Journal of Petroleum Exploration and Production Technology (2020) 10:2465–2482
https://doi.org/10.1007/s13202-020-00887-4

ORIGINAL PAPER - PRODUCTION ENGINEERING

Integrated production optimization of an oil field based on constructing of a proxy model using DOE methods

Mahyar Rajabi-Kochi[1] · **Ehsan Khamehchi**[1]

Received: 18 December 2019 / Accepted: 8 April 2020 / Published online: 15 April 2020

Abstract

An integrated model of the PUNQ-S3 reservoir in the North Sea was constructed. Then, 16 parameters with the highest impact on target functions were selected for sensitivity analysis. The sensitivity analysis was based on the two-level Plackett–Burman design of experimental method. Finally, seven variables with the highest impact on the target functions were selected. Net present value, cumulative oil production, and cumulative water production were three target functions. The proxy model was constructed using the three-level Box–Behnken experimental design for each of the three target functions, taking into account the effect of the variable's interactions on each other. Then the compliance and predictivity of the proxy model for each target function were validated according to the decision variables. In the end, multiobjective optimization was conducted with the aim of maximizing net present value and cumulative oil production and minimizing cumulative water production using a parameter called composite desirability.

Keywords Integrated production model · Optimization · Box–Behnken · Design of experiment · Plackett–Burman · Composite desirability

Abbreviations

DOE	Design of experiments
IPR	Inflow performance relationship
NPV	Net present value
PVT	Pressure–volume–temperature
RSM	Response surface methodology
VLP	Vertical lift performance

List of symbols

ε_i	Residual value
ϕ	Porosity
Q_g	Cumulative gas production
Q_o	Cumulative oil production
W_p	Cumulative water production
d_c	Choke size
d_i	Individual desirability
D_i	Composite desirability
h	Aquifer thickness
i	Interest rate
k	Number of factors
k_x	Horizontal permeability
N	Number of experiments
p	Number of predictors
P_i	Cumulative probability of normal residues
P_{in}	Initial reservoir-aquifer pressure
P_{sep}	Separator operating pressure
SS Error	Residual sum of squares
SS Total	Dependent response for every independent input
T	Target value for target function set by user
U	Limit for target function set by user
w_i	Importance of each factor
y	Value of target function at each step
$\widehat{y_i}$	Value of y_i that line predicted
$\overline{y_i}$	Average value of y_i
Y_i	Actual value of target function
$\hat{Y}_i$	Fitted value of target function

Introduction

Increasing the duration of a reservoir production gradually changes the dynamic characteristics of reservoir, and actual production conditions diverge from the default designing conditions considered for well completion. This leads to a lack of compliance of reservoir production potential with the potential of the well production which often could be seen

✉ Ehsan Khamehchi
khamehchi@aut.ac.ir

[1] Department of Petroleum Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran

in the form of wellhead or downhole flowing pressure drop leading to a reduction in production and eventually the plugging of the well (Shahsavari and Khamehchi 2018). One of the main tasks of petroleum engineers, especially production engineers, is to identify the causes of the problems of such wells and to investigate possible solutions to the problem and restore the well to the production loop (Boyun et al. 2007; Mahdiani et al. 2019). In some cases, it is also necessary to maximize the utilization of a designed operating system. The best solution to do this is to simulate the production system and check the performance of the well flow to minimize energy losses and pressure drop in the produced fluid lines. In this way, in fact, it is possible to separate the production potential capacity of each well and the reservoir through the nodal analysis concept and, by comparing them together, identify the limiting factor of production, and then propose a suitable solution for it to be eliminated (Beggs and Howard Dale 1980). Due to the existence of multiphase fluid in oil production wells and the complexity of mathematical relationships and the high computing volume, the use of a computer program is required to simulate and review diverse production optimization scenarios and demonstrate the importance of optimization in industrial projects (Naderi and Khamehchi 2016) (Hamedi and Khamehchi 2012) (Rashidi et al. 2010) (Vasant et al. 2017). Stoisits were among the first who implemented integrated optimization. In their study, the goal was to provide a method to increase the production flow rate, and variables were including the amount of injected gas allocated to each well and the allocation of any production wells to the surface equipment (Stoisits et al. 2001). In recent years, the construction of an integrated model and, consequently, the optimization of particular functions have been greatly increased by applying many constraints. At the same time, Denney optimized the production rate of each well and the amount of gas allocated to lift produced fluid in two wells using a genetic algorithm (Denney 2003).

Lake et al. (2007) used a quadratic programming method to optimize water injection as a way to increase productivity (Lake et al. 2007). Ehtesham et al. (2011) presented an article using an integrated reservoir and surface equipment model to optimize the performance of five surface separators with two optimized separators (Ehtesham et al. 2011). Codas et al. (2012) also carried out studies on the Urucu oil field in Brazil, succeeded in optimizing the production flow rate as an objective function by using variables of producing gas/oil ratio, water cut and bottom-hole pressure (Codas et al. 2012). Naderi and Khamehchi (2017) also used metaheuristic optimization methods such as genetic and bat algorithms. They chose the net present value (NPV) as the target function and investigated the effect of well location on it (Naderi and Khamehchi 2017).

Marmolejol and Rodriguez used Chambers–Mallows and Stuck algorithm for simulating alpha stable random variables characterizing demand patterns of real electrical systems. The use of Chambers–Mallows–Stuck method for simulating stable random variables provided a new way to generate test systems widely used in power systems research (Marmolejo and Rodriguez 2015).

Experimental designs have been used to the optimization of analytical methods more often than the past. Several benefits could be regarded for these methods like a reduction in the number of tests that need to be done which will lead to reduction in costs and consumption of resources. Meanwhile, these methods permit the construction of mathematical models that allow prediction of the target function value such as Box–Behnken design as well as identifying the most influential factors affecting target function using sensitivity analysis such as a two-level Plackett–Burman method.

In Box–Behnken design, each test can be considered as a combination of a two-level (full or fractional) factorial design with an incomplete block design. In each block, a certain number of factors are put through all combinations for the factorial design, while the other factors are kept at the central values (Box and Behnken 1960). Plackett–Burman designs are highly efficient designs. They use the minimum number of runs to quickly identify the factors with a significant effect on the response. The factors that are identified as important are then investigated more thoroughly in subsequent experiments (Plackett and Burman 1946a, b).

In this paper, an integrated production model was developed based on the available data for the PUNQ-S3 field. Then, using the Plackett–Burman design of experimental method, the sensitivity analysis on the net present value (NPV) target function was performed to determine the most effective factors. By selecting effective parameters and utilizing the Box–Behnken designing as one of the methods for analyzing the response surface, another design was also developed based on three target functions of the net present value, cumulative oil production, and cumulative water production. By conducting experiments and analyzing the results, a high-precision proxy model was generated and then multiobjective optimization by defining a function called composite desirability was performed.

Integrated model description

The field in this study is the PUNQ-S3 field in the North Sea basin. The data file for this field has been prepared from a reservoir study on a real field sample. Six wells are located in different parts of the reservoir, each with a production choke at surface. Figure 1 shows a schematic of an integrated design for wells and surface installations. As shown in Fig. 1, the produced fluid from each well comes together at point A (manifold) and directed toward a separator. In the following, we will

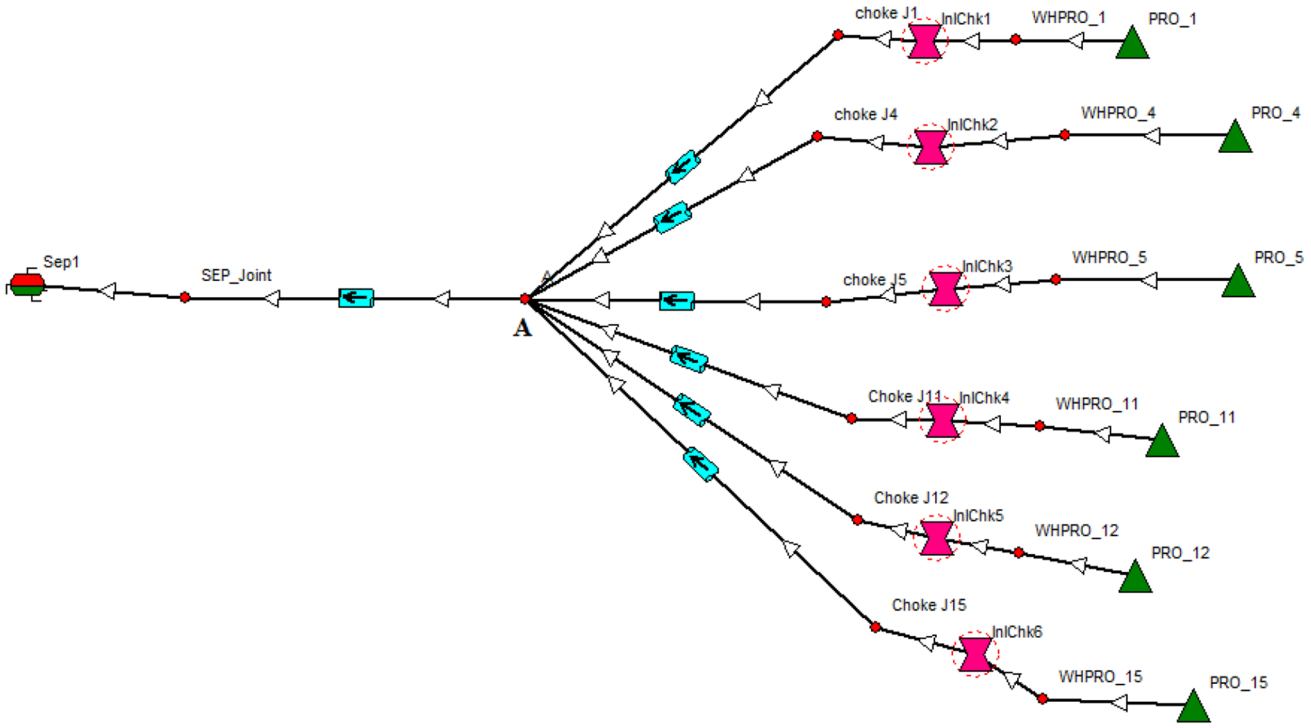

Fig. 1 Schematic diagram of connection of six wells to the manifold and separator

study the construction of an integrated water drive reservoir, well and surface installations model.

