is common to implement more than one method to reduce uncertainty and increase prediction accuracy. Decline curve analysis (DCA) is popular in the industry and is relatively a modest technique. It is one of the oldest methods, which was

developed by Arps [2] and has been utilized even when there were no high-speed processors for handling large amounts of production and pressure data. The empirical Arps equations recognized for different curtailment trends are shown in (1), (2), and (3) [8]. The analysis is based on fitting one of these three empirical relations using measured production data.

For the exponential case ($b = 0$),

$$q = q_i e^{-at}. \quad (1)$$

For the hyperbolic case ($0 < b < 1$),

$$q = \frac{q_i}{(1 + ba_i t)^{1/b}}. \quad (2)$$

For the harmonic case ($b = 1$),

$$q = \frac{q_i}{(1 + a_i t)}, \quad (3)$$

where a is the nominal (instantaneous) decline rate, b is the decline exponent constant, q is the flow rate, and q_i is the initial flow rate

However, it is recognized that the application of DCA is limited to the early production stage of a well, where production is at steady state and the reservoir is under boundary-dominated flow. In addition, rock and fluid compressibility must be low and constant. In cases such as gas reservoir and solution gas derived oil production, where compressibility is higher, the decline is less severe and hence none of the three curtailment trends will fit. An improved DCA which considers reservoir parameters as part of the decline analysis was developed by Fetkovich [4] and is referred to as Fetkovich's type-curve analysis. Fetkovich showed that the value of b was an indication for efficiency of recovery, where zero was the lowest efficiency with small compressibility and 0.5 was the highest efficiency that is achievable from any known recovery mechanism with optimistic relative permeability. Analytical flow equations were applied to generate type curves for transient flow, and they were combined with the empirical decline curve equations documented by Arps [2]. However, Fetkovich's analysis has some notable limitations. It is applicable if the reservoir is homogeneous and has constant thickness, no-flow outer boundary, and constant flowing pressure at the wellbore. In addition, Fetkovich's type-curve analysis assumes a constant productivity index, which implies a straight-line relationship between pressure and flow rate. However, Evinger and Muskat [9] stated that for multiphase flow the relationship between pressure and flow rate is curved. Therefore, Fetkovich's type-curve analysis performs well if only the reservoir is saturated with a single-phase fluid of constant and small compressibility.

On the other hand, numerical reservoir simulation based on first principles provides a more accurate and robust solution to the task of forecasting reservoir performance. Reservoir simulation models are mathematical descriptions of an actual reservoir and are based on material balance, momentum balance, and constitutive equations such as Darcy's Law. Reservoir models are often used to analyze, optimize, and forecast both pressure and saturation dependent terms. Geological, geophysical, petrophysical, well log, core, and fluid data are required to construct reservoir simulation models. However, development of an accurate and representative numerical model is challenging and it involves utilization of a large number of uncertain data. This is because reservoirs are gigantic subsurface systems, which are both heterogeneous and anisotropic. Due to the relatively large size, samples of rock and fluid are collected from only a few locations within the boundary of the reservoir. Statistical techniques are then employed to populate and infer rock and fluid properties for the remaining portion. Moreover, contrary to direct measurement, most of the input data required to build a simulation model are inferred from physical and chemical measurements on rock and fluid samples taken at only few locations of the vicinity of the reservoir. For instance, fluid resistivity measurements are used to determine saturation of water within the reservoir and levels of natural gamma ray and spontaneous potential of a reservoir rock are used to determine the type of rock (lithology) and mineral content. Thus, reservoir property data used in modeling and forecasting are only best estimates and will not be exact

representatives of the reservoir at large. Besides, the history-matching stage by which the inferred reservoir properties are adjusted until the numerical simulation model reasonably mimics the actual reservoir is a formidable task and has substantial limitations [3, 5].

History matching is the process of altering sensitive reservoir model parameters and developing a list of representative numerical models which are capable of reproducing observed production and pressure data [10, 11]. History matching is undertaken via manual alteration of sensitive parameters or automatically by using dedicated multiobjective optimization algorithms [12]. The former is time-consuming and requires expert knowledge and experience on the reservoir under study. In addition, manually adjusting reservoir parameters results in trial-and-error solutions; therefore, it is impossible to use all information; it is not suitable for complex reservoir models; it is time-consuming and it provides suboptimal results. A relatively new technique, which is referred to as assisted history matching (AHM), substantially reduces the time required. In AHM, global optimization algorithms help to reduce a misfit function between the observed and simulated production and pressure data. AHM techniques apply methods that allow for multiple realizations and are able to deal with the uncertainty in an optimal manner.

The Kalman Filter and its variants, such as the Ensemble Kalman Filter (EnKF) and Particle Filter, are currently popular among the research community [13–17]. Though AHM shortened the time taken during the history-matching phase, it is an ill-posed optimization problem, which is known to result in many realizations. Different sets of reservoir parameters may equally reproduce the historical data which results in a different prediction [6]. The so-called best realization could also produce a geologically inconsistent reservoir model [11]. The reader is advised to refer to [3] which has critically summarized the limitations of assisted history matching.

