Bayesian Optimization : A New Sample Efficient Workflow for Reservoir Optimization under Uncertainty

Peyman Kor*,[a], Aojie Hong[a], Reidar Brumer Bratvold[a]

[a]*Energy Resources Department, University of Stavanger, Stavanger, Norway*

Abstract

Field development in E&P requires several level of decision making, where amoung the available options, the one with highest objective is decided. The decisions to make could be well configuration, well types , and well controls. The search to find th eoptimum decison, which called here optimization, a computionally demanding nonlinear problem. More challenging, in teh case of optimization under uncertainty, such problems requires large number of flow simulation given diffrent geological realizations. To alleviate this problem and to provide the fesible framework dealing with computional tcomplexity in essevor optimization, a new sworkflow is presented. The "Bayesian Optimization" in this work efficntley sample from the flow simulation based on the probalistc modeling of the objective function, asissting to perform optimization process with as few as possible flow simulation, without losing the optimum solution. The paper begins with introduction to gaussian process and building a probalistiv view of the objective function as "Surrogate" and then provides the foundation for acquistion function, which decide to the next query point from the exppensive function, here flow simulation. Further, a 1-D probelm will be optimized through the Bayesian optimization which helops to better understading of the workflow. Then, the workflow has been applied to 3D, "Egg model" rto perform optimization in the reaslistic field scenario, considering geological uncertainty. Here, the comparsion of the workflow with other two well-kow algorithm in reservoir optimization literature, namely Genetic Algorithm (GA) and Partcile Swarm Inteligence (PSO) has been performed. The results of the comparison show the the BAyesian Optimization workflow presented here could reach the same global optimum point achieved with GA and PSO, yet reduce computionaly complexity of the optimization 5X, which could b esignificant, in the case of the 3D flow simulation, potentially taking hours of running.

*Corresponding Author

**Equal contribution

Email addresses: peyman.kor@uis.no (Peyman Kor), aojie.hong@uis.no (Aojie Hong), reidar.bratvold.uis.no (Reidar Brumer Bratvold)

1. 1. Introduction:

While developing a field, prediction of reservoir response to change in the variable is an am important factor to have an efficient reservoir management. The field of reservoir engineering is rich in the development and application of full-physics numerical simulations for forward modeling. However, the computational power needed to run such numerical simulators most of the time is huge. Especially, the framework of the Robust Optimization where uncertainty are considered through multiple of geological realization (thousand or multi-thousand), the practical applicability of such forward modeling is considerably limited. To address this challenge, the proxy-modeling for reservoir managemnet has emerged to reduce the cost of forward simulation. The root of this field goes back to the work of the (Bruce, 1943) where the analogy between flow of electricty in tthorugh te electrical device and the response of the reservoir was constructed. (Albertoni and Lake, 2003) Proposed the a Multivariate Linear Regression (MLR) to estimate the production from the well, where it claimed that total production at each is linear combination of injection rates with diffusitivity filter. Building on the work of (Albertoni and Lake, 2003), (Yousef, 2006) proposed a new procedure to quantify communication between wells, where for each injector/producer pair, two coefficients, capacitance to quantify connectivity and time constant to quantify the degree of fluid storage were considered. (Sayarpour, 2008) used superposition in time to analytically solve the numerical integration in CRM. (Zhao et al., 2015)(Zhao et al. 2015) articulated that CRM models are not applicable when production rates change significantly, mainly due to those models neglect to include intearction of production-production and injector-injector pair well. Formulating the model in the framework titled INSIM (interwell numerical simulation model), it was used as efficient replacement to reservoir simulator when history-matching phase rate data or water-cut data obtained during water flooding operations.

Seperatley, in sake of utilization of recent advancemnet in the world of Information technology, couple of reaserach has been don eon the development of “Surrogate Reservoir Models” - (Mohaghegh and Guruswamy, 2006) proposed the workflow for SRM model where Fuzzy Paatern Recognition (FPR) technology was dimensionality reduction, in both static and dynamic parameters of the reservir . Key Performance Indicator (KPI) was used to select the most important variable.- (Sampaio, 2009) tried on use of feed-forward neural networks as nonlinear proxies of reservoir response. A few set of rock and fluid properties were considere a input (porosity, permeability in “x” direction, permeability in “y” direction, and rock compressibility) where in developing the neural network model, only one hidden layer was used. flow proxy modeling to replace full simulation in history matching, and built the proxy with a single hidden layer artificial neural network (ANN). To predict the oil production from the SAGD rocess, (Fedutenko et al. 2014) (Fedutenko et al., 2014)employed the RBF Neural Network to predict the cumulative oil production over the time horizon of

production (10 years) .

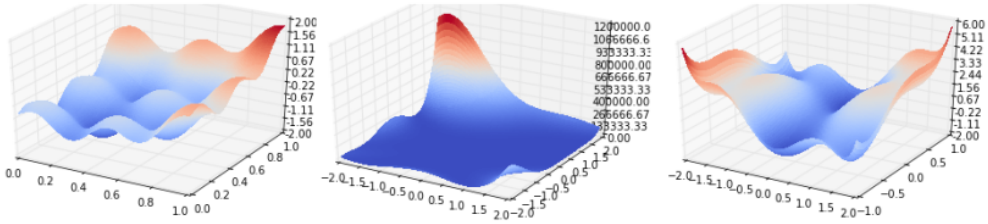
1.1. Survey of suurgate -based papers in onepetro database

2. 2. Problem Statement

Consider a "well behaved" function $f : \chi \to \mathbb{R}$ where $\chi \subseteq \mathbb{R}^D$ is a bounded domain.

$$x_M = argmin\ f(x) \tag{1}$$

reference equation 1



f is explicitly unknown and multimodal. Evaluations of f are expensive. Expensive functions, who doesn't have one?

In reservoir optimization, we can have different objectives: Recovery Factor, Net Present Value (NPV), ...

$$u_M = \underset{u \in \text{constraints}}{\text{argmax}} \quad \text{Objective Func(u)} \tag{2}$$

u is a decision variable to make: could be injection rate, well locations, order of drilling wells.

$$\text{Objective Func(u)} = J(u, G) = \sum_{k=1}^{n_T} \frac{q_o^k(u, G)P_o - q_w^k(u, G)P_{wp} - I^k(u, G)P_{wi}}{(1 + b)^{t_k/D}} \tag{3}$$

$$\text{Objective Func(u)} = \overline{J}(u) = \frac{\sum_{i=1}^{n_e} J_r(u, G_i)}{n_e} \tag{4}$$

2.0.0.1. Example case:.

- Lets' say you are optimizng the well control problem with dimension, $d = 10$
- Number of 3D geological realizations is $n_e = 100$ and running each possible u^*in commercial reservoir simulator, at each realization talks ~ 1 hour. Then time it takes to calculate $\overline{J}(u^*)$ is ~ 100 hours.
- With 6 month CPU running time budget, the total available budget is to run is ~ 50.

- Availablity of only 50 $\overline{J}(u^*)$ evaluation is a small budget, considering the 10 dimentional optimization problem.

Problem $\overline{J}(u)$ is expensive function, meaning the # times we can evaluate it severely limited. Solution Bayesian Optimization (BayesOpt) propose a new workflow to conduct optimization at the small $\overline{J}(u)$ budget, without affecting the optimum solution.

3. 3. Bayesian Optimization Workflow

3.1. Overal View

Workflow to perform global optimization of multimodal black-box functions:

- Step 1. Choose some initial design points and build a probabilistic model over the space of possible objective f, this probabilistic model serves as prior.
- Step 2. Combine prior and the likelihood to get a posterior of probabilistic model over the objective given some observations.
- Step 3. Use the posterior to decide where to take the next evaluation $\mathbf{x}^*$ according to some policy for decision making.
- Step 4. Evaluet the f at $\mathbf{x}^*$ and augment it to the initial data, in step 1.

Iterate between 2 and 4 until the evaluation budget is over. ## 3.1 Gaussian Process

3.1.1. Step 1. Probalistic Model as Prior

3.1.1.1. Gaussian Process (GP). reference to the book (Murphy, 2022)

Key Assumption in (GP) is that: the function values at a set of $M > 0$ inputs, $\mathbf{f} = [\mathbf{f}(\mathbf{x_1}), ..., \mathbf{f}(\mathbf{x_M})]$, is jointly Gaussian, with mean and Covariance

$$(\mu = m(x_1), ...m(x_M)) \sum_{i,j} = \kappa(x_i, x_j) \tag{5}$$

and κ is a positive definite (Mercer) kernel. Suppose we have a initial design points, set $\mathcal{D} = (x_n, y_n) : n = 1 : N$, where $y_n = f(x_n)$ is the noise-free observation of the function evaluated at x.

Now we consider the case of predicting the outputs for new inputs that may not be in $\mathcal{D}$.

$$\mathbf{f}_* = [\mathbf{f}(\mathbf{x_1}), .., \mathbf{f}(\mathbf{x_{N_*}})] \tag{6}$$

By definition of the GP, the joint distribution $p(\mathbf{f_X}, \mathbf{f}|\mathbf{X}, \mathbf{X}_*)$ has the following form:

$$\begin{bmatrix} \mathbf{f_X} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left((\begin{bmatrix} \mu_X \\ \mu_* \end{bmatrix}), \begin{bmatrix} \mathbf{K}_{X,X} & \mathbf{K}_{X,*} \\ \mathbf{K^\intercal}_{X,*} & \mathbf{K}_{*,*} \end{bmatrix} \right) \tag{7}$$

$$\begin{aligned} \mu_X &= [m(x_1), ..., m(x_N)] \\ \mu^* &= [m(x_1^*, ...m(x_N^*))] \end{aligned} \tag{8}$$

$$\begin{aligned} K_{X,X} &= \kappa(X, X; \theta), \qquad size(N \times N) \\ K_{X,*} &= \kappa(X, X_*; \theta), \qquad size(N \times N_*) \\ K_{*,*} &= \kappa(X_*, X_*; \theta), \qquad size(N_* \times N_*) \end{aligned} \tag{9}$$

