Two-Stage Scenario Reduction Process for An Efficient Robust Optimization

S.K. Mahjour\textsuperscript{1}, A.A.D.S. Dos Santos\textsuperscript{1}, M.G. Correia\textsuperscript{1}, D.J. Schiozer\textsuperscript{1}

\textsuperscript{1} CEPETRO/FEM – University of Campinas (UNICAMP)

Summary

Probabilistic approaches for optimization objectives need a large ensemble size to consider uncertainties, which is often computationally expensive. Our proposed method includes two scenario reduction (SR) techniques applied to geostatistical realizations and reservoir simulation models to handle geological and dynamic uncertainties. The goal is to select a subset of simulation models to be used in an efficient robust optimization (RO).

The proposed workflow is summarized in the following sections.

1- Generate total geostatistical (TG) realizations representing grid properties using Latin Hypercube (LH) sampling;
2- Select representative geostatistical (RG) realizations from the TG realizations using an integrated statistical technique named Distance-based Clustering with Simple Matching Coefficient (DCSMC). This section is the first stage of SR;
3- Integrate other uncertainties with the RG scenarios to generate total simulation (TS) models using Discrete Latin Hypercube with Geostatistical models (DLHG);
4- Apply data assimilation process to reduce uncertainty and generate total history-matched simulation (THS) models using a filtering indicator named Normalized Quadratic Deviation with Signal (NQDS);
5- Select representative history-matched simulation (RHS) models from the THS models set using a tool based on a metaheuristic optimization algorithm named RMFinder. This section is the second stage of SR;
6- Perform an RO to maximize NPV as the objective function using the selected RHS models;

The novel SR workflow selects the representative scenarios (RG realizations and RHS models) during two steps: (1) RG selection based on static features before the simulation process and, (2) RHS selection based on simulation-based (dynamic) features after the simulation process. The workflow is applied to a fractured synthetic reservoir model named the UNISIM-II-D flow unit-based.

To check the computational-time and efficiency of the methodology, we compare two candidate production strategies based on (1) five RHS models obtained from the two-stage SR process considering DCSMC and RMFinder techniques (workflow A), and (2) five RHS models obtained from one-stage SR process using the RMFinder method (workflow B). In workflow A, the SR process is performed gradually during two steps while in workflow B, the SR process is applied all at once in one step.

The results show that the distribution of simulation outcomes after RO for the representative scenarios and the total scenarios in workflow A are more similar than workflow B. In addition, the robust production strategy obtained from workflow A is preferred to workflow B because it presents higher chances of high NPV value and lower chances of low NPV value.
Introduction

Well-positions in an oil field have a key role in production performance and financial interests. Defining the location of wells is challenging due to rock-fluid interaction, adjacent wells effects, petrophysical variables, and so on (Janiga et al., 2019). Hence, to overcome the problems and gain maximum economic profits, the optimization of well placement is required (Rahim and Li, 2015). In well placement optimization problems, reservoir flow simulation is normally used to integrate geological (static) and dynamic data, and evaluate the objective functions which are normally related to the economic performance of the field (i.e., the NPV). However, reservoir uncertainties strongly affect the accuracy and reliability of reservoir simulation and optimization outcomes. In the following subsections, we explained the required concepts in the robust well-placement optimization under uncertainties.

Reservoir uncertainties arise when there are some constraints in the understanding of the reservoir properties (Hutahaean et al., 2019). Hence, instead of optimizing a deterministic model, robust well placement optimization is performed to optimize the objective functions over a reservoir model set (Badru and Kabir, 2003; van Essen et al., 2009; Yang et al., 2011 and Chang et al., 2015). During robust optimization, the decision-maker looks for an optimal risk-weighted solution that has good performance for all reservoir models under reservoir uncertainty (Yang et al., 2011).

Reservoir uncertainties can be divided into two groups including (1) geological (static) uncertainties related to geological and petrophysical properties, (2) dynamic uncertainties associated with flow properties, production system accessibility, and oil price fluctuation (Santos et al., 2018a). Static uncertainties in well placement optimization are commonly considered by generating numerous geological realizations (static reservoir models) while dynamic uncertainties are taken into account by building multiple simulation models (dynamic reservoir models). Monte Carlo (MC) and Latin Hypercube (LH) sampling methods are standard tools (Santos et al., 2018b) for generating the geological realization. To combine the static uncertainties with dynamic uncertainties and build the simulation model set, Discretized Latin Hypercube Sampling with Geostatistical realizations (DLHG) has been widely used during the last decade (Almeida et al., 2014; Avansi et al., 2015; Bertolini et al., 2015 and Schiozer et al., 2015).

Representative Scenarios (RSs)

The best way to evaluate and quantify the uncertainty is to generate all possible scenarios including geological realizations and simulation models. However, since the flow simulation and objective function optimization for a large number of scenarios is a very time-consuming task, a smaller subset of scenarios as representatives are normally selected to be used in the well placement optimization process (Shirangi and Durlofsky, 2016). The RSs are a small subset of scenarios that show approximately the features of the full ensemble (Schiozer et al., 2019).