Once history matching is successfully completed, the models are used for production optimization and prediction of future reservoir performance under various development strategies. The process is repeated several times until a viable strategy is realized. For a medium sized reservoir model, each simulation run may require hours to complete. In spite of the demanding effort and time spent on model construction and history matching, the challenge encountered during optimization makes it difficult to exploit the full potential of numerical reservoir simulators. The use of proxy models that mimic the behavior of numerical reservoir models is currently on the rise and significantly eases the optimization stage. Once developed, these proxy models can be evaluated within seconds. Many commercial simulators that generate polynomial proxy models are being developed and used in both industry and academia. In all the simulators, response surface based proxy models are being employed. However, several runs are required to develop a representative proxy model that accurately mimics the response of numerical reservoir models. For instance, if there are ten factors (e.g., water injection wells) and three responses (oil, water, and gas production rates), then the Box-Behnken experiment design will require at least 170 runs. Again, if each run takes

four hours to complete, then 680 computational hours are required to gather data for proxy modeling.

In this paper, we propose system identification based proxy models for optimization of a reservoir model. System identification is the art and science of model building that engages statistical methods to build mathematical models of dynamic systems from measured data. System identification involves proper design of excitation signal, model structure selection, parameter estimation, and validation of the identified model. By properly designing the excitation signal and by selection of an appropriate model structure, it is possible to obtain a representative proxy model using data from a single run of the numerical simulation model. System identification is extremely flexible in a sense that there are several model structures to capture linear as well as nonlinear relationships. In this study, a benchmark reservoir model originally developed to evaluate history-matching techniques is used to validate the efficacy of system identification based proxy models, which have the potential to be used for automatic control and production optimization. The proposed methodology assumes that the numbers of wells and operating conditions except for injection rates remain the same. In cases where infill drilling is required, it is necessary to redo the identification anew.

2. Materials and Methods

2.1. System Identification Based Proxy Models. System identification is a popular and steadily growing modeling approach. It is used to build analytical models of a dynamic system and forecast future behavior easily with less computational effort and time [18–20]. It is commonly used in engineering where prediction is part of an integrated technology such as model based control systems. System identification develops a uniquely structured mathematical model of a system through the application of certain model structure and conventional statistical methods. The number of researches and results reported related to the use of system identification in the oil and gas industry is increasingly growing. Aifa [21] has presented applications of neural network based black-box model, to various problems that arise in the process of understanding the behavior of reservoirs. Sheremetov et al. [22] applied nonlinear autoregressive neural networks with exogenous input to forecast production rate of a naturally fractured reservoir. Olominu and Sulaimon [7] pointed out that autoregressive integrated moving average (ARIMA), which is a popular model structure in time series analysis, has a better efficacy than DCA. In addition, they showed that the model might be useful for short- to mid-long-term forecasting. However, ARIMA model structure does not consider external inputs, such as water injection. As such, the application is limited to primary recovery only. Elmabrouk et al. [23] studied the use of feedforward backpropagation artificial neural network in the prediction of oil production and employed real data from a Libyan oil field to test the hypothesis. The study proves the potential use of advanced empirical models for better prediction of oil production, which is essential for better reservoir management. However, the study does not incorporate the effect of external inputs.

TABLE 1: Polynomial system identification model structures.

	Name
$F(q) = D(q) = C(q) = A(q) = 1$	Frequency impulse response (FIR)
$F(q) = D(q) = C(q) = 1$	Autoregressive with external input (ARX)
$F(q) = D(q) = 1$	Autoregressive moving average with external input (ARMAX)
$D(q) = C(q) = A(q) = 1$	Output error (OE)
$A(q) = 1$	Box-Jenkins (BJ)

System identification is the process of developing mathematical models of a physical system from input and output data [18, 19]. In this study, the system is a reservoir under water injection, the inputs are either water injection rates or well bottomhole pressures, and outputs are production rates. During the lifetime of a reservoir, data such as wellhead flowing pressure, bottomhole flowing pressure, production rates, injection rate, gas/oil ratio, and water cut are recorded regularly to quantify production and evaluate field performance. Therefore, one can develop a dynamic model that assists in decision-making using system identification techniques and those readily available sets of data. In this study, linear polynomial model structures, which are among the many variations of system identification model structures, are used to build a proxy model, which can be used for production optimization. Equation (4) displays the general representation of a linear, dynamic time invariant system [24, 25]:

$$y(n) = q^{-k}G(q^{-k}, \theta)u(n) + H(q^{-k}, \theta)e(n), \quad (4)$$

where

$$G(q^{-1}, \theta) = \frac{B(q)}{F(q)A(q)}, \quad (5)$$

$$H(q^{-1}, \theta) = \frac{B(q)}{F(q)A(q)}.$$

$A(q)$, $B(q)$, $C(q)$, $D(q)$, and $F(q)$ are polynomials with respect to the shift operator q^{-1} , where $q^{-k}u(k) = u(k-1)$. Depending on the nature of these polynomials, different structures can be formulated as shown in Table 1.

