Scientific
Research
Publishing

Stochastic Modeling and Assisted History-Matching Using Multiple Techniques of Multi-Phase Flowback from Multi-Fractured Horizontal Tight Oil Wells

Jesse D. Williams-Kovacs*, Christopher R. Clarkson

Department of Geoscience, University of Calgary, Calgary, Canada
Email: *jdwilli@ucalgary.ca

Abstract

In this paper, the methods developed by [1] are used to analyze flowback data, which involves modeling flow both before and after the breakthrough of formation fluids. Despite the versatility of these techniques, achieving an optimal combination of parameters is often difficult with a single deterministic analysis. Because of the uncertainty in key model parameters, this problem is an ideal candidate for uncertainty quantification and advanced assisted history-matching techniques, including Monte Carlo (MC) simulation and genetic algorithms (GAs) amongst others. MC simulation, for example, can be used for both the purpose of assisted history-matching and uncertainty quantification of key fracture parameters. In this work, several techniques are tested including both single-objective (SO) and multi-objective (MO) algorithms for history-matching and uncertainty quantification, using a light tight oil (LTO) field case. The results of this analysis suggest that many different algorithms can be used to achieve similar optimization results, making these viable methods for developing an optimal set of key uncertain fracture parameters. An indication of uncertainty can also be achieved, which assists in understanding the range of parameters which can be used to successfully match the flowback data.

Keywords

Stochastic Modeling, Assisted History-Matching, Quantitative Flowback Analysis, Rate-Transient Analysis

1. Introduction

In recent years, as a result of low gas prices and relatively high oil prices, many producers have turned their attention to LTO reservoirs as a means of producing commercial wells. Much like shale gas reservoirs, LTO reservoirs are typically very low in permeability and require extensive hydraulic fracturing to allow for commercial production. As a result, operators are seeking new methods to estimate hydraulic fracture properties, particularly early in the well life. Over the past 5 years there have been numerous authors have identified quantitative flowback analysis as a suitable method which in most cases aligns well with more conventional long-term production data analysis (*i.e.* [2]).

Although the majority of the literature has focused on shale gas reservoirs, there has been a substantial amount of research conducted in analyzing flowback from LTO wells. These methods have been applied to LTO plays across North America. A comprehensive literature review was given by [2] and the reader is guided to this work for details. In this paper only the papers relevant to MC simulation and application of evolutionary algorithms to the flowback problem will be discussed.

[3] developed a simple straight-line method for estimating fracture pore volume for shale gas wells that exhibit a period of unit slope which falls between the early transient flow period and the late transient flow period during multi-phase flow. In their relationship, rate-normalized pressure (RNP) is inversely proportional to fracture pore volume and total compressibility. The authors use a combination of the generalized reduced gradient method (GRG) and evolutionary algorithms in order to decouple the involved parameters. The GRG algorithm is used to find a possible optimal combination of the unknown parameters and then the evolutionary algorithm is used to generate the probability density function (PDF) and cumulative distribution function (CDF) associated with the unknown parameters. Although the approach of decoupling parameters is unique, the GRG algorithm has been shown by many authors to get trapped in local optima rather than finding the true global optimum, and therefore, this approach could be strengthened by using a more rigorous algorithm such as a GA which typically will find the near optimal. With the exception of the application of assisted parameter estimation algorithms, this approach is very similar to flowing material balance (FMB) applied by [1], although this analysis focused on single-phase fracture depletion prior to the breakthrough of formation fluids.

Reference [4] developed an approach for analyzing both flowback and long-term online production data from gas condensate wells using numerical simulation combined with a MO GA to derive key fracture and reservoir parameters. The authors rigorously modeled flowback data using a triple-porosity system (matrix, primary hydraulic fractures and induced/natural fractures) using the multiple interacting continua approach (MINC, [5]). The model also includes multiple water trapping mechanisms (permeability jail, capillary pressure and gravity segregation). This is likely the most rigorous method developed to

date, although it is also the most computationally intensive. The model could be more rigorous by including coupled flow-geomechanical simulation, although this would add additional computational intensity.

The base tool used in this work is a modified version of what was developed by [1] which is described in detail in [2]. The objective of the current work is to apply MC simulation both for uncertainty quantification as well as assisted-history matching purposes. Several other SO and MO techniques for assisted history-matching (focusing on evolutionary algorithms) are tested to determine the ability of each algorithm to identify the global minimum and therefore output the most realistic set of key uncertain fracture parameters. Results of the application of a gradient-based algorithm are presented to demonstrate how these techniques are insufficient in optimization of complex problems, unless the initial guess is closer to the absolute minimum than any local minima in the search space. Application of several evolutionary algorithms suggests that these algorithms are useful for this application, assuming a suitable search space is defined. Discussion will also be provided on how to speed up the performance of these algorithms making them more applicable for wide-spread use by industry.

2. Theory and Methods

The analysis procedure used in this work is shown in Figure 1. The emphasis of this work is on the final 2 steps of the analysis procedure in which a variety of algorithms are used in an attempt to find the optimal solution and quantify uncertainty in key fracture parameter estimates. The first three steps, however, will also be demonstrated in the context of the field example, as these steps assist in setting up the search space for the stochastic simulation and assisted history-matching.

Algorithms Used. In this work, six different algorithms were tested for the purpose of uncertainty quantification and assisted history-matching. The methods applied in this paper include: 1) MC MO simulation (Palisade® @RISK™); 2) Microsoft® Excel's SO Gradient-based (GRG2) algorithm (GRG Nonlinear Solver); 3) Microsoft® Excel's SO Evolutionary Solver; 4) Palisade® Evolver's SO GA; 5) GAPS MO GA (based on the NSGA-II-non-dominated sorting genetic algorithm) algorithm; and 6) Palisade® Evolver's SO OptQuest™ algorithm. Each
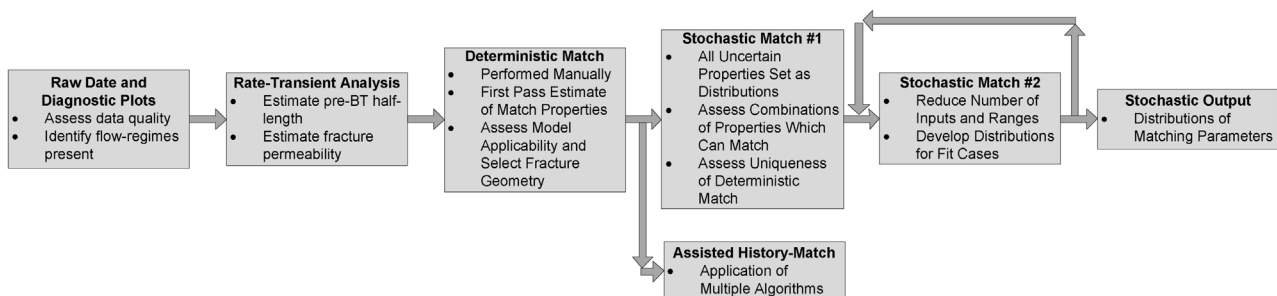
Figure 1. Summary of procedure for analyzing flowback data using deterministic, stochastic and assisted history-matching techniques.

algorithm will be briefly discussed here, with more details available in the literature.

There are two main characteristics of all of these algorithms: 1) the OF; and 2) constraints. The OF is the key parameter that one is attempting to minimize or maximize (minimization in this case). Constraints are relationships which must be satisfied for a solution to be considered acceptable. The OFs used in this work are sum of squares OFs comparing measured water and oil rate with modeled rate. Cumulative production OFs can also be introduced to further constrain the problem. The OFs used in this work will be discussed in the coming sections. Since there are no hard constraints which are applicable to this problem, the only constraints used will be the input ranges of uncertain parameters, which will also be discussed in the coming sections.

Monte Carlo Simulation. Traditional deterministic analysis techniques combine single-point estimates of key input variables to provide a single-point estimate of the result. This type of analysis assumes that the true values of all inputs are known in order to derive an accurate solution. Often these single-point estimates may differ greatly from the actual result and can lead to negative outcomes such as financial loss. In the majority of real-life problems, certainty in all parameters is rarely the case; while some variables may be known precisely or can be estimated with a reasonable degree of accuracy (ex. from lab testing or other methods), others may contain a high degree of uncertainty [6]. Stochastic simulation provides a platform to incorporate the uncertainty of inputs in order to derive a range of possible outcomes. This provides the analyst with vastly more information about the problem and assists in making smart decisions in which both the potential upside and downside are understood. Using these techniques in essence is similar to running hundreds or thousands of what-if scenarios simultaneously, while removing the often time-consuming and/or biased human component. Further, the results are presented in a manner in which they can easily be interpreted. The key components of a stochastic simulation include: 1) defining uncertain variables using a probability distribution; 2) defining key output variables; 3) running a series of simulations using an appropriate sampling technique (MC or Latin Hypercube); 4) developing a distribution of suitable output parameters using an OF. In this work, MC simulations were conducted using Palisade Corporation's @RISK^{TM} add-in for Microsoft$^{®}$ Excel^{TM}. As mentioned previously. MC simulation is conducted in such a way that multiple objectives are considered.

Microsoft$^{®}$ Excel's GRG Non-Linear Solver (*GRG2 Gradient-Based Algorithm*). This technique is based on the Generalized Reduced Gradient 2 (GRG2) algorithm which is an extension of a version of the GRG code developed by [7] and is a SO algorithm. The solver combines a graphical user interface and algebraic modeling language for linear, nonlinear and integer programs and is integrated into the host spreadsheet as closely as possible [8]. These techniques are generally applied to smooth problems (*i.e.* smooth in both the OF and con-

straints), although these methods are often applied unwisely to optimization problems that do not meet the smoothness criteria [9]. These algorithms are "downhill" in nature and therefore tend to get trapped in the closest local minima surrounding the initial guess and struggle to escape these local minima. Application of a good initial guess (*i.e.* the deterministic solution) or the multi-restart technique, in which the algorithm is started from multiple randomly generated starting points, can allow these algorithms to find the absolute minima rather than being trapped in local minima in the vicinity of the starting point.

Microsoft® Excel's Evolutionary Solver. The version of the Evolutionary Solver available in the standard version of Excel™ is a SO algorithm developed by [10]. Although the exact workings of the solver are not available in the literature or from the developer, much like other GAs, this algorithm retains a population of solutions, although this is a steady-state GA (rather than a generational GA) meaning that only one solution is replaced by a better solution at each model iteration. This GA operates on a time-based constraint in which the algorithm has a set amount of time to find a better solution than the existing best solution and outputs a single best solution. As a result of the time-based nature of this GA, the longer maximum time without improvement defined by the user gives the algorithm a better chance of locating the absolute minimum, although the technique is designed to find a "good" solution rather than the optimal solution (the two may be equivalent or at least similar).

As with other GAs there are four main steps applied within the algorithm: 1) Selection; 2) Crossover/Mating; 3) Mutation; and 4) Replacement. Reference [11] offers a variety of other more advanced algorithms which combine the simple GA with classical optimization techniques, other evolutionary algorithms as well as Tabu Search and Scatter Search which may lead to better ultimate solutions. These advanced versions were not tested in this work.

Palisade® Evolver's Genetic Algorithm. Palisade® Evolver's GA is a SO GA which contains 5 potential solving methods: 1) Recipe; 2) Order; 3) Grouping; 4) Budget; and 5) Project. The Recipe method is the default method and is designed to be used when parameter values can be varied independently and can be applied to the majority of optimization problems, especially when the relationship between the adjustable variables are not well understood, or cannot be handled better by one of the other techniques. In this work, the Recipe solving technique is used. The GA used in Evolver™ is unique, much like that used in Microsoft® Excel's Evolutionary Solver, in that it uses a steady-state approach, meaning that only one organism is replaced at a time rather that the entire generation. According to [11], this method has been shown to work as well or better than the generational method, although no evidence is provided in their literature. When comparing the results of Evolver's GA with other GA's that use the generational approach, the number of "equivalent generations" can be set by constraining the number of trails to be equal to the size of the population multiplied by the de-

sired number of generations. Parallelization is also utilized by the program to improve computational efficiency. The same four general steps of a GA are applied in this algorithm as were applied in Excel's Evolutionary Solver. This algorithm is designed to find the global minima. A uniform crossover scheme is used by this algorithm, meaning that half of the parameters of the child come from each parent.

GAPS Multi-Objective Genetic Algorithm (*Based on NSGA-II Algorithm*). GAPS is a MO GA based on the NSGA-II algorithm developed by [12]. As a MO GA, the algorithm is designed to account for objective conflict and yields a Pareto Front of solutions which are all mathematically equivalent. If a single optimal solution is desired, then the user must select the solution along the Pareto Front which is most applicable to the problem being solved. The NSGA-II algorithm is an extension of the NSGA algorithm developed by [13]. The basis of the algorithm is the nondominated sorting procedure, hence the name of the algorithm. The modified algorithm was developed to handle the main criticisms of the original algorithm, and offers the following benefits over the original algorithm:

- Utilizes a faster method for nondominated sorting.
- Preserves elitism, meaning that the best solutions are maintained without modification.
- Incorporates a parameter-less diversity preservation mechanism to replace the need for a sharing parameter, which is the traditional mechanism for maintaining diversity.
- Utilizes parallelization to improve solution speed by allowing calculations to be spread out over multiple processors.

There are two key concepts to the algorithm, being: 1) nondominated sorting; and 2) diversity preservation and follows the same general concept of the other GAs discussed above, although is generational in nature. This algorithm has been shown to work well for three OFs by [5], although it generally begins to fail as further objectives are added in complex problems (four or more in this particular problem). To handle this deficiency, the U-NSGA-III algorithm was developed by [14]. This algorithm uses a continuous single-point crossover scheme, meaning that crossover occurs at randomly chosen points and the two children get the genetic material from either side of the crossover point.