Covariance Kernels	assumeing $h = \|\|x - x'\|\|$
Gaussain	$\kappa(x, x') = \sigma_f^2 exp(-\frac{h^2}{2\ell^2})$
Matern $\mu = \frac{5}{2}$	$\kappa(x, x') = \sigma_f^2(1 + \frac{\sqrt{5}\|h\|}{\ell}\frac{5h^2}{3\ell^2})exp(-\frac{-\sqrt{5}\|h\|}{\ell})$
Matern $\mu = \frac{3}{2}$	$\kappa(x, x') = \sigma_f^2(1 + \frac{\sqrt{3}\|h\|}{\ell})exp(-\frac{-\sqrt{3}\|h\|}{\ell})$
Exponetial	$\kappa(x, x') = \sigma_f^2 exp(-\frac{\|h\|}{\ell})$
Power-Exponetial	$\kappa(x, x') = \sigma_f^2 exp(-(\frac{\|h\|}{\ell})^p)$

$$\kappa(x, x';;\theta) = (1 + \frac{\sqrt{5}|h|}{\theta} + \frac{5h^2}{3\theta^2})exp(-\frac{-\sqrt{5}|h|}{\theta} \tag{10}$$

3.1.1.2. Covariance Kernel, Parameter estimation.

$$p(y|\mathbf{X}, \theta) = \int \mathbf{p}(\mathbf{y}|\mathbf{f}, \mathbf{X})\mathbf{p}(\mathbf{f}|\mathbf{X}, \theta) \tag{11}$$

$$\log p(y|\mathbf{X}, \theta) = \mathcal{L}(\zeta, \sigma_{\mathbf{f}}^{\mathbf{2}}) = -\frac{\mathbf{1}}{\mathbf{2}}(\mathbf{y} - \mu_{\mathbf{X}})^{\intercal}\mathbf{K}_{\mathbf{X},\mathbf{X}}^{-\mathbf{1}}(\mathbf{y} - \mu_{\mathbf{X}}) - \frac{\mathbf{1}}{\mathbf{2}}\mathbf{log}|\mathbf{K}_{\mathbf{X},\mathbf{X}}| - \frac{\mathbf{n}}{\mathbf{2}}\mathbf{log}(\mathbf{2}\pi) \tag{12}$$

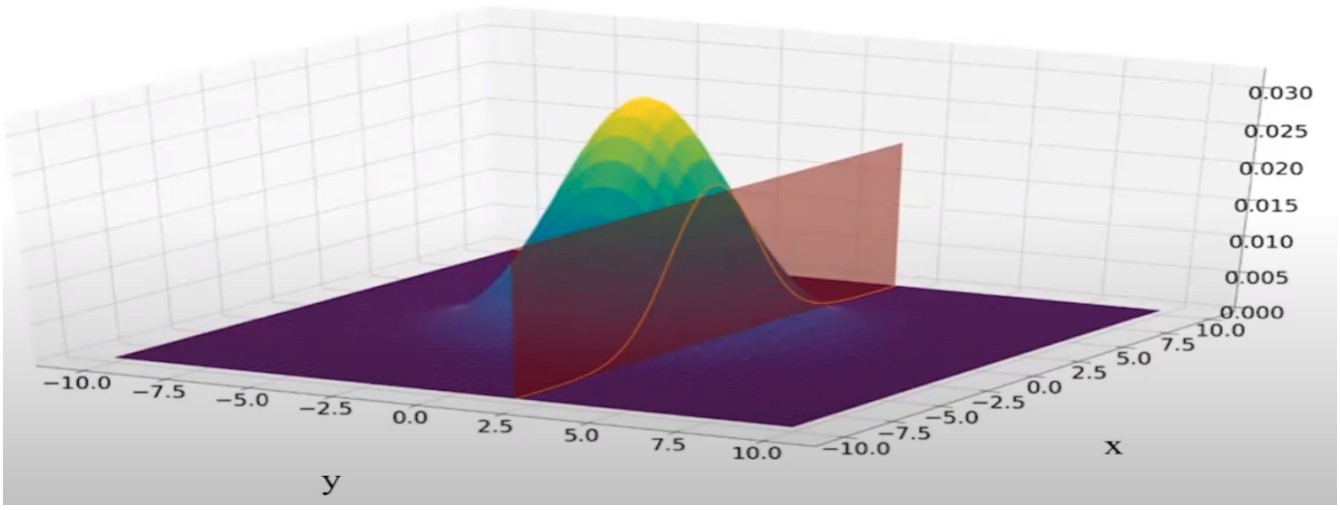
Where the dependence of the $\mathbf{K}_{\mathbf{X},\mathbf{X}}$ on θ is implicit.The gradient-based optimizer is performed in order to:

$$[\zeta^*, \sigma_f^{2*}] = argmax\mathcal{L}(\zeta, \sigma_f^2) \tag{13}$$

However, since the objective $\mathcal{L}$ is not convex, local minima can be a problem, so we may need to use multiple restarts.

3.1.2. Step 2. Posterior of Probabilistic Model

3.1.2.1. Posterior of Gaussain Process, (conditioning on initial data).



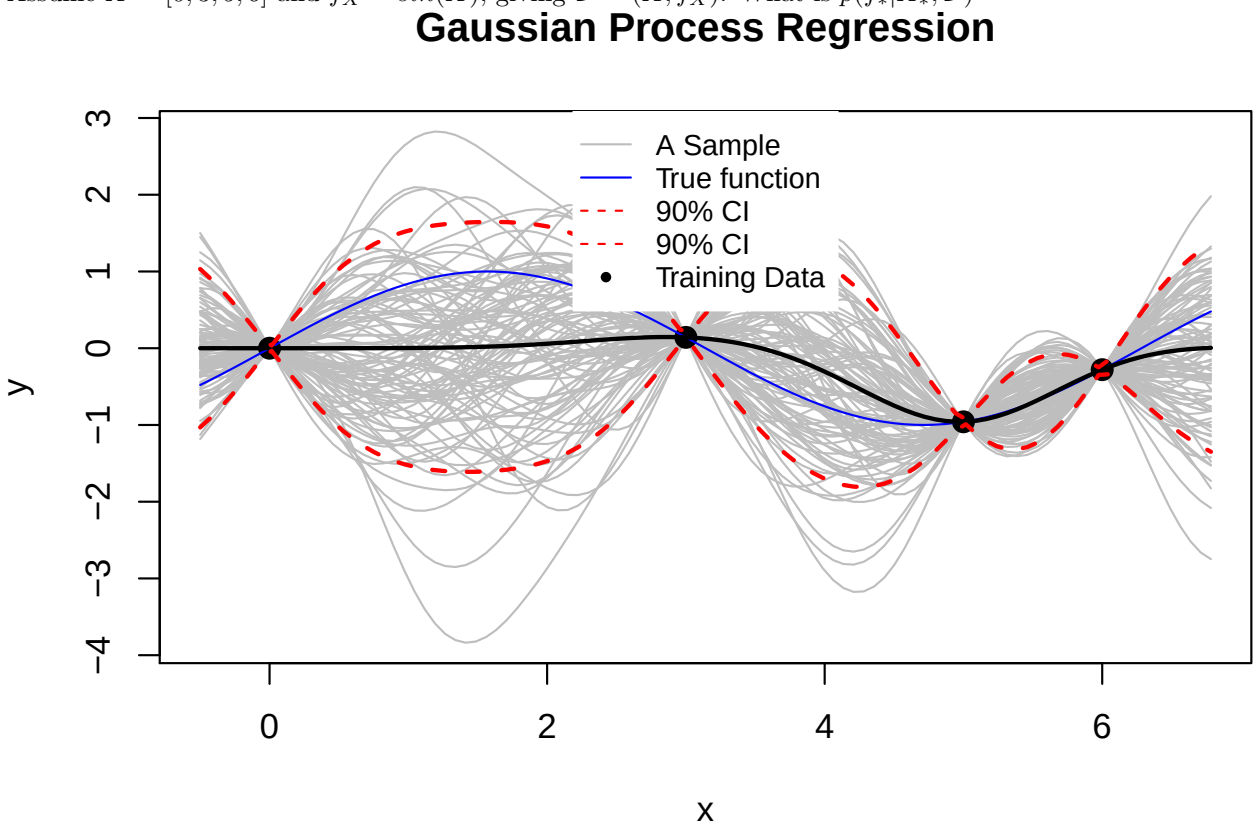
–

$$p(f_*|X_*, \mathcal{D}) = \mathcal{N}(f_*|\mu_*, \sum\nolimits_*) \tag{14}$$

$$\mu_* = m(\mathbf{X}_*) + \mathbf{K}^{\intercal}{}_{\mathbf{X},*}\mathbf{K}^{-1}{}_{\mathbf{X},\mathbf{X}}(\mathbf{f}_{\mathbf{X}} - \mathbf{m}(\mathbf{X}))$$
$$\sum\nolimits_* = \mathbf{K}_{*,*} - \mathbf{K}^{\intercal}_{X,*}\mathbf{K}^{-1}_{X,X}\mathbf{K}_{X,*} \tag{15}$$

