T H E U N I V E R S I T Y O F T U L S A

THE GRADUATE SCHOOL

LIFE-CYCLE PRODUCTION OPTIMIZATION OF STEAM ASSISTED GRAVITY DRAINAGE WITH GAUSSIAN PROCESS REGRESSION PROXY MODEL

by
Yuhao Chen

A thesis submitted in partial fulfillment of

the requirements for the degree of Master of Science in Engineering

in the Discipline of Petroleum Engineering

The Graduate School

The University of Tulsa

2019

T H E U N I V E R S I T Y O F T U L S A

THE GRADUATE SCHOOL

LIFE-CYCLE PRODUCTION OPTIMIZATION OF STEAM ASSISTED GRAVITY DRAINAGE WITH GAUSSIAN PROCESS REGRESSION PROXY MODEL

by
Yuhao Chen

A THESIS

APPROVED FOR THE DISCIPLINE OF

PETROLEUM ENGINEERING

By Thesis Committee

Mustafa Onur, Chair
Ram S. Mohan
Ovadia Shoham

COPYRIGHT STATEMENT

ABSTRACT

Yuhao Chen (Master of Science in Engineering in Petroleum Engineering)

Life-Cycle Production Optimization of Steam Assisted Gravity Drainage with Gaussian Process Regression Proxy Model

Directed by Mustafa Onur

78 pp., Chapter 6: Conclusions

(450 words)

Conventional numerical optimization methods (such as adjoint, SPSA, StoSAG, genetic, etc.) are well established. However, their applicability for solving complex reservoir simulation problems such as the steam assisted gravity drainage (SAGD) process, which involves multiphase thermal flow including convection, conduction, and heat losses to the cap rock and base formation, considered here is limited due to the fundamental challenge of the high computational cost of accurate, high-fidelity reservoir simulators. Instead of using the high-fidelity model, an accurate proxy model which can be used for computationally efficient optimization is preferable. In fact, recent research works conducted at The University of Tulsa Petroleum Reservoir Exploitation Projects (TUPREP) show that using machine learning(ML)-based proxy models, such as the one based on support-vector regression (SVR) are quite promising for efficiently solving the life-cycle optimization, history matching, and uncertainty performance prediction problems for complex reservoir systems and processes.

The objective of this research work is to investigate the efficient estimation of the optimal well-control (or operating) parameters for the life-cycle production optimization problem of SAGD Process by ML-based proxy models. The scope of the research involves a ML-based method to form proxy. This is Gaussian Process Regression (GPR). The cost function to be optimized is the net present value (NPV). The control parameters for the problem of interest

are injection rate and temperature at the horizontal injector, and the bottom-hole pressure at the horizontal producer, given some constraints for these well-controls. Given simulation results, a GPR based proxy is built, and then the optimal control parameters are found by maximizing the NPV based on the proxy as a forward model. The performance of the proxy model is then compared with that of the popular stochastic simplex approximate gradient method (StoSAG).

In this study, we discuss various flow and transport mechanisms that have influence on production from heavy-oil reservoirs by a SAGD process. Then, we present the NPV cost function of interest to be maximized by optimizing the well-control parameters. We present the mathematical details of the GPR method to maximize the NPV cost function. We perform a sensitivity study to understand the effect of various control parameters of interest on production by a SAGD process. For the sensitivity study and generating training sets to be used for a proxy, we use a commercial simulator; Computer Modelling Group's (CMG) STARS thermal simulator. Lastly, we present our results on optimization of production by using a GPR proxy model. Note that a GPR model is a non-parametric kernel-based probabilistic model. The results show that GPR has poor prediction performance when the training set is small, due to the characteristic of the evidence maximization process. However, performance improves quickly when more samples are included in the training process.

ACKNOWLEDGEMENTS

I would like to thank my advisor Prof.Dr. Mustafa Onur for his support and continuous guidance. I am thankful to Zhe Liu who assisted and encouraged me to study the StoSAG, Machine Learning and several computer language. With the help of Prof.Dr. Albert C. Reynolds, I got the opportunity to practice my presentation skill, and presented my work on Advisory Board Meeting (ABM). Thanks to Azad Almazov and Davut Erdem Bircan who helped me a lot on Matlab coding. I also thank Prof.Dr. Ovadia Shoham and Prof.Dr. Ram S. Mohan for serving on my MS thesis committee members.