Noise integrated versions of the model structures listed in Table 1 are also available for system identification. These are special cases of general linear input-output models. They correspond to linear difference equations relating the input to the output under various noise assumptions. For ARX model structure, addition of noise integration leads to an “ARIX” model structure:

$$A(q)y(t) = B(q)u(t - nk) + \frac{C(q)}{1 - q^{-1}}e(t). \quad (6)$$

Similarly, the ARMAX and BJ structures can be enhanced to handle nonstationary disturbances and are referred to as ARIMAX and BJI, respectively.

TABLE 2: Groups and wells control data for water injection case.

Minimum BHP	Individual producer wells control			Minimum Q_O	Individual injector wells control	
	Maximum Q_L	Maximum WC	Maximum GOR		Maximum BHP	Maximum Q_w
280	2000	0.9	200	20	350	7500

TABLE 3: Group wells control for water injection case.

Maximum Q_L	Producers group control		Injectors group control	
	Maximum Q_O	Maximum Q_w	Maximum Q_w	Maximum Q_w
15500	15500	13950	13950	65000

The development of the system identification based models comprises the following key procedures:

- Collection of input and output data from the physical model
- Choice of model structure that represents the physical system
- Estimating unknown model parameters in the selected model structure
- Validating the estimated model using cross validation and/or residual analysis

A detailed discussion on the steps of system identification and model development is available by Ljung [24] and Nelles [25].

2.2. Excitation Input Selection. Selection of a proper excitation input dictates efficacy in system identification. There are several input sequences, which can be used in modeling different complexities. Among them, a pseudorandom binary sequence (PRBS) is very popular and efficient in equally exciting all frequencies. PRBS is a sequence that switches between an upper and a lower bound at varying frequencies. PRBS can also have multiple amplitudes; that is, different upper and lower bounds may be specified at different times for a given input sequence. However, such PRBS inputs are often employed during nonlinear system identification. Since this article studies the application of linear polynomial model structures, the excitation signal used was PRBS that only switches between two amplitudes.

2.3. Model Validation. Model validation constitutes quantifying the deviation exhibited by the estimated model and making sure that all the information contained in the input and output data is exploited. In this article, percentage fit, whiteness test, and crossplot test were used to evaluate the efficacy of the identified model. The mathematical expression of percentage fit is shown in

$$\begin{aligned} \text{percentage fit} &= 1 - \text{NRMSE} \\ &= 1 - \frac{\|y - \hat{y}\|}{\|y - y_{\text{mean}}\|} * 100\%, \end{aligned} \quad (7)$$

where \hat{y} is the predicted output, y is the observed output, y_{mean} is the mean of the observed output, and NRMSE is the normalized root mean square error.

Whiteness test of a residual indicates whether the model has extracted all the information or requires further refining. This is achieved by plotting a histogram of residuals. A good model is one that has a mean value of zero and some variance, indicating that the residuals are uncorrelated. On the other hand, a crossplot analysis is a graphical and visual method of comparing two sets of data whereby the plot of simulated output and validation data was compared graphically and visually, with a 45-degree straight line. The crossplot of a residual is an x - y plot obtained by plotting the simulated output on the x -axis and validation data on the y -axis. A good model will produce an output that gives a perfectly 45-degree straight line when plotted against the validation data set.

2.4. Description of the Case Study Reservoir Model. The reservoir model, which was used to generate the input and output data, is referred to as UNISIM-I-M and is based on a field located in Campos basin in Brazil [27]. The model was originally developed to serve as a benchmark case study on decision analysis regarding reservoir management. For a detailed description of the model, readers are referred to Avansi and Schiozer [26] and Gaspar et al. [27]. Figure 1 shows the porosity distribution while Figure 2 presents the reservoir geometry including the wells placement and vertical permeability distribution of the UNISIM-I-M reservoir model. All the red wells represent producer wells while blue wells represent injector wells. The reservoir volume is meshed into smaller 3D blocks, referred to as grid blocks. Each grid block has a thickness of 50 m, a width of 100 m, and a length of 100 m. There are a total of 93960 such active 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 3200 m and 3000 m, respectively. Initial oil originally in place is calculated to be 175000 m^3 . The initial reservoir pressure was 300 bar. The total numbers of water injectors and producers are 11 and 14, respectively. The original well placement was slightly modified to optimize production; however, production constraints are kept the same as the original. Production and injection well constraints are described in Tables 2 and 3.

3. Result and Discussion

In developing a proxy model, the inputs consisted of water injection rate from individual wells and the output consisted

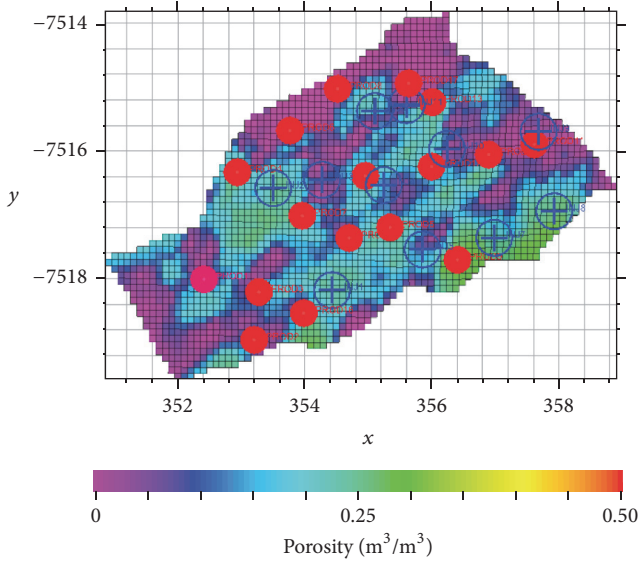


FIGURE 1: Porosity distribution of the first layer of the UNISIM-I-M reservoir model.

