Optimization of SAGD Process Accounting for Geological Uncertainties Using proxy Models

Eugene Fedutenko, Chaodong Yang, Colin Card, Long Nghiem

Computer Modeling Group Ltd.

Summary

Different studies have demonstrated that optimization of SAGD (Steam Assisted Gravity Drainage) operating conditions as well as well locations have a significant potential to improve the economics of SAGD operations. However, in order to capture the variability of optimization parameters of a particular reservoir, such optimization needs to run a large number of simulations, especially when geological uncertainties are taken into account. Running a large number of SAGD simulations is computationally very demanding. Therefore, optimization approaches that allow one to cut down the number of necessary simulations and still find the optimal solutions are highly beneficial.

This paper presents a robust proxy SAGD optimization that takes into account geological uncertainties of the reservoir. It finds optimum operating conditions and well locations (under an assumption of fixed injector / producer vertical separation) by using data interpolation in the N-dimensional parameter space. The approach utilizes the following iterative workflow. First, it ranks all realizations according to the performance of each realization in terms of NPV (Net Present Value). Based on that ranking, a small set of representative realizations are chosen to represent the overall uncertainty of the reservoir. With the selected representative realizations, a robust optimization is performed in the framework of a proxy model. A specified number of initial simulations capturing the parameter variability for selected realizations are launched by using the Latin Hypercube sampling approach. Based on the simulation results, an interpolation proxy model is built. The proxy model is then used to find the optimal parameter values. After that the optimal proxy solutions are verified by running the actual simulations with the corresponding parameter values. Furthermore, the approach allows one to update the proxy interpolator with the optimal simulation results and to iterate over it as many times as necessary. To consider the effect of reservoir uncertainty on optimization results, the workflow utilizes a risk-weighted objective function which is based on the expected value and standard deviation of NPV over the set of the selected realizations. In this study, two types of interpolation proxy models - the second order polynomial and the ordinary Kriging - are investigated in the SAGD optimization process.

The workflow is applied to a typical 3-Well-Pair SAGD model. Both polynomial and Kriging proxy results are compared with the base case (optimization without proxy) results and between themselves. The comparison shows that both polynomial and Kriging proxy optimizations significantly improve the NPV, SOR (Steam-Oil Ratio) and RF (Recovery Factor) outputs of the model. It is also shown that the ordinary Kriging proxy model can find an NPV value that is very close to the optimum value much sooner and, therefore, can significantly cut down the overall simulation time.

Theory and/or Method

Introduction

SAGD is a thermal oil recovery process, which consists of using pairs of parallel horizontal wells drilled near the bottom of the pay (Butler, 1991). The top well is used to inject steam, while the bottom one is used to produce reservoir fluids. The steam injected from the top well rises into the formation, forming an expanding steam chamber around and above the injection well. The rising steam eventually loses its latent heat near the boundary of the steam chamber, heats the oil and allows it to drain to the bottom production well by gravity.

Recent progress in computational hardware and software development has opened new frontiers in reservoir modeling in general and in the simulation of SAGD processes in particular. During the last ten years, a number of simulation studies have been conducted to optimize SAGD operations. Egermann et al. (2001) proposed a method to obtain and maintain an optimized development of the steam chamber throughout the production life of the well pair. Gates et al. (2005) used a genetic algorithm to optimize the cumulative steam-oil ratio (CSOR) by altering the steam injection pressure in a generic two-dimensional McMurray reservoir model. Yang et al. (2009) applied the DECE (Designed Exploration and Controlled Evolution) optimization algorithm to optimize the NPV of a real field three-dimensional reservoir with two SAGD well pairs in the presence of bottom water. Kumar et al. (2010) investigated the impact of geological heterogeneity on SAGD wellbore design and the optimization of the length and positioning of multiple tubular strings as well as the allocation of injected steam among multiple tubing strings. Peterson et al. (2010) conducted an optimization study on the application of a solvent-additive SAGD process to reservoirs with associated basal water.

In order to overcome the limiting factor of simulation time, proxy models have been proposed to solve many optimization problems in field development. Vanegas Prada et al. (2008) discussed the assessment of optimal operating conditions in an Athabasca SAGD process with a polynomial response surface approach. Fedutenko et al. (2011) studied both polynomial and ordinary Kriging proxy workflows in the framework of the 3-Well Pair SAGD model.

These above SAGD optimization studies ignored geological uncertainties of the reservoir and were based on a single realization. Due to the significant impact of reservoir heterogeneity on SAGD performance, the optimal solution obtained based on a single realization may deviate significantly from the actual optimality. To address this problem, Yang et al. (2011) presented a practical robust optimization workflow that significantly reduced computational cost and yet still was able to account for the overall geological uncertainties of the reservoir in the framework of the polynomial response.

The goal of this paper is to expand on the results obtained by Yang et al. (2011) and to present the generic robust optimization workflow that incorporates two different types of the interpolation proxy models: the polynomial and the ordinary Kriging ones. The workflow is applied to a typical 3-Well-Pair SAGD model. Both polynomial and Kriging proxy results are compared with the base case results and between themselves. The comparison shows that both proxy optimizations significantly improve the NPV output of the model. It is also shown that the robust optimization in the framework of ordinary Kriging proxy can be performed with a much smaller number of initial simulations and, therefore, can be used to significantly cut down the overall simulation time.

Robust Optimization

Due to the noisy or sparse nature of seismic data, core samples, and well logs, uncertainty is an inherent characteristic of any geological model. The unique true distribution of reservoir properties is usually unknown. To quantify the uncertainty in a geological model, geo-statistical methods are used to

construct multiple equally probable realizations of reservoir properties. This corresponds to a discretization of the unknown uncertainty space Θ into a finite number of realizations θ_d.

$$\theta_d = \{\theta_1, \theta_2, \cdots, \theta_{N_T}\} \in \Theta, \qquad (1)$$

where N_T is the total number of geo-statistical realizations. Suppose that the performance measure (objective function) of the SAGD process is J, which becomes a function of optimization parameters $\vec{x}$ and geo-statistical realizations θ_d:

$$J = J(\vec{x}, \theta_d) \qquad (2)$$

To account for the uncertainty represented by multiple realizations, the objective function in robust optimization J_{RO} can be defined as (Van Essen et al. 2006):

$$J_{RO} = E_{\theta_d}\left[J(\vec{x}, \theta_d)\right] - r \cdot \Sigma_{\theta_d}\left[J(\vec{x}, \theta_d)\right], \qquad (3)$$

where E_{θ_d} represents the expected value over the discretized uncertainty space θ_d, Σ_{θ_d} is standard deviation, and r is the risk aversion factor. Assuming that the realizations are equiprobable, the robust optimization (RO) objective function can be further written as:

$$J_{RO} = \bar{J} - r \cdot \sqrt{\frac{1}{N_T - 1} \sum_{i=1}^{N_T}(J(\vec{x}, \theta_i) - \bar{J})^2} \quad , \qquad (4)$$

where

$$\bar{J} = \frac{1}{N_T} \sum_{i=1}^{N_T} J(\vec{x}, \theta_i) \qquad (5)$$

The above equations indicate that each RO objective function evaluation requires a simulation run for each realization individually. For $N_T = 100$ realizations, the simulation time will be 100 times longer than the time needed for the optimization based on a single realization. Clearly, it can be costly to use all the N_T realizations in the robust optimization process. On the other hand, randomly choosing a small set of geo-statistical realizations will not accurately represent the uncertainty. For the purpose of selecting representative realizations that span production uncertainty, several ranking methods have

been investigated for the SAGD process (McLennan and Deutsch, 2005; Fenik et al. 2009). In this paper, the set of realizations is ranked with respect to the final NPV produced by the correspondent simulation. After these N_T realizations are reliably ranked, a small set of representative realizations can be chosen for robust optimization:

$$\theta_R = \{\theta_1, \theta_2, \cdots, \theta_{N_R}\} \in \theta_d = \{\theta_1, \theta_2, \cdots, \theta_{N_T}\} \in \Theta \qquad (6)$$

With this small set of representative realizations $\theta_R = \{\theta_1, \theta_2, \cdots, \theta_{N_R}\}$, the robust optimization objective function can be written as:

$$J_{RO} = \bar{J} - r \cdot \sqrt{\frac{1}{N_R - 1} \sum_{i=1}^{N_R} (J(\vec{x}, \theta_i) - \bar{J})^2} \quad \text{where} \quad \bar{J} = \frac{1}{N_R} \sum_{i=1}^{N_R} J(\vec{x}, \theta_i) \quad (7)$$

Robust Optimization Workflow

The workflow can be summarized in the following steps.

1. Use a geo-statistical method to generate N_T (usually≥100) realizations.

2. Rank N_T realizations by applying a reliable ranking method. An accurate ranking method is through direct numerical simulation of all realizations under a reference operating condition and then ranking the realizations according to the calculated performance measures. Note that this number is much smaller than the overall number of simulations that are necessary to perform Robust Optimization and, therefore, this ranking method does not add much to the computational cost of simulations while providing a significant accuracy improvement.