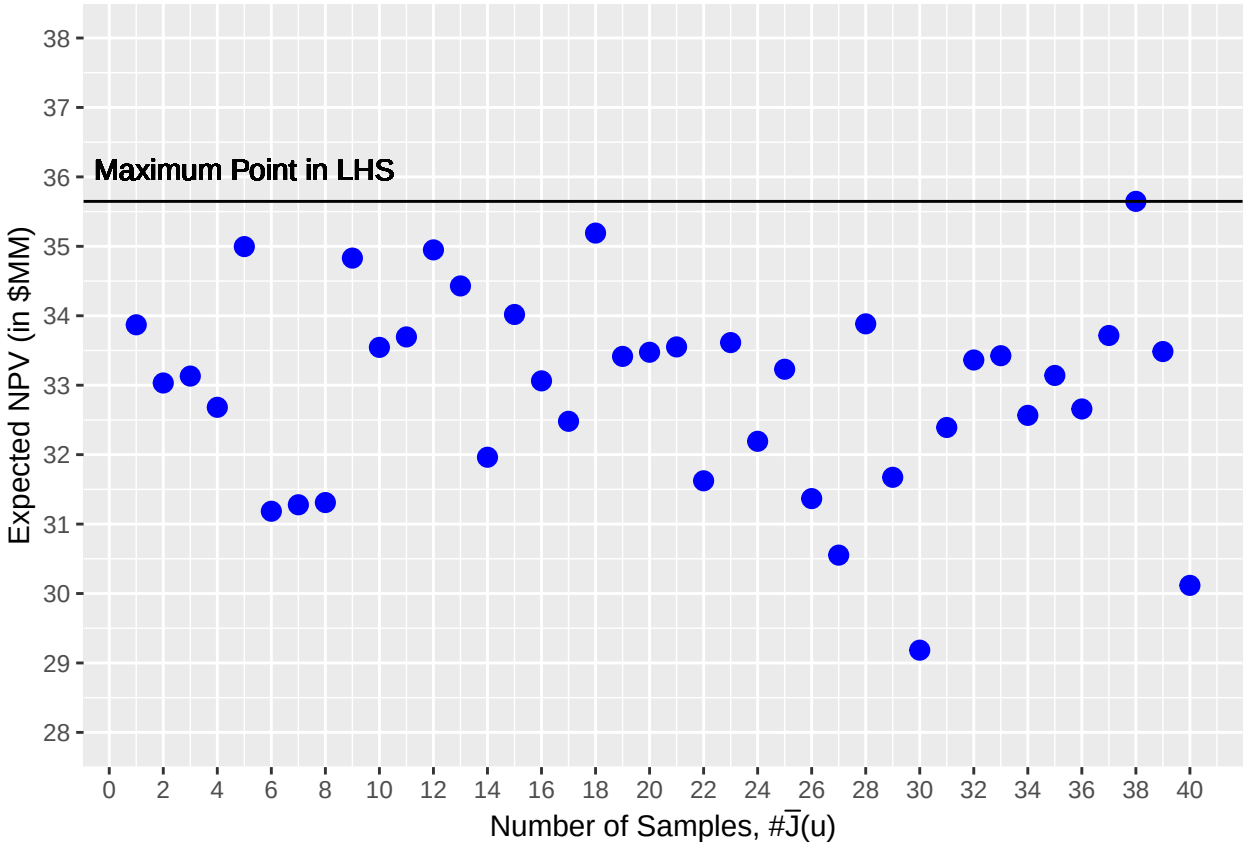
87 *3.1.3. Example of Step.1 and Step.2*

88 Assume $X = [0, 3, 5, 6]$ and $f_X = sin(X)$, giving $\mathcal{D} = (X, f_X)$. What is $p(f_*|X_*, \mathcal{D})$



89

90 Now we sample the point $X = 1$, and add to $\mathcal{D}$



3.1.4. Step.3 Deciding on next $\mathbf{x}^*$ *based on Posterior*

Posterior of the probalistic model quantify the uncertainty over the space of the f. The question is what is the next $\mathbf{x}^*$ to be sampled from the *expensive function*?

Define an utility function to collect new data points satisfying some optimality criterion: optimization as decision making.

There are a few of policies in the literature of Bayesopt, here the *Expected Improvement (EI)* policy will be used.

3.1.4.1. Expected Improvement as Policy for Decision Making. In Expected Improvement (EI) policy choose the next query point as the one which has the highest expected improvement over the space of the *expensive function*

$$utility(x;\theta,\mathcal{D}) = \alpha_{EI} = \int_y max(0, y - f)p(y|x;\theta,\mathcal{D}) \tag{16}$$

$$utility(x;\theta,\mathcal{D}) = \alpha_{EI} = \int_y max(0, y - f)p(y|x;\theta,\mathcal{D})\,dy$$

However, we do not have access to the *expensive function*, f, therefore we replace the f with the best available solution found so far, y^+

$$utility(x;\theta,\mathcal{D}) = \alpha_{EI} = \int_y max(0, y - y^\dagger) p(y|x;\theta,\mathcal{D})\, dy \tag{17}$$

y^+ : The best solution found in the training dataset $\mathcal{D}$

The good news: The analytical form of the utility function is available for the gaussian process

$$\gamma(\mathbf{x}) = \frac{\mu(\mathbf{x};\theta,\mathcal{D}) - y^\dagger}{\sigma(\mathbf{x};\theta,\mathcal{D})} \tag{18}$$

$$utility(\mathbf{x};\theta,\mathcal{D}) = \alpha_{EI}(x;\theta,\mathcal{D}) = (\mu(x;\theta,\mathcal{D}) - y^\dagger)\Phi(\gamma(x)) + \sigma(x;\theta,\mathcal{D})\phi(\gamma(x)) \tag{19}$$

Where $\Phi(.)$ and $\phi(.)$ are CDF and PDF of standard Gaussian distribution.

It is too greedy in the context of the sequential decision making. Therefore, an explorative term is added as explorative" parameter ϵ.

$$\gamma(\mathbf{x}) = \frac{\mu(\mathbf{x};\theta,\mathcal{D}) - y^\dagger - \epsilon}{\sigma(\mathbf{x};\theta,\mathcal{D})} \tag{20}$$

$$\alpha_{EI}(x;\theta,\mathcal{D}) = (\mu(x;\theta,\mathcal{D}) - y^\dagger - \epsilon)\Phi(\gamma(x)) + \sigma(x;\theta,\mathcal{D})\phi(\gamma(x)) \tag{21}$$

3.1.4.2. BO As a "mapping" between two problems. BO is an strategy to transform the problem

$$u_M = \underset{u \in \text{constraints}}{\text{argmax}}\ \overline{J}(u) \tag{22}$$

unsolvabale!

$$u^{next} = \underset{u \in \text{constraints}}{\text{argmax}}\ \alpha_{EI}(u;\mathcal{D}_n,\theta^*) \tag{23}$$

solvabale!

- $\alpha_{EI}(u)$ is inexpensive to evaluate.
- The analytical expression for gradient of $\alpha_{EI}(u)$ is available.
- Still need to find u^{next}, the multi-start BFGS is used for finding u^{next}.

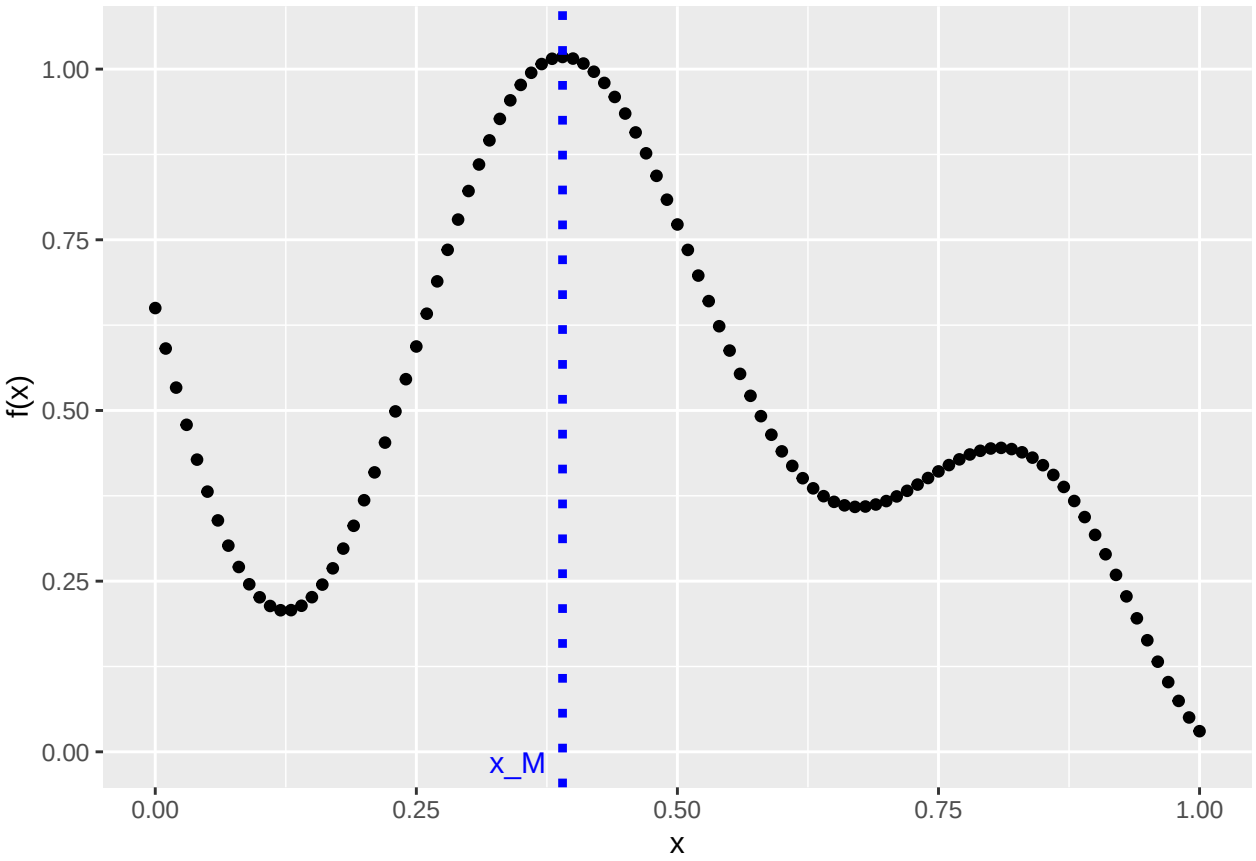


Figure 1: Plot of 1-D equation with blue dash line representing the global optimum

4. 4. Example Cases

4.1. 1-D toy Problem

In this section, a 1-D toy problem is considered to illkuystrate the Bayes Optimization workflow dicussed in the previous section. The 1-D problem was selected sinnce it will help to visuluze all the steps of the workslow making easier explanation of the concepts. Though, it it can bee senn from the next section, the workflow can easily extened to to higer dimetional problem. The *True function* to be optimized in this section has an analytical expression as, given the box constrains:

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & f(x) = 1 - \frac{1}{2}\left(\frac{\sin(12x)}{1+x} + 2\cos(7x)x^5 + 0.7\right) \\ \text{subject to} \quad & 0 \leq x \leq 1 \end{aligned} \tag{24}$$

Since the analytical expression of function available and being 1-D problem, the global optimum of the function had been found at the $x_M = 3.90$. The plot of the function and the optimum point has been shown in the Figure 1

However, it is worth to mention that the analytical expression of objective function in many of real world

problems are not avilable, what is avilable is a *samples* from the objective function. Thefore, in the coming example a few samples are sequentiall drawn from the objective function to resemble the real case scenario. However, we know the global optimum of the objective function in hindsight, just in the case we want to copare the performace of Bayesian optimisation algorithem.