Integrated water drive reservoir model

The PUNQ-S3 is known as a small-size reservoir with nineteen grid blocks along the *x*-axis, twenty-eight grid blocks along the *y*-axis, and five grid blocks along the *z*-axis, with a total of 1761 active grid blocks. Six wells have been drilled in different sections. The reservoir is restricted to two slightly strong aquifers from north and west and two faults from the east and south. There are no injection wells, and the geometry of the reservoir is modeled by a corner point geometry method. The initial reservoir pressure is 234 bar and produces heavy oil with density of 912 kg/m^3. The produced gas/oil ratio for the reservoir oil sample at the surface is 74 m^{3}gas/m^{3}oil, which is expected to increase with time and further pressure drops below the bubble point (Odi 2013). The physical reservoir rock and fluid properties are given in Table 1 (Gu and Dean 2005). The parameters of the table are constant and do not change in the different design of experiments.

well model

Six production wells are located in different parts of the reservoir, named Pro1, Pro4, Pro5, Pro11, Pro12, and Pro15, with no injection well in the reservoir.

Table 1 Reservoir parameters used in reservoir simulation assumed constant during experiments

Reservoir fixed parameters	Value
Initial reservoir pressure (bar$_{absolute}$)	234
Reservoir temperature (°F)	212
Total GOR (Sm^3/Sm^3)	74
Oil density (kg/m^3)	912
Gas-specific gravity	0.86
Water density (kg/m^3)	1000
Rock compressibility at reference pressure = 235 (bara)	0.00045

Calculation of produced fluid properties in the well

The produced reservoir fluid properties (oil + gas) should be specified along the well. For this purpose, the properties of the fluid measured in the fluid properties laboratory (PVT Lab) in various pressure steps are presented in Table 2.

In this paper, these values are adjusted to existing empirical relationships and are used to calculate fluid properties along wells and pipelines. The values of each parameter in Table 2 were subjected to PVT matching, and the values of the shift and multiplier parameters were obtained as statistical correction of empirical relationships. The correlations of Glaso (1980), Standing (1947), Lasater (1958), Vasquez-Beggs (1977), Petrosky et al. (1993), and Al-Marhoun et al.

Table 2 Fluid properties measured in the laboratory at different pressure steps to perform PVT matching

Pressure (Bara)	Solution gas oil ratio (Sm^3/Sm^3)	Oil formation volume factor (m^3/Sm^3)	Oil viscosity (cp)	Temperature (°C)	Bubble point pressure (Bara)
40	11.46	1.064	4.338	100	234.46
60	17.89	1.078	3.878	100	234.46
80	24.32	1.092	3.467	100	234.46
100	30.76	1.106	3.1	100	234.46
120	37.19	1.12	2.771	100	234.46
140	43.62	1.134	2.478	100	234.46
150	46.84	1.141	2.343	100	234.46
160	50.05	1.148	2.215	100	234.46
170	53.27	1.155	2.095	100	234.46
180	56.49	1.162	1.981	100	234.46
190	59.7	1.169	1.873	100	234.46
200	62.92	1.176	1.771	100	234.46
210	66.13	1.183	1.674	100	234.46
220	69.35	1.19	1.583	100	234.46
230	72.57	1.197	1.497	100	234.46
234.46	74	1.2	1.46	100	234.46
250	80	1.198	1.541	100	234.46
300	80	1.194	1.787	100	234.46

(1988) were used in order to obtain the best value for bubble point pressure, oil formation volume factor, and solution gas/oil ratio. In addition, the empirical relationships of Beal et al. (1946), Beggs et al. (1975), Petrosky et al. (1995), and Bergman–Sutton (2006) were used to match experimental data obtained from the measurement of viscosity with empirical relationships. The PVT matching results are shown in Table 3. The following three statistical parameters were compared to determine the best consistency with experimental data for each empirical relationship, which can be seen in Table 3 (1) shift, (2) multiplier, (3) standard deviation.

Shifting data is adding a constant k to each member of a data set, where *k* is a real number. In reality, it is lifting the entire distribution of data points and shifting a distance of *k*.

Multiplier (Scaling): Rescaling data is multiplying each member of a data set by a constant k. Rescaling will change the spread of our data as well as the position of data points. What remains unchanged is the shape of the distribution and the relative attributes of our curve.

Standard deviation shows the overall goodness of fit. The lower the standard deviation, the better the fit (Rietz and Carver 1924).

In order to better comparing the results and to choose the best fit between experimental data and empirical relationships for fluid properties, Fig. 2 is plotted. In Fig. 2a, b, in addition to the main measured data in the laboratory, which are characterized by solid squares, six other graphs are drawn, each of which is presented in legends. In Fig. 2c, in addition to the original data shown with solid squares, five other graphs are also observed in different colors. According to Fig. 2, the results obtained from the Beal et al. correlation for viscosity and the Glaso relationship for gas/oil ratio, oil formation volume factor, and bubble point pressure have the best fitting with the original data. The results of Table 3 also confirm this claim. The best fitting according to the statistical parameters is defined as follows:

A) The closest value of the shift parameter to zero.
B) The closest value of multiplier to one.
C) The lowest value of standard deviation.

All of which agree with two above-mentioned relationships.

Calculation of the produced fluid pressure along the well

Each of the six wells in the PUNQ-S3 reservoir is vertical. In order to calculate the amount of pressure drop in the well, PE2 relationship was used. This relationship is quite suitable for vertical wells. In this empirical relationship, a flowchart was used to determine the flow regime presented by Gould et al. (1974). After determining the flow regime, to calculate the pressure drop in each particular flow regime separately, the empirical relationships were utilized (Griffith and Wallis 1961; Hagedorn and Brown 1965; Duns Jr and Ros 1963).

Table 3 The shift, multiplier, and standard deviation parameters derived from PVT matching of laboratory data with empirical relationships for four parameters: (A) bubble point pressure, (B) solution gas oil ratio, (C) oil formation volume factor, (D) oil viscosity

	Glaso	Standing	Lasater	Vazquez-Beggs	Petrosky et al	Al-Marhoun
(A) Bubble point pressure						
Shift	1.07232	1.18252	1.19508	1.0884	1.13432	1.20157
Multiplier	213.798	452.073	474.397	254.158	358.101	485.622
SD	–	–	–	–	–	–
(B) Solution gas oil ratio						
	Glaso	Standing	Lasater	Vazquez-Beggs	Petrosky et al	Al-Marhoun
Shift	0.85951	0.67556	0.65399	0.83185	0.86034	0.58801
Multiplier	8.14954	17.9431	11.6974	6.8239	-32.6991	40.5153
SD	7.48456	7.41953	11.0101	3.17814	13.6398	14.2669
(C) Oil formation volume factor						
	Glaso	Standing	Lasater	Vazquez-Beggs	Petrosky et al	Al-Marhoun
Shift	0.84175	0.77527	0.79646	0.99881	0.78691	0.87445
Multiplier	0.1628	0.21179	0.18658	-0.040714	0.20822	0.084572
Std deviation	0.0050943	0.0022763	0.0040383	0.00077065	0.0066179	0.0020182
(D) Oil viscosity						
	Beal et al	Beggs et al	Al-Marhoun	Egbogah et al	Bergman	–
Shift	1.3479	1.83515	1.27658	0.24499	7.1	–
Multiplier	− 0.76191	− 0.31628	-0.80494	0.39456	− 2.46917	–
SD	0.020605	0.083944	0.037459	0.060903	0.060903	–

Nodal analysis for connecting the well and reservoir

The coupling of the reservoir and well was performed using the nodal analysis method (Ebrahimi and Khamehchi 2015). In each test design, for each set of well and reservoir settings, the relationship between rate and pressure in the upstream and downstream of the node (connecting grid block between reservoir and well) is written; then, from the intersection point of these two figures, flow rates and operational pressure are given. The pressure at the upstream node (reservoir) is calculated by executing the reservoir simulator and calculating the bottom-hole pressure for a specific production rate. At the downstream of the node, the vertical lift performance (VLP) curves using the empirical relationships in Sect. 2.2.2 have been generated for 20 different liquid flow rate values, 10 different water cut values, 10 different gas/oil ratio values, and 10 different surface pressure. Due to the fluid production rate conditions, water cut, gas/oil ratio, and wellhead pressure, the appropriate VLP chart is selected and intersected with the IPR chart.

Surface facilities model

Chokes are considered as the connection point of well and surface pipeline for each of wells. The well is known as upstream and surface pipe as downstream. The VLP graphs are dominant in the upstream of the choke (as nodes).