In general, two main methodologies for RSs selection were introduced to reduce the number of scenarios for optimization procedure: (1) selecting the representative geological realizations (RGRs) considering static uncertainties (Kang et al., 2019), (2) selecting the representative simulation models (RSMs) considering dynamic uncertainties (Meira et al., 2019). Figure 1 indicates two well placement optimization workflows based on the defined RSs selection techniques. The representative scenario selection steps are shown by the green color in the flow charts.
Figure 1 (a) Representative scenario selection based on geological (static) uncertainties for robust well placement optimization, (b) representative scenario selection based on dynamic uncertainties for robust well placement optimization.

Distance-based Clustering (DBC)

Distance-based clustering (DBC) has been accepted as a simple and applicable technique in several studies for classifying similar scenarios into a number of groups and then select one or more RSs from each group (Caers and Park, 2008; Suzuki and Caers, 2008; Scheidt and Caers, 2009; Haghighat Sefat et al., 2016 and Mahjour et al., 2020a). The principle assumption behind the DBC method is that similar scenarios have similar flow responses, therefore, there is no need to consider all analogous scenarios for the robust optimization purposes (Mahjour et al., 2020b). The outline of DBC includes four stages: (1) measuring a distance, (2) building a matrix, (3) reducing dimensions, and (4) clustering (Kang et al., 2019).

The distance measure is the critical part of DBC to define the dissimilarity of features between two scenarios. Two types of distance are presented: (1) static distance and (2) dynamic distance. Normally, static distance is measured to check the dissimilarity of geological realizations using their geological and petrophysical properties (static properties) such as facies, flow unit, porosity, permeability, etc. A geological realization is made by hundreds to millions of grid cells and each grid cell involves its own static properties (Mahjour et al., 2019). Hence, each group of two geological realizations can be compared grid by grid to define the dissimilarity between them (Mahjour et al., 2020a). Facies and flow units are the most common static properties to be used for calculating the distance between two realizations since the parameters which have important roles in the flow behavior, such as relative permeability and saturation, are determined based on the facies or flow units (Lee et al., 2013; Lee et al., 2016; Kang et al., 2019 and Mahjour et al., 2020b). To obtain the static properties and define the static distances, there is no need to run a flow simulation, while to measure the dynamic distances and obtain dynamic properties (simulation outputs) such as net present value (NPV), oil recovery factor (ORF), and cumulative oil production (Np), as criteria of dynamic dissimilarity (Lee et al., 2011 and Lee et al., 2017), the numerical simulation is essential. Hence, dynamic distance is a tool to check the dissimilarity of simulation models. Furthermore, the process to calculate dynamic distance is significantly slower than the static distance because of the required simulations for determining dynamic parameters (Kang et al., 2019). The static or dynamic distance is defined by a dissimilarity.
distance indicator $S_{ij}$. Static $S_{ij}$ shows the degree of (dis)similarity between two geological realizations $i$ and $j$, and dynamic $S_{ij}$ represents the degree of (dis)similarity between two simulation models $i$ and $j$.

Given the $n$ number of scenarios and the obtained dissimilarity distance indicators between the scenarios, an $n \times n$ dissimilarity matrix is built. The matrix is used as input data for the dimension reduction process (Mahjour et al., 2020b). A single geological realization or simulation model includes numerous grid cells. Therefore, the static and dynamic data are handled in a very high dimensional space which in turn reduces the calculation efficiency (Kang et al., 2019). Dimensionality reduction techniques such as multiple dimensional scaling (MDS) and principal component analysis (PCA) are used as possible solutions to convert high-dimensional data into a 2D Euclidean space in which the features of the data is preserved as much as possible (Jin et al., 2011; Amiran et al., 2015; Chen et al., 2016 and Insuasty et al., 2017). In the 2D Euclidean space, each point shows each scenario and the distance between the two points shows the similarity degree between the two scenarios. Eventually, similar scenarios are classified into the optimal number of groups using different clustering algorithms, and one or more scenarios are selected from each cluster as representative scenarios (Mahjour et al., 2020a). Several clustering algorithms are popular in petroleum engineering: Hierarchical clustering, K-means clustering, K-medoids clustering, and the self-organizing map (SOM) (Mahjour et al., 2020b).

### Mathematical Algorithms

Besides the DBC method, there are some practical solutions to select the RSs based on different mathematical algorithms. Yang et al. (2011) and Yasari et al. (2013) selected a subset of scenarios from a full-set using the traditional ranking of NPV obtained from a base case well position for all scenarios. Sarma et al. (2013) proposed a greedy algorithm named as minimax to closest scenarios to P90, P50, and P10 based on the maximization of the objective-level coverage. Meira et al. (2019) proposed a metaheuristic algorithm using the cross-plots and risk curves of the main well and field objective functions for selecting the RSs. The defined metaheuristic algorithm has been used as an applicable method in some real case studies (Morosov and Schiozer, 2017; Schiozer et al., 2019).