3. Select a small set of representative realizations N_R from the ranked realizations. The number of representative realizations can be determined based on the number of realizations required for reliable estimation of statistics such as the mean and standard deviation of all N_T realizations if normal distribution is assumed.

4. Conduct robust optimization (Van Essen et al. 2006) using the selected N_R representative realizations. The correspondent algorithm is described in detail in the next Section.

5. Review robust optimization results and determine the optimal well location and operating strategy to be used for the reservoir.

6. Quantify the uncertainty of the development plan by applying the optimal well location and operating strategy to N_T realizations. Run simulation using these N_T realizations and obtain N_T predictions from reservoir simulation.

7. Analyze the risk and validate the robustness of the optimal solution of robust optimization. Histogram and cumulative probability distribution plots are utilized in this step.

Optimization Algorithm

In the past decade, tremendous progress has been made in computational hardware and software to improve reservoir simulation speed. However, the availability of computing resources is still seen as a limiting factor for certain workflows of reservoir optimization. The use of proxy models such as polynomial regression models, Kriging models, artificial neural networks, and thin-plate splines has been suggested to reduce the number of required simulations in optimization (Cullick et al. 2006; Zubarev, 2009). This study utilizes optimization algorithms based on second-order polynomial regression and ordinary Kriging models, respectively. The optimization algorithm has four key steps:

1. Experimental design. The purpose of experimental design is to construct combinations of the input parameter values such that the maximum information can be obtained from the minimum number of simulation runs. Latin hypercube design (Cochran et al. 1992), which can handle any number of input parameters with mixed levels, is used in this study.

2. Proxy modeling. In this study two types of proxy models are utilized: the second order polynomial regression model and the ordinary Kriging model. The detailed description of both models is presented below.

3. Proxy-based optimization. Both polynomial regression and ordinary Kriging models are semi-analytical linear models with easily computable derivatives. Therefore, the gradient based optimization methods can potentially be used to quickly obtain the optimal solution. However, due to the simplicity of both models, they may not give accurate predictions for highly nonlinear multidimensional problems. Therefore, the optimal solution obtained based on the gradient of a proxy model may not be the true optimal for the actual reservoir model even though the gradient method is able to successfully locate the optimum of the proxy model. This means that certain suboptimal solutions of the proxy model may become the true optimal solution for the actual reservoir model. To avoid false optimum predictions, both for the polynomial and the ordinary Kriging proxy models, a brute force method is used for the proxy optimization. In the considered context this method randomly selects the specified number N_{BF} (20 million for each iteration in this study) of combinations of operational parameters in the user provided range and verifies whether the selected candidate values provide the maximum value of the objective function. The brute force method takes advantage of the high computational efficiency of both proxy models and prevents the optimization algorithm being trapped in local optimum. However, as the accuracy of the proxy model improves with the increase of the number of the training simulations, using the gradient-based methods in this work flow may become beneficial. The corresponding study that compares the brute force solutions to the ones obtained by the gradient methods and other advanced optimization algorithms will be conducted in the near future.

4. Validation and iteration. $N_V = 10$ best operational solutions (the ones with the largest value of objective function according to the proxy evaluation) are selected of the N_{BF} brute force estimates. For each of them, a reservoir simulation is conducted to obtain the true objective function value for this combination of parameters. Normally, the full physics flow simulation provides different result for the objective function. The difference between the latter and the proxy result defines the proxy prediction accuracy. The accuracy typically improves with the increase of the training data size. To further improve the prediction accuracy of the proxy model at the locations of possible optima, the validated solutions is added to the initial training data set. The updated training data set is then used to build a

new proxy model. With the new proxy model, a new set of possible optimum solutions is obtained. This iteration procedure is continued for a given number of iterations or until a satisfactory optimal solution is found.

In robust optimization, the above-mentioned optimization procedure is carried out over a set of realizations to account for the influence of the uncertainty. Essentially, in order to calculate the robust optimization objective function for a trial solution (i.e., one combination of optimization parameter values), a set of N_R simulation jobs for N_R realizations need to be run. Hence, the total number of simulations required in robust optimization is N_R times larger than nominal optimization which is based on a single realization.

Polynomial Regression Proxy Model

Polynomial regression models have been widely used for the analysis of physical and computer experiments due to their ease of understanding, flexibility, and computational efficiency. The general equation for a second-order polynomial model is:

$$y = a_0 + \sum_{j=1}^{k} a_j x_j + \sum_{j=1}^{k} a_{jj} x_j^2 + \sum_{i<j} \sum_{j=2}^{k} a_{ij} x_i x_j \quad , \tag{9}$$

where y is the response (objective function); x_j are the input variables (optimization parameters); a_0 is the intercept; $a_1, a_2, \cdots, a_k$ are the coefficients of linear terms; a_{jj} are the coefficients of quadratic terms; and a_{ij} are the coefficients of cross (interaction) terms. The coefficients are typically estimated by the least squares fitting of the model (2) to the training data (Myers, Montgomery 2002). As a result, the values of the objective function predicted by this model for the training data, generally speaking, don't match the actual objective function values obtained during the training. In other words, the prediction of the model (9) contains some error not only for the combination of parameters beyond the training data set, but for the training data set itself as well.

A polynomial regression model has an advantage of really quick evaluation (the computational cost defined as an approximate number of atomic computer operations of each estimate is approximately $O(\tilde{M})$, where $\tilde{M}$ is a total number of non-zero coefficients in (9)). At the same time, due to the simplicity of this model, its prediction accuracy is lower than in the case of other models, which can result in slower convergence.

Ordinary Kriging Proxy Model

Kriging model uses the variogram, a function of the distance and direction separating two locations in the parameter space, to quantify the spatial autocorrelation in the data. The variogram is then used to define the weights that determine the contribution of each data point to the prediction of new values at the unsampled locations of the parameter space. There are simple, ordinary and universal Kriging predictors used for the predictions in proxy modeling. Of those, only ordinary Kriging will be considered below, as our study has demonstrated its advantage over the other two in the case of SAGD simulation.

The principle of Kriging method is to estimate the value of the objective function y on a study region noted S at the location $x_p = (x_1, x_2, ..., x_M)$ where the true value is unknown:

$$\tilde{y}(x_p) = \sum_{i=1}^{N} w_i(x_p) y(x_i) \; , \qquad (10)$$

where $\tilde{y}(x_p)$ is the estimator function and $x_1, x_2, ..., x_M$ are the coordinates in the *M*-dimensional parameter space. The region S contains N training data values $y(x_i), i = 1, ..., N$ and the Kriging approach consists of interpolating the value at the unknown location by a weighted summation ($w_i(x)$ weight functions) of the values surrounding the sample points. The central problem in Kriging approach is the determination of the weight functions. It is solved by the minimization of the variance of estimator functions represented by the mean square error as

$$\frac{\partial}{\partial w_i(x_p)} < \left[\tilde{y}(x_p) - y(x_p)\right]^2 > = 0, \qquad (11)$$

where the angular brackets stand for the average over all parameter space. In the considered case of ordinary Kriging the condition (11) must be complemented with the following normalization condition for the weight functions at any location in the parameter space:

$$\sum_{i=1}^{N} w_i(x_p) = 1 \qquad (12)$$

To enforce this constraint, a Lagrange multiplier $\lambda(x_p)$ is introduced. The original problem (11) with the constraint (12) can be then rewritten in the following for

$$\frac{\partial}{\partial w_i(x_p)} < \left[\tilde{y}(x_p) - y(x_p)\right]^2 > + w^T(x_p) I \lambda(x_p) = 0 , \qquad (13)$$

where I is the identity matrix.

In this study we use Kriging as an interpolation method rather than a fitting method, i.e. the interpolated function must match exactly with all training data points. In this case it can be easily shown (Olea, 1999, Press et al., 2007) that the problem (13) for the value (10) is completely solved in the terms of the variogram

$$v(\vec{z}) = \frac{1}{2} < [y(\vec{x} + \vec{z}) - y(\vec{x})]^2 > \qquad , \qquad (14)$$

where $\vec{x}$ and $\vec{z}$ are M-dimensional vectors in operational parameter space. The angle brackets in Eq. (14) stand for the average over all pairs of M-dimensional points in the parameter space of the training data set, $\vec{x}_1$ and $\vec{x}_2$, such that $\vec{x}_1 - \vec{x}_2 \approx \pm\vec{z}$. When obtaining the experimental variogram (14), one

assumes the Euclidean metrics for the vectors $\vec{x}$ and $\vec{z}$. Eq. (14) defines the variogram as a scalar function of a lag vector argument $\vec{z}$. In general case this function is non-omnidirectional, anisotropic with respect to the parameter difference. Normally, the anisotropy factor is completely neglected for high-dimensional cases due to the infeasibility of full-scale multi-dimensional analysis, and the experimental variogram (14) is fitted to some analytical function of Euclidean distance between parameters $\vec{x} + \vec{z}$ and $\vec{x}$. While this approach performs satisfactorily as well, this study tries to improve performance by getting some insight into the inherently anisotropic nature of Eq. (14). It uses an approach that can be considered somewhat intermediate between the full-scale multi-dimensional fitting (which is still infeasible, at least for the limited amount of training data) and the complete anisotropy neglect. Namely, the non-Euclidean distance s in the M-dimensional parameter space is introduced according to the following non-Euclidean diagonal metric relation

$$s^2 = \sum_{\alpha=1}^{M} g_\alpha \Delta x_\alpha{}^2, \qquad (15)$$

where Δx_α stands for the difference between α-components of parameters at the considered points, and g_α are some parameters to be found by fitting of the experimental variogram (14) into some analytical function of non-Euclidean distance defined by (15) under the constraint $g_\alpha \geq 0, \forall \alpha$. The latter constraint guarantees the positive definition of the metrics (15).