Using the various relationships provided, downstream pressure in choke can be obtained. In this paper, the utilized relationships are provided by Perkin (1993), which are known in the choke modeling software as the ELF method. The produced fluid from any well is directed to the manifold by a pipeline. The accumulated fluid in the manifold is terminated through a pipeline to the separator. Specifications for pipelines are shown in Table 4, respectively. Considering the isothermal conditions around the surface pipelines (surrounding temperature 60 °F) and the identical type of the pipes (roughness = 0.0006) and the entirely horizontal pipelines, the PE2 relationships indicated in Sect. 2.2 are used to calculate pressure over the surface pipeline.

Methodology

Sensitivity analysis and verification

In this section, the purpose is to determine the parameters with the greatest impact on the target functions in a production system. Three target functions were selected in this study, which are: (1) net present value (NPV), (2) cumulative oil production, (3) cumulative water production. Given the perspective of this issue, by selecting three functions as target functions, the problem becomes a multiobjective optimization problem. The first step in this path is to analyze the sensitivity and select the parameters with the greatest impact on the target functions in the integrated system. Regarding the complexity of the simultaneous sensitivity analysis on the three target functions, a simplistic assumption is used in this problem, so that only sensitivity is performed on the net

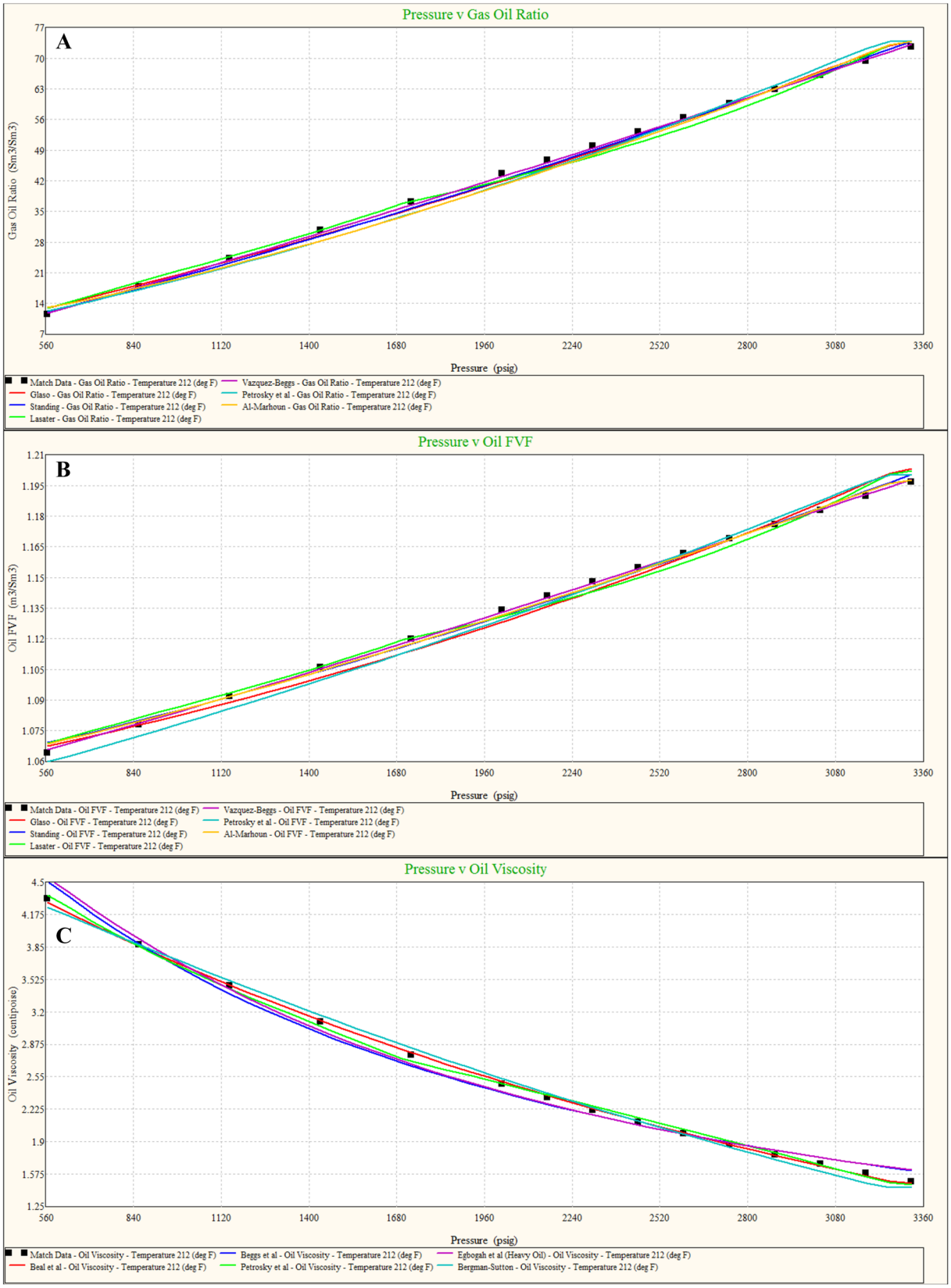

Fig. 2 **a** Compliance of gas/oil ratio with empirical relationships. **b** Matching of the oil formation volume factor with empirical relationship; **c** compliance of the viscosity laboratory data to the empirical relations

Table 4 Length and size of pipelines from each well to manifold and from manifold to separator

Line pipe coordination	Line pipe length (meter)	Line pipe diameter (inch)
Well Pro1-Manifold	768.6	6
Well Pro4-Manifold	178.135	6
Well Pro5-Manifold	432.486	6
Well Pro11-Manifold	326.3	6
Well Pro12-Manifold	1279.83	6
Well Pro15-Manifold	136.9	6
Manifold-Separator	1500	12

present value function. This will not make many errors in the study because the NPV in this study is defined as follows:

$$\mathrm{NPV} = -\mathrm{NPV}_{t=0}(\mathrm{CAPEX}) + \int_0^{T_s} (1+i)^{-t}\left(p_o Q_o + p_g Q_g - c_w W_p - c_{ginj} Q_{ginj} - \mathrm{OPEX}\right)\mathrm{d}t \quad (1)$$

where CAPEX is capital expenditure, OPEX equals operation expenses, c_w is cost per barrel of water produced, c_{ginj} is gas injection cost (ignored in this study), i is interest rate, p_g is gas price per thousand cubic foot, p_o equals oil price a barrel, Q_o is the cumulative amount of oil produced in a time step, Q_g is the cumulative amount of gas produced in a time step, and W_p is the cumulative amount of water produced in a time step.

In Eq. 1, it was assumed that the price per barrel of oil is equivalent to 75 dollars and every 1000 cubic feet of gas would be Sect. 3, and the production of a barrel of water would cost 5 dollars. As can be seen, the NPV itself is expressed as a function of cumulative production of water, oil, and gas, and in other words, it is affected by these three functions.

With a review of previous studies performed in this area, sixteen parameters are selected for sensitivity analysis (Artun 2011) (Avansi 2009) (Shepherd 2009) (Gao 2014). The sixteen selected parameters will include all three water-drive reservoirs, well, and surface systems. In order to perform sensitivity analysis of the net present value function to sixteen selected parameters, a two-level Plackett–Burman test is first designed in which each parameter is set to two upper and lower limit levels. The Plackett–Burman design is a factorial design in two levels and is used to study k factor or variables in two levels. In this method, if the number of experiments is equal to N, the number of factors (k) is obtained from the relation $k = N - 1$ (Plackett and Burman 1946a, b). Given that the number of parameters in this study is sixteen, so the minimum number of experiments will be equivalent to seventeen experiments. However, in order to enhance the statistical population and cover all aspects of the problem, the number of experiments in this study was considered to be 24 tests.

According to the Plackett–Burman method, the order and the procedure for performing the tests are determined. The values of each parameter are set in the designed integrated model, and with the model implementation, the cumulative oil, water, and gas production rate will be obtained. Moreover, using Eq. (1), the net present value will be obtained at each stage of the experiment. These are raw feed results for sensitivity analysis. The method used to perform sensitivity analysis and related parameters is summarized in Table 5. In this study, for the purpose of sensitivity analysis, the goal of fitting a model with high compliance with the NPV target function is based on the sixteen parameters; hence, we start the process with all possible variables. Then, in each step, the variable with the least impact will be eliminated from the model. The process stops when all of the P values of the variables in the model are less than or equal to the alpha to remove. Alpha is a parameter whose value is applied by the user. Asymptotic significance (P value) indicates that the relationship or difference observed in the sample is the result of a chance, and there is no such difference in the society where the sample is selected. In a simpler sense, P value provides us with information about the reality of a result. From technical point of view, the P value is a decreasing index of reliability of a result, and the larger it is, the confidence to the reality of the results reduces. P value shows the probability of an error in accepting the validity of the observed results, although noting that "credibility" means that the observed results represent a

Table 5 Parameters related to sensitivity analysis by two-level Plackett–Burman method

Sensitivity Analysis Responses	Net present value (Million $) Cumulative oil production (BBL) Cumulative water production (BBL)
Confidence level for all intervals	95%
Type of confidence intervals	Tow sided
Box–Cox transformation	No transformation
Stepwise method	Backward elimination
Alpha to remove	0.15

well-known society. Now, "selecting the P value level", the greater the value of which will be rejected as invalid, is contractual and subject to the condition of the problem and is known in our study as Alpha to remove (StatSoft 2006).