### Objective

In this study, a novel robust well placement optimization workflow is proposed based on a two-stage scenario reduction (SR) applied over geological realizations and history-matched simulation models. This can be a practical workflow to perform an efficient well-placement optimization in terms of covering the overall reservoir uncertainty range, reducing the computational time and defining reliable and profitable production strategy under uncertainty. To check the effectiveness of the robust well placement optimization based on two-stage SR, it is compared with a well placement optimization using one-stage SR considering RSMs selection. The study is implemented on a synthetic fracture benchmark case named UNISIM-II-D flow unit-based (Mahjour et al., 2019).

### Methodology

The main base of the methodology is to perform an efficient robust well-placement optimization using a two-stage SR. This is obtained by first selecting the RGRs from the full set of geological realizations considering static uncertainties and then selecting the RSMs from the entire set of simulation models based on dynamic uncertainties. Next, the robust well placement optimization is applied using the selected RSMs to define a production strategy considering well numbers and well locations. Note that it doesn't matter which methods are used for scenario generation, scenario reduction, and robust well placement optimization procedures. The focus of this study is to check the effectiveness of two-stage scenario reduction for well placement optimization looking at the (1) representativeness of selected scenarios using the obtained objective functions from the initial and optimized production strategies, (2) quality of the generated robust strategy, and (3) number of flow simulation runs and computational costs. Figure 2 presents the proposed workflow for robust well placement optimization under dynamic and static uncertainties. Details of each step are explained in the following subsections.
Figure 2 Proposed robust well placement optimization workflow based on two-stage scenario reduction. The scenario reduction steps are shown by the green color in the flow chart.

**Generate Multiple Geological Realizations (GRs)**

A large number of GRs are generated using geostatistical techniques. This critical step needs a multidisciplinary approach to define all possible geological uncertainties related to the facies/flow unit distributions, petrophysical properties, fracture intensities and directions, and so on (Schiozer et al., 2019). Nowadays, LH sampling is more widely-used because the appropriate variables of random sampling and stratified sampling are incorporated (Santos 2018b, and Mahjour et al 2018b).

**Select Representative Geological Realizations (RGRs)**

The RGRs are used to reduce the high number of generated GRs from the previous step. Static distance-based clustering is a practical approach to identify similar GRs and then select the RGRs. In this study, we recommend using Distance-based Clustering with the Simple Matching Coefficient (DCSMC) presented by Mahjour et al. (2020a). The method is fast and reliable to select the RGRs. Note that to select the number of RGRs, two important issues must be jointly considered: (1) the RGRs number should be sufficiently large to maintain the uncertainty space, (2) the RGRs number should be kept limited to reduce the computation cost (Mahjour et al., 2020b). The steps of the DCSMC method are summarized below:

1. Smoothing 3D facies or flow unit (FU) models to remove short-scale elements;
2. Transferring 3D models into 1 D arrays to build a matrix in which each column shows a facies or FU models and each row shows a categorical facies or flow unit value in a grid cell;
3. Computing the simple matching coefficient to measure the dissimilarity distance between any two models;
4. Building the dissimilarity matrix to be used in Multiple Dimensional Scaling (MDS);
5- Applying the MDS method to transfer the models into a 2D Euclidean space in which each point shows each model;

6- Grouping the models into several clusters using a hierarchical clustering method and defying the number of clusters using the Elbow method;

7- Selecting the RGRs from each cluster using the centroid sampling in a way that the closest model to the center of the cluster is sampled.

**Combine Dynamic Uncertainties with the RGRs to Build Simulation Models**

To build the simulation models, dynamic uncertainties are combined with the obtained RGRs from the previous step. Several sampling methods have been introduced in the literature to combine the uncertainties but we recommend using Discretized Latin Hypercube with Geostatistical realizations (DLHG) (Schiozer et al., 2016). The DLHG method allows combining different uncertainty types and it is easy to apply because it does not need the use of proxy models (Santos 2018b). This method has also shown good results in history matching (Bertolini et al., 2015 and Maschio, and Schiozer, 2016) and well placement optimization (von Hohendorff Filho et al., 2016).

**Apply Data Assimilation**

Data assimilation is an approach to filter the models that are well-matched with the obtained production data from the past performance of the reservoir (Santos et al., 2018a and Formentin et al., 2019). The Normalized Quadratic Deviation with Signal (NQDS) is recommended (Avansi and Schiozer, 2015 and Bertolini et al., 2015) as a matching indicator to be used for each well objective function such as Qo, Qw, Qg, and BHP (Schiozer et al., 2019). Normally, the NQDS values between -1 and +1 are considered as an acceptable misfit (Almeida et al., 2014).