Palisade® Evolver's OptQuest Algorithm. OptQuest™ is a "black box" optimizer first developed by [15] to find the global optimum solution. The algorithm is not totally context-independent because the selection of the solution representation gives some information to the optimization algorithm. The model allows the user to represent solutions as a mixture of continuous, discrete, integer, binary, permutation and other specialized variables which provides the optimizer with some information about the system. The software ultimately chooses solvers based on the characteristics of the optimization model (pure or mixed, constrained or unconstrained and deterministic or stochastic). The optimizer is

based in Scatter Search, but also uses the principles of Tabu Search, an Artificial Neural Network and other methods in attempt to derive a global optimum solution more rapidly. Scatter Search is an optimization algorithm comparable to a GA.

3. Field Example

A full analysis demonstrating each step of the analysis procedure (shown in **Figure 1**) will be provided. One MFHW from a 3-well pad in a LTO play in the Western Canadian Sedimentary Basin (WCSB) will be analyzed, although the other two wells on the pad, as well as other wells in the area and other LTO plays in North America, have also been analyzed using this procedure. This well was previously analyzed by [1], but using a deterministic approach. To protect operator confidentiality, well location and reservoir and completion information has been withheld. A summary of the completion and stimulation performed is given below:

- Cased-hole completion.
- Hydraulically fractured with hybrid water fracs in 18 stages using plug and perf technology (single perforation cluster per stage).
- Fracture stages spaced at ~330 ft.
- 1350 STB of fracture fluid and 45 T of proppant pumped per stage.

Assessing the microseismic collected on this well, the assumption of circular bi-wing planar fractures appears to be reasonable and will be used in this analysis. Preceding the flowback data used for this analysis, plugs were drilled out with coil tubing following stimulation, after which the well was placed on flowback monitoring through a test separator. Rate and pressure data was gathered every 15 minutes for approximately 300 hours during flowback following a 12 day shut-in period.

Input common to the different flowback analysis techniques are shown below in **Table 1**. Note that the individual hydraulic fracture width is approximately twice what is expected for a simple bi-wing planar fracture (0.25 in/stage). This is likely due to some fracture complexity (or possibly multi-planar fractures) which could not be resolved at the microseismic level.

Raw Data and Diagnostic Plots. Water, oil and gas rates as well as bottom-hole flowing pressure and gas-oil ratio (GOR) are shown below in **Figure 2(a)**, while water RNP and its derivative (RNP') are shown in **Figure 2(b)**.

From **Figure 2(a)** it can be observed that flowback initiates with more than 2 days of single-phase water production (fracture fluid) prior to the breakthrough of hydrocarbons (formation fluid). Hydrocarbons breakthrough at just after 2 days, with an initial oil rate of ~36 STB/D. For about the first 6 days of hydrocarbon production (8 days total), the GOR is approximately constant at 1250 scf/STB which is equal to the solution gas level. At ~8 days there is a rapid drop in flowing pressure (resulting from a rapid decrease in choke size) below the bubble-point followed by a rapid increase in GOR suggesting a breakthrough of

**Table 1.** Input parameters for flowback field example.

Fracture Properties	Parameter Value
Initial Fracture Pressure (psia)	5000
Initial Water Saturation (%)	100
Fracture Porosity (proppant pack) (%)	31
Fracture Compressibility (psi^{-1})	1×10^{-4}
Number of Hydraulic Fractures	18
Individual Hydraulic Fracture Width (ft)	0.0417
Total Hydraulic Fracture Width (ft)	0.75
Reservoir Properties	**Parameter Value**
Formation Pressure (psia)	3700
Net Pay (ft)	197
Matrix Porosity (%)	4
Initial Mobile Oil Saturation (%)	99
Initial Mobile Water Saturation (%)	1
Formation Compressibility (psi^{-1})	4×10^{-6}
Matrix Permeability (md)	0.0003
Reservoir Temperature (˚F)	140
Fluid Properties	**Parameter Value**
Fracture Water Salinity (%)	50,000
Formation Water Salinity (%)	200,000
Oil Gravity (˚API)	52
Gas-Oil-Ratio (scf/stb)	1250
Bubble-Point Pressure (psia)	2858
Gas Gravity (air = 1)	0.747
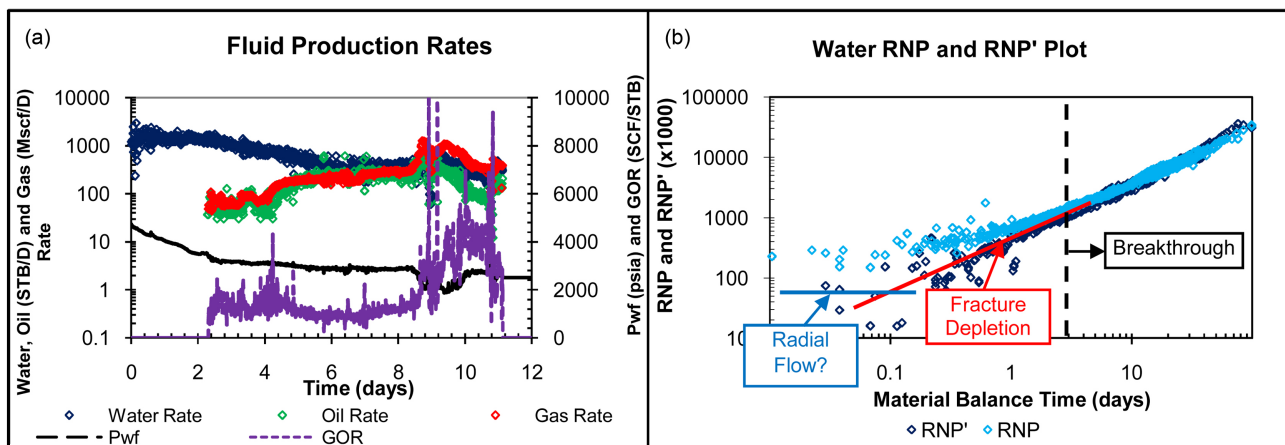
Figure 2. Flowback data: (a) Water, oil and gas rate, as well as bottom-hole flowing pressure and GOR; and (b) Water RNP and RNP'.

gas into the fractures. Therefore, only the first 8 days of production were considered for this analysis as this is the period where production is under two-phase (water + oil) flow in the formation and fractures (the tool cannot currently model three-phase flow). Over the first 8 days of production, water rate and bottom-hole flowing pressure generally decline, while the hydrocarbon rate generally increases following breakthrough as would be expected from a formation with minimal mobile water and constantly decreasing bottom-hole flowing (and fracture) pressure. From Figure 3(b), BBT flow-regimes can be identified using the RNP' curve from the period of single phase production. The first flow-regime interpreted is a short period of transient radial flow within the fractures (0-slope), which appears to last until ~0.1 days of material balance time (MBT), although the data is scarce and noisy during this period making it difficult to conclude this flow-regime identification with a significant degree of certainty. From ~0.1 days to ~3 days of MBT a clear period of fracture depletion (unit-slope) is identified up until breakthrough. After-breakthrough (ABT), the derivative is non-linear in nature, as would be expected as multiple flow-regimes are occurring, although a depletion-like signature remains dominant.

Casing pressures were converted to sandface pressures using a wellbore model, and initial formation pressure was estimated from $p^\star$ obtained from a Diagnostic Fracture Injection (DFIT) test which also yielded the estimate of matrix permeability. The GOR and bubble-point pressure are defined based on PVT analysis of the reservoir fluid from a group of off-setting wells. Initial fracture pressure was determined by a trial and error process conducted by [1] and was maintained for this analysis.

Rate-Transient Analysis of BBT Single-Phase Data. To assist with the history-matching process, rate-transient analysis (RTA) is applied to the flow-regimes identified in Figure 2(b) (repeated in Figure 3(a)). Radial flow analysis, which is shown in Figure 3(b), is used to analyze flow-regime (FR) 1 and estimate fracture conductivity (permeability). Delimiters are shown to highlight the interpreted period of radial flow. Using the slope of the radial flow plot, fracture permeability is estimated to be ~3500 md, assuming a fracture width of 0.5 in/fracture with a small negative skin. The negative skin may be a result of maximum proppant concentration near the wellbore as discussed previously. The FMB, which is shown in Figure 3(c), is used to assess FR 2 and estimate BBT fracture volume and fracture half-length. From the x-intercept of the plot, the total fracture volume is estimated to be approximately 24,000 STB, and the BBT fracture half-length to be ~441 ft per stage, assuming a circular fracture shape and a fracture width of 0.5 in/fracture. From the y-intercept an additional measure of fracture permeability can be derived to be ~3400 md, assuming a fracture width of 0.5 in/stage. The total calculated fracture volume is approximately equal to the total volume injected during the fracture stimulation, suggesting that the majority of pumped fluid has been converted into effective fracture volume (prior to hydrocarbon breakthrough). This conversion percentage is
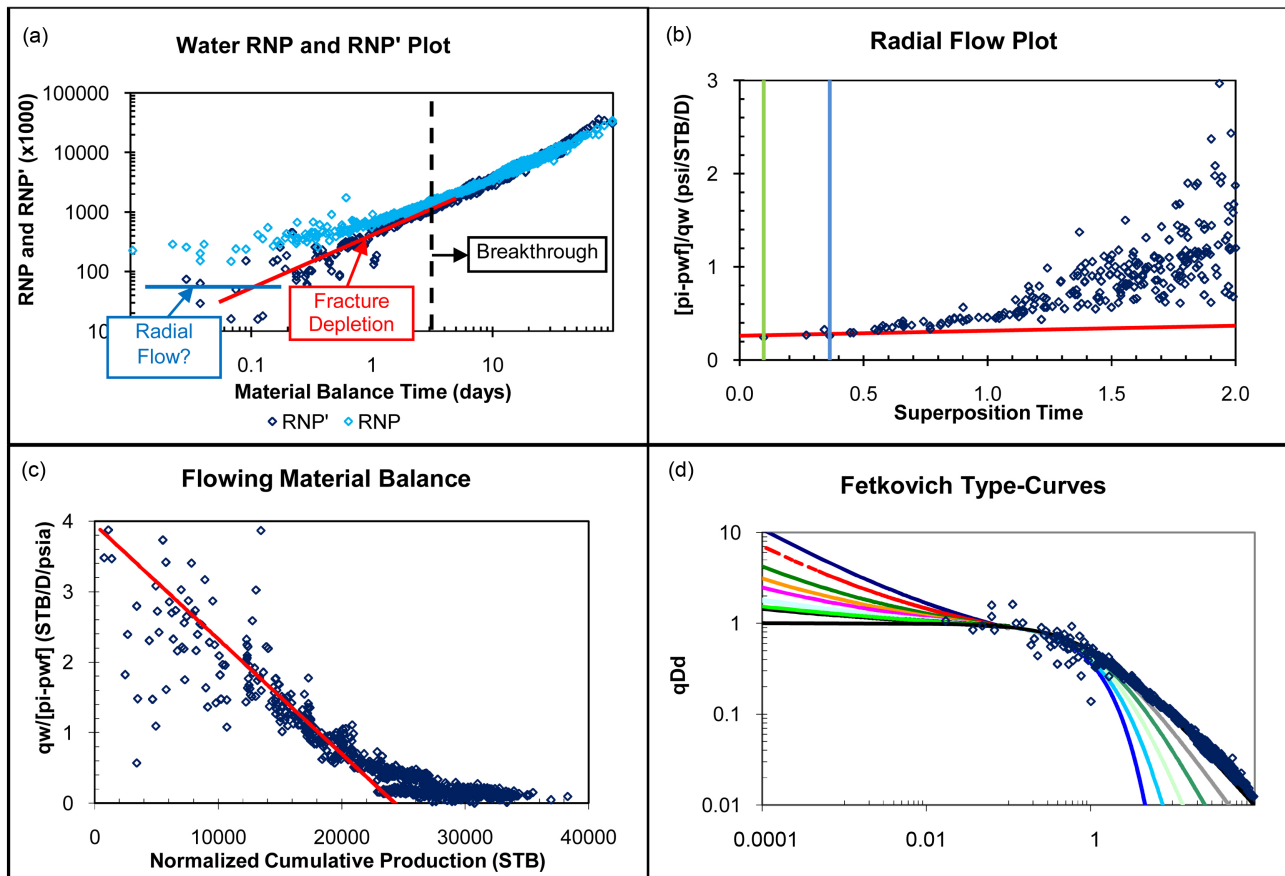
Figure 3. Rate-transient analysis of BBT single-phase data: (a) Water RNP and RNP' plot; (b) Early radial flow analysis; (c) Flowing material balance; and (d) Fetkovich type-curve.

higher than expected even for a well with minimal natural fracturing (as inferred from microseismic and experience in the formation of interest), and may result from the impact of the other two wells being stimulated on the same pad prior to flowback of the well. These values will be used in the deterministic history-matching process. Finally, the fracture parameters estimated from radial flow analysis and the FMB can be confirmed by using the Fetkovich type-curve (Figure 3(d)) which is designed for analyzing radial to boundary-dominated flow behaviour. Because MBT is used in the calculation of t_{Dd}, fracture depletion data falls down the harmonic stem, with a positive deviation indicating the breakthrough of formation fluid.

Parameters estimated from quantitative RTA of this flowback data are provided in Table 2.

Deterministic History-Match. Deterministic history-matching was first conducted to validate the application of the conceptual model to this dataset, confirm selection of a fracture shape and geometry model, and confirm RTA-derived parameters for BBT fracture properties. For this analysis, a circular shape with a single bi-wing fracture being generated from each stage was selected for simplicity, as was done by [1], under the assumption that cylinder radius is equal to fracture half-length.

Table 2. Parameters solved from each BBT RTA technique.

Radial Flow Plot	Parameter Value
Fracture Conductivity, F_{cT} (md-ft)	2625
Fracture Permeability (md)	3500
Flowing Material Balance	Parameter Value
Fracture Fluid-In-Place (STB)	24,000
BBT Half-Length, x_{f_BBT} (ft)	411
Fracture Conductivity, F_{cT} (md-ft)	2550
Fracture Permeability (md)	3400
Fetkovich Type-Curve	Parameter Value
x_f/r_{wa}	~1000
Fracture Conductivity, F_{cT} (md-ft)	2625
Fracture Permeability (md)	3500

The history-matches, guided by the BBT RTA-derived parameters are provided in **Figure 4**. The deterministic history-match was not significantly changed from what was presented by Clarkson *et al.* (2014), although minor improvements were made. Note the deterministic history-match is shown here in black to maintain continuity of this match throughout the paper.