Now, by performing sensitivity analysis, the accuracy of the selected parameters as the most effective parameter on the net present value target function is investigated. For this purpose, the R^2, which is a statistical parameter, is used. Since R^2 is a parameter for assessing regression quality, it may create a question in the reader's mind to be used to validate the sensitivity analysis. In response to this question, as mentioned earlier, by analyzing the results of experiments designed using the Plackett–Burman method, the most effective parameters are selected and a mathematical model is created based on them for predicting the target function. In this section, R^2 represents the degree of compliance of existing data to the fitted math model. Here, the details of the mathematical model are discarded, because the goal is only sensitivity analysis and building the proxy model in Sect. 3.2 will be discussed in detail. The R^2 statistical parameter shows the status of the data around the regression diagram. Its value varies from zero to one hundred percent; 0% indicates the excessive dispersion of data around the fitted line and their significant distances from the average value. 100% indicates the data are close to the average fitted line on the data.

Therefore, the closer the R^2 to 100 is, the higher the regression quality. The value of R^2 in this study was calculated from Eq. 2.

$$R^2 = 1 - \frac{\text{SS Error}}{\text{SS Total}} = 1 - \frac{\sum \left(y_i - \widehat{y_i}\right)^2}{\sum \left(y_i - \overline{y_i}\right)^2} \tag{2}$$

where SS Error is residual sum of squares, SS Total is total sum of squares, y_i is dependent response for every independent input, $\widehat{y_i}$ is value of y_i that line predicted, and $\overline{y_i}$ is average value of y_i. Adjusted R^2 is the modified R^2 parameter which is defined as Eq. 3:

$$R^2_{\text{adj}} = 1 - \frac{\left(1 - R^2\right)(N-1)}{N - p - 1} \tag{3}$$

where p is number of predictors and N is total sample size. If the optional predictor parameter is selected that has slight effect on the target function, the R^2 may be affected, but the adjusted R^2 value will remain constant. So it helps you determine whether you need to add a new predictor or not.

Define proxy and its validation

The response surface methodology, or RSM, is a set of statistical techniques and applied mathematics for constructing empirical models (Mahdiani et al. 2015). The goal in such schemes is to optimize the response (output variable), which is influenced by several independent variables (input variables). Theoretically, RSM was developed for modeling of empirical responses and then led to modeling numerical experiments. An important aspect of RSM is the design of experiments, commonly known as DOE. This strategy was originally developed for fitting experimental models, but can also be used for numerical experiments. The DOE's goal is to select the points where the response should be evaluated (Chatterjee et al. 2015). Selection of testing designs can have a great impact on the accuracy of the estimation and the cost of constructing the surface model (Naderi and Khamehchi 2017). In a traditional DOE, screening tests are carried out in the early stages of the process, when there are a large number of potential project variables that may have small effects or no effect on the response (Box and Draper 1987). To avoid this problem, in step 3.1, sensitivity analysis was performed to identify the variables that have a potential impact on the target function. The result was the selection of a number of affecting parameters on the net present value target function.

In order to build a proxy model, the three-level Box–Behnken design of experimental method (Ferreira 2007) and the analysis of the results are used. The Box–Behnken is a three-level experimental factorial design model and a subfield of the surface response method, which is made up of a combination of two-level factorial and incomplete block design in a special case. In this design, the selected parameters in Sect. 3.1 are set at three levels of low, average, and upper limit. By carrying out experiments in the integrated model, the volumetric flow rates of water, oil and gas are calculated, and then using Eq. 1, the net present value is calculated in each step. The proxy model is constructed by analyzing the response surface, taking into account the linear and quadratic effects of each parameter, as well as the effect of the interaction of the parameters with each other for each target function of cumulative oil production, cumulative water production, and the net present value separately.

Validation and evaluation of the proxy model for each of the target functions were done in three ways. The first method is to use the R-squared parameter, which was previously used for sensitivity analysis verification. In the second method of the proxy validation made for each of the target functions, the fitted model was analyzed and interpreted using the residual value-based curves. The issue that needs to be considered first for each regression is the underlying assumption that analysis is performed based on them. This is a very important point; unfortunately, it is often neglected in analysis and undermines the results of modeling. The underlying assumptions for a regression model are as follows:

1. The error term ε has an average of zero.
2. The error term ε has a constant variance.
3. The error term ε is uncorrelated.
4. The error term ε has a normal distribution.

If the fitting pattern is appropriate, the residues must confirm the above assumptions. It means that because we are fitting a linear model, we assume that the relationship really is linear, and that the errors, or residuals, are simply random fluctuations around the true line. We assume that the variability in the response does not increase as the value of the predictor increases. This is the assumption of equal variance. We also assume that the observations are independent of one another. Correlation between sequential observations, or auto-correlation, can be an issue with time series data that is with data with a natural time ordering.

The residue is the difference between the observed value and the fitted value by the model observed from Eq. 4.

$$\varepsilon_i = Y_i - \hat{Y}_i \tag{4}$$

where ε_i is residual value, Y_i is actual value of target function, and $\hat{Y}_i$ is fitted value of target function. In other words, the residue is a measure of the variability of the response variable that is not expressed by the regression model. Residues can be considered a representative of pattern errors, so any deviation from the four regression assumptions about errors might be seen in the residues. A convenient way to see the regression model's efficiency value for fitting data is to plot the residual curve. Chatterjee et al. analyzed the graphs drawn for regression according to Fig. 3. Figure 3a is a favorable situation in which the variance of the errors is constant. In Fig. 3b, the points are dispersed in a funnel form and result in a nonconstant error of variance. In this case, it is not possible to conduct tests and form confidence intervals, nor to estimate the parameters in the least squares method, and it is necessary to estimate the coefficients using another method. Figure 3c, nonlinear diagram, shows that conversion is required, such as logarithmic or second power transformations, on a predictor variable, or a variable to be added to the proxy model (Chatterjee et al. 2015).

In the third method of verifying the proxy model, the normal probability plot of residuals is used in terms of residual values. Since in the calculation of statistics for constructing a proxy model such as P value for regression tests and also for calculating confidence intervals, the assumption of normal-distributed errors is used, so large deviations from normal distribution can affect the accuracy and validity of the results obtained. Additionally, if the errors follow distributions with narrower or wider sequences than normal distributions, the least square fitting may be sensitive to a small change in the data. A simple way to check the assumption of normalization is to draw a normal probability plot of residuals. If we arrange ε_i as an ascending order and draw ε_i against the cumulative probability of normal residues, which is obtained from Eq. 5, the points must be placed on a straight line.

$$P_i = \frac{i - 0.5}{n}; \quad i = 1;2;3;\ldots \tag{5}$$

The presence of one or more large residues in this graph can be a sign of the existence of distant points, which should be examined precisely. The four different modes in these charts are predictable:

1. A graph with no linear state: This behavior is unusual for the graph.
2. The graph has a linear behavior but one or more points away from the diagram: data out of the range
3. The graph shows the slope change: There is an undefined and unrecognized variable in the function.
4. Linear and straight diagram: normal and ideal mode.

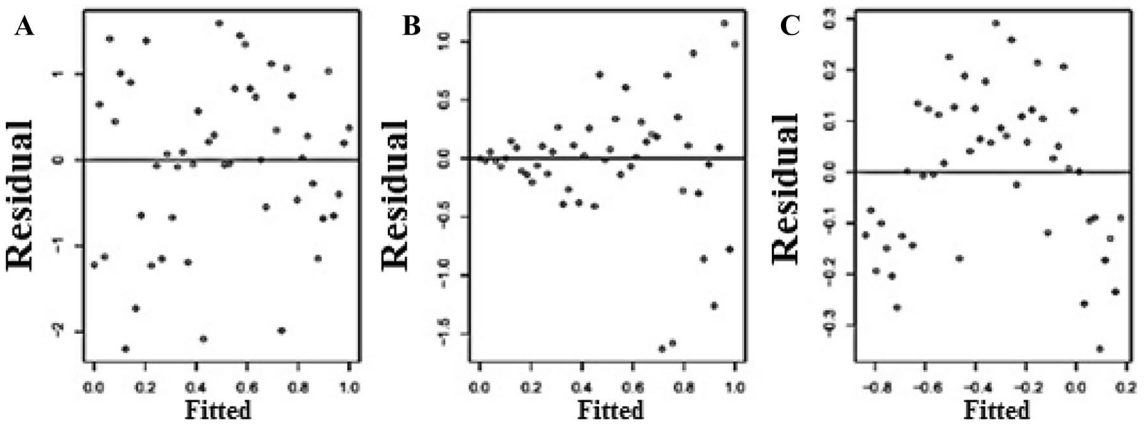

Fig. 3 Interpretation of residual values in a regression and possible states. **a** Ideal mode, **b** nonuniformity mode, **c** nonlinear mode (Chatterjee et al. 2015)

Multiobjective optimization

Obviously, the purpose of this study is to maximize the two functions of net present value and cumulative oil production and minimize cumulative water production. Water production imposes a lot of costs on our manufacturing system, and the aim is to prevent its production.