Thefore, as Figure 2, the 5 sample points, $x = [0.05, 0.2, 0.5, 0.6, 0.95]$ were selected as the initialization of the workflow. In the upper plot, blue lines represnets the samples from posterior of the gaussian model conditioned on the five sample points. The grey area represents the 95% confidence interval while the red curve represents the mean value of the samples (blue lines). The first point to infer from the Figure 2 is there no uncertainty on the sample point. As shown, there is no grey zone on sample point since as was dicussed in the previous section, here we consider the "noise-free" observation. Also, worth to mention that we have wide more uncertainity (wider grey band) in the reas that are more distant from the observation, simply meaning we are less uncertain close to observation points. on the "extrpolation," meaning in the ares outseide of the observation points, the probalistic model shows inetrsting behaviour. on those "exterme" area, the mean curve tend to move toward the mean of all observation points , here around 0, showing the model refelctes the mean-revervion behaviour when it comes extrpolation.

The lower part of Figure 2, shows the plot of utility function at each x values. Worth to note that as the plot suggest, the utility(α_{EI}) function will have the muti-modal structure, meaning in the optimization process multi-start gradient method will be helpful, in order to avoid stock in the local optima. In this work, as was explained in the preious section, the multi-start gradient method was used. The blue dotted line line shows the the x_{next} which is the point where the utility function, is maximum. Then this x_{next} is qured from the real function, and the the pair of $(x_{next}, f(x_{next}))$ is added to the intial data set, $\mathcal{D}$. Going back to the lower figure at Figure 2, the utility has two mode around point $x = 0.5$, say $x_{0.5}^{+}$ and $x_{0.5}^{-}$, however the point $x_{0.5}^{-}$ is selected as the next query point. Readers can be rfereed to the upper plot and it is clear that there is more uncertainity around point $x_{0.5}^{-}$ than $x_{0.5}^{+}$ which given the form of utility function, that is understandable. The utility function always looking for the point that not only maximize the mean value, but also intereded in the points that has higher variannce, which is the case between two points $x_{0.5}^{+}$ and $x_{0.5}^{-}$.

Calling the Figure 2 as the iteration # 1, now we can start sampling sequentially. In the Figure 3 another two iteration has been provided. Wher in each row, the plot on the left represents the posterior on the gaussian condistioning, the right show the utility function. Note that in the Figure 3 all axis labels and legned were not included, to have better visibity. (more info about each plot can ben found in 2) . Interesting to see tat in this example case, at the irteration #2, the workflow query the point $x = 0.385$ which presents the best point so far found through BayesOpt workflow. Thefiore, after just two iteration we are around $\frac{x_{best}}{x_M} = \frac{0.385}{0.390} = 98.7$ of the global optima. Although this is case for 1-D problem, it is clearly showing the stength of the workflow

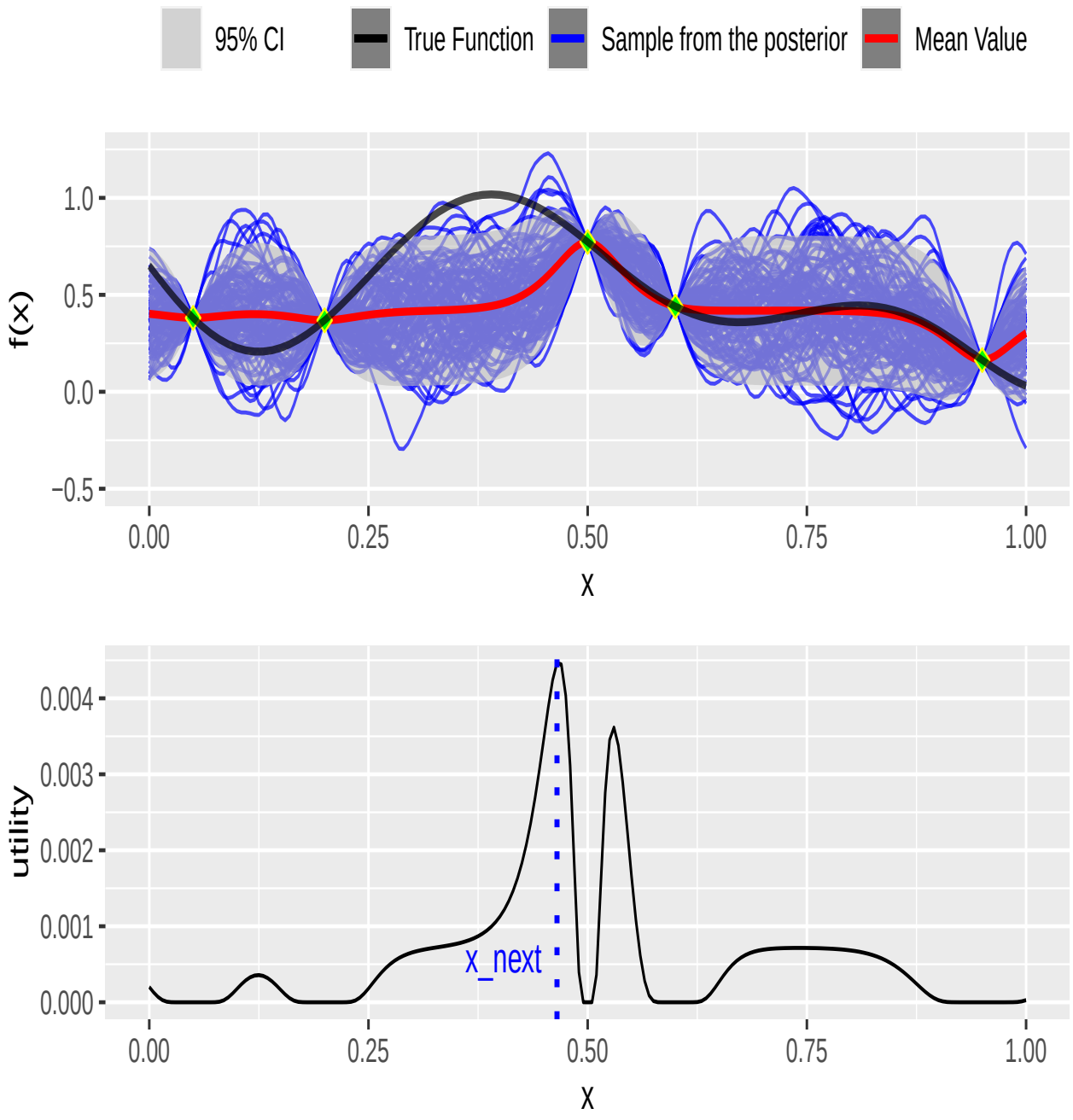


Figure 2: Ite1 - Top: Gaussian posterior over the initial sample points; Lower: Utility function over the x values

to approach the lglobal optima, in as few as possible iteration. In this case after iteration#2, tyhe total number of tim,e that the real function has been quered is $\text{size}(\mathcal{D}) + \text{size}(totaliteration) = 5 + 2 = 7$.

Before going to apply the same workflow at the field scale, the 1-D example presented here offer another useful feature of the Bayesian Optimisation. Looking at 3, we can see that the maximum of utility function is at the iteration # 3 in order of 10^{-6} . That show that after optimization, eve best point to be queried in the next section has a very little utility. So can safely stop the process, since querying points to be sampled from the expensive function has a negligible potential to improve our search in optimization.