**Select Representative Simulation Models (RSMs)**

In this step, the RSMs are selected from the filtered models to be used in the robust optimization process. To do so, dynamic-distance clustering and mathematical algorithms can be applied using the output dynamic data of the simulator. Hence, an initial production strategy needs to generate the dynamic objective functions. Meira et al. (2019) proposed a method named RMFinder as an optimization-based technique to select the RSMs. The method uses a metaheuristic algorithm considering the cross-plots and risk curves of the main output variables. RMFinder technique is easy to use and fast; therefore this technique is recommended in this study. The number of RSMs should be enough to accelerate the simulation runs and maintain the uncertainty space of the full set.

**Apply Robust Well Placement Optimization**

Robust well-placement optimization is applied to maximize the economic objective functions using the RSMs. von Hohendorff Filho et al. (2016) proposed a stochastic approach to perform well-placement optimization using the Iterative Discrete Latin Hypercube (IDLHC) sampling. The posterior frequency distributions of discrete random elements are considered by this method and the uniform objective functions within discontinuous search spaces are maximized (von Hohendorff Filho et al., 2016). The method decreases the search space in each iteration. Hence, in this study, the IDLHC method is recommended to use for robust well-placement optimization. The steps of the IDLHC method are summarized as below:

1. Discretizing each variable in several levels with a uniform probability distribution;

2. Applying the IDLHC method to generate N sample size;

3. Obtaining the objective function values using simulators;
4. Selecting the objective function above the threshold cut percentage $F$ to choose the superior performing samples of $F \times N$ objective functions for updating the level of frequency histogram for the variables;

5. Fitting the level of frequency histogram with the selected elements from samples to generate posterior frequency distribution of the element levels;

6. Producing a new sample set using the distribution of posterior frequency for each variable level;

7. Repeating steps 2 to 6 until convergence criteria are achieved to be the maximum number of iterations or other criteria;

8. Selecting the highest value of the objective function from the samples as the best solution.

**Validation of Results**

To check the effectiveness of the robust well placement optimization based on two-stage SR (Workflow A), we apply well placement optimization using one-stage SR considering dynamic uncertainties as shown in Figure 1b (Workflow B). We compare and evaluate the obtained results from two workflows considering the (1) representativeness of selected scenarios, (2) quality of generated robust strategy and, (3) number of flow simulation runs and computational costs.

**Representativeness of the Selected Scenarios**

Workflows A and B are compared to check the representativeness of the obtained RSs before and after the optimization process. In this case, we first run an initial production strategy with a certain number of wells to generate the defined objective functions aiming to select the RSs and check their representativeness. It is recommended to use five-spot, seven-spot, or nine-spot well patterns as an initial strategy because the wells can affect all parts of the reservoir. We then use the selected RSs for robust well placement optimization and define an optimized production strategy. Eventually, the optimized strategy is applied over the full set of scenarios and RSs to check if the RSs can still preserve the representativeness.

To evaluate the representativeness of RSs, the Cumulative Distribution Function (CDF) curves of the objective functions calculated from the RSs are compared with those from the full scenarios. The non-parametric Kolmogorov-Smirnov (K-S) test is a practical method to compare two CDF curves (Mahjour et al., 2020b). In this test, the maximum absolute difference $D_{max}$ between the obtained CDFs from the RSs and full-set is measured by Equation 1:

$$D_{max} = \max \left| F_{RSS,n}(x) - F_{Full-set,m}(x) \right|$$  \hspace{1cm} \text{Equation 1}

where, $F_{RSS,n}(x)$ is the CDF of the RSs set with $n$ observations and $F_{Full-set,m}(x)$ is CDF of the full-set set with $m$ observations.

If $D_{max}$ is less than $D_{critical}$, the distribution of the two sets is close and the RSs can be the representative of the full-set. $D_{critical}$ with the 0.05 significance level is defined as Equation 2:

$$D_{critical,0.05} = 1.36 \sqrt{\frac{n + m}{n \times m}}$$  \hspace{1cm} \text{Equation 2}

According to Figure 3, the distribution of objective functions of Sets 1 (GRs) with 2 (RGRs) and Sets 3 (total filtered simulation models) with 4 (RSMs) in Workflow A, and Sets 1 (total filtered simulation models) with 2 (RSMs) in Workflow B should be compared before and after the optimization process to check the representativeness of the selected scenarios.
Figure 3 Scenario reduction steps in each workflow shown by green color.

Quality of Generated Robust Strategy

We compare the quality of generated robust strategies obtained from Workflows A and B in terms of reliability and profitability. During this evaluation, the statistical parameters and curves of calculated NPV from the optimized strategies for the history-matched simulation model sets in both workflows are compared.