Proxy models made by analyzing the DOE results of the Box–Behnken experiment are optimized for three target functions, multiobjective optimization, and a parametric definition called desirability. For this purpose, in the first step, for each target function, individual desirability is defined according to the factors affecting them, and for a combination of several target functions, composite desirability is defined. The goal is to maximize the value of desirability. The individual desirability for each target function is defined according to Eq. 6:

$$d_i\left(Y_i(x)\right) = \begin{cases} \left(\frac{Y_i - Y_{\min i}}{T_i - Y_{\min i}}\right)^a & \text{if } Y_{\min i} \le Y_i(x) \le T_i \\ \left(\frac{Y_i - Y_{\max i}}{T_i - Y_{\max i}}\right)^b & \text{if } T_i \le Y_i \le Y_{\max i} \\ 0 & \text{Otherwise} \end{cases} \quad (6)$$

The $Y_{\min i}$ and $Y_{\max i}$ correspond to the desired limits for target function i. In Eq. 6, a and b are user-defined weights which enable the user to determine tighter or wider desirability functions around a target value (T_i) for a response i. Target value for NPV has been set 5,000,000,000 $. This value differs with maximum quantity of target function (i.e., $\mathrm{NPV}_{\max}$). The desirability function method uses RSM approaches to fit polynomials for each response $Y_i(x)$, then substitute the fitted polynomial function in Eq. 6, and substitute the individual desirabilities $d_i(Y_i(x))$ into Eq. 7 at the end. After that, a search-based optimization method is utilized to get x such that $D(x)$ is maximized, because we want $D(x)$ as close to unity as possible. So in this study, we used Hooke–Jeeves method in order to proxy model optimization.

Hence, after defining and obtaining the individual desirability values for each of the target functions proposed in this study, the composite desirability is defined as Eq. 7. The goal of multiobjective optimization is to maximize composite desirability:

$$D_{\mathrm{i}} = \left(\prod \left(d_{\mathrm{i}}^{w_{\mathrm{i}}}\right)\right)^{\frac{1}{w}} \quad (7)$$

As can be seen, the effect of the individual desirability (d_{i}) of each function depends on its importance (w_{i}). In this paper, given that in the sensitivity analysis, the parameters were selected based on the net present value function, and this parameter takes the most importance (value 3). Cumulative oil production parameters and cumulative water production are also of prime importance. Parameters related to optimization are more fully presented in Table 6. The lower or upper limits are set according to the target function and the target values for each function. There is no constraint on the predictor parameters.

Results and discussion

Table 7 indicates sixteen selected parameters with each level up and down. Parameters are selected by searching the literature and from previous studies. The two-level Plackett–Burman method was designed to calculate the net present value based on the sixteen parameters studied, in which Table 8 shows the order and the way of the 24 stages of the experiment. The high level is indicated by the (+) and the lower level with the (−) sign for each parameter. Given the conditions and values stated in Table 8 for each parameter at each stage of experimental design, its value changes in the integrated production model. Then, by setting the values of all sixteen parameters in the integrated model, the model is implemented and first, the cumulative oil production, cumulative gas production, and cumulative water production are obtained in the separator. Then, using Eq. 1 discussed, the net present value (NPV) is calculated. Cumulative water production, cumulative oil production, and cumulative gas production resulting from the implementation of the test at each stage and the resulting net present value are shown in Table 9. The results of Table 9 are the raw feed for sensitivity analysis.

Sensitivity analysis was performed by analyzing the design of a two-level factorial (Plackett–Burman), and taking into account the amount of alpha to remove as 0.15, the sensitivity results were obtained according to Table 10. In this table, seven parameters with the greatest impact on the net present value (NPV) target function are observed.

Table 6 Optimization parameters used in this study, such as the importance, the upper and lower values of each function

Response	Goal	Lower	Target	Upper	Weight	Importance
Cum oil[a]	Maximum	30.1963	54.0221	–	1	1
Cum water[a]	Minimum	–	0.8315	14.7843	1	1
NPV[a]	Maximum	2.2880	4.1165	–	1	3

[a]Cum oil: cumulative oil production/ Cum water: cumulative water production/ NPV: net present value

Table 7 Sixteen selected parameters of the integrated production system with upper and lower levels for each sensitivity test

Parameters	Upper limit	Lower limit
Porosity Multiplier	1	0.8
Net pay thickness	60	40
Kx Multiplier	1	0.8
Kz Multiplier	1	0.8
Wellhead pressure (psi)	400	300
Choke size (in)	1	0.7
Tubing inside diameter (in)	3.85	2.347
Separator pressure (psi)	150	50
Aquifer perm	200	100
Aquifer pressure	300	234
Aquifer compressibility	5×10^{-5}	3.5×10^{-5}
Aquifer porosity	0.4	0.2
Aquifer radius	3500	3000
Aquifer thickness	15	6
Aquifer salt concentration	385	285
Mechanical skin	10	5

As noted earlier, all values of P value in this table are less than the alpha to remove value. The seven parameters that have the greatest impact on NPV include porosity, horizontal permeability, choke size, net pay thickness, separator operating pressure, reservoir-aquifer initial pressure, and aquifer thickness. Meanwhile, reservoir-aquifer initial pressure and net pay thickness had the nearest value to zero which indicates their high importance.

In the next step, the R^2 statistical function is used to validate the sensitivity analysis performed by the Plackett–Burman method, the results of which are shown in Table 11. According to Table 11, the R^2 value is 92.67%, which shows high compliance. But since in this study, the effect of seven independent variables on a target function is being studied, it is better to use the adjusted R^2, because if another predictor is added to the model, which has little effect on the target function, then R^2 changes, but the adjusted R^2 will remain constant. The adjusted R^2 is calculated according to Eq. 3. For sensitivity analysis, 24 samples were selected. Therefore, N in Eq. 3 will be twenty-four. Also, seven predictors were selected with the highest impact on the target function. Therefore, P is equal to seven. So, the adjusted R^2 value is 89.46%, according to Table 11.

In order to construct a proxy model, the seven selected parameters in the sensitivity analysis are set at three levels: low, medium, and high, and the results are shown in Table 12. According to the seven predictors presented in

Table 8 The order and design of the experiment by Plackett–Burman method

	φ	K_x	K_z	P_{wh}	d_{ch}	d_T	P_{sep}	H	K_{aq}	Φ_{aq}	P_{in}	C_{aq}	R_{aq}	h_{aq}	S_{aq}	S
1	− 1	1	− 1	− 1	1	1	− 1	− 1	1	1	− 1	1	− 1	1	1	1
2	− 1	− 1	1	1	− 1	1	− 1	1	1	1	1	1	− 1	− 1	− 1	− 1
3	1	− 1	1	1	1	1	1	− 1	− 1	− 1	− 1	1	− 1	1	− 1	− 1
4	1	− 1	1	− 1	1	1	1	1	1	− 1	− 1	− 1	− 1	1	− 1	1
5	1	1	− 1	1	− 1	1	1	1	1	1	− 1	− 1	− 1	− 1	1	− 1
6	− 1	1	1	− 1	1	− 1	1	1	1	1	1	− 1	− 1	− 1	− 1	1
7	1	− 1	− 1	1	1	− 1	− 1	1	1	− 1	1	− 1	1	1	1	1
8	1	− 1	− 1	1	1	− 1	1	− 1	1	1	1	1	1	− 1	− 1	− 1
9	− 1	1	− 1	1	− 1	− 1	1	1	− 1	− 1	1	1	− 1	1	− 1	1
10	− 1	− 1	1	1	− 1	− 1	1	1	− 1	1	− 1	1	1	1	1	1
11	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1	− 1
12	− 1	1	1	1	1	1	− 1	− 1	− 1	− 1	1	− 1	1	− 1	− 1	1
13	− 1	− 1	− 1	1	− 1	1	− 1	− 1	1	1	− 1	− 1	1	1	− 1	1
14	− 1	− 1	1	− 1	1	− 1	− 1	1	1	− 1	− 1	1	1	− 1	1	− 1
15	− 1	1	− 1	1	1	1	1	1	− 1	− 1	− 1	− 1	1	− 1	1	− 1
16	− 1	− 1	− 1	− 1	1	− 1	1	− 1	− 1	1	1	− 1	− 1	1	1	− 1
17	1	1	1	1	− 1	− 1	− 1	− 1	1	− 1	1	− 1	− 1	1	1	− 1
18	1	1	1	− 1	− 1	− 1	− 1	1	− 1	1	− 1	− 1	1	1	− 1	− 1
19	1	1	− 1	− 1	1	1	− 1	1	− 1	1	1	1	1	1	− 1	− 1
20	− 1	1	1	− 1	− 1	1	1	− 1	1	− 1	1	1	1	1	1	− 1
21	1	− 1	− 1	− 1	− 1	1	− 1	1	− 1	− 1	1	1	− 1	− 1	1	1
22	1	− 1	1	− 1	− 1	1	1	− 1	− 1	1	1	− 1	1	− 1	1	1
23	1	1	− 1	− 1	− 1	− 1	1	− 1	1	− 1	− 1	1	1	− 1	− 1	1
24	1	1	1	1	1	− 1	− 1	− 1	− 1	1	− 1	1	− 1	− 1	1	1

Table 9 The values of target functions for the 24 stages of the Plackett–Burman experimental design (derived from the implementation of the integrated model at each step)