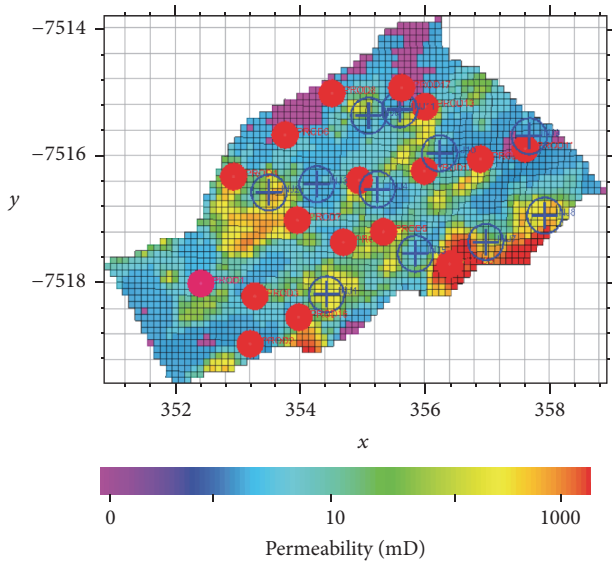


FIGURE 2: Vertical permeability distribution of the first layer of the UNISIM-I-M reservoir model.

of the total oil production rate. The design of a PRBS input was in such a way that each of the multiple inputs has different lower and upper bounds. The lower and upper bounds used to excite the numerical simulation model are shown in Table 4.

The response of the numerical reservoir model is presented in Figure 3. The figure shows that the total number of input and output data is 2807, of which 1500 were used for identification and the next 10 data points were removed to fulfill requirements set by the initial condition of prediction equations. The remaining 1297 data points were utilized for cross validation. Each of the 2807 data points presents the input and output calculated on a daily basis. The numerical simulation run time required to generate the outputs was 4 hours and 53 minutes on Intel® Core™

TABLE 4: Lower and upper bounds of input design for modeling fluid production during water injection.

Well name	Lower bound (Sm ³ /day)	Upper bound (Sm ³ /day)
Well 1	1500	4500
Well 2	1750	4750
Well 3	2000	5000
Well 4	2250	5250
Well 5	2500	5500
Well 6	2750	5750
Well 7	3000	6000
Well 8	3250	6250
Well 9	3500	6500
Well 10	3750	6750
Well 11	4000	7000

i5-3210 M CPU @2.5 GHz Processor with 6 GB RAM. This is considered very large compared to the few seconds required to run polynomial models, which are examined in this study. In addition, when carrying out multiple simulation runs to evaluate reservoir performance, the essence of utilizing reduced model or metamodel as a substitute for numerical reservoir simulator becomes extremely substantial.

3.1. Initial Model Estimation. A simple and yet useful system identification model structure, that is, ARX, and its noise integrated version were used to initiate the system identification process. This allows narrowing the model-order search region and determining time delay. Noise integrated ARX model (ARIX) with both poles and zeros equal to three and a time delay of zero was found to result in a percentage fit of 71.92%. The model was named ARIX (3-3-0). Model structure of ARIX model is

$$A(q)y(t) = B(q)u(t) + \frac{1}{(1-q^{-1})}e(t), \quad (8)$$

where $A(q)$ is a polynomial in the shift operator (q) of order na , $B(q)$ is a polynomial in the shift operator (q) of order nb , $y(t)$ is the output (oil production rate), $u(t)$ is the input (injection rates), na is the number of poles, and nb is the number of zeros.

Figure 4 presents the cross validation result of the ARIX (3-3-0) model. Akaike's final prediction error (FPE) and mean squared error (MSE) exhibited were 0.8157 and 0.8724. The MSE is a measure of the quality of an estimator; it is always nonnegative, and values closer to zero are better. Moreover, according to Akaike's theory, the most accurate model has the smallest FPE. The step response, which illustrates the time behavior of the outputs of ARX (3-3-0) when its inputs change from zero to one in a very short time, is shown in Figure 5. It can be seen that the model has responded without a time delay and has become stable.

