Identification of a Surrogate Model for Economic Performance Prediction of Oil Reservoir Production under Waterflooding Process

Karim Salahshoor[1×] and Seyed Morteza Hoseini[2]

[1] Professor, Department of Instrumentation and Industrial Automation, Petroleum University of Technology, Ahwaz, Iran

[2] M.S. Student, Department of Instrumentation and Industrial Automation, Petroleum University of Technology, Ahwaz, Iran

Highlights

- A framework system theory-based approach modeling for waterflooding process in oil reservoirs is proposed;
- The effect of geological uncertainties on hydrocarbon recovery efficiency is modeled and measured;
- Reservoir management goals can be pursued in the presence of uncertainty based on the obtained model;
- The developed approach has been assessed by being implemented in MRST on the EGG reservoir model.

Abstract

The model-based optimization of the waterflooding process has found significant scope for improving the economic life-cycle performance of oil fields due to geological and economic uncertainties compared to conventional reactive strategies. This paper proposes a new frequency-based system identification method to identify a robust multi-input, multi-output (MIMO) surrogate model for an oil reservoir under waterflooding process so as to describe all the injector-producer relationships. In contrast to the conventional modeling methods, the proposed data-driven modeling approach uses the available injection and production rates as the reservoir input–output data. Meanwhile, it includes a structured-bounded uncertainty model in the form of norm-bounded state-space function blocks to account for uncertainties, facilitating the identified model employed in robust control methodology using linear matrix inequality (LMI) problem formulation so as to eliminate the effect of model uncertainty. The identified MIMO surrogate model is integrated with a desired nonlinear net present value (NPV) objective function in a multi-input, single-output (MISO) system configuration to synthesize a model-based optimization prediction for economical operation and production of oil from oil reservoirs under both geological and economic uncertainties. The introduced approach is implemented on the "EGG model" as a well-recognized three-dimensional synthetic oil reservoir with eight water injection wells and four oil production wells. The results demonstrate that economic performance prediction of the oil reservoir, having an uncertain permeability field, lies in the evaluated bound of the uncertainty model. Waterflooding is a well-known method for increasing oil production. A significant amount of time and effort is required even for high-performance processors to numerically simulate a reservoir with thousands of grid blocks. On the other hand, there is a high uncertainty level in oil reservoir model-based economic optimization due to limited information about

---

× Corresponding author:
    Email: salahshoor@put.ac.ir

geological model parameters. Employing robust control methods can provide robustness for the performance and stability of the control system against model norm-bounded uncertainty. However, in all standard identification methods, it is assumed that the uncertainties in the model can be accommodated in the form of noise. Therefore, the challenge of using the models estimated from the standard identification approach in robust control methods can mainly be considered an essential subject. This paper presents a new frequency-based modeling approach to identify a surrogate model and uncertainty modeling for the waterflooding process with the ease of being employed in robust control methods. A desirable relationship is obtained between the injection rate and the economic production function to model the dynamics of the reservoir using the identification of the surrogate model. Then, the concept of structure bounded uncertainty modeling is presented to describe the model geological uncertainty.

Keywords: Economic, Surrogate model, System identification, Waterflooding

1. Introduction

It is required to increase the recovery of oil from existing reservoirs to fill the current gap in the supply and demand of energy. An oil reservoir has a typical life span of about a few decades, wherein a maximum of three stages of production can take place. Theoretically speaking, 20% to 70% of current procured oil can be recovered using secondary recovery mechanisms (Van den Hof et al., 2009). Waterflooding is the most commonly used secondary oil recovery method in which water is injected into the reservoirs using injection wells to stabilize the pressure in the reservoir and sweep the oil through the porous formation rock into the production wells. Reservoir performance prediction is critical in oil reservoir management strategies to achieve maximum economic benefits in oil reservoirs under the waterflooding process. Decline curve analysis (DCA) and numerical reservoir simulation are classic methods for predicting reservoir performance. These methods have their strengths and weaknesses (Olominu et al., 2014). The DCA is one of the most traditional methods proposed in the work of Arps et al. (1945) for cases where there is no access to high-speed processors for processing vast amounts of production and pressure data.

However, the applications of the DCA are limited to production in stationary states and when the reservoir is under dominant frontier flow, requiring that the compressibility of rocks and fluids be at a constant and insignificant value. This method is inapplicable to high-compressibility cases, such as gas reservoirs and oil extraction using oil-soluble gas. On the other hand, numerical reservoir simulation (Fanchi et al., 2005) is a more accurate and robust solution for predicting the performance of reservoirs. Simulated models of reservoirs provide a mathematical description of natural reservoirs by mass balance, momentum conservation, and the use of principal equations such as Darcy's Law. However, developing such numerical models requires geological, petrophysical, and geophysical data along with charts of wells and information on properties of the desired fluids. Therefore, obtaining an accurate reservoir model is a challenging task accompanied by various and large amounts of uncertain data mainly because reservoirs are extensive systems with heterogeneous and incoherent physical properties. Given the large size of reservoirs, rock and fluid samples are commonly gathered from only certain reservoir parts. Statistical methods are used to predict the properties of rocks and fluids from the remaining parts

of the reservoir. Therefore, data used in modeling and predicting reservoir performance are the best estimates of the actual values. Furthermore, the history matching process, wherein the extracted features of the reservoir are continually adjusted with data until the simulated numerical model shows acceptable patterns similar to natural reservoirs, is a rather tricky procedure with its limitations (Tavassoli et al., 2004; Cancelliere et al., 2011).

Extensive research has been made in reservoir modeling, which shows that selecting the proper method is highly dependent on available data, desired accuracy, and time cost (Abou-Kassem et al., 2001). Reservoir simulation can be used as a proper technique for estimating reservoir performance. A numerical tool or mathematical model is required for predicting the performance of reservoirs to adopt the most suitable strategy for water injection in the secondary stage of oil removal (Fanchi et al., 2005). Recent advances in drilling and reservoir instrumentation have made it possible to use model-based control methods and optimize procedures to develop practical reservoir management strategies to further increase the economic efficiency of oil reservoirs. However, one major challenge in implementing these novel technologies is the issue of geological uncertainty (Alhuthali et al., 2008). Furthermore, the large scale and complexity of hydrocarbon reservoirs have created additional complexity in tools used for controlling these systems (Van den Hof et al., 2009). History matching has been used to address permeability uncertainty to realize different oil reservoir models (Oliver and Chen, 2011).