Number	NPV ($)	Cumulative water production (Million bbl)	Cumulative gas production (Million bbl)	Cumulative oil production (Million bbl)
1	2,108,574,500	3.4403	15,402	27.7276
2	4,080,125,450	6.61721	32,168	53.5561
3	2,203,445,850	2.21703	19,452	28.749
4	2,687,527,170	0.210366	17,723	35.1388
5	2,959,281,850	0.14323	23,166	38.54
6	4,097,234,050	6.16399	28,933	53.8834
7	4,590,532,450	5.88541	31,804	60.3273
8	3,017,833,500	17.9736	26,953	40.3579
9	3,701,677,450	4.82821	28,252	48.5475
10	2,234,650,460	0.786308	16,904	29.1716
11	1,677,118,900	3.15592	17,862	21.8575
12	2,761,095,000	15.1299	24,274	36.8523
13	1,870,314,550	2.67349	14,654	24.5296
14	2,645,106,500	1.3202	21,090	34.5125
15	2,627,477,000	0.0626	19,795	34.2454
16	2,642,213,000	19.3019	22,995	35.5965
17	3,160,551,500	23.1479	28,782	42.5326
18	3,004,116,800	1.13144	25,408	39.114
19	5,400,686,300	9.83144	39,782	71.0733
20	2,855,022,500	28.5872	24,127	39.0077
21	3,817,731,800	1.80254	31,254	49.7731
22	2,713,425,500	12.2561	25,632	35.9708
23	1,828,534,800	2.73684	16,068	23.9202
24	1,929,949,950	3.75861	18,856	25.229

Table 10 The values of *P* value calculated for the seven parameters with the greatest effect on the target function

Predictor	Porosity	*x*-Permeability	Choke size	Separator pressure	Net pay thickness	Aquifer initial pressure	Aquifer thickness
P value	0.016	0.148	0.076	0.032	0.0003	0.0002	0.140

Table 11 Validation of the sensitivity analysis performed on the NPV target function using statistical variables

R^2 (%)	92.67
Adjusted R^2 (%)	89.46
R^2 prediction (%)	83.5

Table 12 High, medium, and low levels of variables required to build a proxy model

Parameters	Low level	Intermediate level	High level
Porosity coefficient	0.8	0.9	1
Horizontal permeability coefficient	0.8	0.9	1
Perforated thickness (m)	40	45	50
Choke size (In)	0.7	0.85	1
Separator operating pressure (Psi)	50	100	150
Aquifer thickness (m)	10	15	20
Initial reservoir-aquifer pressure (Bara)	234	267	300

this study, the Box–Behnken designs 57 experiments to build proxy model based on the predictors. By carrying out each stage of the test, cumulative oil, water, and gas production will be recorded. Then, according to Eq. 1, the net present value is calculated. The results of the target functions at each step of the experiment are presented in Table 13.

The proxy model is presented in terms of first-order and quadratic terms, and the effect of interaction of the parameters with each other in the form of Eqs. 8, 9, and 10 for each target function:

$$\begin{aligned}\mathrm{NPV} = {} & 3.3807 + 0.155\phi + 0.1227k_x + 0.6045h \\ & + 0.0194P_{\mathrm{in}} + 0.1102h_{\mathrm{aqu}} + 0.083d_c \\ & - 0.0628P_{\mathrm{sep}} - 0.2391h^2 - 0.8756P_{\mathrm{in}}^2 \\ & - 0.0756P_{\mathrm{sep}}^2 \end{aligned} \quad (8)$$

$$\begin{aligned}\mathrm{Cumoil} = {} & 44.341 + 2.004\phi + 1.642k_x + 7.688h \\ & + 0.258P_{\mathrm{in}} + 1.54h_{\mathrm{aqu}} + 1.095d_c - 0.903P_{\mathrm{sep}} \\ & - 2.893h^2 - 11.618P_{\mathrm{in}}^2 - 0.933P_{\mathrm{sep}}^2 \end{aligned} \quad (9)$$

$$\begin{aligned}\mathrm{Cumwater} = {} & 4.577 + 0.582k_x - 3.649h - 0.059P_{\mathrm{in}} \\ & + 1.234h_{\mathrm{aqu}} - 1.375P_{\mathrm{sep}} + 2.892h^2 - 3.518P_{\mathrm{in}}^2 \end{aligned} \quad (10)$$

where ϕ is porosity, k_x is horizontal permeability (md), h is perforated thickness, h_{aqu} is aquifer thickness, P_{in} is initial reservoir-aquifer pressure, P_{sep} is separator operating pressure, and d_c is choke size.

It should be noted that in the above equations the predictor parameters are encoded. The values of the input parameters must be numerically between the two lower and upper levels in the range $[-1, +1]$. Meanwhile, the values of h and d_c are different for each well, but since the equation is encoded, when, for example, the value + 1 is assigned to each of the two parameters, it means that the upper level of that parameter in each well should be used in the equation. In addition, interaction effect of the two parameters has no significant effect on the target functions and did not appear in the proxy model. The seven parameters are therefore independent of each other and have their own separate effect on target functions.

Validation and evaluation of the above equations were done in three ways due to adapt them to the data used to construct the proxy and to check the prediction power of the testing data. The first method is to use the R-squared parameter, which was previously used for sensitivity analysis verification. Table 14 shows the results of validation using R^2. These results are examined in two aspects.

Firstly, for two cumulative oil production and net present value target functions, R^2 values are above 98%. Compliance between proxy model and test data is high. High adaptation is the result of the correct selection of predictive parameters. In other words, proxy results indicate the accuracy and validity of the sensitivity analysis.

Secondly, for the cumulative water production target function, the R^2 ratio is less than the other two target functions (about 88%). The degree of compliance is acceptable since it should be assumed that the uncertainty about the water-drive parameters is high as the main cause of water production.

In the second verification method, the fitted model was analyzed and interpreted using the residual values-based

Table 13 The results of 57 test designs by the Box–Behnken method to build a proxy model

Order	NPV	Cumulative water	Cumulative gas	Cumulative oil
1	2.41425	14.7843	18,697	32.4277
2	2.78974	19.8993	22,365	37.6286
3	2.31683	0.8315	18,448	30.2086
4	3.66763	8.3028	26,436	48.3978
5	3.44200	3.0292	24,283	45.1239
6	3.11383	6.4066	23,417	41.0081
7	2.86404	21.5509	23,410	38.6875
8	3.67794	6.3411	26,800	48.3900
9	3.50730	2.8765	25,151	45.9497
10	3.23601	2.5433	23,882	42.3611
11	2.56034	1.1682	18,829	33.4626
12	2.70579	1.1005	18,751	35.4005
13	3.34455	4.6911	24,420	43.9300
14	2.49197	0.8328	18,539	32.5402
15	3.08835	4.0524	23,675	40.5011
16	2.40864	11.1429	19,788	32.0666
17	2.78833	14.2248	21,880	37.2508
18	3.38778	6.4424	24,692	44.6122
19	2.78686	1.1283	20,625	36.4083
20	2.33025	0.8915	16,851	30.4554
21	3.27798	6.9057	23,336	43.2334
22	4.02328	3.5831	27,294	52.7908
23	3.24755	3.9293	24,919	42.5658
24	3.05250	3.3460	22,238	40.0335
25	2.58035	13.6208	21,045	34.4709
26	2.28800	6.4639	18,531	30.1963
27	2.68534	13.7455	20,699	35.8930
28	3.78709	2.6343	27,610	49.5657
29	2.55545	1.1098	18,205	33.4184
30	3.36452	5.6534	26,255	44.1869
31	4.36079	6.9539	31,265	57.3569
32	3.66206	4.9970	25,213	48.1521
33	2.51916	0.9275	18,892	32.8949
34	3.70104	3.1079	25,447	48.5365
35	2.00398	1.6907	15,739	26.2029
36	2.64911	6.7906	21,874	34.8992
37	3.53592	4.4234	24,996	46.4407
38	4.11655	4.3725	28,917	54.0221
39	2.90862	0.6950	20,414	38.0114
40	2.48293	8.1756	19,341	32.8771
41	3.15945	2.7426	22,088	41.4253
42	2.49197	0.8328	18,539	32.5402
43	2.12895	2.9937	17,029	27.9044
44	2.69449	0.2048	18,048	35.2183
45	3.25756	3.5010	23,030	42.7463
46	3.77814	4.2396	27,670	49.5511
47	2.78686	1.1283	20,625	36.4083
48	3.47636	3.3736	25,608	45.5521

Table 13 (continued)

Order	NPV	Cumulative water	Cumulative gas	Cumulative oil
49	2.49197	0.8328	18,539	32.5402
50	2.51916	0.9275	18,892	32.8949
51	4.26476	6.2645	29,996	56.0812
52	2.33025	0.8915	16,851	30.4554
53	2.31683	0.8315	18,448	30.2086
54	2.48960	1.1777	18,701	32.5252
55	3.36515	4.6714	25,574	44.1571
56	2.40545	1.1098	18,205	31.4184
57	2.95941	2.1972	22,100	38.7213

Table 14 The values of statistical parameters indicating the conformity of the data with the proxy model and the degree of compliance with the prediction data

R^2 (%)	Adjusted R^2 (%)	Predicted R^2 (%)
Cumulative water production		
91.05	89.56	87.08
Cumulative oil production		
98.86	98.57	98.14
Net present value		
98.86	98.56	98.11

curves. Figure 4 shows the status of the residue diagram for each of the three target functions in this study. Comparison of these diagrams with Fig. 3 makes it clear that Fig. 4a, c are corresponding graphs of net present value and cumulative oil production target function, respectively, which are symmetrical compared to the midline drawn in each graph and most similar to Fig. 3a. Figure. 4b, as seen in the R^2 test, has a slight deviation in the results of the proxy model and actual values, but the difference is not significant (according to the R^2 corresponding to this model). Therefore, an effective parameter on cumulative water seems to be ignored. But the model constructed for NPV and cumulative oil production is highly adapted and the residual-based graph is well distributed.

In the third method of proxy model verification, the normal residue probability chart is used in terms of residual values. Figure 5a–c shows linear behavior, so that points are scattered around the line drawn on the graph. The straight line diagram shows the normal and ideal state of the proxy model. In Fig. 5b, most points are located around the drawn line, but there are three points far from the straight line. According to the results of the Chatterjee et al. study, the presence of data out of the defined range for the proxy is deduced according to the shape of the graph.