3.2. Identification of a More Complex and More Accurate Model. Based on the analysis of the ARIX (3-3-0) model, more complex model structures were examined by

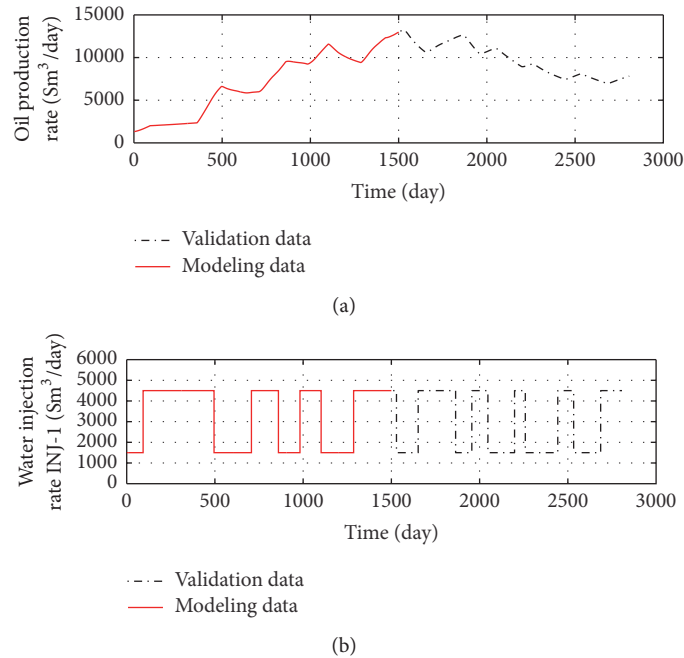


FIGURE 3: Oil production rate (a) and water injection rate of well INJ-1 (b), used during modeling and validation of linear time invariant model structures.

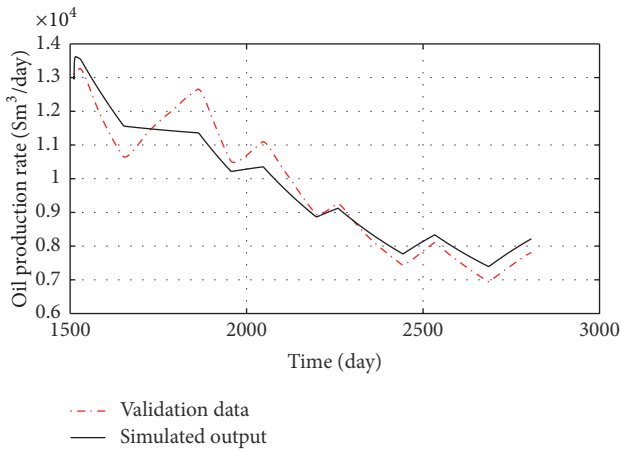


FIGURE 4: Validation of the ARIX (3-3-0) model for MISO system.

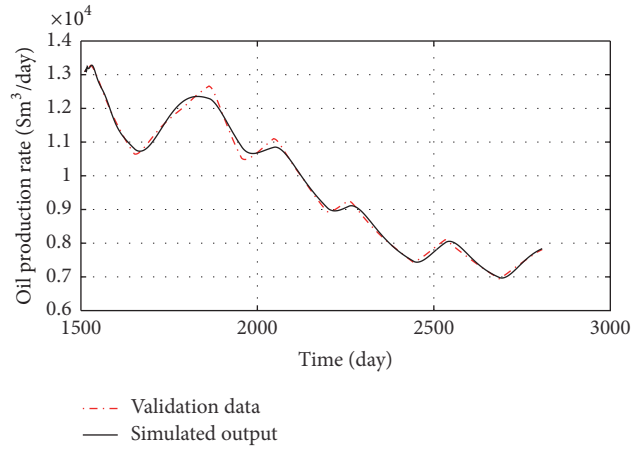


FIGURE 6: Comparison of validation data and BJI (4-4-4-4-2) model's simulated output.

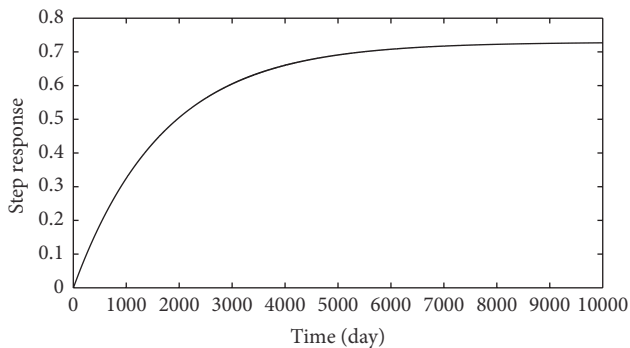


FIGURE 5: Step response of ARIX (3-3-0) model.

introducing a variation in model order and time delay close to the ARIX (3-3-0) model. The best model that resulted in the best percentage fit with the validation data set was found to be Box-Jenkins model with noise integration (BJI). The structure of the BGI model is

$$y(t) = \left[\frac{B(q)}{F(q)} \right] u(t) + \left[\frac{C(q)}{D(q)(1-q^{-1})} \right] e(t). \quad (9)$$