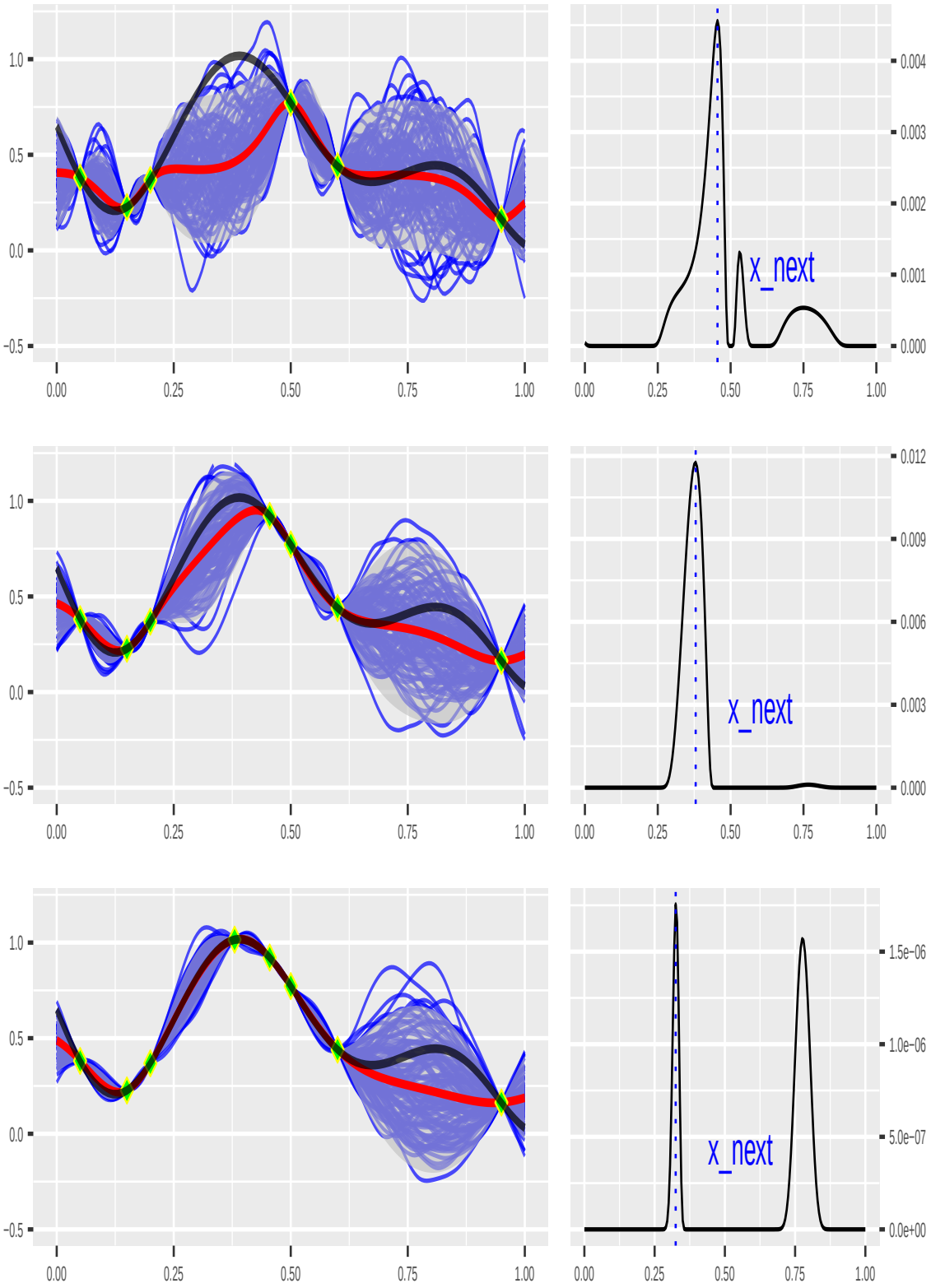


Figure 3: Gaussian posterior of over the initial sample points

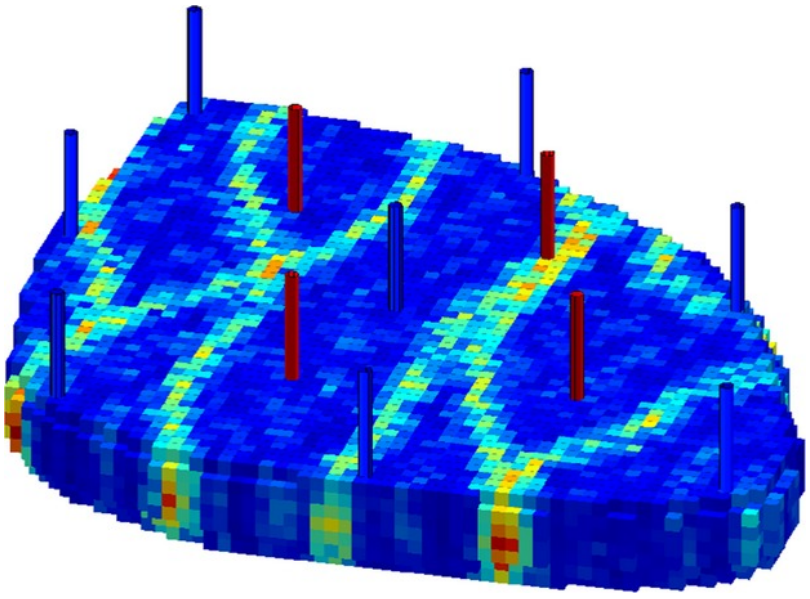


Figure 4: Well locations in Egg model, blue ones are injection, the red producers

4.2. Field Scale

4.2.1. Synthetic 3D Reservoir Model

In this section, the BayesOpt workflow is applied to the synthetic 3D reservoir model. The trough introduction of the model and gelogical describtion can be found in (Jansen et al., 2014) . Known as "Egg Model" it has a geology of channelized depositional system.

The 3D model has eight water injectors and four producers wells shown in Figure 4. The has a geological realizations of patterns of highly permeable channels which are described by 100 equi-probable geological realizations, three of which are illustrated in left side of Figure 5.(Hong et al., 2017).

Relative permeabilities and the associated fractional flow curve of the model have shown in right side of Figure 5 .All the wells are vertical and completed in all seven layers. Capillary pressure is ignored. The reservoir rock is assumed to be incompressible. The model has a life-cycle of 10 years. Here, the injection rate to be maiintaned over life-cycle of reservoir is going to be optimized. Thus, given eight injection wellls, the optimizatijon workflow has the eight dimnetions.However, the optimization in not unbounded, the water can be adjusted from 0 to 100 m3/day, making the box-constrain optimization. The injectors are operated with no pressure constraint, and the producers are under a minimal BHP of 395 bars without rate constraint.

4.2.2. Well Control Optimization

Reviewing the equation raised in the section 3, here the goal is robust optimization of the field , given geological realizations as follow:

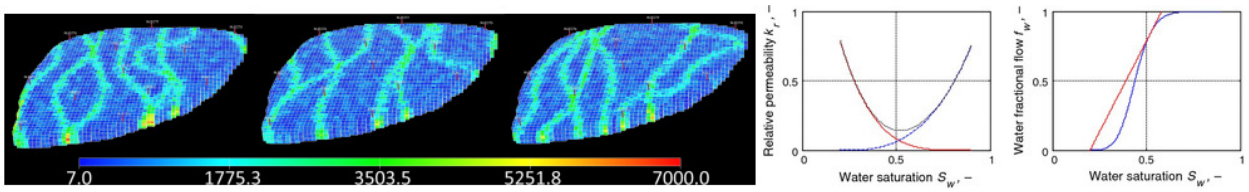


Figure 5: Left: Three geological realizations of the 3D model; Right: Rel perm and fractional flow curve

$$\text{Objective Func(u)} = \overline{J}(u) = \frac{\sum_{i=1}^{n_e} J_r(u, G_i)}{n_e} \tag{25}$$

Equation (25)

u is Injection rate for the each injection well, therefore the control vector, to be optimizaed in this case is defined as:

$$u = [u_{inj1}, u_{inj2}, u_{inj3}, u_{inj4}, u_{inj5}, u_{inj6}, u_{inj7}, u_{inj8}]^{\intercal} \tag{26}$$

As the (25) suggest, the $\overline{J}(u)$ need some parameters to be defined. The oil price (P_o), water production cost (p_{wp}) and water injection cost (P_{wi}) in $dollar/m^3$ has been provided in the Table 1. Also, in this work the cash flow is disconted daily and the discount factor is avilable in the 1. We would like to note that in this work due to avoid further computional burden in optimization process, 10 realizations of the egg model has been considered, therefore $n_e = 10$ in Equation (25).

Table 1: Required Parameters needed for calculation of Expected NPV

Item	Pric	Items	Value
P_o	315	b	8%
P_wp	47.5	D	365
P_wi	12.5	n_e	10

4.2.3. BayesOpt Workflow

As it was discussed, the starting point of the BAyesOpt workflow is to randomly sample the initial data pairs $\mathcal{D}$ which is used to build the Gaussian model of the response surface to the input variables. In this work, forty samples fom the Latin hyper cube sampling (LHS) method were drawn. The LHS is prefred in this work to Monte Carlo since it provides the stratifcation of the CDF of each variable, leading to better coverage of the input variable space. The Figure 6 show the results of the $\overline{J}(u)$ for each sample from LHS. Also, The maximum $\overline{J}(u)$ found from sampling has been shown with blue line. Setting the specific seed number (since LHS is in itself is random process), we get the max NPV aciehved here was 35.65$\$MM$. Looking at Figure 6

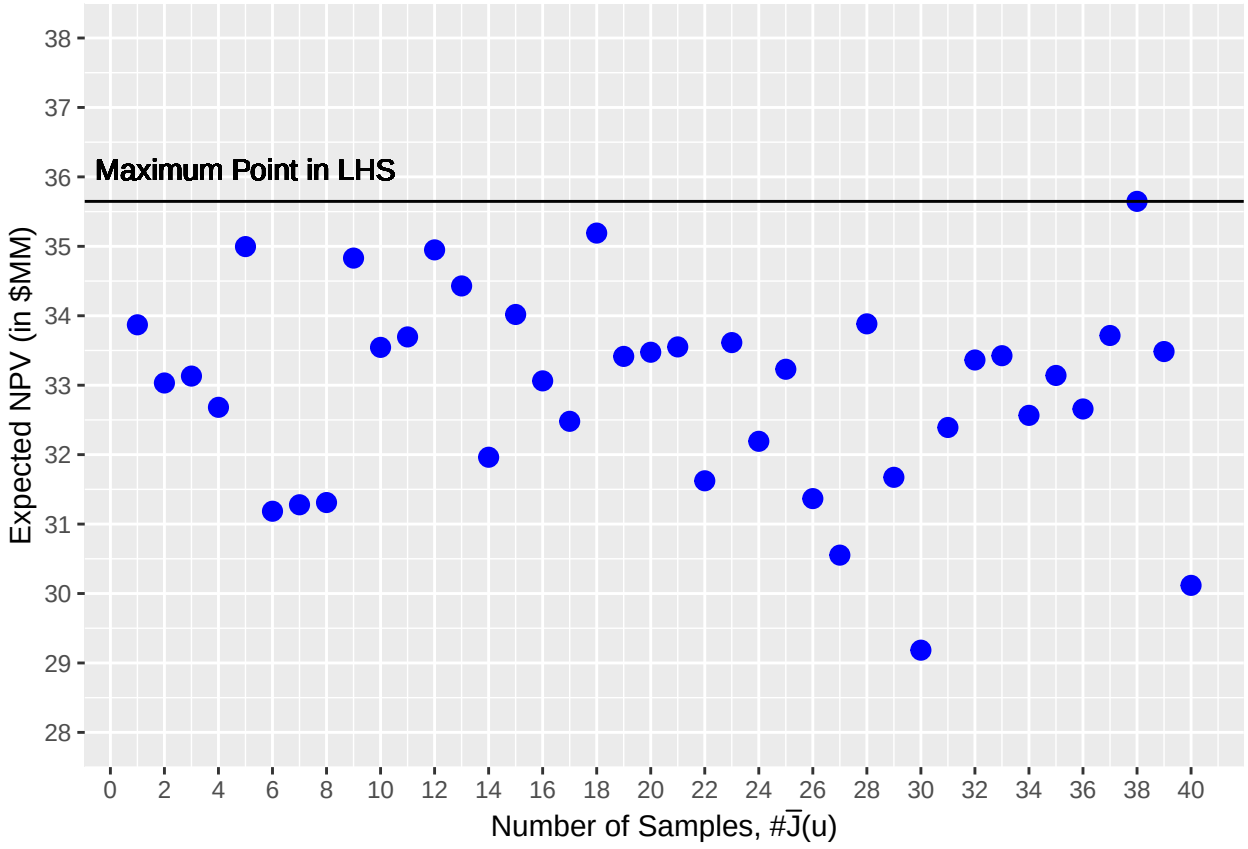


Figure 6: Expected NPV as result of forty sampling from LHS

it is worth to mention that random sampling like the LHS is not helpful to consistently approach the global optimum point, and there is a need for efficient workflow to find the optimum point while using the a few as possible sampling from real function.