It has been found out that, for the base case considered below, the Ordinary Kriging algorithm based on the metrics (15) provides better optimal solutions than the corresponding algorithm based on Euclidean distance (the latter corresponds to the case $g_\alpha = 1, \forall \alpha$).

Interpolation is performed according to the following algorithm:

1. Estimate the experimental variogram from the input training data according to Eq. (14). As a result, obtain a scalar function of M-dimensional vector defined by the differences of parameter values.

2. Fit the obtained experimental variogram into the power function of the non-Euclidean distance in the M-dimensional parameter space s as $v(s) = as^\delta$, where s is defined according to the Eq. (15), and a, δ and g_α are some parameters to be found during the fitting under the constraint $g_\alpha \geq 0, \forall \alpha$. It has been found out that, for the case considered, such power function provides better goodness-of-fit than exponential and Gaussian functions.

3. Define two vectors of length $N+1$

$$\vec{X} = (y(x_1),..., y(x_N),0) \qquad (16)$$

$$\vec{Y} = (v_{p1},..., v_{pN},1), \qquad (17)$$

where $v_{pi} = v(|x_p - x_i|)$, x_p being a point at which an interpolated objective function $y(x_p)$ is sought after.

4. Define an $(N+1) \times (N+1)$ symmetric matrix

$$V = \begin{pmatrix} v_{1,1} & \cdots & v_{1,N} & 1 \\ \cdots & \cdots & \cdots & \cdots \\ v_{N,1} & \cdots & v_{N,N} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}, \qquad (18)$$

where $v_{i,j} = v(|x_i - x_j|)$ and i and j are training data input points.

5. Find the sought after ordinary Kriging interpolation as

$$\widetilde{y}(x_p) = \vec{Y}^T V^{-1} \vec{X} \qquad (19)$$

Note that in this approach it is possible to perform the *LU* decomposition of the variogram matrix *V* and then find the vector $V^{-1}\vec{X}$ by the mean of forward-backward substitution only once, at the model set-up time. Then the computational cost of each interpolation will be somewhat $O(N)$, where *N* is a size of the training data. Note that *N* normally can be of the order of a few hundreds, while the number of non-zero coefficients in (7) is typically of the order of a few tens. This is why the evaluation time for the Polynomial proxy is typically an order of magnitude smaller than for the Kriging proxy. However, as it will be shown below, this shortcoming is compensated by higher prediction accuracy, resulting in faster convergence to the optimal solution.

Examples

Model Description

We consider a SAGD example of a 2-dimensional reservoir model, consisting of 3 well pairs, over a time horizon of 10 years. The model dimension is 498m×100m×69m (I×J×K). It is modeled with 17,181 grid blocks of dimension 2m×100m×1m (DI×DJ×DK). The Advanced Process and Thermal Reservoir Simulator STARS 2009.11 was used to model the SAGD process. The average reservoir initial pressure is 2400 kPa and the temperature is 14°C. The reservoir has a pay zone of about 30m thick plus a top and bottom water zone. The pay zone has an initial oil saturation ranging from 0.5 to 0.95. The top water zone is about 19m thick with an initial water saturation ranging from 0.15 to 1.00. The bottom water zone is about 20m thick with its initial water saturation equal to 1.0.

The reservoir is highly heterogeneous with high and low permeability zones. To capture geological uncertainties of the permeability and porosity distribution, a set of 100 geological realizations of the reservoir was generated using a Gaussian geo-statistical simulation method. Since these realizations

are based on the same set of variograms and no production data is available to reduce the uncertainty, all these realizations are considered to be equiprobable. Fig.1 shows the horizontal permeability field for realization 011. The horizontal permeability range is 0.65~15803 md. The permeability field can be characterized by high permeability zones with permeability above 500 md and low permeability zones with permeability less than 2 md. A constant vertical to horizontal permeability ratio of 0.7 is applied to the entire model.
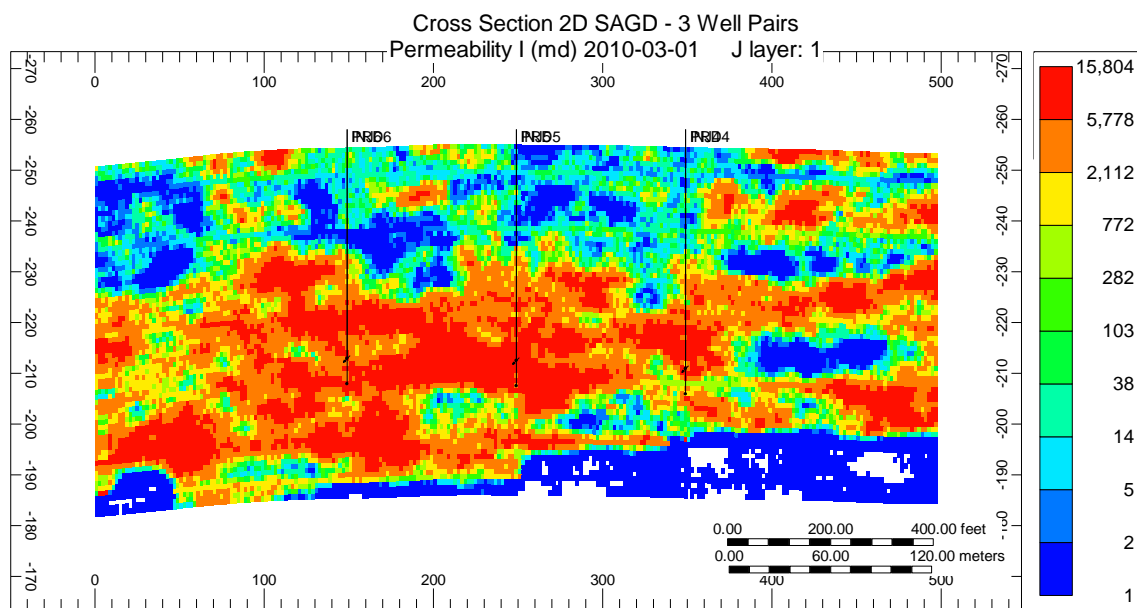
Fig. 1. Permeability field of realization 011 out of a set of 100.

Objective Function

The objective is to maximize the simple net present value (NPV) which takes into account oil revenue, steam injection cost, and well capital cost over a time horizon of 10 years. The objective function is thus given by:

$$J = \sum_{k=1}^{K} \frac{\Delta t_k [u_o \cdot q_{o,k} - u_s \cdot q_{s,k}]}{(1+D)^{t_k}} - C_{well} \quad , \tag{20}$$

where k is time step, K is the total number of time steps, Δt_k is the time interval of time step k in (day), t_k is the cumulative time until k in (year), D is yearly discount rate (1/year), u_o is the unit value of oil revenue ($/bbl), $q_{o,k}$ is oil production rate (bbl/day), u_s is steam injection cost ($/bbl), $q_{s,k}$ is steam injection rate (bbl/day), and C_{well} is well capital cost. In this study, the NPV calculation is based on the following assumptions: yearly discount rate $D = 10\%$, bitumen price $u_o = \$50/bbl$, steam injection cost $u_s = \$8/bbl$ of cold water equivalent, well capital $C_{well} = \$8,000,000$ per well pair, and total well length is 1000m (all rates are multiplied by 10 because the model only represents a 100m slice).

Optimization Parameters

Due to the permeability and porosity heterogeneity and presence of top and bottom water zones, it is expected that the placement of SAGD well pairs and their operations are critical for proper steam chamber development and the ultimate performance of the recovery process. To optimize the objective function (NPV) of the process, we consider the location of SAGD well pairs and well operating conditions as optimization parameters. For well placement, each SAGD well pair is allowed to move horizontally and vertically, but the injector-producer vertical distance is kept constant at 5m. For steam injection wells, both maximum bottom-hole pressure and maximum steam injection rate constraints are applied. Due to the low injectivity caused by the low bitumen viscosity, the maximum bottom-hole pressure is the active constraint on steam injection. Therefore, the maximum bottom-hole pressure is chosen as the optimization parameter for the injectors. Furthermore, to accommodate different steam requirements during different phases (start up, ramp up, plateau, and wind down) of the SAGD process, the maximum bottom-hole pressure for steam injection is allowed to change every two years. For SAGD production wells, minimum bottom-hole pressure, maximum liquid rate, and maximum steam

rate (water component of the gas phase) constraints are applied. Because controlling the right amount of steam production is one of the keys to prevent energy loss and to increase energy efficiency of the SAGD process, the maximum steam production rate is selected as the optimization parameter for the producers. Table 1 summarizes the optimization parameters considered in this study.