However, regarding the problems of classic modeling of reservoirs, it is essential to find a simple, fast, and accurate method for modeling oil reservoirs under waterflooding. Fortunately, recent studies have focused on this issue. For instance, Sayyafzadeh et al. (2011) proposed a method for showing the relationship between the rate of different wells and the rate of a production well using a set of transform functions. A data assimilation algorithm can be introduced in a history matching scheme to achieve complete predictions through the modeling process (Emerick and Reynolds, 2013). However, the obtained history-matched models are complex, and the developed models cannot be used in robust production performance prediction and advanced oil reservoir management due to unbounded uncertainty intervals. Further, the reservoir models that incorporate highly nonlinear dynamic equations, having numerous parameters and states, make it impractical to realize efficient reservoir production management (Chen and Hoo, 2013). Tafti et al. (2013) proposed an AutoRegressive model with eXogenous input using an experimental design approach based on the physical properties of reservoirs. Rezapour et al. (2015) compared the prediction accuracy of various linear models in oil reservoirs under waterflooding, and Horoofar et al. (2016) also applied a system identification approach using a linear structure for modeling the SPE10 oil reservoir.

Recent advances in excavation drilling and reservoir instrumentation have made it possible to use model-based control methods and optimize procedures to develop executable practical reservoir management strategies to further increase the economic efficiency of oil reservoirs. However, the large scale and complexity of such hydrocarbon reservoirs have created additional complexity in tools used for controlling these systems (Van den Hof et al., 2009). Furthermore, one major challenge in implementing these novel technologies is the issue of geological uncertainty (Alhuthali et al., 2008).

Despite the advantages of the proposed studies in the simulation of oil reservoirs under waterflooding, all the studies used a linear structure to model the procedure, which is inapplicable when considering the nonlinear dynamic time-variant behavior of reservoirs during all stages of operation. For example, the nonlinear and dynamic behavior of reservoirs is highlighted during the first stages of production, and linear-based predictions will not provide reliable results. Moreover, one of the main challenges in the closed-loop management of reservoirs is the high uncertainty caused by limited knowledge of the model parameters, which prevents the potential advantages of the economic performance prediction of the

closed-loop reservoir management from manifesting and can risk going beyond predicted economic costs.

This study proposes a new identification method for modeling waterflooding process in oil reservoirs. The geological uncertainty is included using a structure-bound uncertainty model with the system identification method. System identification is the science and art of creating mathematical models for dynamic systems using observed input-output data (Ljung et al., 2010). System identification involves the proper design of excitation signals, selection of model structure, estimation of model parameters, and evaluation of the identified model. This modeling tool is highly flexible considering that it contains various structures for expressing different linear and nonlinear relationships. For the first time, this study proposes using surrogate model system identification and structure uncertainty modeling in state-space block formulation to account for the dynamics of waterflooding process in oil reservoirs. The proposed mechanisms can be used in reservoir management and advanced control and optimization strategies to achieve the main objective of improving the economic performance of an oil reservoir during the productive life span of oil reservoirs.

Section 2 of this study includes theoretical and background information on the proposed strategy. Section 3 shows the stages of developing the proposed method along with the obtained outcomes of model performance, and Section 4 demonstrates the results and conclusions.

2. Background materials

2.1. System identification

System identification (SI) is a well-known method for modeling an unknown system for optimization and control. The main procedure focused in SI modeling is typical because the unknown system is considered a black box with specified inputs and outputs. Then, the dynamic mathematical model between any system input and output of the mathematical dynamic unknown systems can be determined using statistical methods and the recorded system input-output data. The ultimate objective is to obtain an appropriate dynamic mapping between the relevant system input and output to correctly predict the output of the system under different conditions. Therefore, after training the identified model, it can be used for prediction, control, and optimization purposes (Ljung et al., 2010).

Equation-based methods for modeling oil reservoirs are commonly used in commercial and numerical simulators to model oil reservoirs. However, most of the information produced by commercial simulators can be regarded as redundant from the actual waterflooding procedure point of view. During waterflooding, the performance of oil reservoirs can be affected by certain input–output variables such as the rate of water injection in injection wells and water and oil production rates from production wells. Therefore, it is not required to model every existing phenomenon in oil reservoir; instead, it is only essential to determine the external model relationship between water injection rate as the input and water and oil production rates as the output relative input–output parameters variables. For this purpose, an oil reservoir model can be structured and configured in a multiple-input multiple-output dynamic model structure. This external input–output dynamic perspective of an unknown system is commonly used for modeling purposes in control theory. Therefore, it is essential to determine the accessible input and output ports of an oil reservoir. System identification procedure can be exercised to identify all the relevant external input–output dynamic models. In other words, the input and the output of the system must be accessible. In this study, water injection rates from injection wells and bottom hole pressure are typically used as the inputs for modeling waterflooding in oil reservoirs.

a. Surrogate model

Surrogate modeling is an engineering tool used for calculating a $y = f(x)$ mapping that hides the physical relationship between the x input vector (x) and the y input (y) as the output scalar (Forrester et al., 2008). Surrogate models, otherwise known as meta-models, are compressed analytical models used to describe the behavior of multi-input multi-output systems based on a limited number of simulations with a high computational load. The main idea behind creating surrogate models is to calculate a secondary function for each specified input value to estimate the desired model using the observed data without feed-forward simulations. The performance of the surrogate models depends on the input values and model structure. These models imitate the complex behavior of simulated models and can be used for optimization, sensitivity analysis, and control, requiring thousands of simulation iterations. It is required to calculate the desired output or the value of the secondary function corresponding to the input vector to create a surrogate model. The main objective of this study regarding reservoir management is to identify a robust surrogate model for an oil reservoir to be integrated with a net present value (NPV) function to control the rates of water and oil production from production wells for reservoir management objectives. Therefore, this study uses a net present value (NPV) function as the output for the proposed system. The NPV is an economic function used to evaluate the economic performance of a production procedure. In the oil industry, economic performance is analyzed by considering the profit of produced oil concerning production costs, including the cost of water injection and purification of produced water. The NPV function can be described by:

$$J_k = \frac{-\sum_{i=1}^{Ninj} r_{wi}(q_{wi,i})k + \sum_{j=1}^{Nprod}[-r_{wp}(q_{wp,j})kr_{op}(q_{op,j})k]}{(1+b)^{\frac{t_k}{dt}}} \tag{1}$$

where $q_{wi,i}$, $q_{wp,j}$, and $q_{op,j}$ are the rates of injected water, produced water, and produced oil, respectively. r_{wi}, r_{wp}, and r_{op} indicate the costs of water injection, water purification, and oil price, respectively. b, dt, tk, *Ninj*, and *Nprod* denote the discount rate, time of sampling, the moment of time, and the number of injection and production wells, respectively.