Having the initial data found through LHS, we can build the probalistic model of the reposnse surface and sequentially sample from the *expensive-to-evaluate* function. Unfortunately, win this section we can not plot the posterior of the probalistic model, condition on the above forty LHS samples, due being the space is eight-dimetional, and hard to visulize. The Figure 7 shows the expected NPV found after ten sequential sampling resulted from the BayesOpt workflow. Readers are refreed to this point that in the figure, not all red points are increasing and some points are lower than previous points. The reason for this behaviour is the nature of BayesOpt algorith. We can suggest that in the points that has lower expected NPV from the previous, we may reached the lower optimum point, but those points helped us to decrease the uncertainity, which is helpful for the further sampling. We can see that after just ten evaluation of the expenside function (here it means finding the expected NPv from running 10 geological realization using flow simulation) we reach the new record Expeted NPV of $max\overline{J}(u) = 36.85\$MM$.

Now, as we explained in the 1-D section, the plot of the utility at each iteration could provide some useful

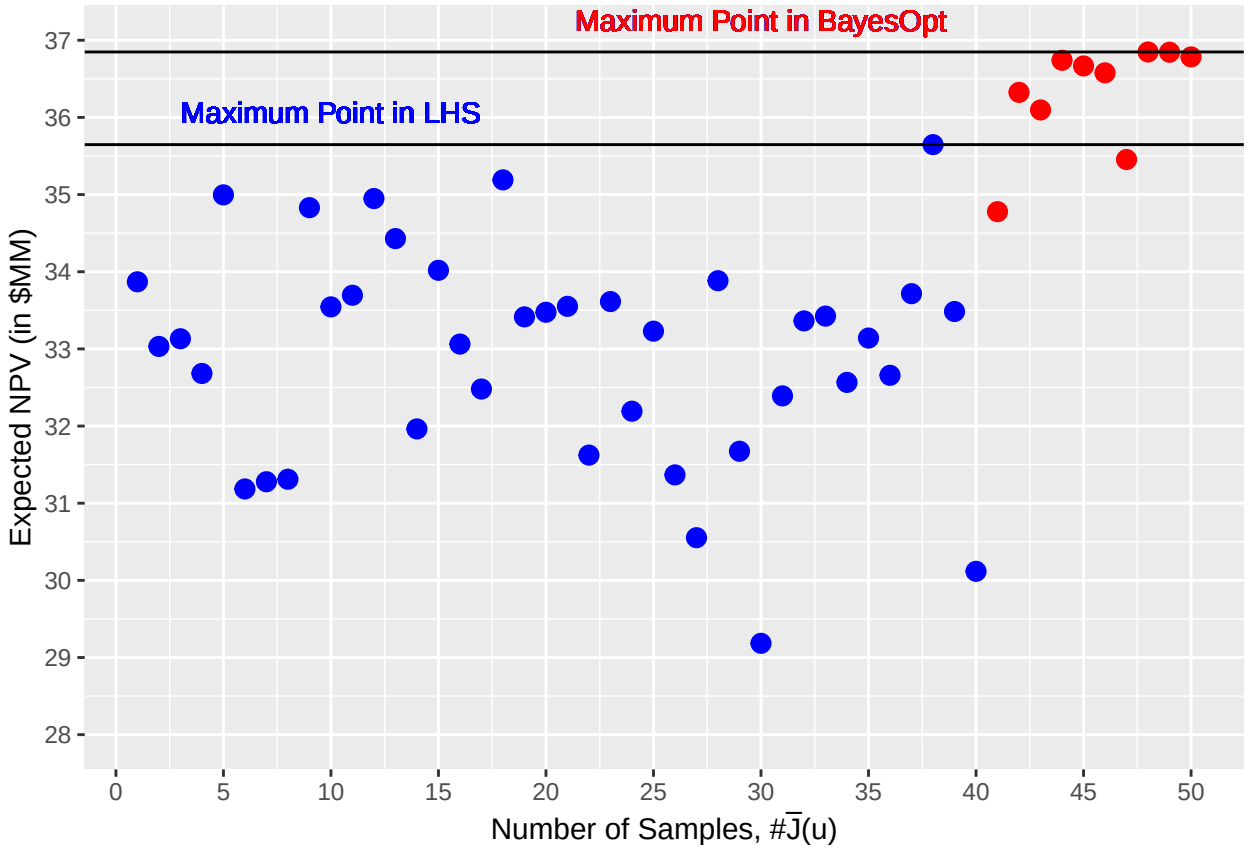


Figure 7: Blue points represnts the sample from LHS, red points represents the samples from the BayesOpt Workflow

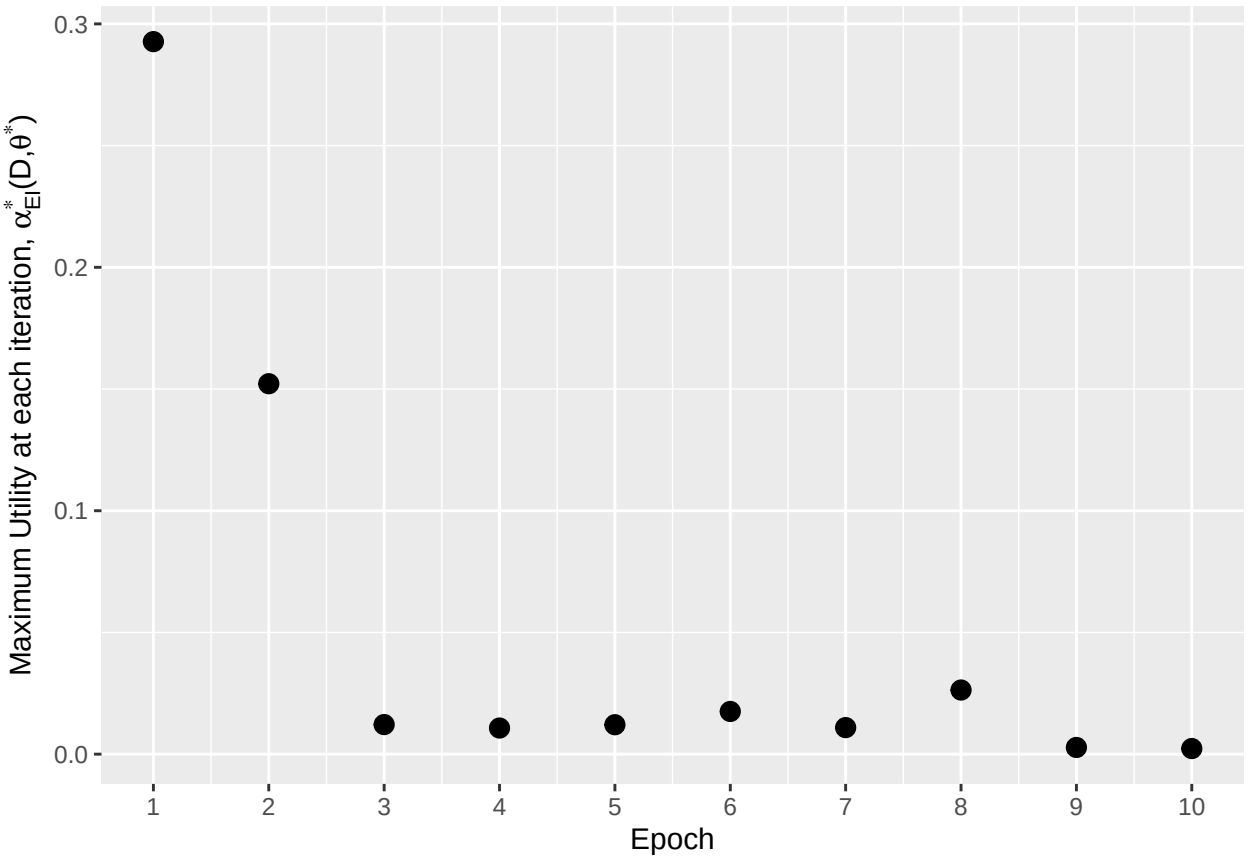


Figure 8: Maximum utility at each iteration, after running L-BFGS-B to find the u with max utility, α^*_{EI}

information about the optimization process. The Figure 8 plots the $\alpha^*_{EI}(\mathcal{D}, \theta^*)$ (Equation (23))versus the ten iteration in this work. In fact the notaion α^*_{EI} means the optimum of the $\alpha_{EI}(u; \mathcal{D}, \theta^*)$ after running multi-start (1000)- L-BFGS-B on all u values. Now, we can see that in the figure the α^*_{EI} is decreasing going toward the zero. It can be inferred from this trend that, we are going out of the *good* u values to be sampled from the expensive function, can be intepreted that we are in the vicinity of global optima, if we see after several iteration still α^*_{EI} is less than 10^-6. Repeat the Optimization, three times, in different initial design points. BayesOpt Applied to Reservoir Case: Final Solution u, in three different initial design:

Repeat the Optimization, three times, in different initial design points

BayesOpt Applied to Reservoir Case:

Final Solution u, in three different initial design:

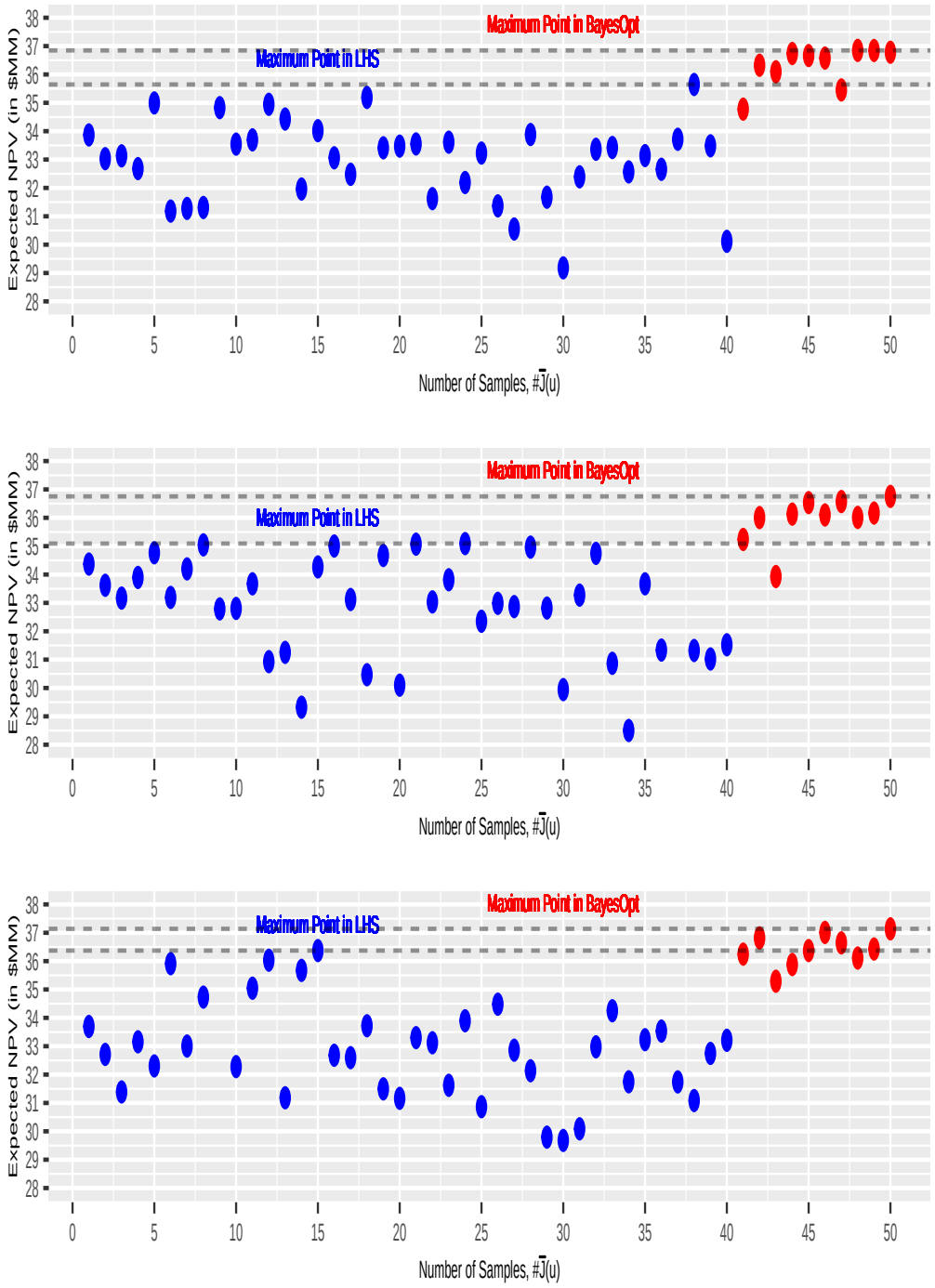
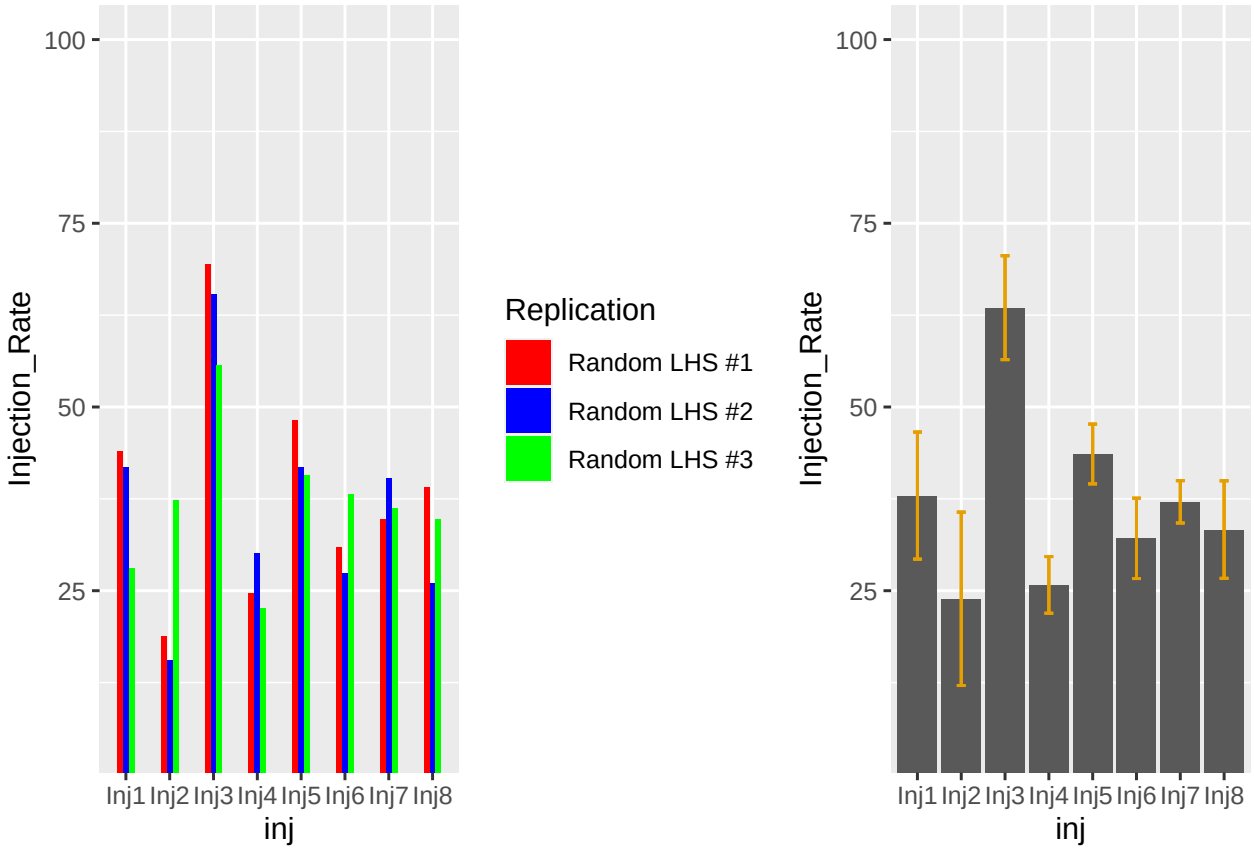


Figure 9: BayesOpt workflow applied to Syntetic 3D model, in three different initialization



4.3. BayesOpt VS. with other Global Optimization Algorithms

4.3.1. Particle Swarm Optimization (PSO)

- de Brito, D. U., & Durlofsky, L. J. (2021). Field development optimization using a sequence of surrogate treatments. Computational Geosciences, 25(1), 35-65.

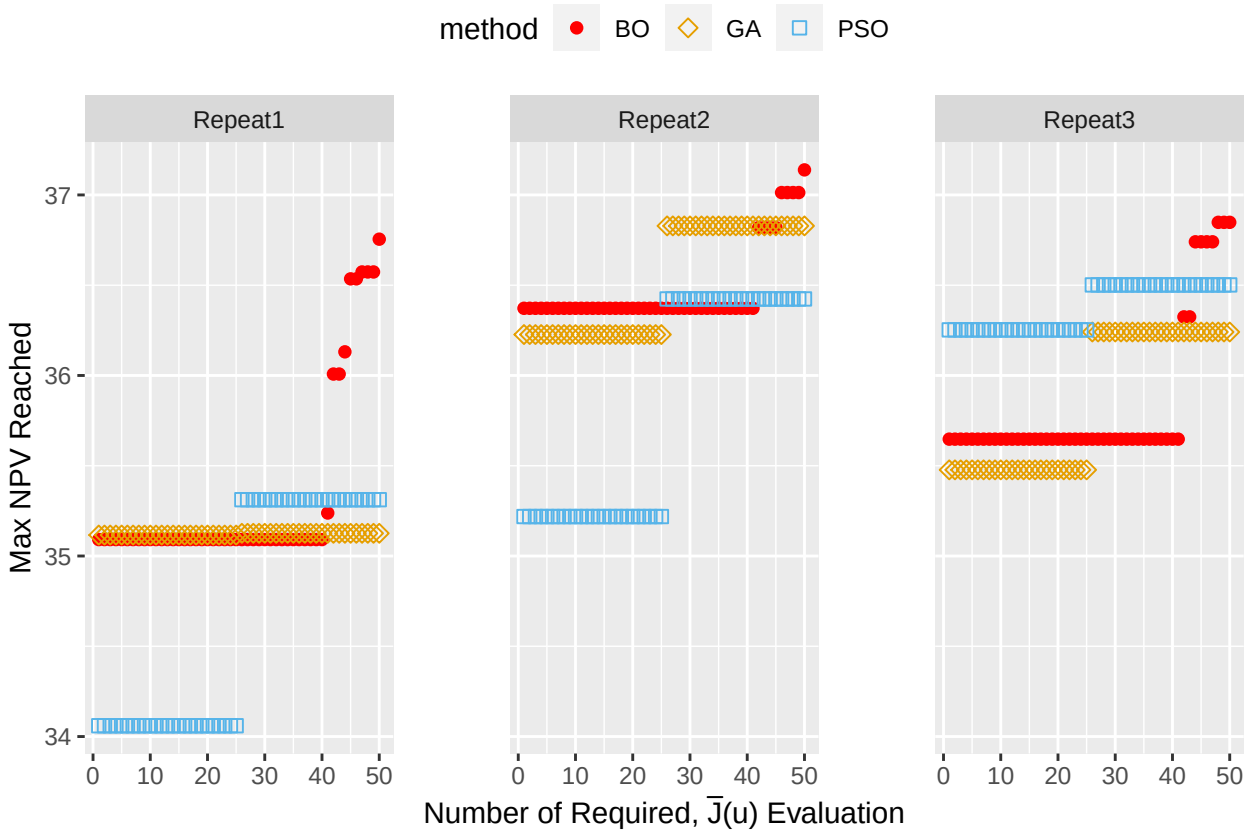
- Jesmani, M., Bellout, M. C., Hanea, R., & Foss, B. (2016). Well placement optimization subject to realistic field development constraints. Computational Geosciences, 20(6), 1185-1209.