Number of Flow Simulation Runs and Computational Costs

To check and compare the computational time of the whole process in each workflow, the computational time for each part of the Workflows A and B including scenario generation, scenario reduction, data assimilation, and well placement optimization are measured.

Application

In this study, a synthetic Benchmark case named the UNISIM-II-D flow unit-based model is used. All geologic and simulation data of the model including well and platform data, deterministic economic variables, and uncertainty elements for generating the scenarios are fully explained by Mahjour et al. (2019). The simulation model includes 95000 grid cells (41000 active cells) and the average grid size is 100 × 100 × 8m. We define five-spot well patterns (28 vertical production wells and 28 vertical injection wells) as an initial production strategy under water flooding as the recovery mechanism (Figure 4). The final simulation time (forecast) is 30 years and some field and well objective functions are considered to analyze the results.
Results

In this section, we check the effectiveness of the proposed two-stage scenario reduction for the robust well placement optimization (Workflows A) and then compare it with the one-stage scenario reduction aiming the robust well placement optimization (Workflow B).

According to Workflow A (Figure 3), we first used the results of the study of Mahjour et al. (2019) to generate 200 GRs (Set 1). They considered different types of geological uncertainties to generate 200 GRs using the LH method considering the UNISIM-II-D flow unit-based model. Mahjour et al. (2020a) then selected 18 RGRs (Set 2) from 200 GRs using a combination of statistical methods including filtering, simple matching, multidimensional scaling, and clustering. In their studies, the elbow method was applied to select the optimal number of representative realizations (18 RGRs). Note that during the RGRs selection, the static distance between GRs was considered without needing the simulation runs. To validate the results, we check the uncertainty space of some field and well objective functions using the K-S test. The objective functions are:

Field: Net Present Value (NPV), cumulative oil production (Np), Oil Recovery Factor (ORF), Oil-In-Place (OIP), Water-In-Place (WIP), cumulative Water Production (Wp), cumulative Water Injection (Wi);

Wells: Well Economic Indicator (WEI) of producers and injectors (Botechia et al., 2013).

We simulate the Sets 1 and 2 in Workflow A under static uncertainties (without considering the dynamic uncertainties) using the initial strategy (vertical five-spot well patterns) to obtain the objective functions. According to Figures 5 to 7, the $D_{max}$ between the field and well objective functions of two sets is less than $D_{critical}$ (0.3) therefore; the uncertainty range between the sets is identical and the first step of the scenario reduction process in Workflow A is properly carried out. During this step, 81 percent of total 200 GRs have been decreased.

Next, Mahjour et al. (2020b) conducted the DLHG sampling to generate 54 simulation models based on 18 RGRs. The number ratio of the RGRs and the simulation models is 1-3. They proved that the defined ratio (1-3) is enough to cover the static and dynamic uncertainty space. Subsequently, they applied the NQDS as a filtering indicator to reduce the uncertainty according to 516 days of production history data for one vertical production well. 16 out of 54 models recorded good match due to the short history period without water production. Hence, 16 filtered simulation models (Set 3) is selected to be used in the second step of the scenario reduction process in the Workflow A. To select the RSMs (Set
4), we use the RMFinder technique based on the multiple risk curves and cross-plots for four objective functions: NPV, Np, Wp, and ORF. RMFinder is an ease-of-use method that can select the RSMs automatically using the simulation-based outputs. Thus, we simulate the models in Set 3 under static and dynamic uncertainties using the initial strategy during 30 years of production forecast and the defined objective functions are considered to be used as input data of RMFinder technique. The required data for running the simulation is fully-explained by Mahjour et al. (2020b). In RMFinder method, the number of the RSMs is defined by the user (Meira et al., 2019). In this study, five RSMs are selected from 16 models, as Set 4. To check the representatives of the selected RSMs, we compare the distribution of defined well and field objective functions between Sets 3 and 4 using the K-S test. According to Figures 8 to 10, $D_{\text{max}}$ between the field and well objectives of two sets is less than $D_{\text{critical}}$ (0.7) therefore; the uncertainty range between the sets is identical and the second step of the scenario reduction process in Workflow A is properly performed. During this step, 69 percent of total 54 history-matched simulation models have been decreased.

Figure 5 $D_{\text{max}}$ between the field objective functions of Set 1 and Set 2 in Workflow A ($D_{\text{critical}}$=0.3). The objective functions are obtained from initial strategy under static uncertainty.

Figure 6 $D_{\text{max}}$ between the WEI of Set 1 and Set 2 for injection wells in Workflow A ($D_{\text{critical}}$=0.3). The objective functions are obtained from initial strategy under static and dynamic uncertainties.
**Figure 7** $D_{\text{max}}$ between the WEI of Set 1 and Set 2 for production wells in Workflow A ($D_{\text{critical}}=0.3$). The objective functions are obtained from initial strategy under static and dynamic uncertainties.