b. Linear system identification

Most characteristic features of a system can be extracted from the linear time-invariant (LTI) model of the system because such systems can be described using linear differential equations. These features include impulse response, convolution, double, stability, and scaling, and the characteristics of a time-invariant linear system cannot generally be applied to nonlinear systems. However, according to the law of simplicity, the first solution for identifying nonlinear systems is to use a linear model structure to preserve such characteristic features for unknown systems.

Transfer functions are external mathematical representations of the dynamic relationships between the individual inputs and outputs of a system. In a linear time-invariant system, transfer functions can be defined as the ratio of the Laplace transform of the output to the input, which is, in fact, the Laplace transform of the system impulse response of the system (Equation (2)).

$$Y(s) = G(s)U(s) \tag{2}$$

The transfer function of a system is a fractional function where the numerator and denominator are polynomials (Equation (3)).

$$G(s) = \frac{b_m s + \cdots + b_0 s}{a_n s + \cdots + a_0 s} = \frac{N(s)}{D(s)} \tag{3}$$

The roots of $N(s)$ and $D(s)$ are called the zeros and poles of the system, respectively. Poles and zeros are generally complex numbers that describe the dynamics of the system (Ljung et al., 2010).

After determining the relevant input–output pairs for the unknown system, system identification can be carried out based on a proper transfer function model structure that can optimally describe the performance of an oil reservoir. Prior to estimating the model parameters, the model order should be selected. For this purpose, fitness and simplicity are usually used as two main principles to determine an order for the system model. According to the fitness principle, model structure and its order should be chosen so that the sum of squared errors, i.e., $\hat{S} = \sum_{i=1}^{N} \hat{e}_i^2$, is minimized where the prediction error is defined as:

$$\hat{e}_i = y_t - \hat{y}_t \tag{4}$$

$$\hat{y}_t = u_t^T \hat{\theta} \tag{5}$$

The fitness principle seeks to create a relationship between the gathered system input and output variables based on the minimum error. The principle of simplicity ensures that in choosing between two optional model orders n_1 and n_2 ($n_2 < n_1$), the simpler model structure should be selected if it is sufficiently accurate because working with lower-order models is far simpler than high-order models. The final choice for the order of the system is highly dependent on the final application objective of the system modeling. For example, if the objective is to find faults in the system, a highly accurate model is required. However, if the objective is to design a controller for the system or to perform an optimization, a lower-order model will suffice. Accordingly, trial and error can obtain the proper order for an oil reservoir under the waterflooding process.

2.2. Uncertainty modeling

Uncertainty is an inevitable part of modeling real-world control systems. Generally, uncertainty falls into two categories:

- Disturbance signals;
- Dynamic deviations;

Disturbance signals involve a disturbance in both input and output of the unknown system deviations in sensors and operators. In experimental conditions, dynamic deviations contribute to the difference between the mathematical and the actual dynamic models. A mathematical model of practical systems is, in fact, always an estimate of the actual dynamic system. Common sources of these differences include dynamics not correctly modeled (usually in high frequencies), neglected nonlinear dynamics in modeling, effects of deliberate model order reduction in order, and changes in system parameters due to factors such as environmental changes and worn-out equipment. Such errors in modeling can affect the stability and performance of the system. This section will investigate how to model uncertainty in system identification to be later used in analyzing the robustness of the identified model in applications such as the design of controllers. Equation (4) describes a linear continuous-time system (Mao et al., 1998).

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{6}$$

$$y(t) = C(t)x(t) + D(t)u(t)$$

This system can be modeled as an uncertain linear system using Equation (5) as described below:

$$\dot{x}(t) = \big(A + \Delta A(t)\big)x(t) + (B + \Delta B(t))u(t)$$
$$y(t) = \big(C + \Delta C(t)\big)x(t) + \big(D + \Delta(t)\big)u(t) \tag{7}$$

where $x \epsilon R^n$ is the state vector, $y \epsilon R^p$ denotes the measured output, and $u \epsilon R^m$ indicates the system input control vector. A, B, C, and D are constant matrices with adequate dimensions. Uncertainty was modeled using the following structure bound formula below:

$$\Delta A(t) = M_A F(t) N_A \tag{8}$$

where $M_A \epsilon R^{n \times k}$ and $N_A \epsilon R^{l \times n}$ are both known, and $F(t) \epsilon R^{k \times l}$ is unknown with an upper bound of $F(t) * F(t) \le I$; I is the identity matrix (Mao et al., 1998). Similarly, we can have:

$$\Delta B(t) = M_B F(t) N_B$$
$$\Delta C(t) = M_C F(t) N_C \tag{9}$$
$$\Delta D(t) = M_D F(t) N_D$$

The systematic algorithm for modeling uncertainty using a structure bound method can be described as follows:

Calculating the maximum and minimum of all variable elements in the matrices:

$$a_{ijmin} \le a_{ij}(t) \le a_{ijmax}$$
$$b_{ijmin} \le b_{ij}(t) \le b_{ijmax}$$
$$c_{ijmin} \le c_{ij}(t) \le c_{ijmax}$$

Calculating the average of all variable elements in the matrices:

$$a_{ijm} = \frac{a_{ijmin} + a_{ijmax}}{2}$$
$$b_{ijm} = \frac{b_{ijmin} + b_{ijmax}}{2}$$
$$c_{ijm} = \frac{c_{ijmin} + c_{ijmax}}{2}$$

Calculating the maximum and minimum values of elements in $\Delta a_{ij}(t)$, $\Delta b_{ij}(t)$, and $\Delta c_{ij}(t)$:

$$h_{Aij} = \frac{a_{ijmax} - a_{ijmin}}{2}$$
$$h_{Bij} = \frac{b_{ijmax} - b_{ijmin}}{2}$$
$$h_{Cij} = \frac{c_{ijmax} - c_{ijmin}}{2}$$

We place all the constant elements in matrices $A(t)$, $B(t)$, and $C(t)$ along with the average values in matrices A, B, and C, respectively.