Ranking Of Realizations

This study uses direct numerical simulations of the SAGD process as the ranking method. This method works by applying the base case well location and operating strategy (i.e., the base case values listed in Table 1) to the 100 realizations and running 100 simulation jobs. The realizations are then ranked by the calculated NPV value from low to high. There are two reasons for choosing this method: (1) The ranking results will be the most accurate and therefore the geological uncertainties can be accurately represented by a small set of realizations; (2) Compared to multiples of hundreds of simulation runs required by robust and nominal optimization, the 100 simulation jobs needed for the most accurate ranking of realizations are insignificant.

Fig.2 shows the histogram of NPV resulted from the base well location and operating strategy applied to the 100 realizations. The NPV values of 100 realizations range from 94.0 to 168.0 M$. The P50 (median) NPV value is 133.3 M$. As indicated by the symmetry of the histogram and the way the tails of the distribution taper off, the histogram in Fig.2 suggests a bell-shaped normal distribution model. This can be further explained by Fig. 3 which is the normal probability plot for the 100 NPV values. As shown in Fig. 3, the points in the graph are approximately close to a straight line which indicates that the population is normally distributed.
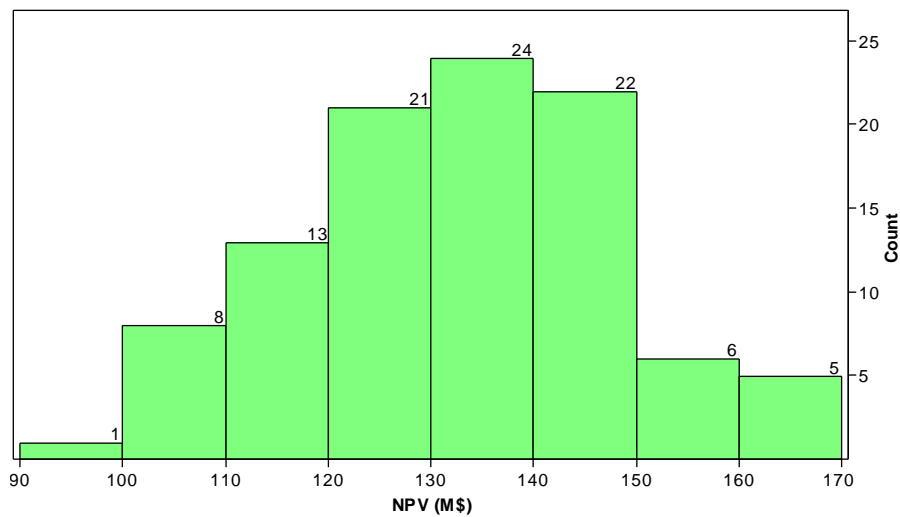
Fig. 2. Histogram of NPV resulting from the base well location and operating strategy
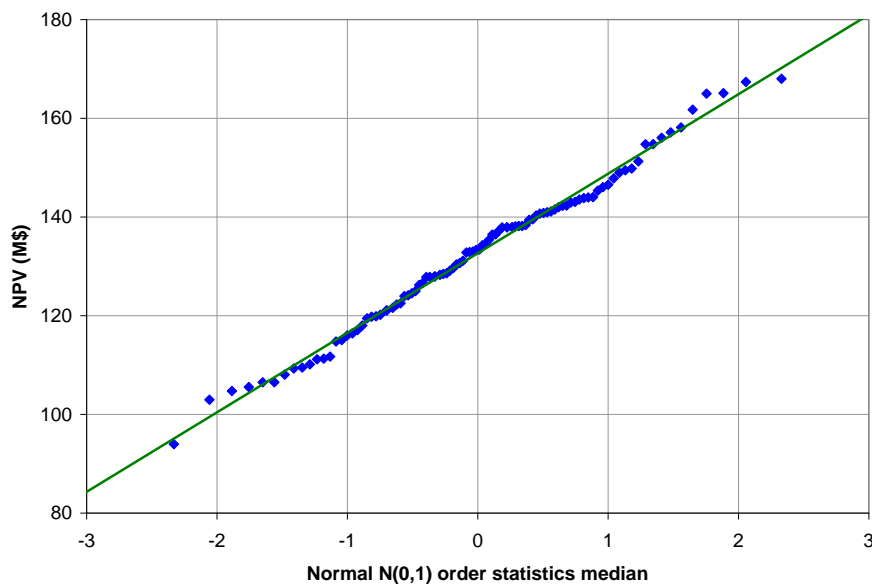

Fig. 3. Normal probability plot of NPV resulting from the base well location and operating strategy.

To select representative realizations for robust and nominal optimization, the 100 realizations are ranked by the NPV values from low to high. As listed in Table 2, the 9 realizations that are ranked as 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 90th NPV are the representative realizations for robust optimization. The 50th NPV realization (#11) is used for nominal optimization. Table 3 compares the mean and standard deviation of NPV for the entire 100 realizations and the selected 9 representative realizations. The comparison shows that the mean is almost identical while the standard deviations are close to each other. This indicates that the 9 representative realizations are sufficient for estimating reliable statistics of the entire 100 realizations

Robust Optimization Results

Robust Optimization was carried out using the nine representative realizations (selected from the initial list of 100 realizations as described in the previous Section) listed in Table 2 for the following numbers of initial runs: $N = 30, 60, 180$, i.e. the initial experimental design is a Latin hypercube design with 30×9, 60×9, and 180×9 simulation jobs, respectively. Risk averted NPV which is the objective function of robust optimization was calculated with a risk aversion factor $r = 0.5$. The robust optimization processes have a total number of 230, 260 and 380 robust optimization solutions which means that a total of 230×9=2070, 260×9=2340, and 380×9=3420 runs are made for each experimental design (including initial sampling runs and optimization iteration runs but excluding the 100 runs for ranking the models), respectively. The run progresses of optimization are shown in Figs. 4- 9 both for polynomial and Kriging proxy models. In robust optimization, each symbol in the run progress plot represents a robust optimization solution which consists of 9 simulation jobs for 9 realizations (one job per one realization). After the initial exploration, 20 iterations of proxy-based optimization are performed. Each iteration consists of three steps: (1) Build a proxy model using the training data; (2) Find possible optimum solutions using the proxy model; (3) Validate the possible optimum solutions. In each iteration, 10 possible optimum solutions are validated by running actual simulations. Therefore, a total of 20×10=200 solutions (in addition to the initial 30, 60, and 180 sampling solutions) are validated, for each experimental design, and the true objective function values (i.e., calculated from actual simulation results not proxy model predictions) of these solutions are shown in Figs. 4 – 9. Here the Job ID means a sequential number of any successful (not failed due to any reason) simulation with unique combination of operational parameters. It can be seen from all these figures that the maximum value of NPV obtained with the initial experimental design solutions is smaller than the final solution. Thus, as more and more solutions are added to the training data set, the proxy model becomes more accurate and leads to a better optimum. The red dots represent the optimal solutions lying within a 2% difference

gap from the best solution (in terms of the total NPV). The green line represents the moving average value of total NPV.