**Figure 8** $D_{\text{max}}$ between the field objective functions of Set 3 and Set 4 in Workflow A ($D_{\text{critical}}=0.7$). The objective functions are obtained from the initial strategy under static and dynamic uncertainties.

**Figure 9** $D_{\text{max}}$ between the WEI of Set 3 and Set 4 for injection wells in Workflow A ($D_{\text{critical}}=0.7$). The objective functions are obtained from the initial strategy under static and dynamic uncertainties.
The objective functions are obtained from the initial strategy under static and dynamic uncertainties.

According to Workflow B, we applied the results of the study of Mahjour et al. (2019) to generate 200 GRs. Mahjour et al. (2020b) then used the DLHG method to build 600 simulation models combining 200 generated GRs and other dynamic uncertainties. The number ratio of the GRs and the simulation models is 1-3. Next, they performed the NQDS method to select 182 history-matched simulation models from 600 simulation models. Subsequently, we run simulations using the initial strategy on 182 history-matched simulation models (Set 5) to obtain the defined objective functions: NPV, Np, Wp, and ORF. The objective functions finally are used in RMFinder as input data to apply scenario reduction process and select the RSMs. We define the same number of RSMs that we have selected in Workflow A to compare and evaluate the results of both workflows. Thus, five RSMs (Set 6) are selected from Set 5. To check the representatives of the selected RSMs, we compare the distribution of defined well and field objective functions between Sets 5 and 6 using the K-S test. According to Figures 11 to 13, $D_{\text{max}}$ between the field and well objectives of two sets is less than $D_{\text{critical}}$ (0.6) except Wells i13 and i14 in Figure 12. Therefore the uncertainty range between the sets is identical and the scenario reduction step in Workflow B is well-performed. During this step, 97 percent of total 182 history-matched simulation models have been reduced.

Although Workflow B reduces the number of scenarios successfully, the value of $D_{\text{max}}$ for the defined objective functions in Workflow B is more than Workflow A. It means that the differences of the CDF curves of objective functions between the RSMs and the full set in workflow B are more than Workflow A. Hence, the representativeness quality of obtained RSMs from Workflow A is higher than Workflow B. This is because the number of selected RSMs in Workflow B (five RSMs) is not enough to cover the uncertainty space of the full set of history-matched simulation models (182 models) for all defined objective functions while the number of selected RSMs in Workflow A (five RSMs) is sufficient to cover the uncertainty range of the full set of history-matched simulation models (16 models). Therefore, if we gradually reduce the number of scenarios in two different steps, the uncertainty space will be effectively maintained by very few numbers of RSMs which can accelerate the optimization process dramatically.
Next, robust well placement optimization is performed using the obtained five RSMs from each workflow. In this study, we perform the IDLHC method to maximize the Expected Monetary Value.
(EMV) as an objective function shown the average NPV of the RSMs. During the robust optimization process, all RSMs are evaluated simultaneously, so that at the end of the process, an optimized strategy is obtained according to the EMV. The IDLHC optimization method is set with 10 iteration with value N=100 evaluations of the objective function, totaling 5000 samples, and F=0.2. The water flooding is considered as the recovery mechanism. Figure 14 shows the calculated EMV values from the IDLHC using five RSMs in Workflow A. The objective function values rise with each iteration. The simulation runs are parallelized for reducing the computational time. The maximum value of EMV is achieved from iteration 10 (US$ 2.45 million).

![Figure 14 Obtained EMV values from IDLHC using the RSMs of Workflow A (EMV Max. = US$ 2.45 million).](image)

Figure 15 represents the wells map of the optimized strategy according to the obtained maximum EMV from Workflow A (Strategy A). The number of wells is 27 including 10 vertical production wells and 7 vertical injection wells.

![Figure 15 Optimized production strategy based on the obtained RSMs from Workflow A (Strategy A).](image)

Figure 16 represents the calculated EMV values from IDLHC using five RSMs in Workflow B. The maximum value of EMV is achieved from iteration 10 (US$ 2.46 million).

![Figure 16](image)
Figure 16 Obtained EMV values from IDLHC using the RSMs of Workflow B (EMV Max. = US$ 2.46 million).

Figure 17 indicates the wells map of the optimized strategy according to the obtained maximum EMV from Workflow B (Strategy B). The number of wells is 27 including 10 vertical production wells and 7 vertical injection wells.

Figure 17 Optimized production strategy based on the obtained RSMs from Workflow B (Strategy B).

To check the representativeness of the RSs after optimization process, we apply the optimized production strategies on the scenario sets of each workflow in a way that Strategy A is used to generate the objective functions in Workflow A and Strategy B is used to achieve the objective functions in Workflow B. Hence, we simulate the Sets 1 and 2 of Workflow A under static uncertainties (without considering the dynamic uncertainties) using the Strategy A and we then apply the K-S test. According to Figures 18 to 20, $D_{\text{max}}$ between the field and well objective functions of two sets is less than $D_{\text{critical}}$ (0.3). Thus, the uncertainty range between the sets is identical, and the representativeness of the scenarios is preserved after optimization process during the first step of the scenario reduction process in Workflow A.
Figure 18 $D_{\text{max}}$ between the field objective functions of Set 1 and Set 2 in Workflow A ($D_{\text{critical}}=0.3$). The objective functions are obtained from Strategy A under static uncertainty.