From **Figure 4** it can be seen that the deterministic history-match to water is very good throughout the eight days of flowback modeled, while the hydrocarbon match is very good for the first 6.5 days at which point it begins to underestimate production. After conducting the first 3 stages of the analysis procedure, the remainder of the paper will focus on the stochastic simulation and assisted history-matching component of the procedure which is the main contribution of this work.

The key history-match parameters are given in **Table 3** (assumed to be fracture permeability, cylinder radius (fracture half-length) BBT and ABT of formation fluid, breakthrough pressure and the Corey relative permeability exponents for oil and water in the fractures). Matrix relative permeability coefficients were assumed to be 2 for oil and water formation fluid, although due to the very high mobile oil saturation, these curves have minimal impact on the analysis.

From **Table 3**, it can be seen that the fracture half-length decreases following the breakthrough of formation fluids, as expected. A ~7% decrease in drainage radius (~4% decrease in effective half-length) was observed and this was applied during stochastic simulation and assisted history-matching to reduce the number of uncertain parameters. In many cases the reduction in half-length as a result of breakthrough can be significantly larger. Further, breakthrough pressure was set slightly higher than formation pressure to account for the supercharge effect associated with high-rate injection. Matrix relative permeability exponents were set to two for both water and oil throughout.

Stochastic Simulation and Assisted History-Matching. In this section, the results from multiple stochastic and assisted history-matching techniques will be
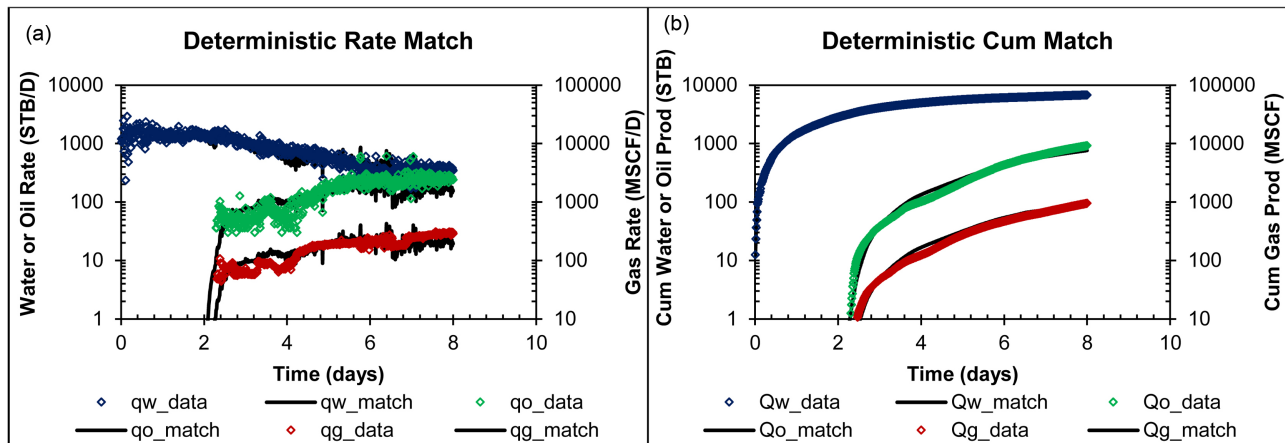
Figure 4. Deterministic flowback match: (a) Water, oil and gas production rates; and (b) Cumulative water, oil and gas.

Table 3. Key history-match parameters for flowback simulation.

History-Match Parameter	Parameter Value
Fracture Permeability (md)	3500
BBT Drainage Area (Ac)	14
BBT Half-Length, x_{f_BBT} (ft)	441
ABT Drainage Area (Ac)	13
ABT Half-Length, x_{f_ABT} (ft)	425
Breakthrough Pressure (psi)	3825
Corey Oil Exponent—Fracture, n'	1.2
Corey Water Exponent—Fracture, m'	5.0

discussed. As mentioned previously, the algorithms used include: 1) MC simulation (Palisade® @RISK™); 2) Microsoft® Excel's Gradient-based (GRG2) algorithm (GRG Nonlinear Solver); 3) Microsoft® Excel's Evolutionary Solver; 4) Palisade® Evolver's GA; 5) GAPS MO GA (based on the NSGA-II) algorithm; and 6) Palisade® Evolver's OptQuest™ algorithm. The results of each individual technique will be discussed followed by a comparison of the results of each of the techniques. As discussed previously, only the first 8 days of flow data were analyzed because, during this flow period, the flowing pressure remains above the bubble point, and therefore only two-phase flow exists in the matrix and fractures. During this period, the GOR is also relatively constant, as would be expected for flow above the bubble point.

Monte Carlo Simulation. As discussed previously, stochastic history-matching can be a multi-step process, with multiple refinement stages. For example, two sets of MC simulations were conducted in the presented example. Following the first stage, inputs including fracture compressibility and matrix properties were held constant for the final set of simulations, which will be discussed here. Further refinement stages could be conducted using information from past runs to adjust input distributions and increase the number of success cases. The parameter distributions for the first refinement stage are shown below in **Table 4**.

Table 4. Input distributions/ranges for Monte Carlo simulation and assisted history-matching.

Uncertain Parameter	Distribution Type	Low Value	High Value
Fracture Permeability (md)	Uniform	3000	4000
BBT Drainage Area (Ac)	Uniform	10	14*
BBT Half-Length, $x_{f\,BBT}$ (ft)	Uniform	372	441*
Breakthrough Pressure (psi)	Uniform	3700	4100
Corey Oil Exponent—Fracture, n'	Uniform	1	3
Corey Water Exponent—Fracture, m'	Uniform	1	10

*Upper bound on fluid in place given assumptions of fracture shape, width and porosity.

Because enough data was not available to construct proper input distributions, uniform distributions were used for each parameter between a reasonable low and high value. In some cases, the high and low value were constrained by physical limits (*i.e.* the upper limit of half-length, the lower limit of n' and m' and breakthrough pressure) whereas other limits were set at a reasonable range and then adjusted following the screening stage of iterations. The input parameter ranges were also further constrained by the initial screening phase of simulations (500,000 iterations with significantly wider parameter ranges), as well as reasonable limits on the uncertain parameters. The initial screening phase was used to rule out outlier matches which occurred with minimal frequency. The same limits were used for the application of the assisted-history matching techniques, which will be discussed in the following section. This analysis is comparable to that conducted by [16], although the number of uncertain parameters was reduced from 10 to 5, placing the focus on the most important parameters. Using this approach allows reasonable coverage of the sample space with a smaller number of iterations. [17] conducted significantly less simulations than would be needed to cover the sample space, although the purpose of that work was to demonstrate the purpose and application of MC simulation to history-matching flowback data from MFHWs.

The following objective functions (OFs) were used in either the MC simulations, assisted history-matching algorithms or both. The OFs take the form of sum of squared residuals for the rate and cumulative production of the water and oil phases. Because the well is flowed above the bubble point throughout the analysis period of the flowback, and the GOR is relatively constant at approximately the solution gas level, the gas phase is not considered and is effectively lumped in with the oil phase.

Water Rate OF:

$$OF_{q_w} = \sum_{i=1}^{n} \left(q_{w,data} - q_{w,sim} \right)^2 \tag{1}$$

where, n is the number of data points collected during the portion off the flowback data being analyzed for each phase.

Oil Rate OF:

$$OF_{q_o} = \sum_{i=1}^{n} \left(q_{o,data} - q_{o,sim} \right)^2 \tag{2}$$

Cumulative Water OF:

$$OF_{Q_w} = \sum_{i=1}^{n} \left(Q_{w,data} - Q_{w,sim} \right)^2 \tag{3}$$

Cumulative Oil OF:

$$OF_{Q_o} = \sum_{i=1}^{n} \left(Q_{o,data} - Q_{o,sim} \right)^2 \tag{4}$$

Summed Rate OF:

$$OF_{q_t} = \sum_{i=1}^{n} \left[w_w \left(q_{w,data} - q_{w,sim} \right)^2 + w_o \left(q_{o,data} - q_{o,sim} \right)^2 \right] \tag{5}$$

where,

$$w_w + w_o = 1 \tag{6}$$

The summed rate OF given by Equation 5 is used for the SO algorithms. There are however two issues associated with using summed OFs: 1) objective conflict leading to erroneous results; and; 2) weighting can have a significant impact on the results of the algorithm. In this work, a 1:1 weighting was used for direct comparison to the MO algorithms (equivalent to using $w_w = w_o = 0.5$).

Alternate criteria, such as those applied by [16] [17], could also be used if a reasonable baseline deterministic match is not available. In these two papers, the authors used an R^2 value of each phase (in terms of rate) greater than 0.9 (as suggested by [18]) and a total cumulative production for each phase within 10% (as suggested by [19]) for each phase to determine a successful match. This approach was shown to be relatively successful in the studies by [16] [17]. However, some successful solutions in both cases were found where a high R^2 value existed but the match was poor, due to the inherent nature of R^2 as a match fit indicator, especially when dealing with highly nonlinear problems with a large number of data points. For the current study, 100,000 iterations are conducted, and fairly strict criteria were enforced to obtain a successful match, including the following (solution has to be better than the deterministic solution for both phases):

- $OF_{q_w} < OF_{q_w,deterministic}$

- $OF_{q_o} < OF_{q_o,deterministic}$

- $OF_{Q_w} < OF_{q_w,deterministic}$

- $OF_{Q_o} < OF_{q_o,deterministic}$

Using all four of these criteria, only 79 matches were found (~0.1%), while if only the two rate criteria were used, as is usually done with the assisted history-matching techniques, ~10× the number of matches were found (~1%). Based on these results, it is clear that the random behaviour of MC simulation is not particularly efficient in finding optimal solutions, making the use of modern assisted history-matching techniques desirable, particularly when a deterministic solution is not available. For the remainder of this paper, only the two rate OFs will be used in the application of all algorithms, because adding more OFs can often cause these algorithms to converge slowly and creates further objective conflict, potentially leading to finding a less desirable solution. The parameters

for the deterministic match, as well as the best 5 matches (in order of increasing OF value) when considering the summation of the two OFs, are shown below in Table 5 along with the mean, standard deviation and P10/P90 ratio of the 861 acceptable matches.

The history-match associated with the top 10 MC simulation history-matches as well as the deterministic history-match are shown in Figure 5.

For later comparison to the assisted history-matching results, the average of the top five iterations were assumed to represent the best solution found using MC simulation, as each of these solutions have a total OF within 1% of each other. The values for each of the uncertain parameters are provided in Table 6. The average summed rate OF is ~14% lower than the deterministic solution.

Table 5. Stochastic history-match parameters.

Uncertain Parameter	Deterministic	Match 1	Match 2	Match 3	Match 4	Match 5	Mean	Std. Dev.	P10/P90
Fracture Permeability (md)	3500	3225	3184	3167	3159	3216	3425	254.41	1.23
BBT Drainage Area (Ac)	14	13.82	13.73	13.97	13.78	13.67	13.52	0.37	1.07
BBT Half-Length, x_{f_BBT} (ft)	440	438	436	440	437	435	432	6.01	1.04
ABT Drainage Area (Ac)	13	12.83	12.75	12.97	12.79	12.69	12.54	1.41	1.04
ABT Half-Length, x_{f_ABT} (ft)	425	422	420	424	421	420	417	5.79	1.04
Breakthrough Pressure (psi)	3825	4093	4096	4058	4099	4089	4024	56.75	1.04
Corey Oil Exponent—Fracture, n'	1.20	1.44	1.44	1.35	1.42	1.51	1.40	0.12	1.26
Corey Water Exponent—Fracture, m'	5.00	4.81	5.04	5.31	6.10	5.25	5.31	1.41	2.03
Total Objective Function (millions)	41.5	35.4	35.4	35.5	35.6	35.7	39.2	0.14	1.10

Table 6. Average values of top 5 Monte Carlo simulations.