Figs. 4-5 show that at the high size of the initial Latin hypercube design (N =180) robust optimizations with both Polynomial and Ordinary Kriging proxies produce very similar optimum results. However, the ordinary Kriging proxy produces twice as many optimal solutions as the polynomial one (36 optimal solutions for the former vs. 17 optimal solutions for the latter). It means that the ordinary Kriging proxy can provide more operational flexibility as some optimal solutions, while being very close in the terms of total NPV, may be very different from the standpoint of feasibility of their operational parameters. Such difference can be a result of more complex shape of Kriging Response Surface as compared with the Polynomial Response Surface. The latter has only one minimum while the former one has many local minima with different operational parameters.
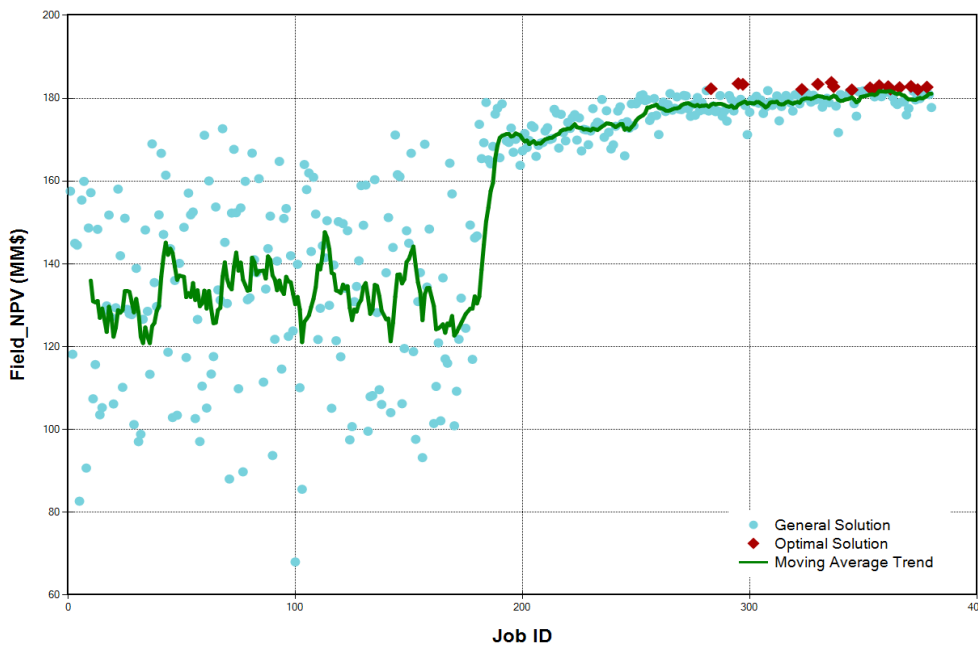
Fig. 4. NPV Progress for RO with Polynomial Proxy and 180 Initial LHD Samples
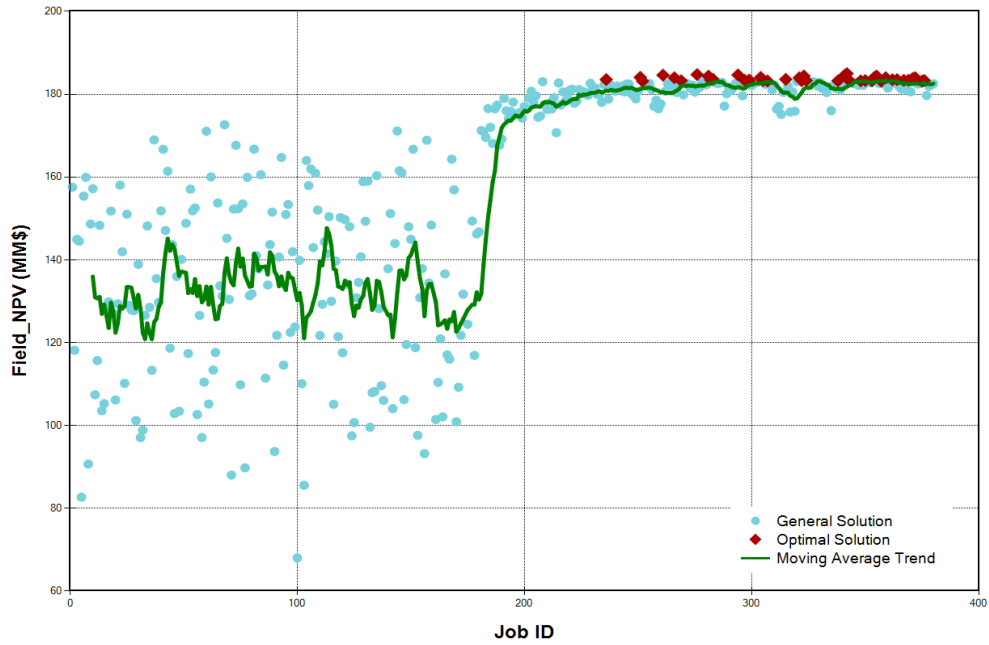
Fig. 5. NPV Progress for RO with Ordinary Kriging Proxy and 180 Initial LHD Samples
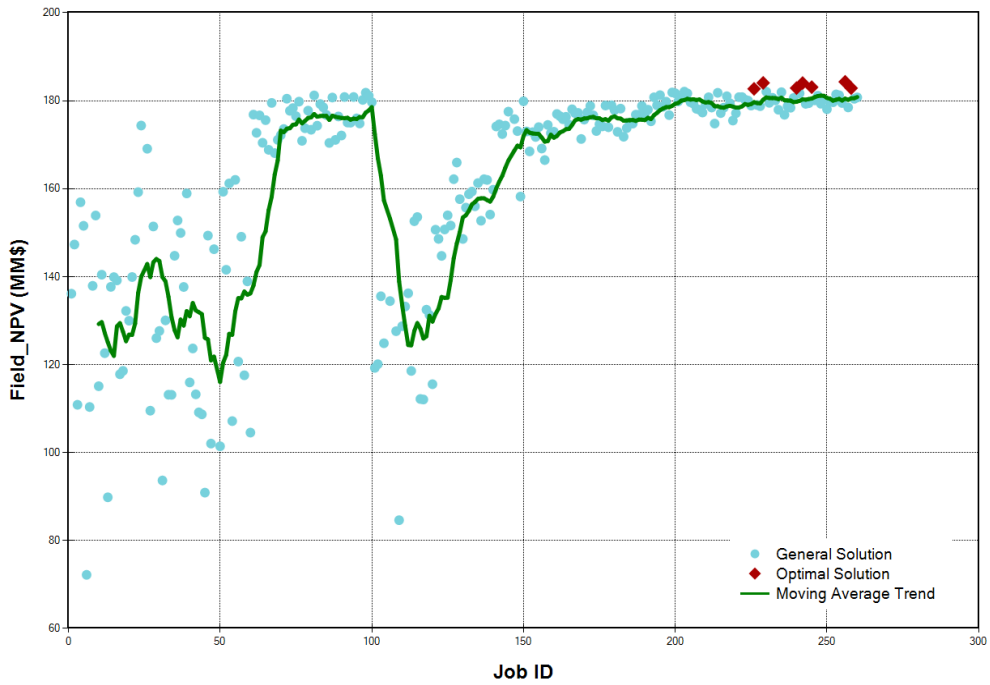
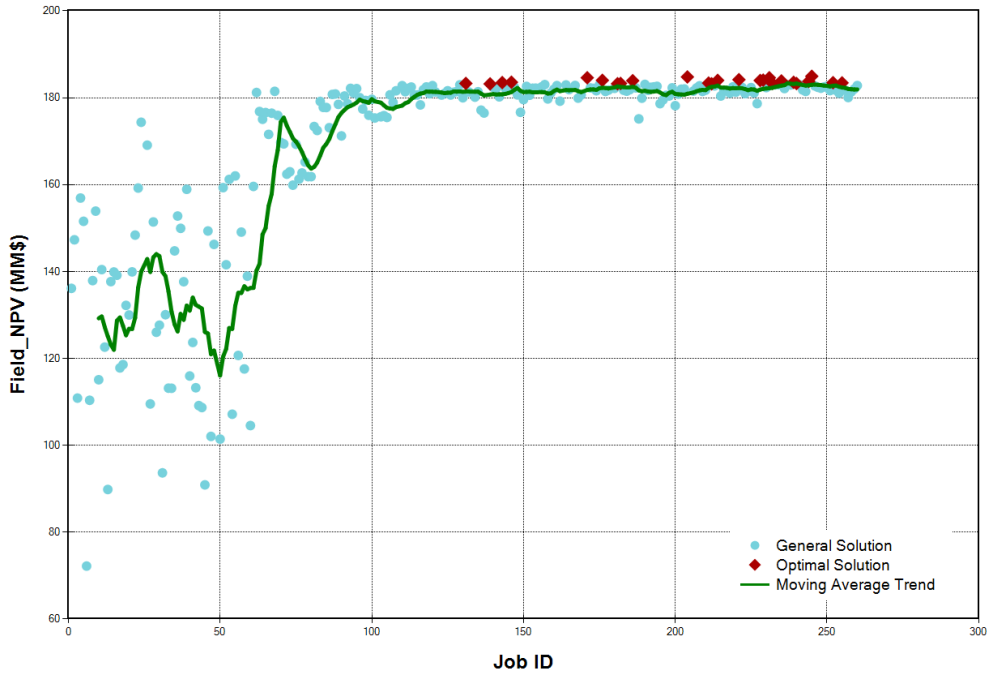Fig.6. NPV Progress for RO with Polynomial Proxy and 60 Initial LHD Samples

Fig. 7. NPV Progress for RO with Ordinary Kriging Proxy and 60 Initial LHD Samples

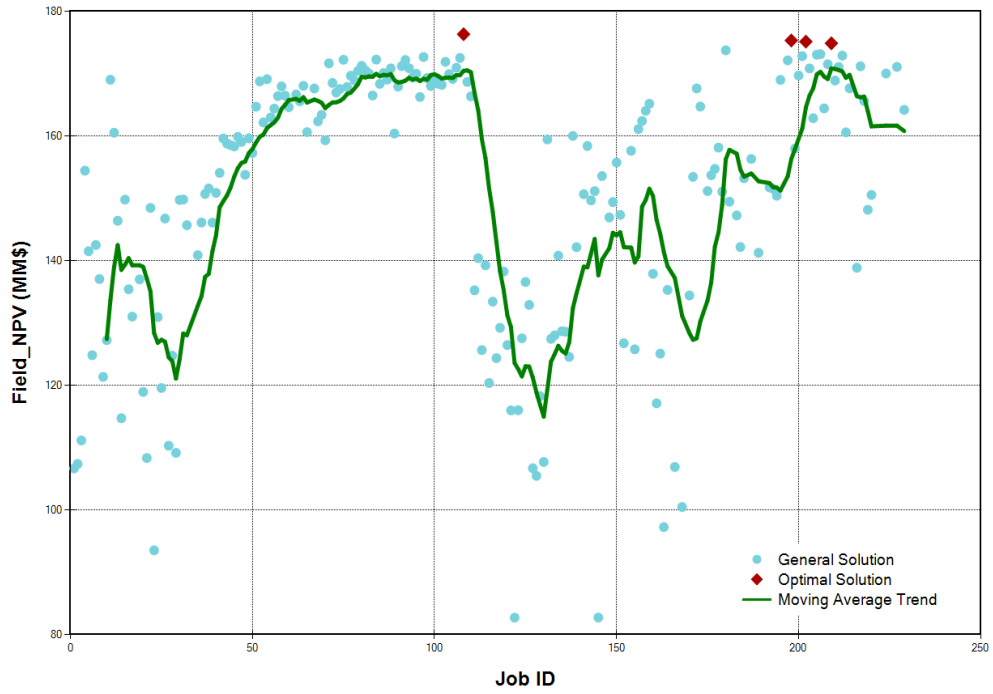Fig. 8. NPV Progress for RO with Polynomial Proxy and 30 Initial LHD Samples