The number of poles and zeros of the deterministic part was found to be four for all the inputs. The stochastic part of the BJI model structure that gave the best performance was also of order four in both poles and zeros. The time delay is found to be two. The model can be referred to as BJI (4-4-4-4-2). Figure 6 presents a graphical comparison

of simulated output and measured validation data. It can be seen that the simulated output has closely reproduced the validation data. The percentage fit, FPE, and MSE exhibited were 94.01%, 0.207, and 0.5632, respectively. The estimated polynomial functions of the BJI (4-4-4-2) model are

$$\begin{aligned}
B1(q) &= 0.0003314q^{-2} - 0.0009479q^{-3} \\
&\quad - 0.001406q^{-4} - 0.001088q^{-5} \\
B2(q) &= -0.0003488q^{-2} - 1.278e - 05q^{-3} \\
&\quad - 0.00049q^{-4} - 0.0001398q^{-5} \\
B3(q) &= -0.0003505q^{-2} - 2.031e - 05q^{-3} \\
&\quad + 0.0001999q^{-4} + 0.000143q^{-5} \\
B4(q) &= 0.003034q^{-2} + 0.001483q^{-3} + 0.001071q^{-4} \\
&\quad + 0.0009769q^{-5} \\
B5(q) &= -0.0001085q^{-2} + 0.000201q^{-3} \\
&\quad + 0.0003909q^{-4} + 0.0003442q^{-5} \\
B6(q) &= -0.001266q^{-2} - 0.000845q^{-3} \\
&\quad - 0.0005904q^{-4} - 0.0008796q^{-5} \\
B7(q) &= -0.000173q^{-2} + 8.793e - 05q^{-3} \\
&\quad + 0.0002124q^{-4} + 0.0001657q^{-5} \\
B8(q) &= -0.0001733q^{-2} + 8.761e - 05q^{-3} \\
&\quad + 0.0002121q^{-4} + 0.0001654q^{-5} \\
B9(q) &= -0.0001734q^{-2} + 8.75e - 05q^{-3} \\
&\quad + 0.000212q^{-4} + 0.0001653q^{-5} \\
B10(q) &= -0.0000.3q^{-2} + 1.986e - 05q^{-3} \\
&\quad + 0.0001798q^{-4} + 0.0001357q^{-5} \\
B11(q) &= -0.0002676q^{-2} + 2.000e - 05q^{-3} \\
&\quad + 0.0002q^{-4} + 0.0001157q^{-5} \\
C(q) &= 1 - 0.01847q^{-1} - 0.2401q^{-2} - 0.1495q^{-3} \\
&\quad - 0.03079q^{-4} \\
D(q) &= 1 - 1.85q^{-1} + 1.009q^{-2} - 0.2766q^{-3} \\
&\quad + 0.1215q^{-4} \\
F1(q) &= 1 - 1.937q^{-1} + 1.131q^{-2} - 0.3102q^{-3} \\
&\quad + 0.1171q^{-4}
\end{aligned}$$

$$\begin{aligned}
F2(q) &= 1 - 1.936q^{-1} + 1.131q^{-2} - 0.3103q^{-3} \\
&\quad + 0.1166q^{-4} \\
F3(q) &= 1 - 1.936q^{-1} + 1.131q^{-2} - 0.3097q^{-3} \\
&\quad + 0.1175q^{-4} \\
F4(q) &= 1 - 1.937q^{-1} + 1.131q^{-2} - 0.3097q^{-3} \\
&\quad + 0.1177q^{-4} \\
F5(q) &= 1 - 1.936q^{-1} + 1.131q^{-2} - 0.3098q^{-3} \\
&\quad + 0.1174q^{-4} \\
F6(q) &= 1 - 1.936q^{-1} + 1.131q^{-2} - 0.3097q^{-3} \\
&\quad + 0.1176q^{-4} \\
F7(q) &= 1 - 1.937q^{-1} + 1.123q^{-2} - 0.3103q^{-3} \\
&\quad + 0.117q^{-4} \\
F8(q) &= 1 - 1.937q^{-1} + 1.123q^{-2} - 0.3103q^{-3} \\
&\quad + 0.117q^{-4} \\
F9(q) &= 1 - 1.937q^{-1} + 1.113q^{-2} - 0.3103q^{-3} \\
&\quad + 0.117q^{-4} \\
F10(q) &= 1 - 1.937q^{-1} + 1.113q^{-2} - 0.3101q^{-3} \\
&\quad + 0.1172q^{-4} \\
F11(q) &= 1 - 1.922q^{-1} + 1.123q^{-2} - 0.3101q^{-3} \\
&\quad + 0.1165q^{-4}.
\end{aligned} \tag{10}$$

In addition, the step response of the identified model is shown in Figure 7. It can be observed that the response of model BJI (4-4-4-2) is stable after an abrupt change in input. Knowing how the system responds to a sudden input is important because large and possibly fast deviations from the long-term steady state may have extreme effects on the prediction ability of the model. This will in turn influence decisions.

3.3. Residual Analysis. Two types of residual analysis were carried out: a whiteness test and crossplot analysis. The whiteness test was undertaken by plotting a histogram of the residuals, which were calculated by differencing simulated output and validation data. Figure 8 presents the histogram of the residuals. The mean and sample variance of the residual distribution are 0.0873 and 1272, respectively. According to the whiteness test, a good model that has exploited most of the information, which is inherently contained in the data, will have a white noise. Since the mean is almost zero, it can be considered that the model has captured most of the information in the input and output data.