If the results of the three validation methods are compared and summarized, it concludes that the proxy model for the two functions of cumulative oil production and net present value has a high accuracy and acceptable conformance with

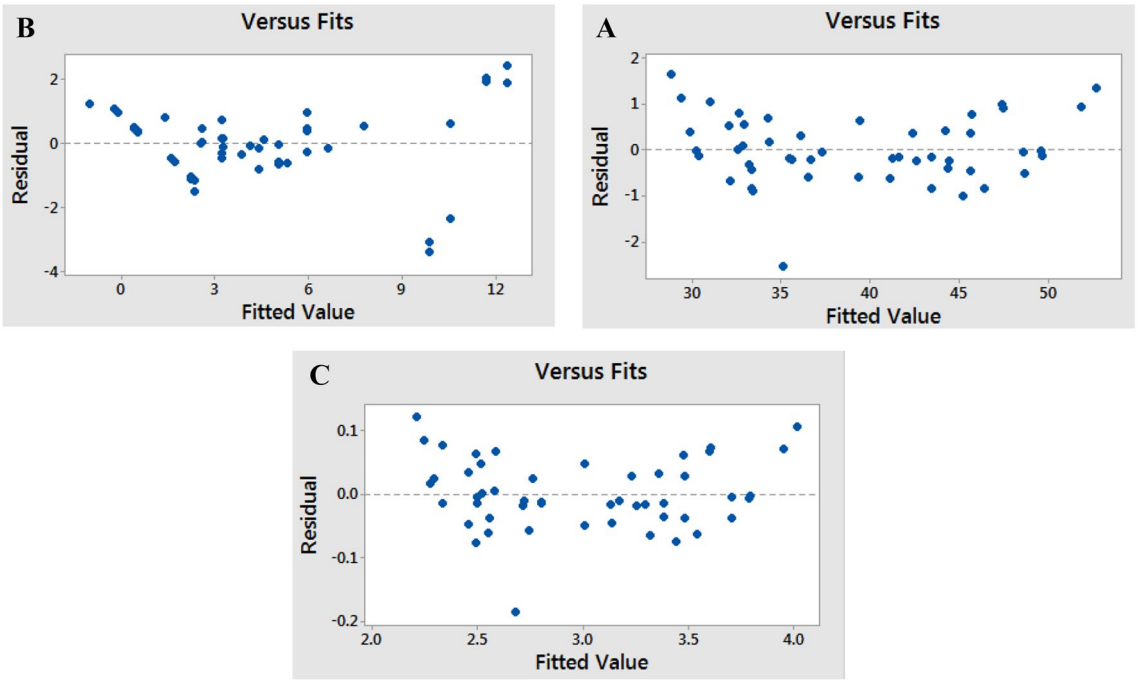

Fig. 4 Residue diagram status for three target functions **a** cumulative oil production, **b** cumulative water production, **c** net present value

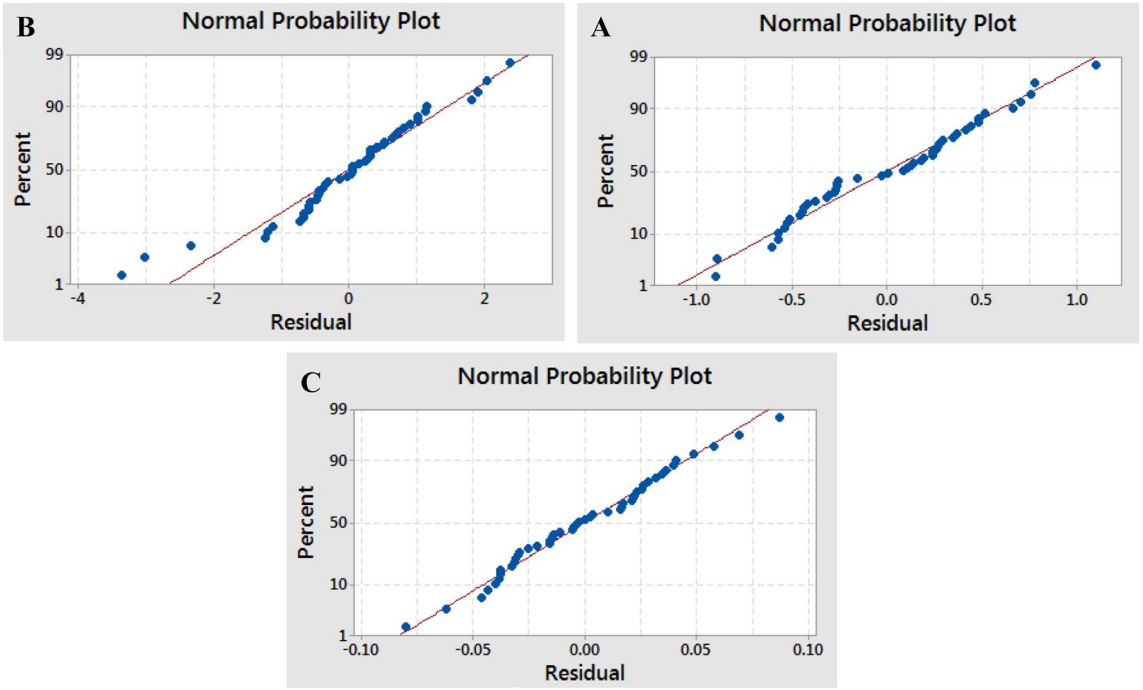

Fig. 5 The diagram of probability of normal residues based on ascending order for **a** cumulative oil production, **b** cumulative water production, **c** net present value

the data from the integrated model, and the seven parameters obtained from the sensitivity analysis have the ability to predict the values of the target functions. However, the proxy model made for the cumulative water production target function has a roughly ninety percent R-squared; according to the validation charts, it seems that by adding a parameter related to the amount of water production and aquifer physical properties to the proxy model, it can improve data compliance. In other words, the effect of a parameter is ignored.

By calculating the individual desirability values followed by the composite desirability value, a multiobjective optimization is performed. The results are in accordance with Table 15.

Net present value has the highest individual desirability among the three target functions. The main reason for this issue should be found in the proxy model for each of the functions. In the net present value mathematical model, three parameters have appeared with a second order: the initial reservoir-aquifer pressure, the separator pressure, and the perforated interval height. Meanwhile, with the change in the perforated interval height in the interval defined between the two upper and lower levels, the net present value function behaved quite descending. However, the other two parameters showed nonlinear behavior and their optimal values were calculated according to Table 15.

The net present value has a higher importance than two other target functions ($w=3$). Therefore, in order to achieve the highest amount of composite desirability, it is reasonable that the individual desirability value of NPV would be set at its highest ($d=1$).

Only two parameters of initial reservoir-aquifer pressure and aquifer thickness showed a nonlinear effect on composite

Table 15 Multiobjective optimization results using a proxy model made for each of the target functions

Parameters	Values	Parameters	Values
Porosity	1	Cumulative oil in optimum situation	53.4933
Horizontal permeability	1	Cumulative water in optimum situation	2.16898
Perforated thickness	1	Net present value in optimum situation	4.11653
Initial reservoir-aquifer pressure	− 0.0303030	Individual desirability of cum oil in optimum situation	0.97781
Aquifer thickness	− 0.494949	Individual desirability of cum water in optimum situation	0.90414
Choke size	1	Individual desirability of cum NPV in optimum situation	1.0000
Separator pressure	1	Composite desirability	0.9757

desirability. The behavior of each of the parameters affecting the target functions in terms of the individual desirability and composite desirability is shown in Fig. 6. The behavior corresponds to the degree of parameters in the proxy model, because reservoir-aquifer pressure and aquifer thickness appeared in a second-order proxy model. However, other parameters have a linear relationship with the target functions, and their interaction with each other has no significant effect on the target function and the optimal values. By optimizing the multiobjective problem, optimal conditions were obtained for control variables. The value of the composite desirability in optimal mode, compared to the nonoptimal state of the current production of the reservoir, is 21% higher.

Lyu et al. (2019) aimed to provide an efficient method to optimize the well configuration in the fractured reservoir. They have proposed a method for the optimization of multilateral well trajectories and determination of the suitable well configuration in the fractured reservoir with the objective function being the maximum profits and minimum costs. The effectiveness and accuracy of the method have been verified through a series of benchmark tests, but they have not considered the water and gas production; meanwhile, they have not conducted the sensitivity analysis to identify the most influential parameters.

In addition, integration of well and surface facility development is crucial to guarantee the selection of the optimum configuration for oil and gas field developments. It is an important step in early planning to minimize investment and reduce the planning period. So Almedallah et al. have presented an integrated model that combines surface and subsurface infrastructure design. The model uses a Monte Carlo Markov chain method to explore the field connectivity, coupled with an optimization routine based on a detailed cost model to determine facility placement (Almedallah and Walsh 2019). But the main differences between this paper and previous works are:

Sensitivity analysis which examine the contribution of each parameter on target function.
Conduction multiobjective optimization by means of composite desirability function evaluation.

Conclusions

In this study, using Plackett–Burman method, a sensitivity analysis was performed to identify the most influential parameter on the net present value of the PUNQ-S3 oil field. Among the 16 parameters, seven effective ones were selected based on the comparison of the *P* value of each predictor with alpha to remove.