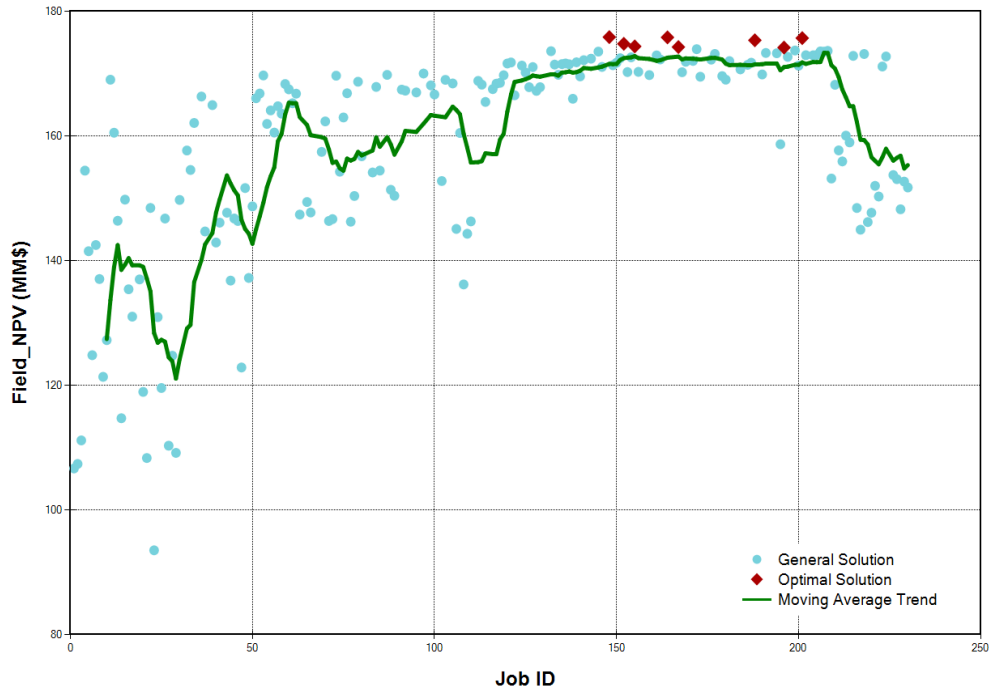Fig. 8. NPV Progress for RO with Polynomial Proxy and 30 Initial LHD Samples

Fig. 9. NPV Progress for RO with Ordinary Kriging Proxy and 30 Initial LHD Samples

Figs. 6 – 9 show that, as the size of the initial Latin hypercube design goes down, the prediction quality of both proxy models immediately after the initial training decreases as well, but with different rates. At $N = 60$ the ordinary Kriging proxy produces results that are comparable to the case of $N = 180$, and it begins to generate results that are very close to the optimal at the JobID = 100, while it takes the polynomial proxy about a hundred more validation points to get close to the optimal solution in this case. Therefore, the ordinary Kriging proxy model allows one to significantly cut down both the size of the initial experimental design and the number of validation jobs and, as a result, significantly reduces the overall simulation time without significant sacrifice of the quality of the optimal results. Figs. 5 – 6 show that in the case of $N = 30$ both proxy models produce the optimal results which are $5-6M lower than the results of $N = 60$, 180 runs.

Figs. 6 – 9 also show that the number of optimal solutions produced by the ordinary Kriging proxy keeps significantly exceeding the number of optimal solutions produced by the polynomial proxy for any considered size of the initial Latin hypercube design. Figs. 10 – 13 show what that difference means from the standpoint of such important operational parameter as injection pressure. Figs. 10 – 11 present the time dynamics of the injection pressure for optimal solutions found by the polynomial and the ordinary Kriging proxies, respectively, in the case of the initial Latin hypercube design of size 60. Figs. 12 – 13 present the corresponding dynamics in the case of the initial Latin hypercube design of size 30. One can see that, despite the fact that most of optimal solutions for the injection pressure overlap in time, the ordinary Kriging proxy does provide more operational choices in the case of the initial Latin hypercube design of size 60. Despite their quantitative difference operational strategies seem to be quite similar for both proxies in this case. In the case of the initial Latin hypercube design of size 30, the situation is quite opposite. Figs. 12 – 13 show that the variety of actual operational scenarios is, in fact, approximately the same, however it is worth noting that polynomial and ordinary Kriging proxies have found two quite different operational strategies in this case.
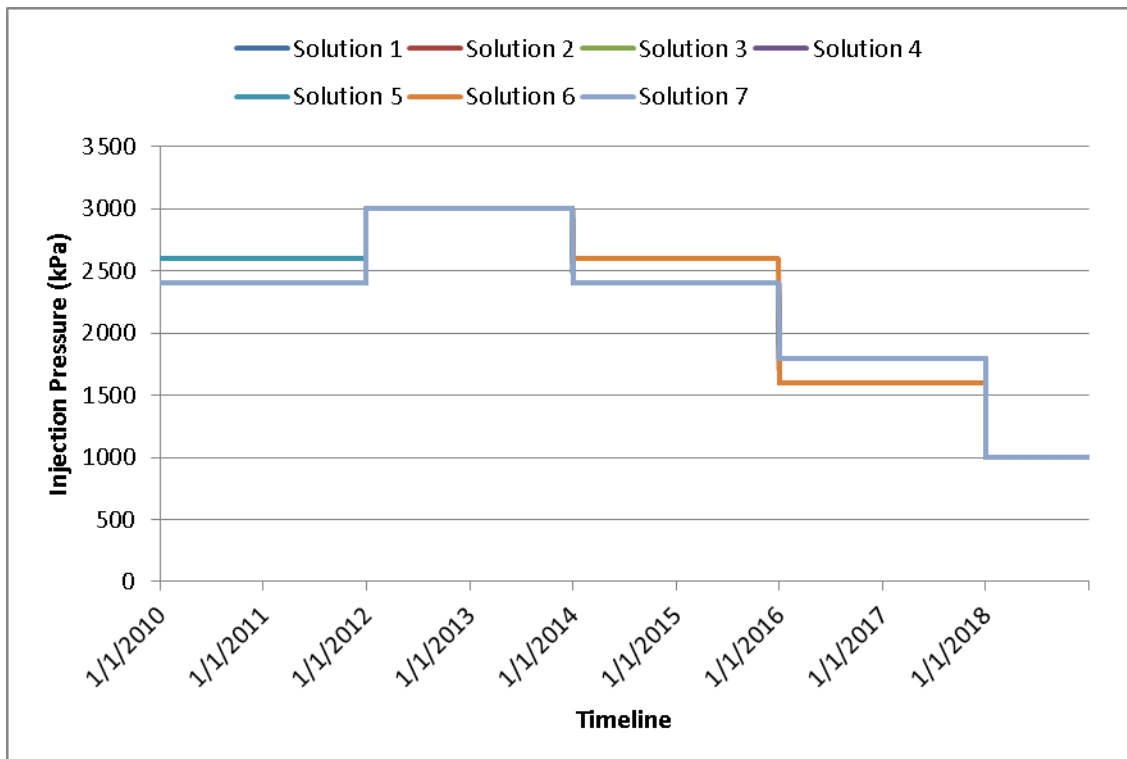
Fig. 10. Injection Pressure Dynamics for Optimal Solutions with Polynomial Proxy and 60 Initial LHD Samples
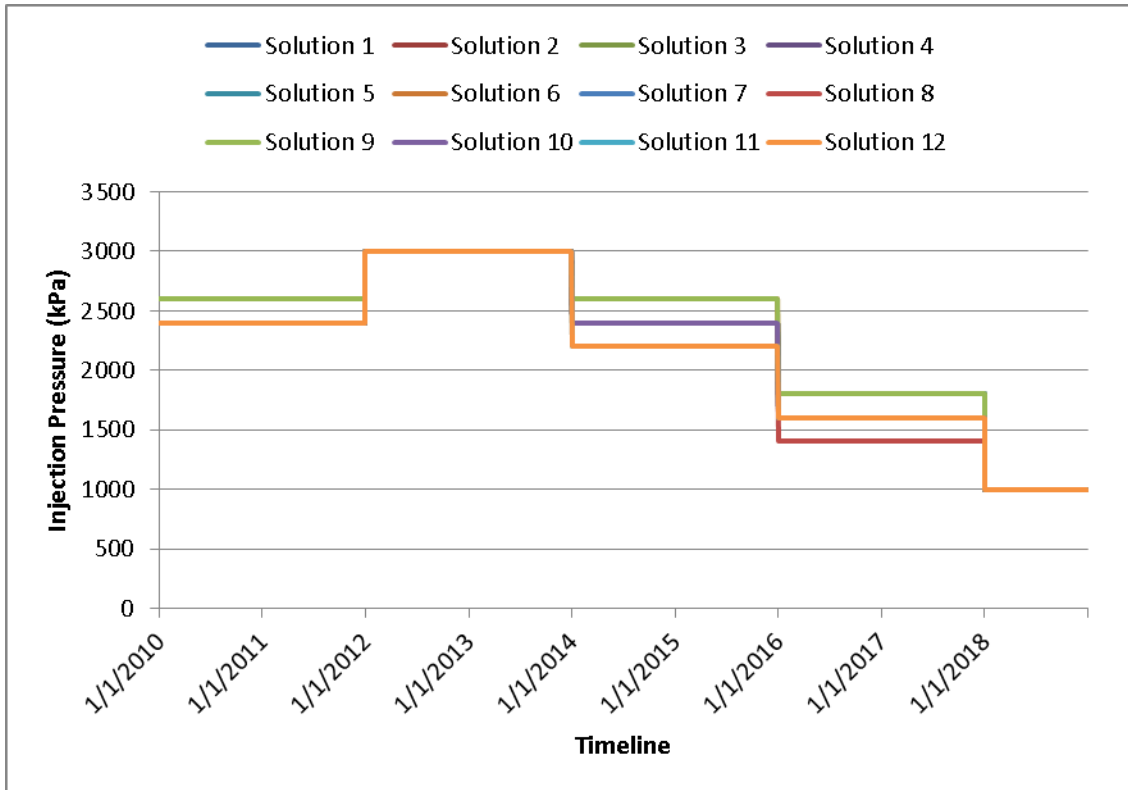