We determine the constant matrices in the uncertainty model of $(M_A, N_A, M_B, N_B, M_C, \text{and } N_C)$ according to the indefinite band:

$$\Delta A(t) = M_A F(t) N_A$$
$$\Delta B(t) = M_B F(t) N_B$$
$$\Delta C(t) = M_C F(t) N_C$$
$$-1 \le F(t) \le 1$$

2.3. Reservoir model

The Egg model was used in this study to produce the required input and output data. The egg model was developed as part of Martin Zendolit's doctoral thesis. Later on, Zendolit et al. (2007) published the first paper on this model. The Egg model represents a three-dimensional synthetic reservoir model which contains 100 permeability realizations of an oil reservoir with 8 water injection wells and 4 oil production wells. The Egg model has been used in various studies to simulate a two-phase current (water–oil). Since the model lacks water and gas caps, the initial mechanism production rate is insignificant, and production typically occurs by waterflooding using eight injection wells and four production wells (Figure 1). Approximately 100 permeability scenarios have been carried out to show the geological uncertainty in oil reservoirs. Unfortunately, details on confining the Egg model parameters have not been the same in every research. The differences were mainly due to fluid parameters and network cell sizes.

Additionally, the parameter configurations in these studies have not been well documented, which makes it difficult and almost impossible to reproduce similar numerical results. The configurations for this study were set according to the standard model proposed by Johnson et al. (2014). Table 1 lists the parameters of the standard model.
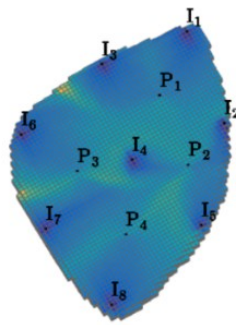

Figure 1
Locations of both injection and production wells used in the reservoir model.

Table 1
Standard model parameters and fluid properties.

Symbol	Variable	Value	SI unit

h	Grid-block height	4	m
$\Delta x, \Delta y$	Grid-block length/width	8	m
Φ	Porosity	0.2	-
C_o	Oil compressibility	1.0×10^{-10}	Pa^{-1}
C_r	Rock compressibility	0	Pa^{-1}
C_w	Water compressibility	1.0×10^{-10}	Pa^{-1}
μ_o	Oil dynamic viscosity	5.0×10^{-3}	Pa s
μ_w	Water dynamic viscosity	1.0×10^{-3}	Pa s
k_{ro}^0	End-point relative permeability, oil	0.8	-
k_{rw}^0	End-point relative permeability, water	0.7	-
n_o	Corey exponent, oil	4	-
n_w	Corey exponent, water	3	-
S_{or}	Residual-oil saturation	0.1	-
S_{wc}	Connate-water saturation	0.2	-
P_c	Capillary pressure	0.0	-
P_R	Initial reservoir pressure (top layer)	40×10^6	Pa
$S_{w,0}$	Initial water saturation	0.1	-
P_{bh}	Production well bottom-hole pressures	39.5×10^6	Pa
r_{well}	Well-bore radius	0.1	m

3. Data production

Modeling in system identification is based on the identification test. If this test is not carried out correctly, even the most complex identification algorithms will not produce a valid system model for the desired procedure application. On the other hand, if the system input and output data result from an appropriate identification test, the least-squares method can be used as the simplest statistical method in system identification which commonly satisfies modeling requirements.

The required data for system identification can include data gathered from actual reservoirs or data procured using an excellent commercial simulator such as Eclipse. The reservoir model structure can be represented by is a typical sample of a MIMO system. This study employed a basic system identification strategy for MIMO systems. In other words, to identify the dynamic model relationship between each system input and output pair at any moment in time, an excitation signal was applied to the desired system input, while the other system inputs remained fixed. Using a proper excitation signal can enhance the performance quality of system identification. In other words, using rich input signal excitation data increases the accuracy of the procedure identified model. Therefore, the unknown system must be simulated so that all its dynamic modes in all the frequency ranges are excited. For this purpose, a persistently exciting signal must be applied to the input of the unknown system. Therefore, a pseudorandom binary signal (PRBS) is selected in this study as the excitation signal that covers a vast spectrum of frequencies through trial and error. However, in actual-world circumstances, such a random signal cannot be applied as the input to an oil reservoir because the control valves cannot perform such high-frequency commands.

Consequently, a suitable excitation signal with a limited frequency range must be used to satisfy the operational constraints for the proposed unknown system oil reservoir. This signal is then placed on top of the optimum input signal in the steady-state reservoir conditions to improve economic performance. The optimum input can be obtained using a model-based optimization procedure. Figures 2 and 3 show the water injection rate in injection well No. 1 and the amount of produced water and oil from the production wells, while inputs from the other injection wells are kept fixed in their steady-state conditions. As discussed in the previous sections, the desired output or value of the NPV secondary