Figure 19 $D_{\text{max}}$ between the WEI of Set 1 and Set 2 for injection wells in Workflow A ($D_{\text{critical}}=0.3$). The objective functions are obtained from strategy A under static and dynamic uncertainties.

Figure 20 $D_{\text{max}}$ between the WEI of Set 1 and Set 2 for production wells in Workflow A ($D_{\text{critical}}=0.3$). The objective functions are obtained from strategy A under static and dynamic uncertainties.

Next, we simulate the models in Set 3 and Set 4 in Workflow A using Strategy A under static and dynamic uncertainties. According to Figures 21 to 23, $D_{\text{max}}$ between the field and well objectives of
two sets is less than $D_{critical}$ (0.7) therefore; the uncertainty range between the sets is identical and during the second step of the scenario reduction process in Workflow A, the representativeness of the scenarios is maintained after optimization process. From the result, it is shown that two-stage scenario reduction process is able to properly preserve the representativeness of the obtained RSs before and after robust well placement optimization.

**Figure 21** $D_{max}$ between the field objective functions of Set 3 and Set 4 in Workflow A ($D_{critical}$=0.7). The objective functions are obtained from strategy A under static and dynamic uncertainties.

**Figure 22** $D_{max}$ between the WEI of Set 3 and Set 4 for injection wells in Workflow A ($D_{critical}$=0.7). The objective functions are obtained from strategy A under static and dynamic uncertainties.

According to Workflow B, we run simulations using strategy B on the models in Set 5 and Set 6. **Figures 24 to 26** show that $D_{max}$ between some field and well objectives of two sets are more than $D_{critical}$ (0.6). Hence the representativeness of the selected RSs using Workflow B is not well-preserved after the optimization process. This may be due to an excessive reduction in the number of scenarios in one step based on Workflow B so that 97 percent of scenarios were reduced in Set 5.
Figure 23 $D_{\text{max}}$ between the WEI of Set 3 and Set 4 for production wells in Workflow A ($D_{\text{critical}} = 0.7$). The objective functions are obtained from strategy A under static and dynamic uncertainties.

Figure 24 $D_{\text{max}}$ between the field objective functions of Set 5 and Set 6 in Workflow B ($D_{\text{critical}} = 0.6$). The objective functions are obtained from Strategy B under static and dynamic uncertainties.

Figure 25 $D_{\text{max}}$ between the WEI of Set 5 and Set 6 for injection wells in Workflow B ($D_{\text{critical}} = 0.6$). The objective functions are obtained from Strategy B under static and dynamic uncertainties.
As a measure of the quality of the optimized strategies A and B, we apply Strategy A on the models of Set 3 in Workflow A, and Strategy B on the models of Set 5 in Workflow B. The models in Set 3 and Set 5 are the history-matched models in Workflow A and Workflow B, respectively.

Table 1 compares the mean and median of NPV values for the Set 3/Strategy A and Set 5/Strategy B. It represents the NPV mean and median in Set 3/Strategy A are more than Set 5/Strategy B. Hence from the results, it can be said that Strategy A is more profitable than Strategy B based on the defined sets. Furthermore, the NPV mean and median for Set 3/Strategy A, and the NPV mean and median for Set 4 (RSMs in Workflow A)/Strategy A are almost identical which means that five RSMs are sufficient for covering the uncertainty range of the full set in Workflow A. However, the NPV mean and median for Set 5/Strategy B is not close to the NPV mean and median for Set 6 (RSMs in Workflow B)/Strategy B which shows that five RSMs are not enough to support the uncertainty space of the full set in Workflow B during the robust optimization process.

In a cross-validation step, we apply Strategy B on the models of Set 3, and Strategy A on the models of Set 5. Table 2 compares the mean and median of NPV values for the Set 5/Strategy A and Set 3/Strategy B. It represents the NPV mean and median in Set 5/Strategy A are more than Set 3/Strategy B, therefore; according to this analysis, Strategy A is more profitable than Strategy B based on the defined sets. Moreover, the NPV mean and median for Set 3/Strategy B, and the NPV mean and median for Set 4/Strategy B are almost identical which means that five RSMs are sufficient for covering the uncertainty range of the full set in Workflow A even if the production strategy has been changed. However, the NPV mean and median for Set 5/Strategy A is not close to the NPV mean and median for Set 6/Strategy A which shows that five RSMs are not enough to support the uncertainty space of the full set in Workflow B.