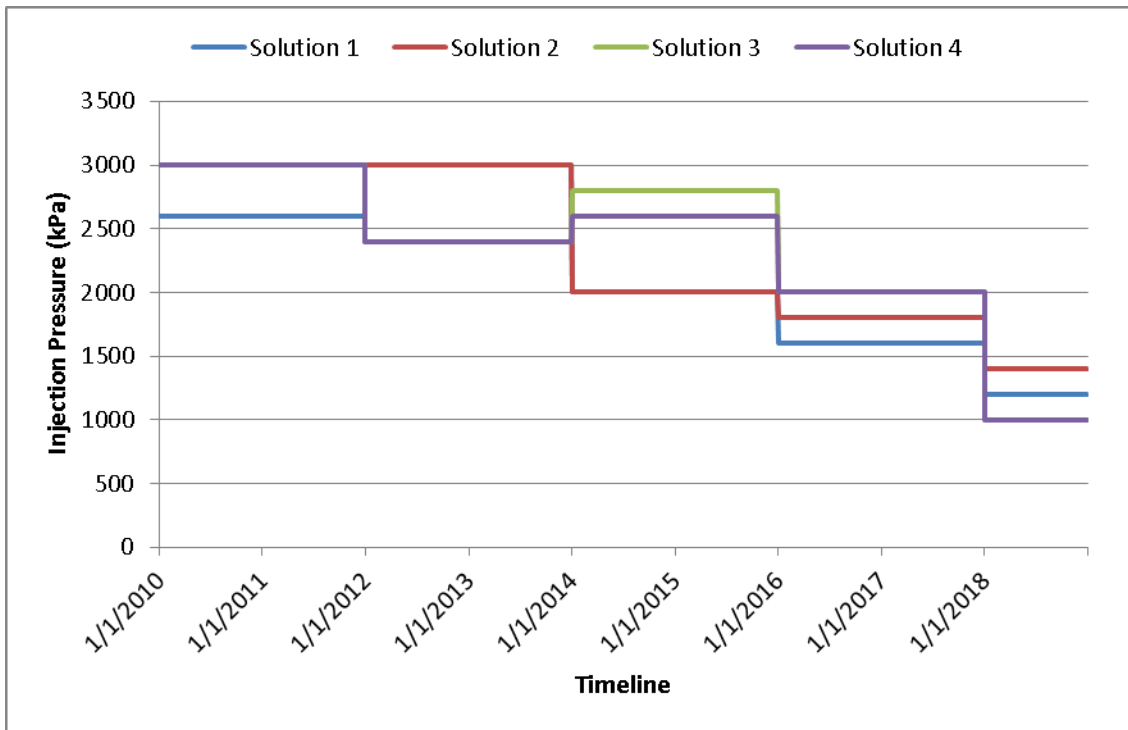Fig. 11. Injection Pressure Dynamics for Optimal Solutions with Ordinary Kriging Proxy and 60 Initial LHD Samples

Fig. 12. Injection Pressure Dynamics for Optimal Solutions with Polynomial Proxy and 30 Initial LHD Samples
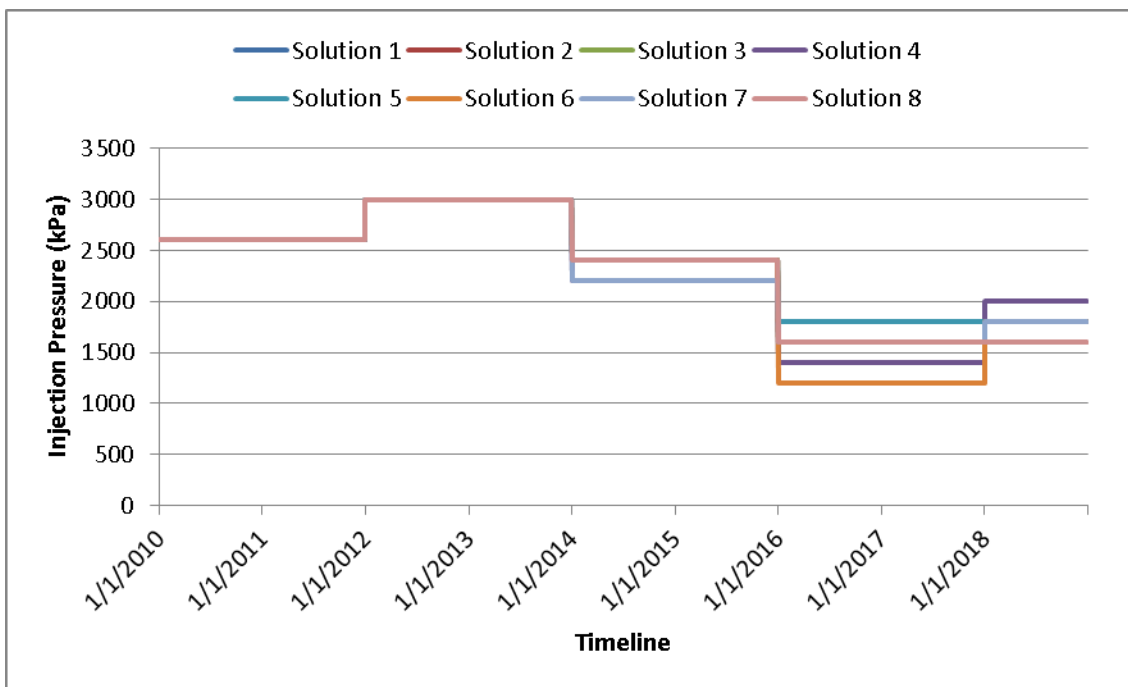
Fig. 13. Injection Pressure Dynamics for Optimal Solutions with Ordinary Kriging Proxy and 30 Initial LHD Samples