4.3.2. Genetic Algorithm Optimization (GA)

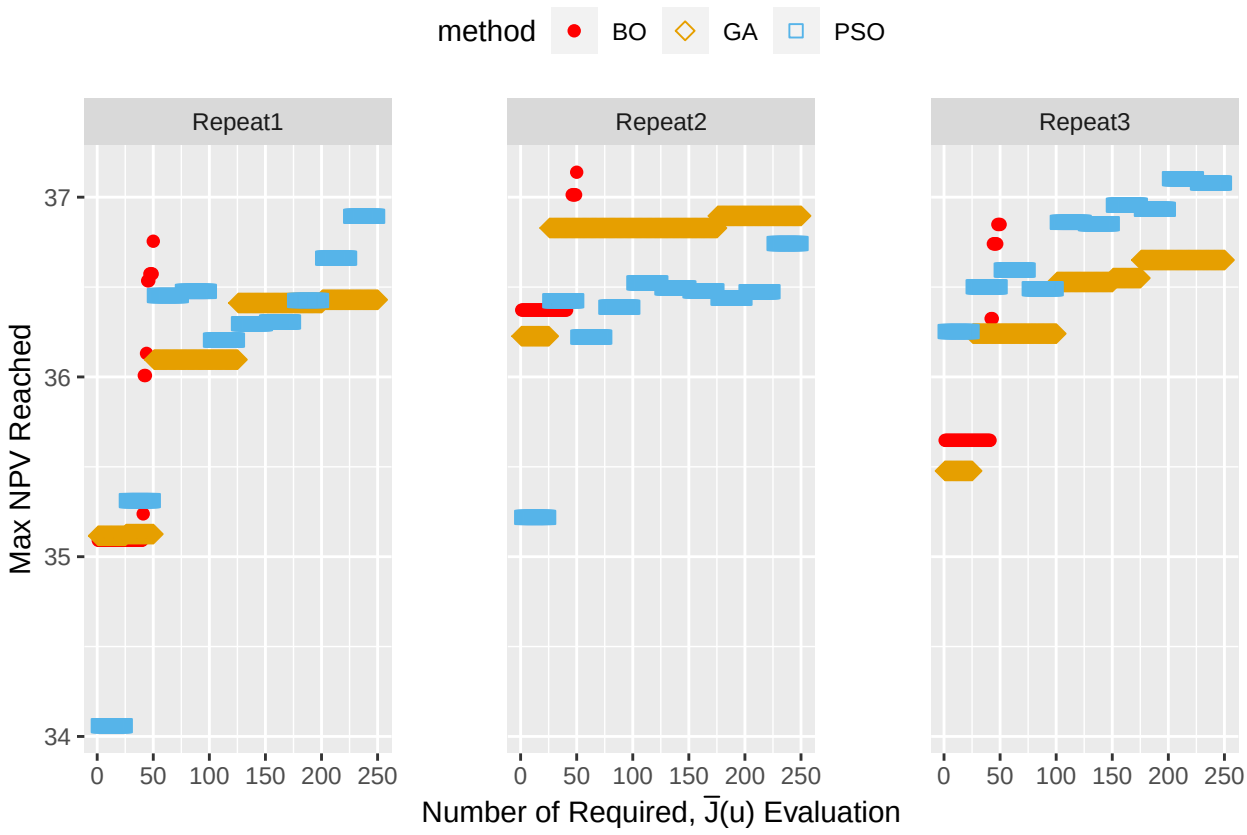
- Chai, Z., Nwachukwu, A., Zagayevskiy, Y., Amini, S., & Madasu, S. (2021). An integrated closed-loop solution to assisted history matching and field optimization with machine learning techniques. Journal of Petroleum Science and Engineering, 198, 108204.
- Bukhamsin, A. Y., Farshi, M. M., & Aziz, K. (2010, January). Optimization of multilateral well design and location in a real field using a continuous genetic algorithm. In SPE/DGS Saudi Arabia Section Technical Symposium and Exhibition. Society of Petroleum Engineers.

BayesOpt VS. with other Global Optimization Algorithms

Fixed Reservoir Simulation Budget (N=50)



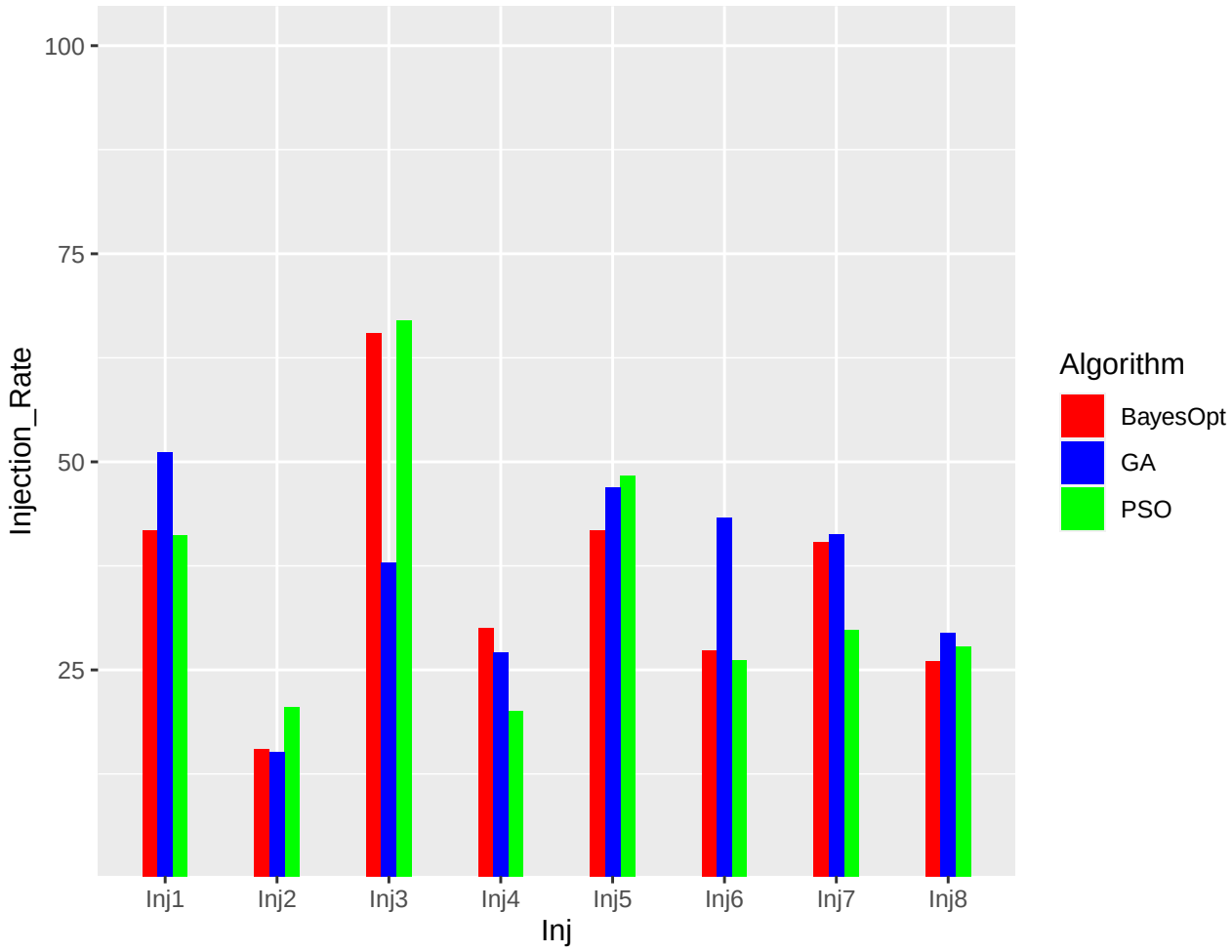
BO: Reservoir Simulation Budget (N=50), GA, PSO: Reservoir Simulation Budget (N=250)



BayesOpt VS. with other Global Optimization Algorithms

Table Summarizing Comparison of BayesOpt, PSO, GA

Comparing the Final Solution u of the Opt algorithms. . . (the Median Replication was used)


100
75
50
25
Injection_Rate
Inj1
Inj2
Inj3
Inj4
Inj5
Inj6
Inj7
Inj8
Inj
Algorithm
BayesOpt
GA
PSO

251

5. 5. Conclusing Remarks

6. Acknowledgements

This work received support from the Research Council of Norway and the companies AkerBP, Wintershall–DEA, Vår Energy, Petrobras, Equinor, Lundin, and Neptune Energy, through the Petromaks–2 DIGIRES project (280473) (http://digires.no). We acknowledge the access to Eclipse licenses granted by Schlumberger.
Here are two sample references: (Dirac, 1953; Feynman and Vernon Jr., 1963).

References

Albertoni, A., Lake, L.W., 2003. Inferring Interwell Connectivity Only From Well-Rate Fluctuations in Waterfloods. SPE Reservoir Evaluation & Engineering 6, 6–16. doi:10.2118/83381-PA

Bruce, W.A., 1943. An Electrical Device for Analyzing Oil-reservoir Behavior. Transactions of the AIME 151, 112–124. doi:10.2118/943112-G

Dirac, P.A.M., 1953. The lorentz transformation and absolute time. Physica 19, 888–896. doi:10.1016/S0031-8914(53)80099-6

Fedutenko, E., Yang, C., Card, C., Nghiem, L.X., 2014. SPE Heavy Oil Conference-Canada. SPE, Calgary, Alberta, Canada, p. D021S008R001. doi:10.2118/170085-MS

Feynman, R.P., Vernon Jr., F.L., 1963. The theory of a general quantum system interacting with a linear dissipative system. Annals of Physics 24, 118–173. doi:10.1016/0003-4916(63)90068-X

Hong, A.J., Bratvold, R.B., Nævdal, G., 2017. Robust production optimization with capacitance-resistance model as precursor. Computational Geosciences 21, 1423–1442. doi:10.1007/s10596-017-9666-8

Jansen, J.D., Fonseca, R.M., Kahrobaei, S., Siraj, M.M., Essen, G.M.V., Hof, P.M.J.V. den, 2014. The egg model – a geological ensemble for reservoir simulation. Geoscience Data Journal 1, 192–195. doi:10.1002/gdj3.21

Mohaghegh, S.D., Guruswamy, S., 2006. Development of Surrogate Reservoir Models (SRM) for Fast-Track Analysis of Complex Reservoirs. San Antonio, Texas, U.S.A, p. 9.

Murphy, K.P., 2022. Probabilistic machine learning: An introduction. MIT Press.

Sampaio, T.P., 2009. An Application of Feed Forward Neural Network as Nonlinear Proxies for the Use During the History Matching Phase 11.

Sayarpour, M., 2008. Development and application of capacitance-resistive models to water/CO? floods (PhD thesis).

Yousef, A.A., 2006. A Capacitance Model To Infer Interwell Connectivity From Production- and Injection-Rate Fluctuations 17.

283 Zhao, H., Kang, Z., Zhang, X., Sun, H., Cao, L., Reynolds, A.C., 2015. SPE Reservoir Simulation Symposium.
284 Society of Petroleum Engineers, Houston, Texas, USA. doi:10.2118/173213-MS