Figure 27 compares the obtained CDF curves of NPV values from history-matched model sets with different optimized strategies A and B. The CDF of the Sets with Strategy A (Set 3/Strategy A and Set 5/Strategy A) are at the right side of the CDF of the sets with Strategy B (Set 3/Strategy B and Set 5/Strategy B) which means that most of the models simulated with Strategy A have higher values of NPV than the models simulated with strategy B. Figure 28 shows the box plot of the NPV values of different history-matched model sets and optimized strategies. From the Figure 28, the uncertainty range in Set 3 is lower than Set 5 and the quality of Strategy A is higher than Strategy B. Hence, the comparison work indicates that the quality of the obtained RSMs in Workflow A is better than those from Workflow B for the robust well placement optimization since Strategy A is more reliable and profitable than Strategy B.


<table>
<thead>
<tr>
<th></th>
<th>Total models</th>
<th>NPV mean (Million US$)</th>
<th>NPV median (Million US$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full set in Workflow A</td>
<td>16</td>
<td>2.44</td>
<td>2.48</td>
</tr>
<tr>
<td>(Set 3/Strategy A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSMs in Workflow A</td>
<td>5</td>
<td>2.45</td>
<td>2.46</td>
</tr>
<tr>
<td>(Set 4/Strategy A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full set in Workflow B</td>
<td>182</td>
<td>2.37</td>
<td>2.39</td>
</tr>
<tr>
<td>(Set 5/Strategy B)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSMs in Workflow B</td>
<td>5</td>
<td>2.46</td>
<td>2.62</td>
</tr>
<tr>
<td>(Set 6/Strategy A)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 27** CDFs of obtained NPV values from history-matched model sets with the different optimized strategies A and B.

According to Table 3, the computational time for selecting the representatives in Workflow A is significantly lower than Workflow B (nearly half). Hence, two-stage scenario reduction is preferred for the robust well placement optimization in terms of the cost of computations. The optimization steps in both workflows took around 17 days and 16 hours using five RSMs. Note that the computational cost and the number of simulations for the robust optimization step in Workflow B will be higher if further RSMs are selected to cover the uncertainty range of the full set.
**Table 3** Computational time for each step of robust well placement optimization.

<table>
<thead>
<tr>
<th>Workflow Step</th>
<th>Computational time/ Workflow A</th>
<th>Computational time/ Workflow B</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRs generation</td>
<td>7 hours, 25 minutes</td>
<td>7 hours, 25 minutes</td>
</tr>
<tr>
<td>RGRs selection</td>
<td>5 hours, 25 minutes</td>
<td>------</td>
</tr>
<tr>
<td>DLHG sampling and data assimilation</td>
<td>5 minutes</td>
<td>45 minutes</td>
</tr>
<tr>
<td>RSMs selection</td>
<td>2 hours, 15 minutes</td>
<td>20 hours, 25 minutes</td>
</tr>
<tr>
<td>Total time</td>
<td>15 hours, 10 minutes</td>
<td>28 hours, 35 minutes</td>
</tr>
</tbody>
</table>

**Conclusions**

This paper presented a novel scenario reduction workflow for robust well placement optimization applied to a synthetic fracture benchmark case named UNISIM-II-D flow unit-based. This could gradually and significantly reduce the number of scenarios (geological realizations and simulation models) during two stages. The fundamental basis of the workflow was (1) selecting the RGRs considering static uncertainties and, (2) selecting the RSMs considering dynamic uncertainties. To check the effectiveness of the proposed workflow, it was compared to a one-stage scenario reduction workflow based on RSMs selection (without considering the RGRs selection) for well configuration. From the results of the two-stage scenario reduction workflow, we got conclusions that five RSMs are able to be used in robust well placement optimization under water flooding mechanism while the overall uncertainty range of the reservoir is still preserved. To evaluate the uncertainty space of representative scenarios and the full set, the generated distribution of some field and well objective functions from initial and optimized strategies are considered.

The comparison work also showed that the outcomes of the proposed workflow are superior to the one-stage scenario reduction workflow for the robust optimization problems in terms of preserving overall uncertainty space, reducing computational cost, and defining a reliable and profitable production strategy.

This work can be extended to contain more studies including the evaluation of the presented workflow in different recovery mechanisms (polymer-flooding, WAG, etc.) during the optimization process as well as performing a comparison work between the proposed workflow and one-stage scenario reduction based on RGRs selection (Figure 1a).
Acknowledgements

We gratefully acknowledge the support of PETROBRAS, EPIC (Energy Production Innovation Center), hosted by the University of Campinas and sponsored by FAPESP – São Paulo Research Foundation (grant number 2017/15736-3), Equinor Brazil, ANP (Brazil’s National Agency of Oil, Natural Gas and Biofuels), UNISIM and Energi Simulation. The authors are also grateful to Schlumberger Information Solution for the use of Petrel® and to Computing Modeling Group for the use of Builder CMG®.

References