function corresponding to the applied input must be calculated to build the surrogate model. The NPV function was used as the secondary function in this study.
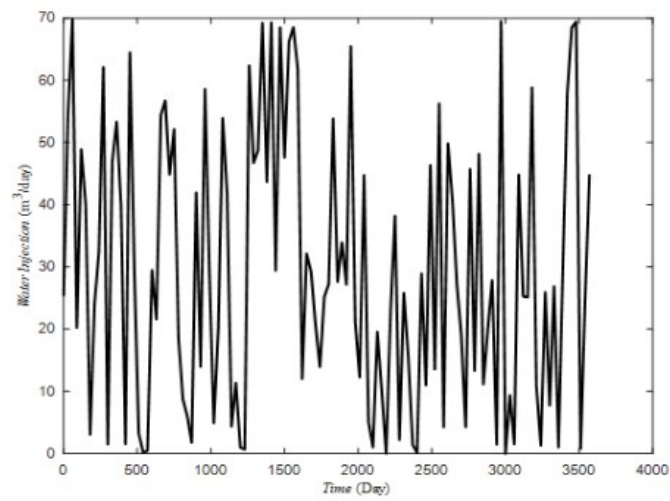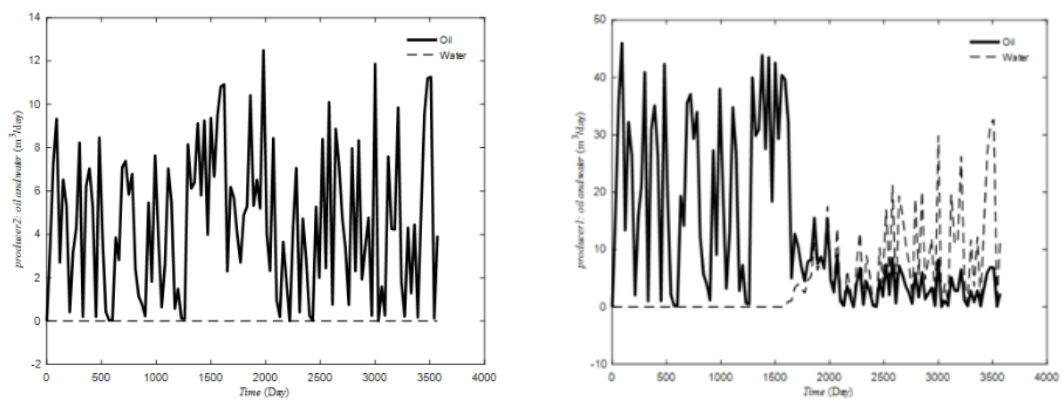

Figure 2
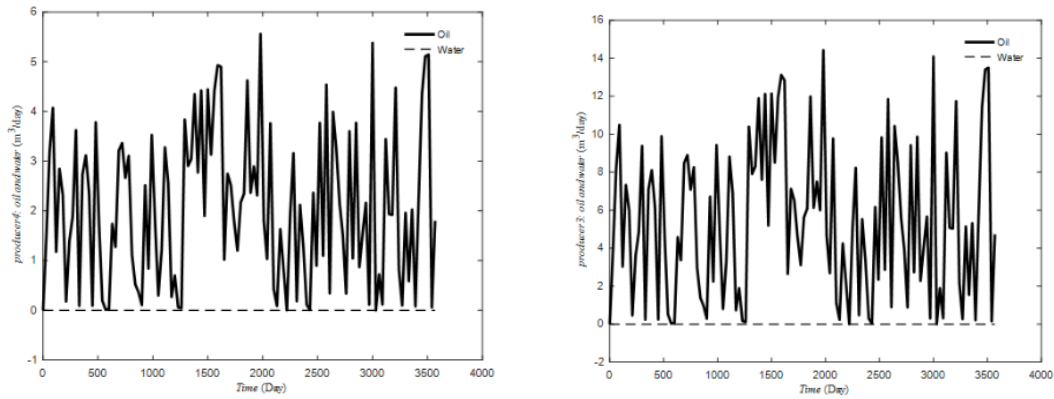Water injection profile for injection well No. 1.

Figure 3
Profiles for produced water and oil.

Table 2 shows the values for these NPV function parameters. Without hindering the totality of the subject and for convenience, uncertainty was modeled by simultaneously applying the desired input for five different instances of permeability. After using injection and production data for the first permeability case in Equation (1), the desired data for identifying the surrogate model is obtained. Figure 4 shows this concept. Input and output are simultaneously recorded according to Figures 4 and 5.

Table 1
The NPV function parameters.

Parameter	Value	Unit
b	0	-
r_{op}	80	$\$/m^3$
r_{wp}	10	$\$/m^3$
r_{wi}	5	$\$/m^3$

$$J_k = \frac{-\sum_{i=1}^{Ninj} r_{wi}\left(q_{wi,i}\right)_k + \sum_{j=1}^{Nprod}\left[-r_{wp}\left(q_{wp,j}\right)_k + r_o q_{op,j}\right]}{(1+b)^{\frac{t_k}{\tau}}}$$

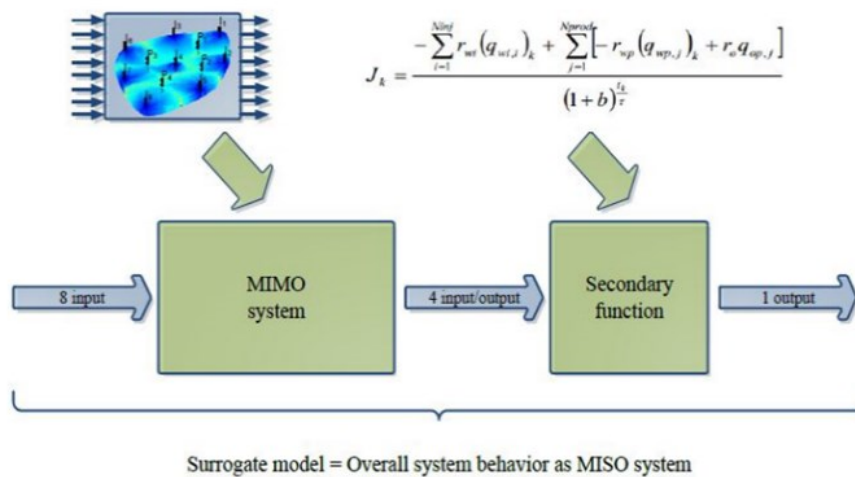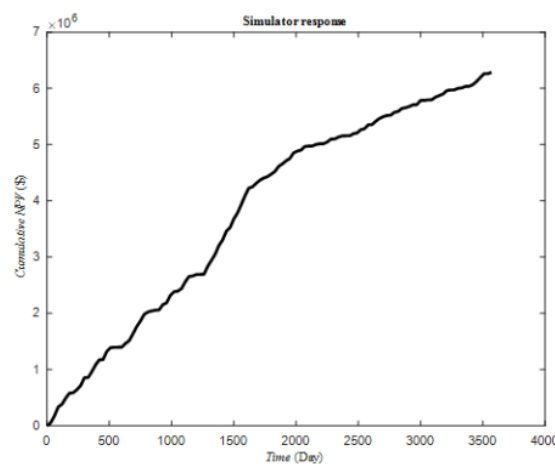Surrogate model = Overall system behavior as MISO system

Figure 4
Block diagram for calculating the desired output (Zanbouri and Salahshoor, 2018).