| Uncertain Parameter	Top 5 Average
Fracture Permeability (md)	3190
BBT Drainage Area (Ac)	13.78
BBT Half-Length, x_{f_BBT} (ft)	437
ABT Drainage Area (Ac)	12.80
ABT Half-Length, x_{f_ABT} (ft)	421
Breakthrough Pressure (psi)	4087
Corey Oil Exponent—Fracture, n'	1.43
Corey Water Exponent—Fracture, m'	5.30
Total Objective Function (millions)	35.4
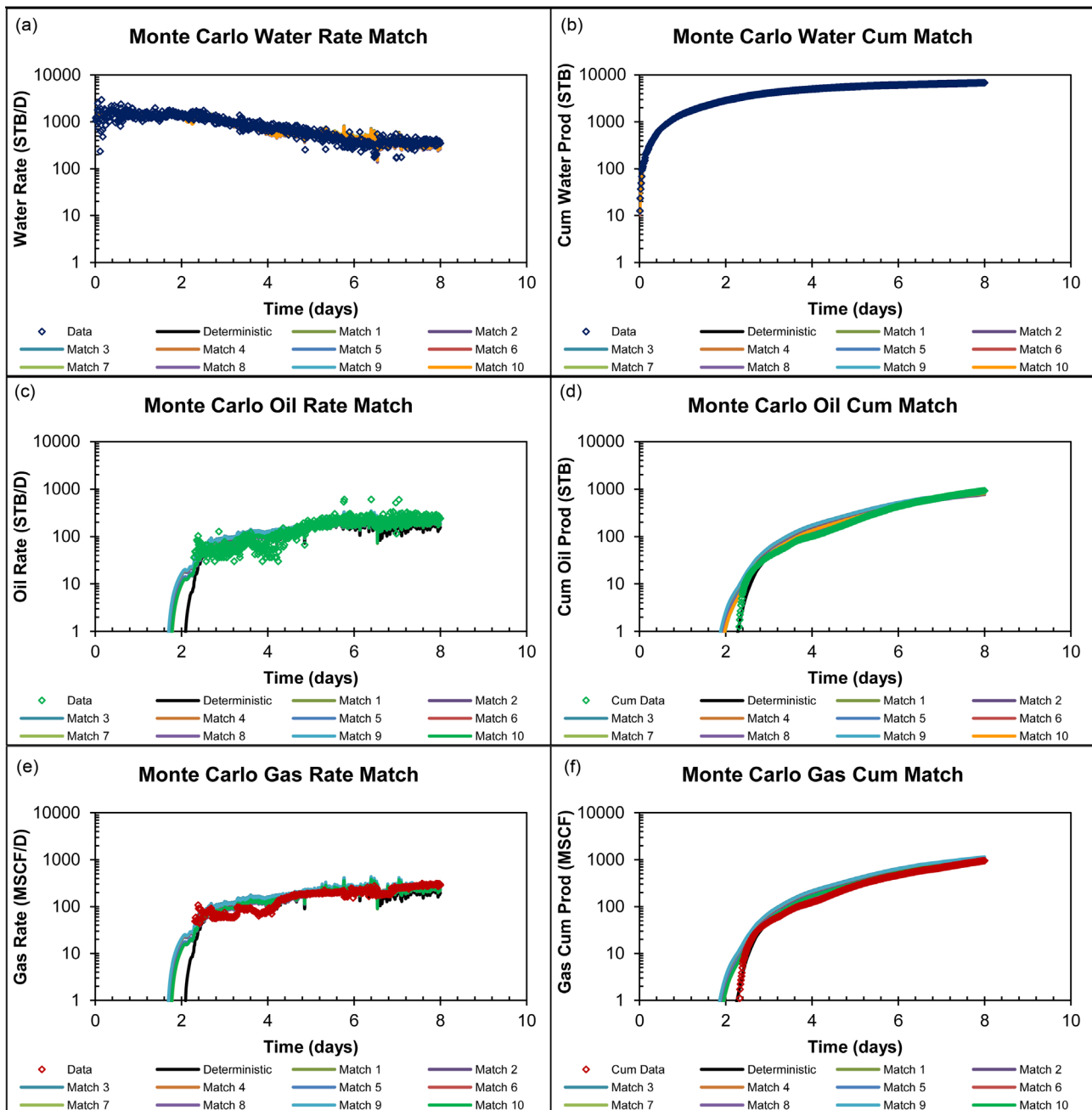
Figure 5. Stochastic flowback history-match: (a) Water production rates; (b) Cumulative water production; (c) Oil production rates; (d) Cumulative oil production; (e) Gas production rates; and (f) Cumulative gas production.

The parameter distributions generated from the stochastic history-matching exercise are provided in **Figure 6**. R^2 values are shown to indicate the lognormal nature of the output distributions, and the deterministic and mean values are provided for reference. Note that the parameter distribution for BBT drainage area is not provided because the focus is on the half-length calculated from the drainage area (using the assumed shape and geometry constraints), since half-length is one of the key parameters controlling long-term production of the well.

From **Table 5** and **Figure 6** it can be seen that the range of values used to match the flowback data is fairly small (P10/P90 ≤ 2). Further it can be seen that the average values of the 861 successful matches are within 5% of the deterministic match, with the exception of the relative permeability exponents (n' = 17% and m' = 6%), which are sensitive to small changes due to the small magnitude of their values. Some key results to point out are as follows:

- Fracture permeability covers the entire input distribution, suggesting that fracture permeability may fall outside the search space. However, very few matches beyond the selected range were found during the screening phase.
- Breakthrough pressures, including the deterministic match, are greater than
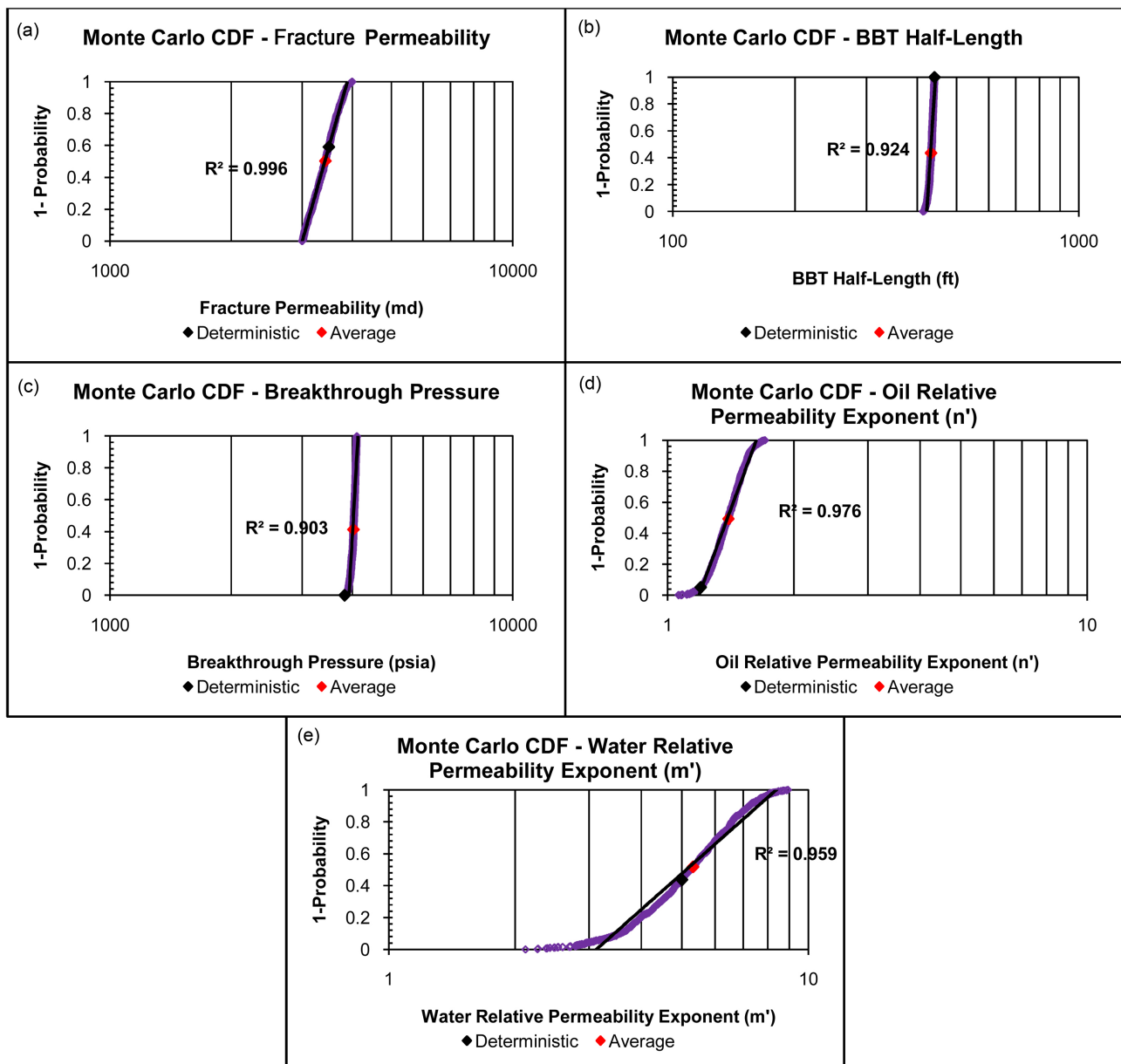
Figure 6. Stochastic flowback history-match parameter distributions: (a) Fracture permeability; (b) Breakthrough pressure; (c) BBT half-length; (d) Corey oil relative permeability exponent, n'; (e) Corey water relative permeability exponent, m'.

reservoir pressure estimated from DFIT analysis. Values (other than the deterministic match) also fell near the upper limit of 4100 psia, suggesting a better match to the available data could be achieved using a breakthrough pressure of 4100 psia. However, experimentation with a variety of fracture parameters suggested that breakthrough pressures greater than 4100 psia led to model breakthrough significantly BBT in the actual data, which led to setting the upper limit at the selected value. 4100 psia still yields a breakthrough earlier than the data—however, the fact that flow initiates at ~36 STB/D, which is higher than other similarly completed wells on the same pad, suggests that some early-time hydrocarbon data may not have been recorded. An improved late-time match was observed with earlier breakthrough. Overall the results suggest a near fracture supercharge of up to 10% following an 11 day shut-in between stimulation and the onset of flowback in which the bridge plugs were milled out. The supercharge has been shown to be much higher in some formations depending on factors such as the stimulation pumped, shut-in time between stimulation and flowback and other reservoir and fluid properties.

- BBT half-length values fell in the range of 413 to 441 ft suggesting that a BBT half-length less than 400 ft is unlikely and that a high degree of fracture efficiency was achieved.

- Fracture relative permeability exponents to oil fall in a tight band between 1.1 and 1.6, suggesting minimal potential variability in this parameter.

- Fracture relative permeability exponents to water are far less constrained than those to oil falling between 2.1 and 9, although are significantly higher than those to oil. This has been observed in nearly all wells analyzed using these methods. In this case it can be seen that the values between the P10 (3.5) and P90 (7.2) follow a lognormal distribution and yield a P10/P90 ratio of ~2.20% of the solutions fell outside this range and may be considered as outliers.

Assisted History Matching. In addition to the MC simulation approach demonstrated above, five assisted history-matching algorithms were applied in an attempt to find the best possible history-match to the same flowback data set discussed above. Two types of algorithms were used in this analysis (gradient-based and evolutionary) with a total of five techniques being tested: 1) Microsoft® Excel's SO Gradient-based (GRG2) algorithm (GRG Nonlinear Solver); 2) Microsoft® Excel's SO Evolutionary Solver; 3) Palisade® Evolver's SO GA; 4) GAPS MO GA (based on the NSGA-II) algorithm; and 5) Palisade® Evolver's SO OptQuest™ algorithm. For this analysis, the lower and upper bounds given above in Table 4 are used as constraints on the algorithms, and no further constraints were applied. The same initial guesses for the uncertain parameters (Table 7) are used to seed each of the algorithms, although many evolutionary algorithms do not require an initial guess as they generate an initial population based on the constraints in the uncertain parameters. Most available evolutionary algorithms are implemented in a way that the initial guess will be a member of the

Table 7. Initial guesses for assisted history-matching parameters.

Uncertain Parameter	Value
Fracture Permeability (md)	3100
BBT Drainage Area (Ac)	13.96
BBT Half-Length, x_{f_BBT} (ft)	440
ABT Drainage Area (Ac)	13
ABT Half-Length, x_{f_ABT} (ft)	424
Breakthrough Pressure (psi)	3900
Corey Oil Exponent—Fracture, n'	1
Corey Water Exponent—Fracture, m'	1

first population. As was discussed previously, and as will be demonstrated below, the initial guess is critical to achieving good results from the GRG algorithm because these algorithms will tend to find the closest local minima in the OF (downhill nature of the algorithm). The initial guesses were selected based on the following criteria:

- Fracture permeability—RTA of early-time flowback data suggested a maximum fracture permeability of ~3500 psia (as was used in the deterministic history match) and therefore a slightly lower value was selected for this application.
- Breakthrough pressure—DFIT analysis suggested an initial reservoir pressure of ~3700 psia and therefore a 200 psia supercharge effect was assumed (~5%).
- Drainage area—set based on results of the FMB in the deterministic analysis.
- Relative permeability exponents—straight-line relative permeability curves were assumed as may be expected for homogeneous perfectly planar fractures under ideal flowing conditions.

The results of each algorithm will be discussed, followed by a comparison of the results of each algorithm, as well as the average results of the top five MC simulations.

Microsoft® Excel's GRG Nonlinear Solver. As discussed previously, the initial guess is critical to the quality of the result using this type of algorithm due to the "downhill" nature of the algorithm and tendency to get trapped in local minima. This impact will be demonstrated in this section. Due to the deficiencies of this algorithm in solving complex problems with multiple minima, a poor result was expected using the initial guesses shown in Table 7. However, it was interesting to determine whether a reasonable quality initial guess (*i.e.* the deterministic solution) could be used to converge on the absolute minimum and ultimately find an optimized solution. This is of interest because these algorithms run very quickly compared to GAs due to their simplistic nature, and are built directly into Microsoft® Excel™, allowing for fast and simple application following the deterministic history-matching exercise. To test the capacity of the algorithm for solving this problem, two runs were completed. In the first run the initial guesses shown in Table 8 were used, and in the second run, the deterministic solution

Table 8. Initial guesses and final solutions for Attempt #1 using EXCEL's GRG non-linear solver.

Uncertain Parameter	Initial Guess	Final Solution
Fracture Permeability (md)	3100	3000
BBT Drainage Area (Ac)	14	14
BBT Half-Length, $x_{f,BBT}$ (ft)	441	441
ABT Drainage Area (Ac)	13	13
ABT Half-Length, $x_{f,ABT}$ (ft)	425	425
Breakthrough Pressure (psi)	3900	3911
Corey Oil Exponent—Fracture, n'	1	1
Corey Water Exponent—Fracture, m'	1	1
Total Objective Function (millions)	42.8	

(see Table 2) was used. In both cases the algorithm converged to a solution very quickly, given the speed of the tool being used, suggesting very few iterations were required to locate a minima, although the exact iteration count is not provided by the standard version of Excel™ unless the algorithm is stopped at each iteration (which was not done in this case). The initial guess and final solution for the two sets of input parameters are provided in Table 8 and Table 9, with the solutions being compared to actual data and the deterministic history-match in Figure 7.