Fig.14 shows the optimum well location obtained from robust optimization. The grid property shown in the figure is water saturation. The black circles represent the well location of the base case. The red circles represent the optimal well location in the case of Polynomial Regression proxy model. The yellow circles represent the optimal well location in the case of Ordinary Kriging proxy model. The figure shows that for both proxy models the wells are positioned slightly higher in the optimal solution than the base case. The optimal horizontal location of well pair 5 is exactly the same both for Polynomial Regression and Ordinary Kriging models. For well pair 4 and 6, Polynomial Regression and Ordinary Kriging proxy models have found different optimal horizontal well locations.
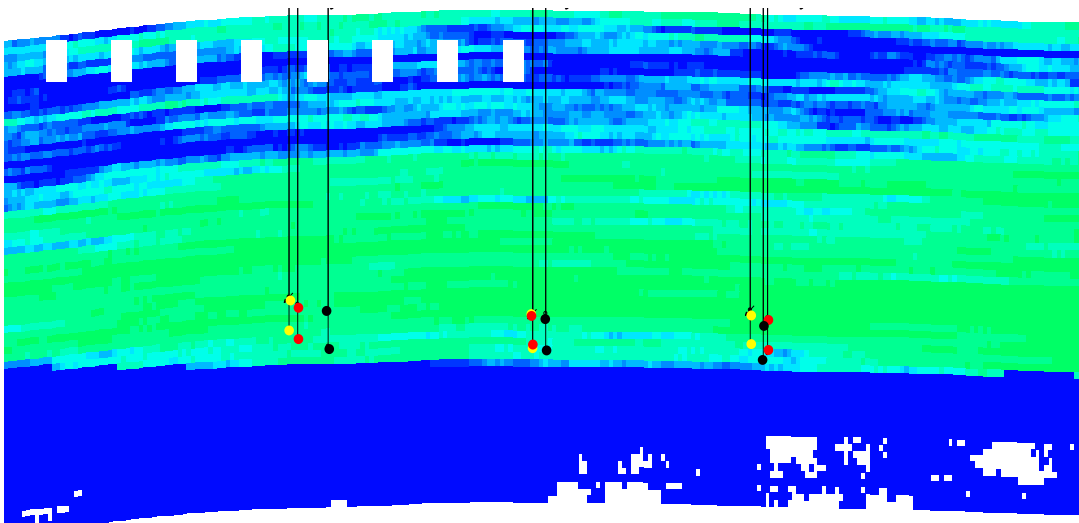
Well Pair 6 Well Pair 5 Well Pair 4

Fig.14. Well-Pair Locations for Base Case (black), Polynomial (Red), and Kriging (Yellow) Proxies

Fig.15 shows the optimal cumulative SOR obtained at the high size of the initial Latin hypercube design (N =180). It is obvious that in both cases of Polynomial Regression and Ordinary Kriging proxies, the cumulative SOR can be significantly decreased by the Robust Optimization.
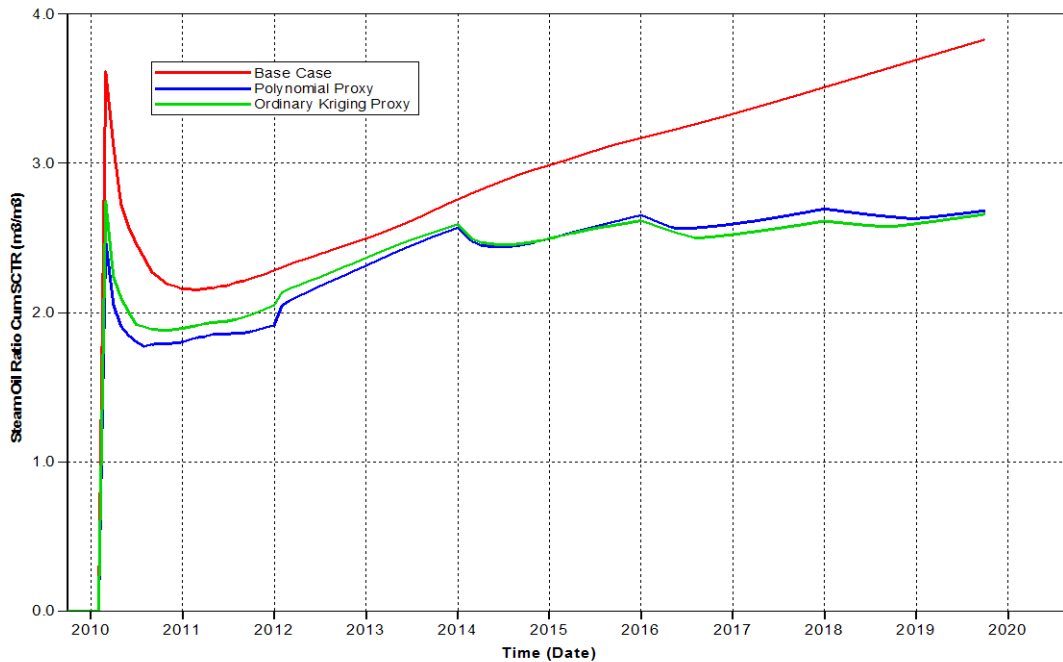
Fig.15. Optimal Cumulative SOR Dynamics (180 LHD Samples)

Conclusions

This paper presents a practical robust optimization workflow that can significantly reduce computational cost and yet still be able to account for the overall geological uncertainties of the reservoir. The key technique of the workflow is (1) ranking the entire set of geological realizations, (2) selecting a small set of representative realizations for robust optimization, and (3) building the proper proxy model for the training data. The results following from the case study point out that

1. Robust optimization using 9 representative realizations is able to account for the uncertainty of 100 realizations. Robustness of such optimization is validated by applying the obtained optimal well location and operating strategy to the full set of 100 realizations. The results show that the proposed robust optimization procedure is able to find an optimal risk weighted solution that gives good performance for any realization of the uncertainty in the given set.

2. Both Polynomial Regression and Ordinary Kriging proxies can significantly improve NPV and cumulative SOR of a 3-Well-Pair model.

3. Both proxy models can be effectively used in the optimization process by brute-force search of optimal solutions.

4. Ordinary Kriging proxy model provides more accurate prediction, faster finds an optimal solution, and finds more optimal solutions than Polynomial Regression one as a result of more complex form of the Kriging Response Surface

Acknowledgements

References

1. Butler, R.M. 1991. Thermal Recovery of Oil and Bitumen, Prentice-Hall Inc., Englewood Cliffs, NJ.

2. Cochran, W.G., and Cox, G.M. 1992. Experimental Designs, John Wiley & Sons, Inc., 2nd Edition, New York.

3. Cullick, A.S., Johnson, D., and Shi, G. 2006. Improved and More-Rapid History Matching With a Nonlinear Proxy and Global Optimization, paper SPE 101933, presented at the SPE Annual Technical Conference and Exhibition.

4. Egermann, P., Renard, G., and Delamaide E. 2001. SAGD Performance Optimization through Numerical Simulations: Methodology and Field Case Example, paper SPE 69690, presented at the SPE International Thermal Operations and Heavy Oil Symposium.

5. Fedutenko, E., Yang, C., Card, C., and Nghiem, L. 2011. Optimization of SAGD Process with Proxy Models: Case Study of a 3-Well-Pair Model, Proceedings Of World Heavy Oil Congress, Edmonton, p. 533-537.

6. Fenik, D.R., Nouri, A., and Deutsch, C.V. 2009. Criteria for Ranking Realizations in the Investigation of SAGD Reservoir Performance, paper 2009-191, presented at the Canadian International Petroleum Conference (CIPC).

7. Gates, I.D., Kenny, J., Hernandez-Hdez, I.L., and Bunio, G.L. 2005. Steam-Injection Strategy and Energetics of Steam-Assisted Gravity Drainage, paper SPE 97742, presented at the SPE/PS-CIM/CHOA International Thermal Operations and Heavy Oil Symposium.

8. Kumar, A., Oballa, V., and Card, C.C. 2010. Fully-Coupled Wellbore Design and Optimization for Thermal Operations, paper SPE 137427, presented the Canadian Unconventional Resources & International Petroleum Conference.

9. McLennan, J.A., and Deutsch, C.V. 2005. Ranking Geostatistical Realizations by Measures of Connectivity, paper SPE 98168, presented at the SPE/PS-CIM/CHOA International Thermal Operations and Heavy Oil Symposium.

10. Myers, R.H., Montgomery, D.C. 2002. Response Surface Methodology. Process and Product Optimization Using Designed Experiments, John Wiley & Sons Inc., NY.

11. Olea, R.A. 1999. Geostatistics for Engineers and Earth Scientists, Springer, Norwell, MS.

12. Peterson, J., Riva, D., Edmunds, N., and Solanki, S. 2010. The Application of Solvent-Additive SAGD Processes in Reservoirs With Associated Basal Water, paper SPE 137833, presented at the Canadian Unconventional Resources & International Petroleum Conference.

13. Press, W.H., Teukolsky,S.A., Vetterling,W.T. 2007. Numerical Recipes 3rd Edition, Cambridge University Press.

14. Van Essen, G.M., Zandvliet, M.J., Van den Hof, P.M.J., Bosgra, O.H., and Jansen, J.D. 2006. Robust Waterflooding Optimization of Multiple Geological Scenarios, paper SPE 102913, presented at the SPE Annual Technical Conference and Exhibition.

15. Vanegas Prada,J.W., Cunha, L.B. 2008. Assessment of Optimal Operating Conditions in a SAGD Project by Design of Experiments and Response Surface Methodology, Petroleum Science and Technology, Vol. 26, No. 17, 2095.

16. Yang, C., Card, C., and Nghiem, L. 2009. Economic Optimization and Uncertainty Assessment of Commercial SAGD Operations, Journal of Canadian Petroleum Technology, Vol. 48, No. 9.

17. Yang, C., Card, C., Nghiem, L., and Fedutenko, E. 2011. Robust Optimization of SAGD Operations under Geological Uncertainties, paper SPE 141676, presented at the SPE Reservoir Simulation Symposium, Woodlands.

18. Zubarev, D.I. 2009. Pros and Cons of Applying Proxy-Models as a Substitute for Full Reservoir Simulations, paper SPE 124815, presented at the SPE Annual Technical Conference and Exhibition.