The mathematical model was constructed and verified to predict net present value, cumulative oil production, and

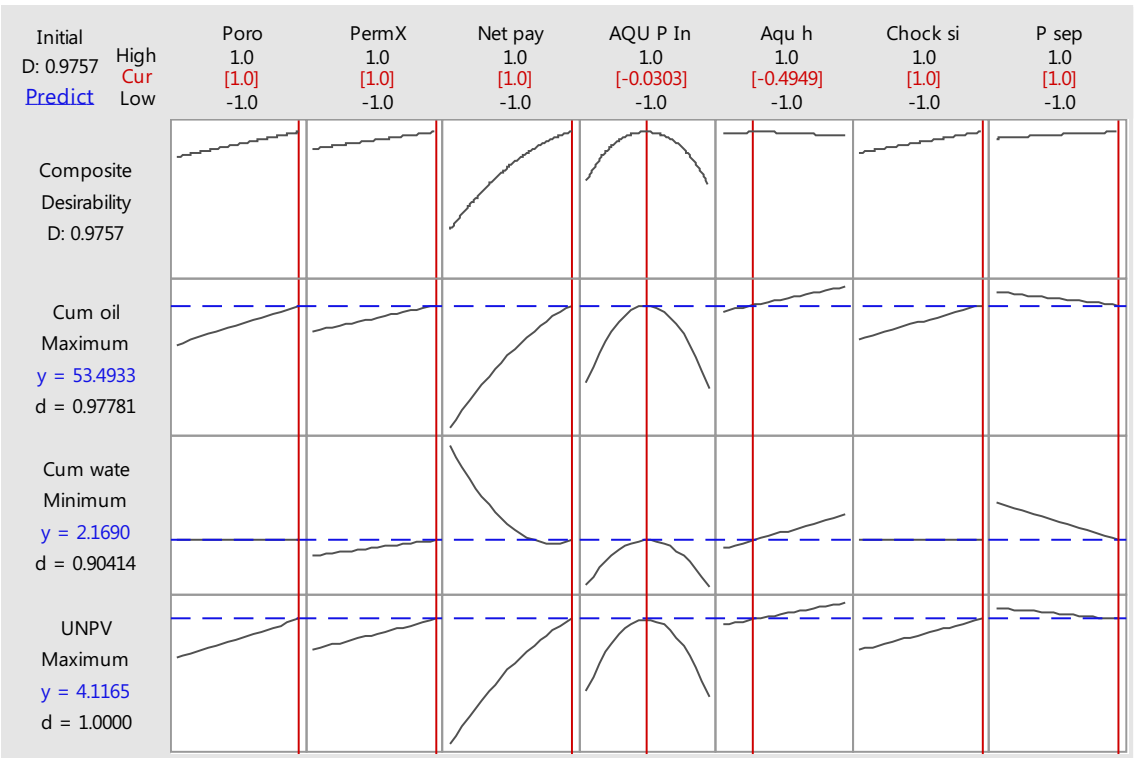

Fig. 6 The impact of each effective seven parameters on target functions and composite desirability

cumulative water production of the field based on the seven selected parameters. The accuracy of constructed model was well over 90%, which indicates great compliance with test data.

By optimizing the multiobjective problem, optimal conditions were obtained for control variables. The value of the composite desirability in optimal mode, compared to the nonoptimal state of the current production of the reservoir, is 21% higher.

References

Al-Marhoun MA (1988) PVT correlations for Middle East crude oils. J Pet Technol 40(05):650–666

Almedallah MK, Walsh SD (2019) Integrated well-path and surface-facility optimization for shallow-water oil and gas field developments. J Petrol Sci Eng 174:859–871

Artun E et al (2011) Development of universal proxy models for screening and optimization of cyclic pressure pulsing in naturally fractured reservoirs. J Nat Gas Sci Eng 3(6):667–686

Avansi GD et al. (2009) Assisted procedures for definition of production strategy and economic evaluation using proxy models. In: 71st European association of geoscientists and engineers conference and exhibition

Beal C (1946) The viscosity of air, water, natural gas, crude oil and its associated gases at oil field temperatures and pressures. Trans AIME 165(01):94–115

Beggs HD (1980) Production optimization. Society of Petroleum Engineers

Beggs HD, Robinson JR (1975) Estimating the viscosity of crude oil systems. J Petrol Technol 27(09):1–140

Box GE, Behnken DW (1960) Some new three level designs for the study of quantitative variables. Technometrics 2(4):455–475

Box GEP, Draper NR (1987) Empirical model-building and response surfaces. Wiley, New York, p 424

Boyun G, William C, Ghalambor A (2007) Petroleum production engineering: a computer-assisted approach

Chatterjee S, Hadi AS (2015) Regression analysis by example. Wiley, Hoboken

Codas A, Campos S, Camponogara E, Gunnerud V, Sunjerga S (2012) Integrated production optimization of oil fields with pressure and routing constraints: the Urucu field.". Comput Chem Eng 46:178–189

Denney D (2003) Optimizing production operations. J Petrol Technol 55(03):59–60

Duns Jr. H, Ros NC (1963) Vertical flow of gas and liquid mixtures in wells. In: Proceedings of the 6th world petroleum congress, Frankfurt am Main, Germany, pp 22–106

Ehtesham MH, Alberto M, Ahmed A-S (2011) Facilities planning using coupled surface and reservoir simulation models. In: SPE/DGS Saudi Arabia section technical symposium and exhibition. Society of Petroleum Engineers

Ferreira SLC et al (2007) Box-Behnken design: an alternative for the optimization of analytical methods. Anal Chim Acta 597(2):179–186

Gao M (2014) Reservoir and surface facilities coupled through partially and fully implicit approaches. Diss

Glaso O (1980) Generalized pressure-volume-temperature correlations. J Petrol Technol 32(05):785–795

Gould TL, Rasin Tek M, Katz DL (1974) Two-phase flow through vertical, inclined, or curved pipe. J Petrol Technol 26(08):915–926

Griffith P, Wallis GB (1961) Two-phase slug flow. J Heat Transf 83:307–318

Gu Y, Oliver S, Dean. (2005) History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter. SPE J 10(02):217–224

Hagedorn AR, Brown KE (1965) Experimental study of pressure gradients occurring during continuous two-phase flow in small diameter vertical conduits. J Pet Technol 4:475–484

Hamedi H, Khamehchi E (2012) A nonlinear approach to gas lift allocation optimization with operational constraints using particle swarm optimization and a penalty function. Pet Sci Technol 30(8):775–785

Lake WL, Ximing L, Edgar TF, Al-Yousef A, Morteza S, Daniel W (2007) Optimization of oil production based on a capacitance model of production and injection rates. In: Hydrocarbon economics and evaluation symposium. Society of Petroleum Engineers

Lasater JA (1958) Bubble point pressure correlation. J Petrol Technol 10(05):65–67

Lyu Z, Song X, Geng L, Li G (2019) Optimization of multilateral well configuration in fractured reservoirs. J Pet Sci Eng 172:1153–1164

Marmolejo JA, Rodriguez R (2015) Fat tail model for simulating test systems in multiperiod unit commitment. Math Probl Eng

Naderi M, Khamehchi E (2016) Nonlinear risk optimization approach to water drive gas reservoir production optimization using DOE and artificial intelligence. J Nat Gas Sci Eng 31:575–584

Naderi M, Khamehchi E (2017) Well placement optimization using metaheuristic bat algorithm. J Petrol Sci Eng 150:348–354

Odi U (2013) Optimal process design for coupled CO_2 sequestration and enhanced gas recovery in carbonate reservoirs. Diss

Perkins THOMASK (1993) Critical and subcritical flow of multiphase mixtures through chokes. SPE Drill Complet 8(04):271–276

Petrosky Jr. GE, Farshad FF (1993) Pressure-volume-temperature correlations for Gulf of Mexico crude oils. In: SPE annual technical conference and exhibition. Society of Petroleum Engineers

Petrosky Jr. GE, Farshad FF (1995) Viscosity correlations for Gulf of Mexico crude oils. In: SPE Production Operations Symposium. Society of Petroleum Engineers

Plackett RL, Burman JP (1946) The design of optimum multifactorial experiments. Biometrika 33(4):305–325

Plackett RL, Burman PJ (1946) The design of optimum multifactorial experiments. Biometrika 33(4):305–325

Rashidi F, Khamehchi E, Rasouli H (2010) Oil field optimization based on gas lift optimization. ESCAPE20 6.

Rietz HL, Carver HC (1924) Handbook of mathematical statistics, vol 3. Houghton Mifflin, Boston

Shepherd M (2009) Factors influencing recovery from oil and gas fields, pp 37–46

Standing MB (1947) A pressure-volume-temperature correlation for mixtures of California oils and gases. Drilling and Production Practice. American Petroleum Institute

StatSoft, Inc. (2006). Electronic statistics textbook. Tulsa: OK: Electronic Statistics Textbook.

Stoisits FR, Kelly DC, Donald MJ, Micheal DM (2001) Petroleum production optimization utilizing adaptive network and genetic algorithm techniques. U.S. Patent 6,236,894 issued May 22

Sutton RP, Bergman DF (2006) Undersaturated oil viscosity correlation for adverse conditions. In: SPE annual technical conference and exhibition. Society of Petroleum Engineers

Vasant P, Litvinchev I, Marmolejo-Saucedo JA (2017) Modeling, Simulation, and Optimization. Springer, New York

Vazquez M, Beggs HD (1977) Correlations for fluid physical property prediction. In: SPE annual fall technical conference and exhibition. Society of Petroleum Engineers.