Figure 5
The NPV cumulative profile calculated by simulating the oil reservoir using the MRST simulator.

4. Results

The MSRT simulator was used to implement the proposed algorithms (Lie et al., 2014). This simulator software provides facilities for controlling each well by altering injection rate and bottom-hole pressure. Since the ultimate objective of modeling oil reservoirs under the waterflooding process is to control the reservoir output and optimize reservoir performance, water injection rate in injection wells can be considered manipulative variables. In addition, bottom-hole pressure has been assumed as fixed, following operational standards. Furthermore, the output of the procedure configuration identified in Figure 4 was chosen as the cumulative NPV value for the relative oil reservoir .

This section investigates the identification of the surrogate model using a system identification approach. Numerous models that have been obtained from particular varying permeability scenarios were then obtained to model uncertainty. For this purpose, the system identification approach was repeated for different permeability scenarios. The identified model is then evaluated in regards to validity and prediction efficiency. Finally, bound structure modeling of uncertainty is discussed in detail.

4.1. Model estimation

Transfer functions can also be used to describe the frequency response of a system, which represents how a system responds to signals with different frequencies. Figure 6 shows the general structure of the surrogate model that needs to be identified.
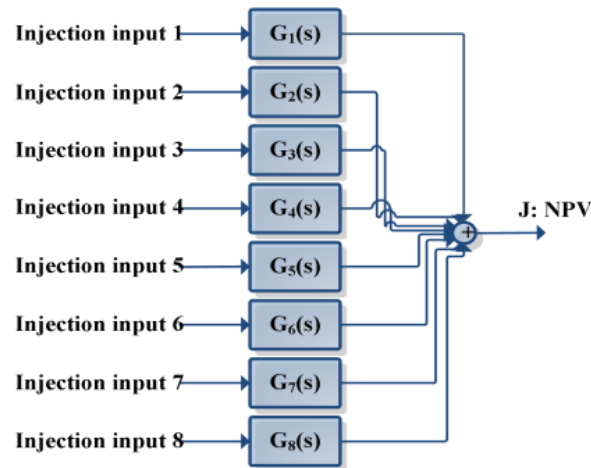
Figure 6
The general structure of the MISO surrogate model.

The transfer function between each input–output pair in Figure 4 can now be obtained using the system identification toolbox in MATLAB software. For this purpose, an excitation signal was applied to one input at a time, while the remaining inputs were kept fixed in their steady-state conditions. Table 3 shows the identified MISO model structure for the first permeability scenario in terms of transfer functions. The obtained consistency terms in Table 3 represent that the identified transfer function is entirely accurate. This measure indicates the accuracy of each individual identified frequency response model, which is confirmed by comparing the identified frequency response plot with the actual frequency response data in terms of mean squared error (MSE), being automatically calculated by MATLAB for validation purposes.

Table 3
The identified MISO surrogate model for permeability scenario of well No. 1.

No	Model	Consistency
1	$G_1(s) = \dfrac{89.14}{s + 0.000204}$	97.08%
2	$G_2(s) = \dfrac{88.36}{s + 0.0004144}$	96.01%
3	$G_3(s) = \dfrac{83.37}{s + 0.0001411}$	98.06%
4	$G_4(s) = \dfrac{85.54}{s + 0.00013}$	96.32%
5	$G_5(s) = \dfrac{92.31}{s + 0.0002211}$	95.11%
6	$G_6(s) = \dfrac{83.76}{s + 0.0001017}$	98.13%
7	$G_7(s) = \dfrac{84.28}{s + 0.0001027}$	97.79%
8	$G_8(s) = \dfrac{89.1}{s + 0.0001637}$	95.85%

4.2. Evaluating the identified model

Certain assumptions have been made during modeling using the system identification approach. One of these assumptions is that the system is linear and may not be valid or right or wrong under different conditions. For this reason, the obtained identified model must be evaluated. If evaluation results prove the model validity, the model can be used; otherwise, the entire identification procedure must be

inspected to find the root of the problem. Both simulation and statistical analysis techniques have been used to evaluate the proposed model in this study.

A new input was applied to the identified model using simulation, and the real and predicted outputs were compared. The simulation approach works on a graphical basis, where one can look at a graph to observe the difference between predicted and accurate outputs and decide on the reliability of the model. However, the analysis can also be carried out quantitatively. For this purpose, the sum of squared errors for different models is obtained, and the identified model with the slightest error is selected as the desired model.

One statistical analysis method used in this paper to validate the identified model is the correlation test. Ideally, prediction errors of a dynamic model must not depend on the input or the remnants of the previous stages. Since the number of samples in this study is limited, the excellent correlation value for nonzero points is not zero, demanding further investigation on whether these nonzero values are due to the limited number of samples or a result of the correlation with remnant errors and whether the identification process is complete. It is only required to calculate the autocorrelation and the cross-correlation functions using the following equations to find the answer to these questions. If the values of these functions reside in the optimal region, the identified model is valid.

$$R_e(\tau) = \frac{1}{N}\sum_{t=1}^{N} e(t)e(t-\tau) \quad \tau = 1,2,\dots \tag{11}$$

$$R_{u\varepsilon}(\tau) = \frac{1}{N}\sum_{t=1}^{N} u(t)\varepsilon(t-\tau) \quad \tau = 1,2,\dots \tag{12}$$