From Table 7 and Table 8 it can be seen that the final solutions yield a 4% higher summed rate OF than the deterministic solution when using the initial guesses shown in Table 6, while the optimal solution using the deterministic match as the initial guess yields a 16% reduction in the summed rate OF which equates to an ~2% improved result over any of the MC simulations. From Figure 7 it can be seen that Attempt #1 yields a very poor history-match, especially to the hydrocarbon phases, while Attempt #2 yields an excellent history-match to all three phases. As will be seen in the coming sections, the gradient solver replicated the results of the evolutionary algorithms when seeded with the deterministic history-match as an initial guess. This suggests that no local minima exist between the deterministic solution and the absolute minimum. Further, this type of method may be used to optimize a history-match once a reasonable deterministic solution is found, although additional testing would be required to further substantiate this claim.

Microsoft° Excel's Evolutionary Solver. In this section the results of the Excel's Evolutionary Solver will be demonstrated. This Solver algorithm is the first of two SO GAs which will be tested in this work. This solver also uses several classical optimization methods to attempt to improve upon the solutions found by the GA, thus making it a hybrid GA. As discussed previously, details of how

Table 9. Initial guesses and final solutions for Attempt #2 using EXCEL's GRG non-linear solver.

Uncertain Parameter	Initial Guess	Final Solution
Fracture Permeability (md)	3500	3102
BBT Drainage Area (Ac)	14	14
BBT Half-Length, x_{f_BBT} (ft)	441	441
ABT Drainage Area (Ac)	13	13
ABT Half-Length, x_{f_ABT} (ft)	425	425
Breakthrough Pressure (psi)	3825	4099
Corey Oil Exponent—Fracture, n'	1.2	1.45
Corey Water Exponent—Fracture, m'	5	5.61
Total Objective Function (millions)	34.7	

Excel's Evolutionary Solver works are not readily available and very little assistance was provided by the developer to help understand exactly which techniques are employed. Given that this is a time-based, rather than a generation-based algorithm, the exact number of iterations conducted is unknown, although the algorithm converged significantly faster than the other algorithms tested, suggesting that significantly fewer 10,000 iterations were conducted in finding the best solution. The input parameters used for this algorithm are shown below in Table 10. A mutation rate of 15% was selected for all GAs, as this was preprogrammed into the version of the GAPS algorithm which was used in this work. A population size of 100 was used for all of the population-based algorithms based on the suggestions of Kanfar and Clarkson (2016). Max time without improvement was set to a high value to give the algorithm sufficient time to search for a better solution given the calculation speed of the spreadsheet-based tool used in this work (~30 seconds/iteration). Based on the values provided in Table 10, and the approximate run speed of the spreadsheet, ~300 iterations were allowed to find an improved solution, alternatively the algorithm will terminate once the maximum change of the combined OF falls below 0.01%. In this case it is unclear whether the algorithm terminates based on the time or convergence criteria. The same convergence criteria were used with Palisade's algorithms, which will be discussed in a coming section.

As with other SO algorithms, a single best solution is found by the algorithm. The parameters resulting from the optimization are found in Table 11 and the resulting combined OF is ~16% lower than that of the deterministic match.

Palisade® Evolver's Genetic Algorithm. In this section the results of Palisade® Evolver's SO GA will be demonstrated. Evolver™ is the second SO GA used in this work. Much like Excel's Evolutionary Solver, Evolver™ uses a steady-state approach, which the company has found to work as well or better than the generational approach. Further, given that this is a proprietary commercial tool, details on the exact workings of the algorithm are not readily available. Based on the information provided by the developer, the algorithm operates in a manner
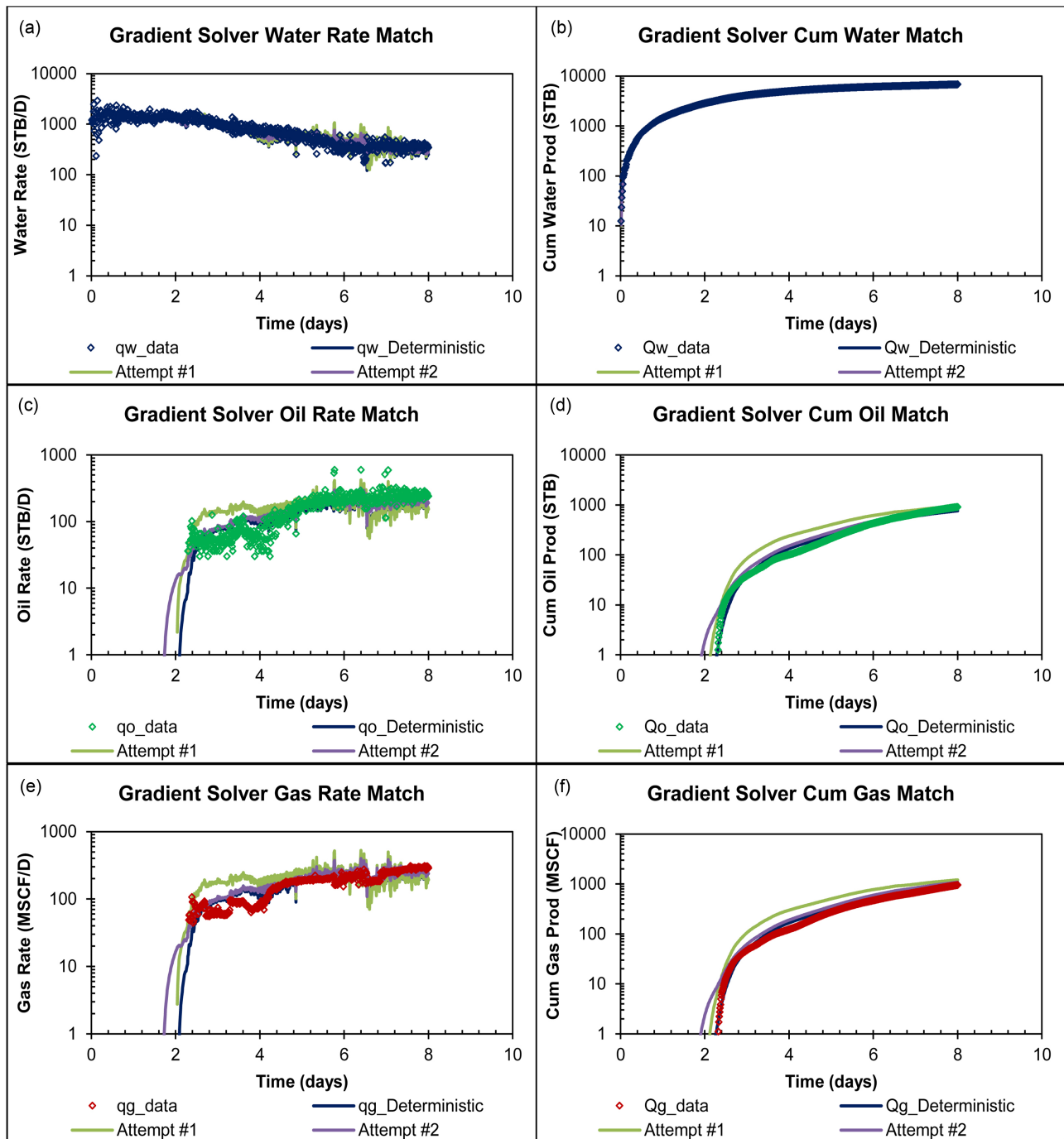
Figure 7. Gradient solver flowback history-match: (a) Water production rates; (b) Cumulative water production; (c) Oil production rates; (d) Cumulative oil production; (e) Gas production rates; and (f) Cumulative gas production.

Table 10. Input parameters for excel's single-objective evolutionary solver.

Mutation Rate (%)	15
Population Size	100
Max Time Without Improvement (s)	10,000
Convergence Criteria for Max Change (%)	0.01

Table 11. Optimal match parameters for excel's single-objective evolutionary solver.

Uncertain Parameter	Value
Fracture Permeability (md)	3156
BBT Drainage Area (Ac)	14
BBT Half-Length, x_{f_BBT} (ft)	441
ABT Drainage Area (Ac)	13
ABT Half-Length, x_{f_ABT} (ft)	225
Breakthrough Pressure (psi)	4098
Corey Oil Exponent—Fracture, n'	1.46
Corey Water Exponent—Fracture, m'	5.66
Total Objective Function (millions)	34.7

comparable to a basic GA, although several specialty operators are included to improve the results of the algorithm. The algorithm is trial-based rather than generational-based, and therefore to mimic the generational approach used by the GAPS MO GA, 10,000 trials were conducted (equivalent to 100 generations with a population of 100). A convergence criteria for maximum change in the OF is also used as an input for termination of the algorithm, although this was not achieved. The input parameters used for this algorithm are shown below in Table 12. A cross-over rate of 50% is used as this is the default setting in the program, meaning that each child receives half of its genes from each parent. Changing the cross-over rate could significantly impact algorithm performance and can be changed during an optimization run.

As with other SO algorithms, a single best solution is found by the algorithm. The parameters resulting from the optimization are given in Table 13, and the resulting combined OF is ~16% lower than that of the deterministic match.

The best solution was found in the 7455th trial, although only 245 trials were required to get within less than 1% of the best solution, suggesting that significantly fewer trials could have been run for this particular scenario. Fewer trials, however, would limit the search extent of the algorithm, which may lead to poor results in some cases, as the majority of the early trials produce significantly higher OF numbers. Figure 8(a) shows the average and minimum OF for the 100 equivalent generations (100 trials is equal to 1 generation). From the average curve the steady-state nature of the algorithm becomes apparent. Unlike a generational GA, where you would expect to see the generational average decrease over time, in this case the average decreases for approximately 15 equivalent generations before beginning to fluctuate between 37 million and 47 million for the remaining 85 equivalent generations. From the minimum curve it can be seen that a value within 1% of the minimum is found (within the third equivalent generation) and remains relatively constant for the remainder of the equivalent generations. This result can be seen by plotting the algorithm's improvement progress Figure 8(b). The logarithmic x-axis is used to better show the optimization progression. Lastly, Figure 8(c) provides the OF value per trial, and

Table 12. Input parameters for excel's single-objective evolutionary solver.

Crossover Rate (%)	50%
Mutation Rate (%)	15
Population Size	100
Number of Iterations	10,000
Convergence Criteria for Max Change (%)	0.01

Table 13. Best match parameters for excel's single-objective evolutionary solver.

Uncertain Parameter	Value
Fracture Permeability (md)	3102
BBT Drainage Area (Ac)	14
BBT Half-Length, x_{f_BBT} (ft)	425
ABT Drainage Area (Ac)	13
ABT Half-Length, x_{f_ABT} (ft)	4099
Breakthrough Pressure (psi)	14
Corey Oil Exponent—Fracture, n'	1.45
Corey Water Exponent—Fracture, m'	5.61
Total Objective Function (millions)	34.7
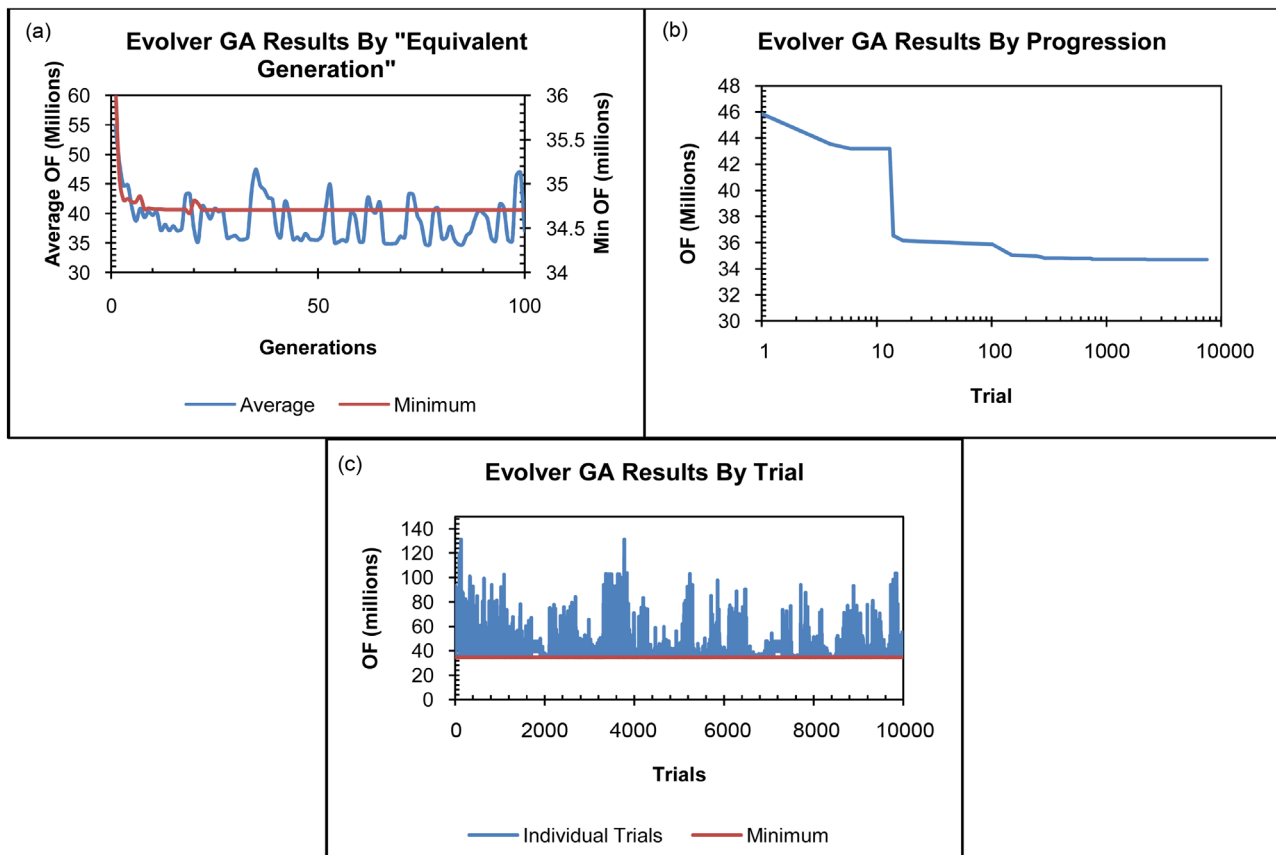
Figure 8. Evolver™ SO GA results: (a) By equivalent generation, showing both the equivalent generation average and minimum; (b) By progression step; and (c) By trial, also showing the minimum achieved value.

it can again be seen that values approaching the minimum are found quite quickly and continue to be found throughout the remainder of the optimization.