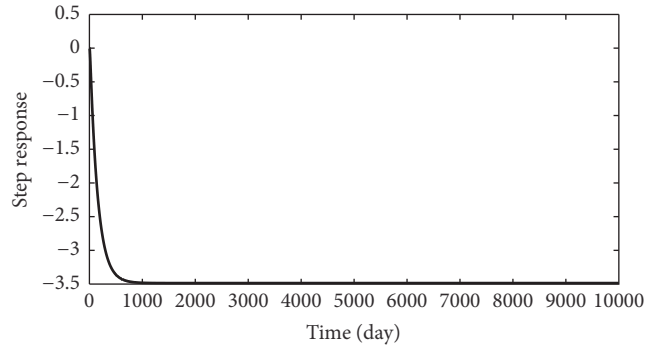


FIGURE 7: Step response of model BJI (4-4-4-4-2).

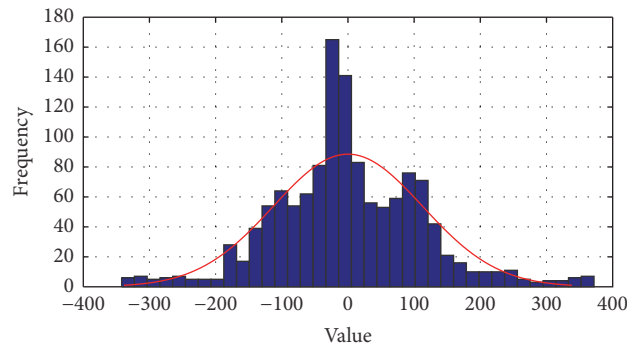


FIGURE 8: Histogram of residuals for model BJI (4-4-4-4-2).

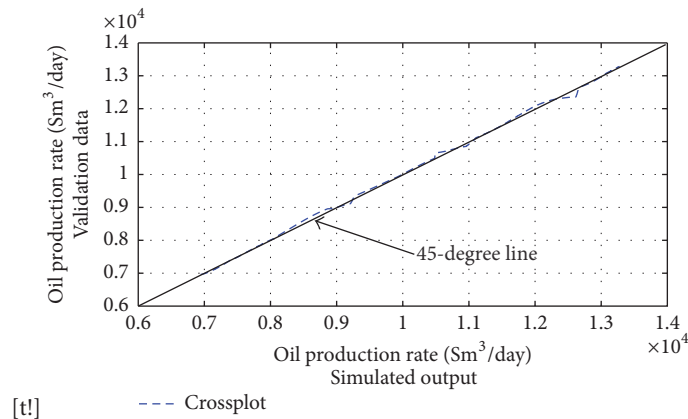


FIGURE 9: Crossplot of simulated output and validation data for model BJI (4-4-4-4-2).

In addition, a crossplot of simulated output versus validation data was plotted to allow further analysis on the identified model. The data set was concatenated and sorted in descending order based on the simulated output before it was plotted on x - y plane. A good model will have its crossplot aligned with the 45-degree line. A model that underestimates the output will have its crossplot below the 45-degree line and a model that overestimates the output will have its crossplot above the 45-degree line. Figure 9 presents a crossplot of model BJI (4-4-4-4-2). It can be seen that the model neither underestimates nor overestimates during the

validation period. In other words, the prediction ability of the identified model is satisfactory.

4. Summary and Conclusion

Numerical reservoir modeling and simulation has many applications in the upstream oil and gas industry. However, due to the large computational time required, full exploitation of its potential use remains challenging. Currently, the use of response surface based proxy models is becoming a common trend. Nowadays, several simulation software packages

incorporate a module for proxy modeling. While such proxy models are useful and significantly reduce optimization time, their development still requires several simulation runs. In this study, a system identification based proxy model that only requires a single run and properly designed excitation signal was proposed. A case study that involves water injection was used to investigate and establish a procedure for application of system identification techniques in reservoir simulation. Injection flow rates were designed in the form of a pseudorandom binary sequence that is known to excite all frequencies equally and allow successful identification. Input and output data were collected at a frequency of 1/day and were divided into modeling and validation sets. The modeling set was used to construct a proxy model. The validation set was used for cross validation and residual analysis. Residual analysis and crossplot test together with the calculation of percentage fit have shown that the identified model was able to accurately mimic the performance of the numerical reservoir simulator while predicting total oil production during oil recovery by water injection. In conclusion, with proper selection of input sequence and model structure, system identification can deliver a proxy model with just one simulation run.