The identified model for the first permeability scenario using methods is shown in Figures 7 and 8. Figure 8 consists of an auto-correlation function plot (the first from the left) and eight cross-correlation function plots. The same approach can be used for other scenarios.
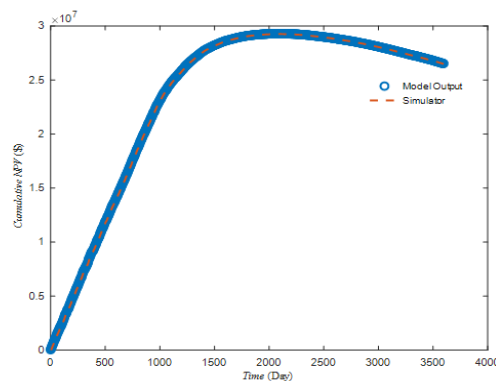

Figure 7

Comparison of the actual output with the model output for the first realization of the unseen input.
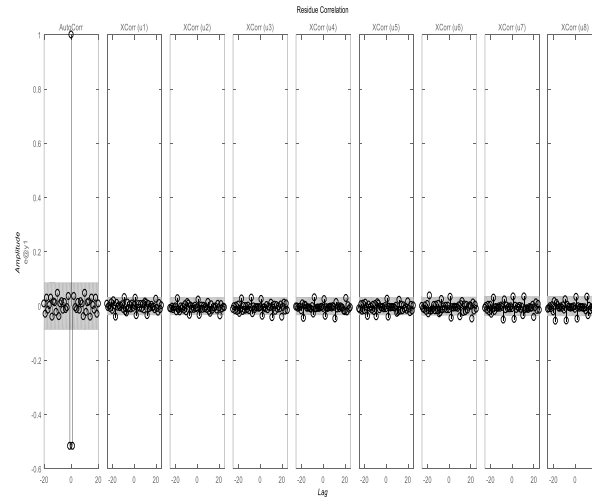
Figure 8
The diagram of auto-correlation and cross-correlations functions for the identified model.

The evaluation results for the identified models in Figures 7 and 8 show that the system identification has been very efficient, leading to the accurate models listed in Table 3 that can be used in oil prediction, designing controllers design, and other applications.

4.3. Modeling uncertainty using structure-bound approach

As discussed in the previous sections, uncertainty is an inevitable part of modeling oil reservoirs. This section models the geological uncertainty of reservoirs using a structure-bound approach. One of the advantages of this approach is that it can transform the robust control problem into a linear matrix inequality (LMI) and can eliminate the effects of uncertainty on NPV control.

According to the explanations given in Section 2.2 for modeling uncertainty, the range of changes in variable values in the oil reservoir model matrix must be determined. For this purpose, five permeability scenarios were considered, and the range of the rate of changes in uncertain parameters corresponding to the identified models for each of the scenarios was determined. Tables 4 and 5 present the range of changes in uncertain parameters.

Table 4
The maximum and minimum values for the variable elements in matrix A.

A	Minimum	Maximum
a_{11}	`–3.2730e–04	–2.0447e–04
a_{22}	–4.1440e–04	–3.3420e–04
a_{33}	–1.9563e–04	–1.4115e–04
a_{44}	–1.2996e–04	–6.3110e–05
a_{55}	–2.2114e–04	–1.5980e–04
a_{66}	–1.0165e–04	–3.3299e–05
a_{77}	–1.5363e–04	–1.0268e–04
a_{88}	–2.8491e–04	–1.6368e–04

Table 5
The maximum and minimum values for the variable elements in matrix C.

C	Minimum	Maximum
c_{11}	11.1419	11.4683
c_{12}	11.0448	12.4077
c_{13}	10.4213	10.5943
c_{14}	10.0801	10.6923
c_{15}	10.9128	11.5386
c_{16}	9.7436	10.4702
c_{17}	10.5353	10.9524
c_{18}	11.1378	12.2015

Following the systematic algorithm for modeling uncertainty in Section 2.2, the structure-bound terms can be determined by placing the constant elements and average of variable elements in matrices A and C. By obtaining the indefinite band $\Delta A(t)$ and $\Delta C(t)$, we obtain (Section 2.2):

$$\dot{x}(t) = \big(A + \Delta A(t)\big)x(t) + Bu(t)$$
$$y = \big(C + \Delta C(t)\big)x(t)$$
$$\Delta A(t) = M_A F(t) N_A \tag{13}$$
$$\Delta C(t) = M_C F(t) N_C$$

Table 6
The indefinite band for matrix A.

$\Delta A(t)$	Value	$M_A N_A$
$\Delta A(t)_{11}$	$M_{A1} F(t) N_{A1}$	6.1414e–05
$\Delta A(t)_{22}$	$M_{A2} F(t) N_{A2}$	4.0102e–05
$\Delta A(t)_{33}$	$M_{A3} F(t) N_{A3}$	2.7244e–05
$\Delta A(t)_{44}$	$M_{A4} F(t) N_{A4}$	3.3427e–05
$\Delta A(t)_{55}$	$M_{A5} F(t) N_{A5}$	3.0670e–05
$\Delta A(t)_{66}$	$M_{A6} F(t) N_{A6}$	3.4176e–05
$\Delta A(t)_{77}$	$M_{A7} F(t) N_{A7}$	2.5472e–05
$\Delta A(t)_{88}$	$M_{A8} F(t) N_{A8}$	6.0612e–05

Table 2
The indefinite band for matrix C.

$\Delta C(t)$	Value	$M_C N_C$
$\Delta C(t)_{11}$	$M_C F(t) N_{C1}$	0.1632
$\Delta C(t)_{22}$	$M_C F(t) N_{C2}$	0.6815
$\Delta C(t)_{33}$	$M_C F(t) N_{C3}$	0.0865
$\Delta C(t)_{44}$	$M_C F(t) N_{C4}$	0.3061
$\Delta C(t)_{55}$	$M_C F(t) N_{C5}$	0.3129
$\Delta C(t)_{66}$	$M_C F(t) N_{C6}$	0.3633
$\Delta C(t)_{77}$	$M_C F(t) N_{C7}$	0.2086
$\Delta C(t)_{88}$	$M_C F(t) N_{C8}$	0.5318

One of the advantages of the proposed method in this study compared to a recent study of Zanbouri et al. (2018) regarding the modeling uncertainty issue is that the obtained structure-bound terms can easily

convert the problem of robust control into a linear matrix inequality problem to facilitate implementation of the problem in a robust control context, eliminating the effect of uncertainty.