GAPS Multi-Objective Genetic Algorithm. In this section, the results of the only MO GA tested will be demonstrated. This is the GAPS algorithm developed by Mohammed Kanfar for the Tight Oil Consortium at the University of Calgary, and is based on the NSGA-II algorithm as discussed previously. The benefits of using MO algorithms were discussed previously, so in this section, the focus will solely be on the results of the algorithm. Although it is common practice in the application of GAs to run half the number of generations as the population size, in this application an equal number of generations and populations were conducted to allow the algorithm to "dig deeper" towards an absolute minimum. Note that larger population sizes allow the algorithm to explore further in the search space. The impact of running more generations will be discussed below. The algorithm was run with 100 generations with populations of 100 following the recommendations of [5], who ran 50 generations with populations of 100 in a similar application using numerical rather than analytical simulation (for a total of 10,000 runs). The input parameters used for this algorithm are shown given in Table 14.

As is the case with all MO Gas, the final generation does not converge to a single solution, but instead converges to a Pareto Front of nondominant mathematically-equivalent solutions. In this case, the Pareto Front is convex in nature, which suggests that two phase rate objectives are conflicting (a straight-line would suggest non-conflicting OFs). To converge on a single best solution, the solutions were filtered, removing solutions that have an OF higher than a certain threshold (with the threshold being continuously reduced until only several solutions remained around the corner point of the Pareto Front), and then visual inspection was used to pick the final solution. There are currently no methods available in the literature for selecting the single best solution, and therefore an approach similar to that used by [5] was used in this application due to the similarity of the problems. The evolution of the Pareto Front from generation to generation will first be investigated. The advancing Pareto Front is shown in Figure 9. Figure 9(a) provides the final generation, along with the other generations being shown in groups of ten generations. From this plot it can be seen that there is a large amount of scatter in the first 10 generations, but by the second ten generations convergence on the ultimate Pareto Front begins. For the Pareto Front, a semi-log presentation was chosen with the oil rate OF being on a log scale, while the water rate OF is plotted on a Cartesian scale, as this was found to best demonstrate the results in this case (however a Cartesian plot will

Table 14. Input parameters for GAPS multi-objective genetic algorithm.

Mutation Rate (%)	15
Population Size	100
Number of Generations	100
Number of Iterations	10,000
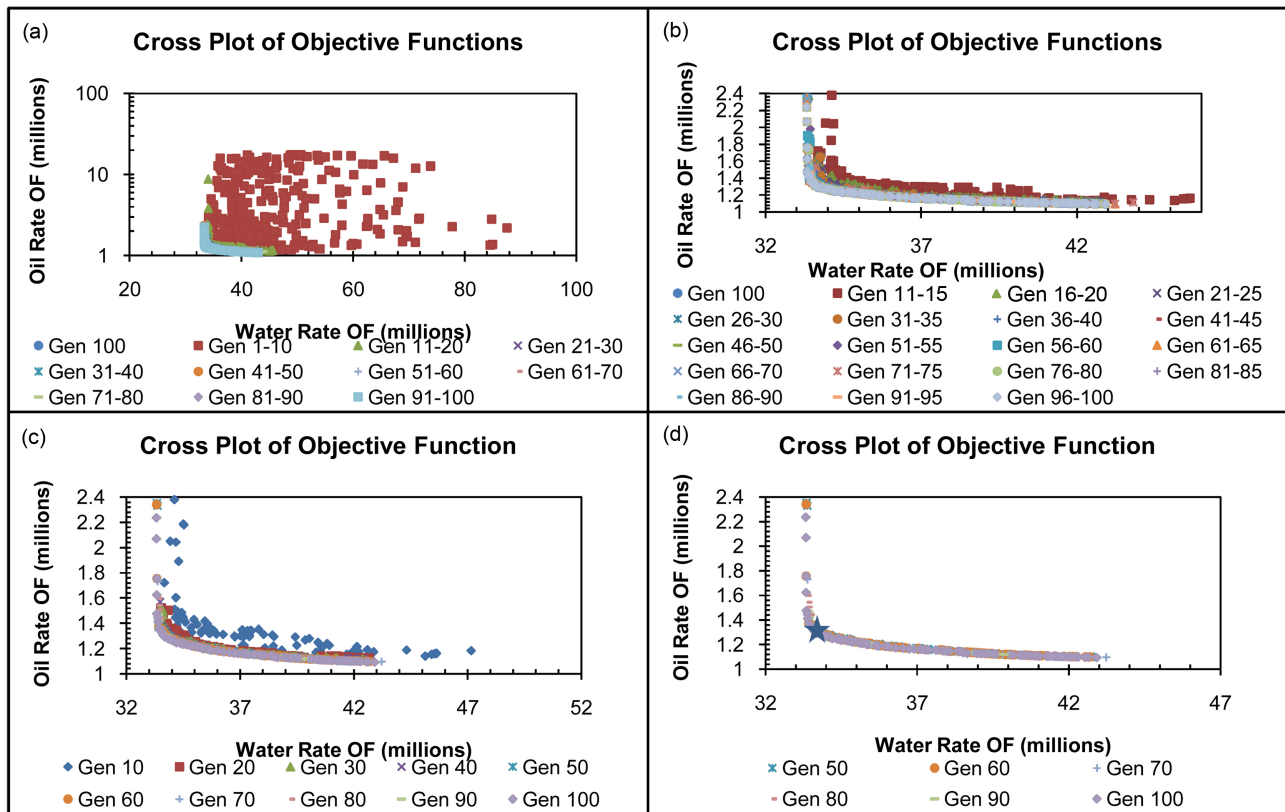
Figure 9. Pareto diagram for flowback history-match: (a) Generations grouped into sets of 10 generations showing significant scatter in the first 10 generations; (b) Generations grouped into sets of 5 generations starting at generation 11; (c) Every 10th generation to show advancement of Pareto Front over time; (d) Every 10th generation between Generation 50 and Generation 100 to demonstrate convergence on the ultimate Pareto Front. The single best solution is shown with a star.

be used in **Figure 9(d)**). In **Figure 9(b)**, the first ten generations are eliminated and the remaining 90 generations are broken into groups of 5. Form this plot it can be seen that there is consistent improvement for approximately 50 generations prior to converging on the ultimate Pareto Front. This can also be seen in **Figure 9(c)**, which shows every tenth generation starting at Generation 10. Finally, in **Figure 9(d)**, the final 50 generations are shown and it can be seen that there is no obvious improvement beyond 50 generations and therefore the population to generation ratio of two used by [4] as well as many others when applying GAs is suggested for future applications of this algorithm. This will reduce runtime by 50% without having a significant impact on the final generation results. The single best solution selected using the method described above is shown with a star in **Figure 9(d)**, and it can be seen that this point is near the corner point of the Pareto Front, suggesting relatively equivalent trade-off between the two objectives. From **Figure 9** it can also be seen that the average value of the water rate OF is ~20× greater than that for the oil rate OF. This is due to the fact that water rates are much larger than the oil rates, and therefore visually similar deviations in rate will be approximately an order of magnitude higher. This information could have been used with the SO algorithms to reduce potential bias towards achieving a better water than hydrocarbon history-match,

although an equal weighting appears to still be effective for this problem based on the results of the proceeding and following sections.

Next, generation 100 will be investigated in greater detail, focusing primarily on the extent of variability in the key parameter estimates during this final generation. The parameters corresponding to the best match and the average, standard deviation and P10/P90 ratio for Generation 100 are given in Table 15. The best match leads to a summed OF which is ~16% lower than that of the deterministic match.

Palisade Evolver's OptQuest^{TM} Algorithm. In this section the results of Palisade Evolver's OptQuest^{TM} will be demonstrated. Much like Evolver's GA, this is a SO algorithm. This algorithm has its basis in Scatter Search which draws many similarities to GAs, although also includes integer programming, Tabu Search and an Artificial Neural Network to improve its results and efficiency, as discussed previously. The algorithm is trial-based much like Evolver's GA. In this case, since there is no basis for comparison of the algorithm, the maximum number of trials was set to a very large value (100,000) allowing the convergence criteria for maximum change in the OF to control the termination of the optimization. The input parameters used for this algorithm are shown below in Table 16.

As with other SO algorithms a single best solution is found by the algorithm. The parameters resulting from the optimization are found in Table 17 and the resulting combined OF is ~16% lower than that of the deterministic match.

In this particular case, 33,756 trials were required to reach the set criteria, although a value with a combined OF within 1% of the optimal value was found in 13,421 trials which equates to a ~60% reduction in optimization time, although many significantly higher OF values were found in the final 20,000 trials. To

Table 15. GAPS multi-objective genetic algorithm generation 100 results.

Uncertain Parameter	Optimal Match	Average	Std. Dev.	P10/P90
Fracture Permeability (md)	3136	3,584	290	1.24
BBT Drainage Area (Ac)	13.98	13.99	0.21	1.00
BBT Half-Length, $x_{f\,BBT}$ (ft)	440	440	0.20	1.00
ABT Drainage Area (Ac)	12.99	12.99	0.01	1.00
ABT Half-Length, $x_{f\,ABT}$ (ft)	424	424	0.20	1.00
Breakthrough Pressure (psi)	4099	4093	0.01	1.00
Corey Oil Exponent—Fracture, n'	1.45	1.50	0.03	1.05
Corey Water Exponent—Fracture, m'	5.87	7.89	1.57	1.57

Table 16. Input parameters for palisade's single-objective genetic algorithm.

Maximum Number of Iterations	100,000
Convergence Criteria for Max Change (%)	0.01

Table 17. Best match parameters for evolver's single-objective OptQuest algorithm.

Uncertain Parameter	Value
Fracture Permeability (md)	3148
BBT Drainage Area (Ac)	14
x_{f_BBT} (ft)	441
ABT Drainage Area (Ac)	13
x_{f_ABT} (ft)	425
Breakthrough Pressure (psi)	4099
Corey Oil Exponent—Fracture, n'	1.46
Corey Water Exponent—Fracture, m'	5.65
Fracture Permeability (md)	34.7

allow comparison with the GAs, the results were filtered into "equivalent generations" of 100 trials. Figure 10(a) provides the average and minimum OF for the 338 "equivalent generations" (the 338th "equivalent generation only contains 56 trials"). From the average curve, the differences between OptQuest's performance and a generational GA become apparent. Unlike a generational GA, where one would expect to see the generational average go down over time, in this case the average decreases for approximately 10 "equivalent generations" prior to stabilization with four groups of "equivalent generations" with significantly higher values which occur when the algorithm tries radically different areas of the search space. This is characteristic of a Scatter Search Algorithm which utilizes Tabu Search and an Artificial Neural Network to stop the algorithm from going back to areas of the search space which either have, or are expected to yield inferior solutions. From the minimum curve, it can be seen that a value within 1% of the minimum is found in the 133rd "equivalent generation" and remains relatively constant for the remainder of the "equivalent generations". This result can be seen by plotting the algorithms improvement progress which is shown in Figure 10(b). Lastly, Figure 10(c) shows the OF value per trial, and it can again be seen that values approaching the minimum are found quite quickly and continue to be found throughout the remainder of the optimization. The same impact can be seen as in the "equivalent generation" case in Figure 10.

Summary of Results. In the previous sections the results of several techniques including: 1) Deterministic Analysis; 2) MC MO simulation (Palisade® @RISK™); 3) Microsoft® Excel's SO Gradient-based (GRG2) algorithm (GRG Nonlinear Solver); 4) Microsoft® Excel's SO Evolutionary Solver; 5) Palisade® Evolver's SO GA; 6) GAPS MO GA (based on the NSGA-II) algorithm; and 6) Palisade® Evolver's SO OptQuest™ algorithm were discussed individually. In this section the results of the different techniques will be compared. In Figure 11 the history-match to both the water and oil phases is shown for each of the techniques, while in Figure 12 the key parameters and total OF's are provided. The results from Excel's GRG Nonlinear Solver have not been included s this required
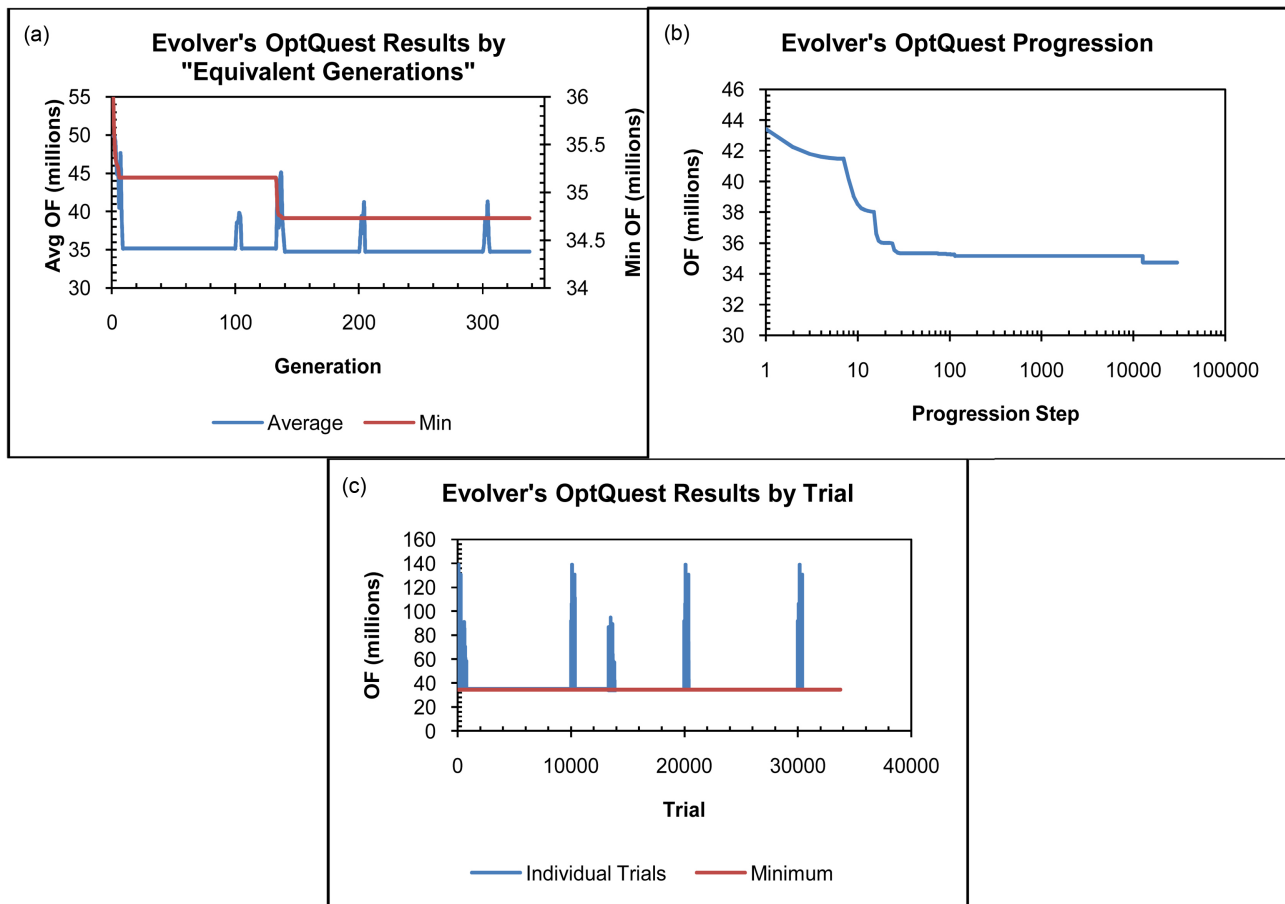
Figure 10. Evolver's SO OptQuest results: (a) By equivalent generation showing both the equivalent generation average and minimum; (b) By progression step; and (c) By trial also showing the minimum achieved value.