Nomenclature

Q_L :	Liquid flow rate
Q_O :	Oil flow rate
DCA:	Decline curve analysis
Q_W :	Water flow rate
WC:	Water cut
BHP:	Bottomhole pressure
GOR:	Gas-to-oil ratio
NRMSE:	Normalized root mean square error
PRBS:	Pseudorandom binary sequence
MISO:	Multiple input-single output.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

References

- [1] P. Obidike, C. Emelle, and R. Vlaardingerbroek, "Integrated production forecasting for business planning," in *Proceedings of the 38th Nigeria Annual International Conference and Exhibition*, Nigeria, August 2014.
- [2] J. Arps, "Analysis of decline curves," *Transactions of the AIME*, vol. 160, no. 1, pp. 228–247, 1945.
- [3] M. Cancelliere, F. Verga, and D. Viberti, "Benefits and limitations of assisted history matching," *Offshore Europe*, 2011.
- [4] M. Fetkovich, "Decline curve analysis using type curves," *Journal of Petroleum Technology*, vol. 32, pp. 1065–1077, 1980.
- [5] Z. Tavassoli, J. N. Carter, and P. R. King, "Errors in history matching," *SPE Journal*, vol. 9, no. 3, pp. 352–361, 2004.
- [6] Y. Tomomi, "Non-uniqueness of history matching," in *Proceedings of the SPE Asia Pacific Conference on Integrated Modelling for Asset Management*, Society of Petroleum Engineers, 2000.
- [7] O. Olominu and A. A. Sulaimon, "Application of time series analysis to predict reservoir production performance," in *Proceedings of the SPE Nigeria Annual International Conference and Exhibition*, Society of Petroleum Engineers, Dallas, Tex, USA, 2014.
- [8] H. C. Slider, *Worldwide Practical Petroleum Reservoir Engineering Methods*, PennWell Books, 1983.
- [9] H. Evinger and M. Muskat, "Calculation of theoretical productivity factor," *Transactions of the AIME*, vol. 146, no. 1, pp. 126–139, 1942.
- [10] H. B. Crichlow, *Modern Reservoir Engineering: A Simulation Approach*, Prentice Hall, 1977.
- [11] D. S. Oliver, A. C. Reynolds, and N. Liu, *Inverse Theory for Petroleum Reservoir Characterization and History Matching*, Cambridge University Press, 2008.
- [12] R. Rwechungura, M. Dadashpour, and J. Kleppe, "Advanced history matching techniques reviewed," in *Proceedings of the 17th Middle East Oil and Gas Show and Conference (MEOS '11)*, pp. 1729–1747, Manama, Bahrain, September 2011.
- [13] A. A. Emerick and A. C. Reynolds, "Combining the ensemble Kalman filter with Markov Chain Monte Carlo for improved history matching and uncertainty characterization," *SPE Journal*, vol. 17, no. 2, pp. 418–440, 2012.
- [14] J. Ping and D. Zhang, "History matching of channelized reservoirs with vector-based level-set parameterization," *SPE Journal*, vol. 19, no. 3, pp. 514–529, 2014.
- [15] Y. Chen and D. S. Oliver, "History matching of the norne full-field model with an iterative ensemble smoother," *SPE Reservoir Evaluation & Engineering*, vol. 17, no. 2, pp. 244–256, 2014.
- [16] A. A. Emerick and A. C. Reynolds, "History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations," *Computational Geosciences*, vol. 16, no. 3, pp. 639–659, 2012.
- [17] M. Tarrahi, E. Gildin, J. Moreno, and S. Gonzales, "Dynamic integration of DTS data for hydraulically fractured reservoir characterization with the ensemble Kalman filter," in *Proceedings of the SPE Energy Resources Conference*, pp. 822–835, June 2014.
- [18] B. George, *Time Series Analysis: Forecasting & Control*, Pearson Education, Uttar Pradesh, India, 3rd edition, 1994.
- [19] M. Nerlove, D. M. Grether, and J. L. Carvalho, *Analysis of Economic Time Series: A Synthesis*, Academic Press, 2014.
- [20] A. Zellner and F. Palm, "Time series analysis and simultaneous equation econometric models," *Journal of Econometrics*, vol. 2, no. 1, pp. 17–54, 1974.
- [21] T. Aifa, "Neural network applications to reservoirs: physics-based models and data models," *Journal of Petroleum Science and Engineering*, vol. 123, pp. 1–6, 2014.
- [22] L. Sheremetov, A. Cosultchi, J. Martínez-Muñoz, A. Gonzalez-Sánchez, and M. A. Jiménez-Aquino, "Data-driven forecasting of naturally fractured reservoirs based on nonlinear autoregressive neural networks with exogenous input," *Journal of Petroleum Science and Engineering*, vol. 123, pp. 106–119, 2014.
- [23] S. Elmabrouk, E. Shirif, and R. Mayorga, "Artificial neural network modeling for the prediction of oil production," *Petroleum Science and Technology*, vol. 32, no. 9, pp. 1123–1130, 2014.
- [24] L. Ljung, *System Identification: Theory for the User*, PTR Prentice Hall Information and System Sciences Series, Prentice Hall, Englewood Cliffs, NJ, USA, 1987.
- [25] O. Nelles, *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*, Springer, Berlin, Germany, 2001.

- [26] G. D. Avansi and D. J. Schiozer, "UNISIM-I: synthetic model for reservoir development and management applications," *International Journal of Modeling and Simulation for the Petroleum Industry*, vol. 9, no. 1, pp. 21–30, 2015.
- [27] A. T. Gaspar, A. Santos, C. Maschio, G. Avansi, and D. Schiozer, "UNISIM-IM: study case for management variables optimization of reservoir exploitation strategy," *International Journal of Modeling and Simulation for the Petroleum Industry*, vol. 9, pp. 1–7, 2015.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