5. Conclusions

This study proposes a novel strategy to develop a robust surrogate model in MIMO structures. It incorporates uncertainty modeling in terms of state-space structure-bound terms for oil reservoirs under waterflooding process. For this purpose, a frequency response-based system identification model is proposed for modeling waterflooding oil reservoirs leading to the identified MISO model represented in Table 3 for the first permeability scenario. The modeling procedure starts by initially designing an appropriate excitation signal and applying it through an injection well. The production data can be acquired from the economic NPV function, facilitating the generation of the input–output data required during the identification experiment. Having selected a suitable order, the model structure is then selected following the principles of simplicity and fitness, and the model is estimated using the MATLAB system identification toolbox. Finally, simulation and remnant analysis are used to evaluate the validity of the identified model. This procedure is performed for five different permeability scenarios for a set of indefinite models. The evaluation results show that the identified model is adequately valid for control and optimization applications. Following that, an uncertainty border estimation approach is presented for linear identification models to enable the model uncertainty to be expressible by state-space model block configurations in Tables 6 and 7 in terms of matrices A and C. For this purpose, uncertainty modeling is interpreted in the form of identifying various models for different permeability scenarios and determining the range of changes of indefinite parameters based on the identified model. The indefinite bands are then determined using the uncertainty modeling algorithm. Therefore, the identified oil reservoir model, integrated with an economic NPV function, can emulate a model-based optimization prediction scheme for the economic operation and production of an oil reservoir under uncertainties. Meanwhile, the obtained state-space uncertainty block configurations demonstrate that the identified model can be applied to robust control strategies using linear matrix inequalities for reservoir management objectives. The impact of oil reservoir uncertainty on total economic performance can then be compensated.

Nomenclatures

ARX	Autoregressive model with exogenous input
bh	Bottom-hole
C	Compressibility (Pa^{-1})
K_r	Relative permeability
MIMO	Multi-input multi-output
MISO	Multi-input single-output
MRST	MATLAB reservoir simulation toolbox
NPV	Net present value
O	Oil
P	Pressure (Pa)
PE	Persistently exciting
PRBS	Pseudorandom binary signal
r	Rock
R	Reservoir

S	Saturation
SI	System identification
W	Water
μ	Viscosity

References

A Tafti, T., Ershaghi, I., Rezapour, A., and Ortega, A., Injection Scheduling Design for Reduced-order Waterflood Modeling, in Proceeding of, Society of Petroleum Engineers, SPE-165355-MS, 2013.

Abou-Kassem, J., Ertekin, T., and King, G, Basic Applied Reservoir Simulation, Society of Petroleum Engineers, 2001.

Alhuthali, A. H. H., Datta-Gupta, A., Yuen, B. B. W., and Fontanilla, J. P, Optimal Rate Control Under Geologic Uncertainty, in Proceeding of, Society of Petroleum Engineers, SPE-188368-MS, 2008.

Arps, J. J., Analysis of Decline Curves, Transactions of The AIME, Vol. 160, No. 01, p. 228–247, 1945.

Cancelliere, M., Verga, F., and Viberti, D., Benefits and Limitations of Assisted History Matching, in Proceeding of, Society of Petroleum Engineers, SPE-146278-MS, 2011.

Fanchi, J. R., Principles of Applied Reservoir Simulation: Gulf Professional Publishing, 2005.

Forrester, A., and Keane, A., Engineering Design via Surrogate Modelling: A Practical Guide: John Wiley & Sons, 2008.

Hourfar, F., Moshiri, B., Salahshoor, K., Zaare-Mehrjerdi, M., and Pourafshary, P., Adaptive Modeling of Waterflooding Process in Oil Reservoirs, Journal of Petroleum Science and Engineering, Vol. 146, p. 702–713, 2016.

Jansen, J., Fonseca, R., Kahrobaei, S., Siraj, M., Van Essen, G., and Van Den Hof, P., The Egg Model– A Geological Ensemble for Reservoir Simulation, Geoscience Data Journal, Vol. 1, No. 2, p. 192– 195, 2014.

Lie, K.-A., an Introduction to Reservoir Simulation Using MATLAB: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST), SINTEF ICT, May 2014.

Ljung, L., Perspectives on System Identification, Annual Reviews in Control, Vol. 34, No. 1, p. 1–12, 2010.

Mao, X., Koroleva, N., and Rodkina, A., Robust Stability of Uncertain Stochastic Differential Delay Equations, Systems and Control Letters, Vol. 35, No. 5, p. 325–336, 1998

Olominu, O., and Sulaimon, A. A., Application of Time Series Analysis to Predict Reservoir Production Performance, in Proceeding of, Society of Petroleum Engineers, 2014.

Rezapour, A., Ortega, A., and Ershaghi, I., Reservoir Waterflooding System Identification and Model Validation with Injection/Production Rate Fluctuations, in Proceeding of, Society of Petroleum Engineers, SPE-174052- MS, 2015.

Sayyafzadeh, M., Pourafshary, P., Haghighi, M., and Rashidi, F, Application of Transfer Functions to Model Water Injection in Hydrocarbon Reservoir, Journal of Petroleum Science and Engineering, Vol. 78, No. 1, p. 139–148, 2011.

Tavassoli, Z., Carter, J. N., and King, P. R., Errors in History Matching, SPE Journal, Vol. 9, No. 03, p. 352–361, 2004.

Van Den Hof, P. M., Jansen, J.-D., Van Essen, G., and Bosgra, O. H., Model-based Control and Optimization of Large Scale Physical Systems-Challenges in Reservoir Engineering, in Proceeding of, IEEE, p. 1–10, Xlii-Li, 2009.

Zanbouri, H., and Salahshoor, K., Development of Robust Surrogate Model for Economic Performance Prediction of Oil Reservoir Production Under Waterflooding Process, Journal of Petroleum Science and Engineering, Vol. 165, p. 496–504, 2018.

Zandvliet, M., Bosgra, O., Jansen, J., Van Den Hof, P., and Kraaijevanger, J., Bang-Bang Control and Singular Arcs in Reservoir Flooding, Journal of Petroleum Science and Engineering, Vol. 58, No. 1, p. 186–200, 2007.