manipulation of the initial guess to achieve an acceptable history-match (although its key match parameters will be discussed below).

From **Figure 12** it can be seen that the deterministic match matches the early-time hydrocarbon production better than the other algorithms, although it provides a far less superior late time match to the two hydrocarbon phases. The late-time hydrocarbon match can be improved further by increasing breakthrough pressure, although it was determined that this leads to premature breakthrough by the model and therefore an upper limit of 4100 psia was enforced. It can also be seen that the water match from each of the solving techniques is similar. Further the hydrocarbon rate profiles for all but the deterministic history-match look similar. Additional key observations based on **Figure 12** are presented below:

- Fracture permeability ranges from 3102 md to 3190 md with the lowest value coming from Evolver's GA and the highest coming from the average of the top five MC simulations. Each of the algorithms finds a fracture permeability ~350 md lower than the deterministic match (~10% difference). The percent variability from the five algorithms is approximately is ~2.5% when compared to the deterministic match.
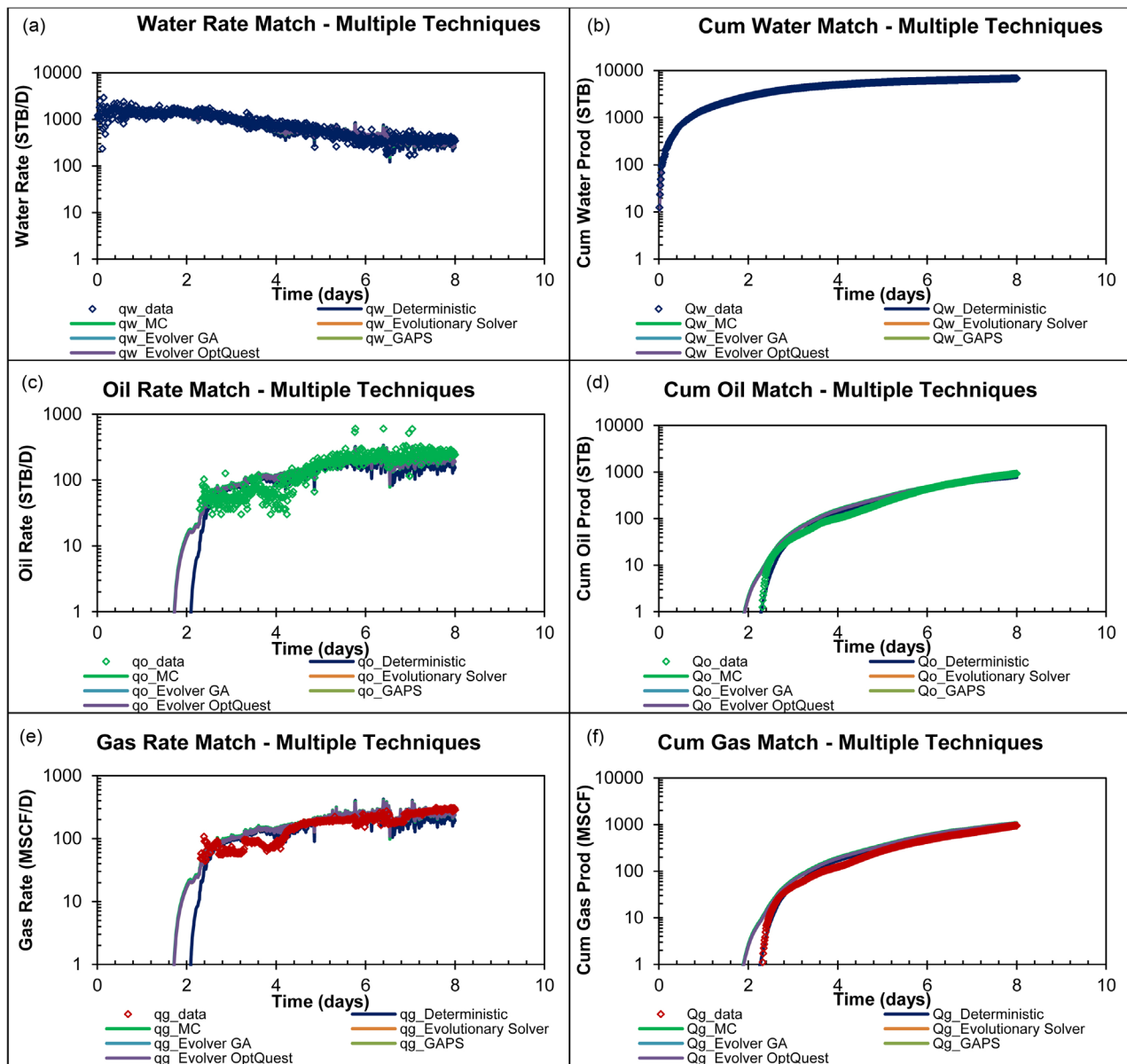
Figure 11. Flowback history-match using different algorithms: (a) Water rate match; (b) Cumulative water production match; (c) Oil rate match; (d) Cumulative oil production match; (e) Gas rate match; and (f) Cumulative gas produced match.

- Breakthrough pressure approaches the upper limit for each of the five algorithms and is significantly higher than the deterministic match (~7%). An earlier breakthrough yields a better late-time oil match, which is where oil rates are highest and therefore have greatest potential to add to the OF value. This is also the piece of data were the deterministic solution deviates most from the measured data. As mentioned previously, a breakthrough pressure of greater than 4100 psia leads to premature breakthrough, although also yields a better late-time history-match. A breakthrough pressure of 4100 psia suggests a 10% supercharge of the formation directly surrounding the fractures which results from pumping the fracture at significantly higher pressures than formation pressure (mini water flood effect).
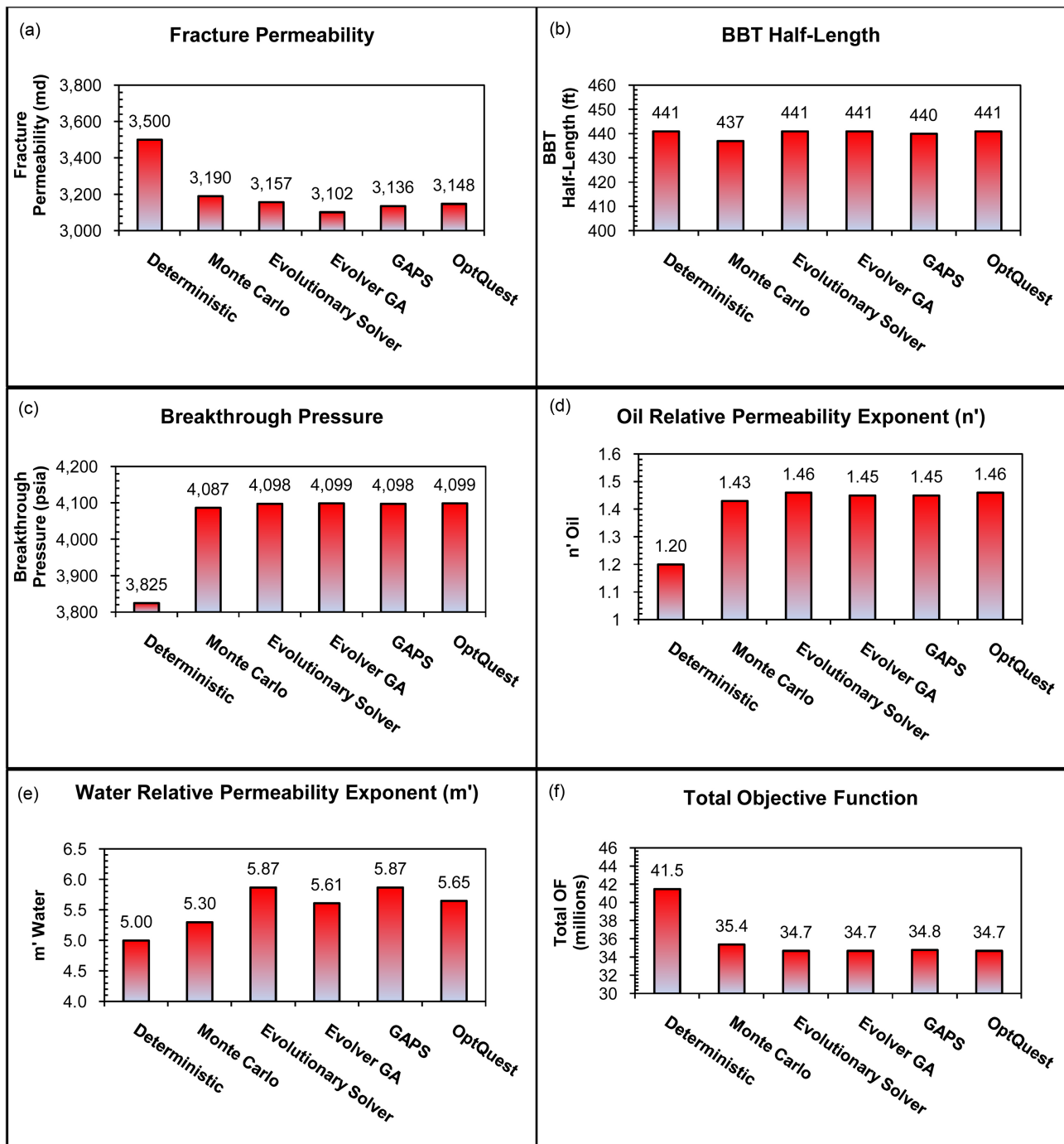
Figure 12. Flowback key history-match parameters found using different algorithms: (a) Fracture permeability; (b) BBT half-length; (c) Breakthrough Pressure; (d) Oil relative permeability exponent, *n*'; (e) Water relative permeability exponent, m; and (f) Total OF value.

- BBT half-length is nearly constant using each of the six techniques, ranging from 437 - 441 ft which is to be expected given the rather definitive results of the FMB shown above. The deterministic history-match used the same BBT half-length as the four main assisted history-matching techniques.
- Oil relative permeability exponent shows almost no variability from the five

algorithms ranging from 1.43 - 1.46. This is ~20% higher than the value used in the deterministic history-match.

- Water relative permeability exponent shows slightly more variability from each of the five algorithms, ranging from 5.30 - 5.87. Each of the algorithms predicted a water exponent exceeding that of the deterministic history-match by an average of ~4%. The percent variability from the five algorithms is approximately is ~11.4% when compared to the deterministic match.

- The total OF for the four assisted history-matching algorithms was nearly identical, ranging from 34.7 - 35.4 million. The average of the top five MC simulations was ~2% higher than the other assisted history-matching techniques. The five different algorithms improved the total OF from 14.7% - 16.4%, although this suggests that the deterministic match still falls within the ±20% range often accepted in industry in this particular case.

The above results demonstrate that each of the algorithms find a very similar optimal value for each of the key parameters suggesting that this likely represents the global optimum. After reviewing the total OF, it is clear that there is significant benefit to applying these algorithms once bounds on key parameters can be estimated. Another interesting observation is that the deterministic history-match yielded values within 10% of the optimal values for three out of the five uncertain parameters. The only exceptions are the relative permeability exponent to oil water, which varied by ~20% and ~11% respectively. This higher differential can be attributed to the low values of these exponents, making them particularly sensitive when calculating percent difference (although the absolute value was within 0.25 and 0.66 of the average optimal values respectively).

Based on the results shown above, it would be expected that application of Excel's GRG Non-Linear Solver with the use of multi-restart mode would likely yield the same results. This was not tested to its full extent in this analysis, although using the deterministic history-match as an initial guess led to similar parameters as those solved by the other algorithms. This result suggests that the multi-restart method would likely be successful in this problem and also demonstrates that there are no local minima between the deterministic match and global optimum.

4. Discussion

The basis of this work is the tool developed by [1] for analyzing multi-phase (water, oil and gas) flowback data from MFHWs following hydraulic fracture stimulation to estimate key fracture properties such as effective fracture half-length and fracture permeability. The base tool, with the modifications discussed previously, was then used to conduct a deterministic history-match. Following the deterministic history-match, MC simulation was used to determine the variability in the key history-match parameters which can be used to effectively match the data (rate OF for oil and water lower than the deterministic match). Once this analysis was conducted, the results of the best MC simulations

were compared to the results of several assisted history-matching techniques in an attempt to find the global optimum (which corresponds to the "true" fracture parameters, assuming the model and other hard inputs are correct). Algorithm complexity varied from Excel's GRG Non-Linear Solver, which is based on GRG, to SO and MO GAs, and an algorithm known as OptQuest™ which combines several optimization techniques into a single algorithm. It was demonstrated that each technique could essentially locate the same optimal set of parameters, suggesting that this corresponds to the absolute minimum rather than a local minimum, which in turn led to a significant improvement in history-matching over the deterministic analysis. Despite the versatility of the methods described, there are several areas which warrant further discussion.

Two of the biggest challenges when applying MC simulation and other assisted history-matching techniques are: 1) selecting which variables to consider unknowns; and 2) developing an input distribution for the unknowns. These methods are typically most successful and converge faster when the number of inputs is limited to the minimum possible number with the smallest range to minimize the search space for the algorithm. In the case of flowback analysis, there are many uncertain inputs making this a difficult problem to solve using these methods, and therefore it is important to select the most important parameters as uncertain (*i.e.* fracture half-length and conductivity), while assuming that some less critical inputs that are constant (*i.e.* initial fracture pressure and fracture porosity). The next challenge is developing an input distribution for the uncertain parameters (particularly for MC simulation). In an ideal scenario, the input distributions can be developed from existing data allowing for greater precision and ultimately better output results, although this requires a significant amount of analogous data. For some scenarios, such as history-matching long-term production from wells with a significant number of analogs which have all been analyzed, this is feasible. Further, in many cases parameters such as matrix permeability have been demonstrated extensively in the literature to show a lognormal distribution. Unfortunately this is not the case with flowback analysis, where the data set is generally limited, or in many cases non-existent, due to the very new nature of industry interest in analyzing this data and lack of widespread (although rapidly growing) application. For example, the basic techniques used in this work have been applied by several companies including in an SPE paper written by [20]. Due to a lack of data for developing an input distribution, a simple uniform distribution was used in this work for the 5 selected input parameters, where the distribution range was limited as much as possible using available offset analysis as well as the deterministic history-match. As these methods gain further traction in industry, and are applied to more wells, developing better input distributions will likely be possible making the application of these methods more versatile.

Another challenge is determining an acceptable number of iterations (*i.e.* the number of generations and population size in the GAPS algorithm) to allow achieving reasonable results while minimizing run time to make the application

of the techniques to a large number of wells more feasible. In this work, the purpose was to demonstrate the applicability of the different techniques used, and therefore run time was not a consideration, although this will become more important as these techniques continue to gain traction in industry. In this case, other than the Excel™ Solver methods, each technique required multiple days of run time making the techniques not practically applicable to a large number of wells. Further, the tool is still in the research phase, and could be made significantly more efficient (~1 iteration per second comparable to other similar commercial tools), which would also help to significantly reduce run time. Trying to determine an acceptable number of iterations is an area of future work which will require application to more than the several wells which have been analyzed using these techniques.

In this paper, it was demonstrated that Excel's GRG Non-Linear Solver is highly ineffective when a relatively generic initial guess is used, as this algorithm will find the closest local minima to the initial guess. When the deterministic solution was used as the initial guess, the algorithm converged to parameters similar to the other techniques applied, suggesting there is no local minima between the deterministic solution and the optimal solution. This may not always be the case, and in some applications a complete deterministic analysis may not be conducted prior to applying an assisted history-matching algorithm. The convergence speed of this algorithm makes it ideal, although its application clearly has limitations. One solution is to apply the multi-restart techniques discussed previously, where the algorithm will run for a series of different initial guesses in attempt to find the global minima. It is likely that a substantial number of restarts would be required to find the optimal solution for the flowback problem, and therefore extensive testing would be required before confidently applying this technique and determining how its run time compares to the other algorithms tested. The standard version of Solver available in Excel™ does not offer a multi-restart option, although the developer of this solver (Frontline Solver's) offers more advanced versions which include this option as well as further improvements and additional algorithms.

In this work, six techniques were applied for assisted history-matching purposes. These methods were selected as they were either developed within the research group (GAPS algorithm) or commercially available from reputable vendors that are used extensively in industry (Microsoft® and Palisade®).

Future work on this topic will focus on the application of addition algorithms available both commercially and in the literature. Testing of further MO GA's would be of particular interest as they overcome the biggest challenge of SO algorithms which were the primary focus of this work. Testing further techniques is warranted, seeking algorithms which converge faster and/or are potentially more effective in consistently finding the optimal solution. A detailed investigation, using multiple examples (both simulated and field examples) will also be conducted to determine the number of iterations required to achieve the desired

result. This will allow for a better comparison of both convergence speed and accuracy which was not directly addressed in this paper.

5. Conclusions

In this work, several algorithms were tested for the purpose of uncertainty analysis and assisted history-matching of flowback data. In previous work, [16] [17] applied MC simulation to the same data set, although less iterations were conducted with significantly more input parameters and with wider bounds, bringing the results into question. The GAPS algorithm has also been tested by [4] for history-matching three-phase flowback with a numerical simulator and [3] attempted to use a combination of algorithms to try to decouple parameters in one of their analysis tools, although more rigorous methods could have been applied. Other than these limited studies, uncertainty quantification and assisted history-matching have not been investigated for application to the flowback problem. The main conclusions of this study are as follows:

- MC simulation can effectively be applied for both uncertainty quantification and assisted-history matching, assuming enough trials are conducted to effectively cover the search space. For practical application, this limits the number of uncertain parameters and the distribution range for these parameters.

- As anticipated, application of a gradient-based algorithm was not successful unless a very good initial guess was provided. This is due to the nature of the algorithm limiting its application in the absence of using the multi-restart feature.

- Each of the techniques tested (excluding Excel's GRG Non-Linear Solver), including both SO and MO techniques, was able to converge to a very similar optimal solution, suggesting that they were likely finding the global optima. There are often problems associated with applying SO algorithms to MO problems due to competing objectives, although this issue did not appear to arise in the analyzed well. It was demonstrated that each of these techniques provided a significant improvement in history-match quality over a single deterministic analysis, although deterministic history-matching is useful in determining which parameters should be considered uncertain and constraining the range of these uncertain parameters.

- Further testing is warranted to determine the wide-spread applicability of these techniques, and to reduce run time making the application more desirable for industry applications.

- Additional algorithms should be investigated for a larger number of wells to determine which techniques are most applicable to the flowback problem. Specifically, testing additional MO algorithms would be desirable as these algorithms tend to better represent the problem. It is possible that MO algorithms other than the GAPS algorithm could provide better results for flowback analysis.

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Clarkson, C.R., Qanbari, F. and Williams-Kovacs, J.D. (2014) Innovative Use of Rate-Transient Analysis Methods to Obtain Hydraulic-Fracture Properties for Low-Permeability Reservoirs Exhibiting Multi-Phase Flow. *The Leading Edge*, **33**, 1108-1122. https://doi.org/10.1190/tle33101108.1

[2] Williams-Kovacs, J.D. (2017) Quantitative Analysis of Multi-Phase Flowback from Multi-Fractured Horizontal Wells. Ph.D. Thesis, University of Calgary, Calgary, Alberta.

[3] Ezulike, D.O., Dehghanpour, H., Virues, C.J., Hawkes, R.V. and Jones Jr., R.S. (2016) Flowback Fracture Closure: A Key Factor for Estimating Effective Pore Volume. *SPE Reservoir Evaluation & Engineering*, **19**, 567-582. https://doi.org/10.2118/175143-PA

[4] Kanfar, M.S. and Clarkson, C.R. (2016) Reconciling Flowback and Production Data: A Novel History Matching Approach for Liquids Rich Shale Wells. *Journal of Natural Gas Science and Engineering*, **33**, 1134-1148. https://doi.org/10.1016/j.jngse.2016.04.042

[5] Pruess, K. (1985) A Practical Method for Modeling Fluid and Heat Flow in Fractured Porous Media. *Society of Petroleum Engineers Journal*, **25**, 14-26. https://doi.org/10.2118/10509-PA

[6] Palisade Corporation (2015) @RISK™ User's Guide: Risk Analysis and Simulation Add-In for Microsoft® Excel™. Palisade Corporation, Ithaca, NY.

[7] Lasdon, L.S., Waren, A.D., Jain, A. and Ratner, M. (1978) Design and Testing of a Generalized Reduced Gradient Code for Nonlinear Programming. *ACM Transactions in Mathematical Software*, **4**, 34-50. https://doi.org/10.1145/355769.355773

[8] Arun, C. and Tayo, O.A. (2014) Design and Optimization of Single Span Fixed-Feet Portal Frame Using Generalized Reduced Gradient. *Civil and Environmental Research*, **6**, 20-29.

[9] Laguna, M. (2011) OptQuest: Optimization of Complex Systems. OpTek Systems Inc., Greenville, SC.

[10] Frontline Solvers (2016) Optimization and Simulation User Guide. Frontline Systems Inc., Incline Village, Nevada.

[11] Palisade Corporation (2015) Evolver™ User's Guide: The Genetic Algorithm Solver

for Microsoft® Excel™. Palisade Corporation, Ithaca, NY.

[12] Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. (2002) A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, **6**, 182-197. https://doi.org/10.1109/4235.996017

[13] Srinivas, N. and Deb, K. (2007) Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evolutionary Computation*, **2**, 221-248. https://doi.org/10.1162/evco.1994.2.3.221

[14] Seada, H. and Deb, K. (2015) U-NSGA-III: A Unified Evolutionary Algorithm for Single, Multiple, and Many-Objective Optimization. Evolutionary Multi-Criterion Optimization. 8*th International Conference*, *EMO* 2015, Guimarães, Portugal, 29 March-1 April 2015, 34-49.

[15] Glover, F., Kelly, J.P. and Laguana, M. (1996) New Advances and Applications of Combining Simulation and Optimization. *Proceedings of the* 28*th Conference on Winter Simulation*, Coronado, CA, 8-11 December 1996, 144-152. https://doi.org/10.1145/256562.256595

[16] Williams-Kovacs, J.D. and Clarkson, C.R. (2013) Stochastic Modeling of Two-Phase Flowback of Multi-Fractured Horizontal Wells to Estimate Hydraulic Fracture Properties and Forecast Production. *SPE Unconventional Resource Conference*, The Woodlands, TX, 10-12 April 2013, SPE-164550-MS. https://doi.org/10.2118/164550-MS

[17] Williams-Kovacs, J.D. and Clarkson, C.R. (2013) Stochastic Modeling of Multi-Phase Flowback from Multi-Fractured Horizontal Tight Oil Wells. *SPE Unconventional Resources Conference*, Calgary, Alberta, 5-7 November 2013, SPE-167232-MS.

[18] Oudinot, A.Y., Koperna Jr., G.J. and Reeves, S.R. (2005) Development of a Probabilistic Forecasting and History Matching Model for Coalbed Methane Reservoirs. 2005 *International Coalbed Methane Symposium*, Tuscaloosa, Alabama, 16-20 May 2005.

[19] Roadifer, R.D., Farnan, R.A., Crabtree, B.J., Raterman, K.T. and Moore, T.R. (2003) History Matching (Reservoir Parameter Estimation) for Coalbed Methane Wells via Monte Carlo Simulation. 2003 *International Coalbed Methane Symposium*, Tuscaloosa, Alabama, 5-9 May 2003.

[20] Cugnart, R., Rasoanaivo, S., Williams-Kovacs, J.D., Raverta, M. and Clarkson, C.R. (2017) Early Time SRV Characterization through Flowback Analysis: Application of Clarkson/Williams-Kovacs Technique to Vaca Muerta. *SPE|AAPG|SEG Unconventional Resources Technology Conference*, Austin, TX, 24-26 July 2017, URTEC-2689844-MS. https://doi.org/10.15530/urtec-2017-2689844

Nomenclature

Abbreviations

ABT = After-breakthrough

BBT = Before-breakthrough

CDF = Cumulative distribution function

DFIT = Diagnostic fracture injection test

FMB = Flowing material balance

FR = Flow-regime

GA = Genetic algorithm

GOR = Gas-oil ratio

GRG = Generalized reduced gradient

LTO = Light tight oil

MBT = Material balance time

MC = Monte Carlo

MINC = Multiple interacting continua approach

MO = Multi-objective

NSGA = Nondominated sorting genetic algorithm

PDF = Probability density function

PVT = Pressure-Volume-Temperature

RNP = Rate-normalized pressure

RNP' = Rate-normalized pressure derivative

RTA = Rate-transient analysis

SO = Single-objective

Field Variables

F_c = Fracture conductivity, md-ft

m' = Corey water relative permeability constant for the fractures, dimensionless

n' = Corey oil relative permeability constant for the fractures, dimensionless

p = Pressure, psia

p_{wf} = Sadface flowing pressure, psia

p^* = Extrapolated initial reservoir pressure, psia

q_o = Oil production (surface) flowrate, STB/D

q_w = Water production (surface) flowrate, STB/D

Q_o = Cumulative water production (surface), STB

Q_w = Cumulative production (surface), STB

r_{wa} = Apparent wellbore radius ($r_{wa} = r_w e^{-s}$), ft

w = Objective function weighting factor, dimensionless

x_f = Fracture half-length, ft

Dimensionless Variables

t_{Dd} = Dimensionless decline time

Subscripts

ABT = After-breakthrough

BBT = Before-breakthrough

BT = Breakthrough

D = Dimensionless variable

f = Fracture

o = Oil

T = Total

w = Water

wf = Sandface