I extend my thanks to the Computer Modeling Group (CMG) for providing the simulator licenses and patiently replying to my question.

I sincerely thank my parents who provided me the opportunity to study in The University of Tulsa and gave me their unconditional love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

There are large deposits of heavy-oil including bitumen resources in the world, which make up about 70% of the total oil resources. Heavy-oil is classified as any oil having an API gravity between 10° and 20° and viscosity greater than 100 and up to 1×10^6 cp. All methods to produce heavy-oil involve some form of thermal process that aims to reduce oil viscosity by introducing heat to the reservoir. Steam assisted gravity drainage (SAGD) is an effective thermal recovery method for heavy oil and bitumen extraction. Generally, a pair of horizontal wells is used in a SAGD process. In a typical SAGD approach, initially, there is a preheating period, in which steam is circulated in both wells to mobilize oil around wells through heat conduction. After the preheating period, steam is injected into a horizontal well located directly above a horizontal producer helping to displace heated oil. As steam enters the reservoir, it heats the reservoir fluids and surrounding rock. Hot oil and condensed water drain through the force of gravity to a production well at the bottom of the formation. SAGD maximizes the role of gravity forces during steam flooding of heavy oils.

1.1 Literature Review

Heavy oil is a complex amalgam with compounds of high molecular weight. For millions of years, hydrocarbon deposits have been degraded by algae to form heavy oil, which has resulted in the loss of their lighter hydrocarbon fractions - the heavier fractions remaining. Heavy oil is the liquid petroleum with API gravity between 10° and 20° and by definition, typically, heavy oil has a viscosity of 100 cp or greater [31]. Heavy oil usually contains significantly higher levels of asphaltenes. Extra heavy oil is commonly referred to as a heavy oil with viscosity, which is a characteristic of defining the flow, exceeding 10,000

cp. Heavy oil ranges from the oil itself to the asphalt - or tar sands “overweight oil" - actually embedded in the sand. The tar sands are less than 75 meters deep and are mined as "solid" rather than in liquid form. More than 90% of the heavy oil and bitumen (oil sands) in the world are deposited in Canada and Venezuela. As Canada's conventional crude oil resources continue to decrease, further development of heavy oil and oil sand in-situ recovery technologies is critical to meeting Canada's current and future energy needs [33]. The scale of the heavy oil resources is huge and the technology has made great progress. One method is to mine the shallow Athabasca deposit, which has become the main source of Canadian oil. Other new tar sands mining projects are underway. Although mining overcomes the problem of delivering oil to oil wells, it requires large amounts of solids and sludge to treat. Therefore using mining method on the shallowest deposits is economical. However, the main part of Canadian oil sands resources is too deep to be obtained through mining. The challenges to efficiently produce heavy oil is given in the following: 1) to find and define suitable reservoirs; 2) to create conditions for oil to flow at economic rates; 3) to drain the reservoir systematically to obtain high recoveries. There are several ways to achieve these challenges: 1) Cold production using vertical wells and horizontal wells; 2) Stimulation by wellbore heating; 3) Cyclic steaming using conventional wells; 3) Steam Assisted Gravity Drainage (SAGD); 4) Steam and Gas Push (SAGP); 5) Cyclic Steaming with horizontal wells; 6) Vapour Extraction (Vapex).

As the availability of conventional crude oil has declined, it developed several ways to improve the recovery of the reservoir, which called enhanced oil recovery (EOR). The most important way that can be used on heavy oil reservoir is thermal recovery, which contains steam stimulation, steam flooding, hot water flooding, and in-situ combustion.

Shell discovered the steam stimulation process in Venezuela accidentally, when it produced heavy crude oil from steam flooding in the Mene Grande field near the eastern shore of Lake Maracaibo. During steam stimulation, steam is injected into the reservoir for several weeks at a rate of approximately 1000 bbl/day. The well is refluxed and subsequently pumped. If in a suitable application, the production of oil is rapid and the process is effective

in the early cycles. It can also be used to produce very viscous oils for oil sands if the steam pressure is high enough to destroy the reservoir and thus allow for injection. For this operation, the steam pressure is required to be about 1 psi per foot depth to overcome the in-situ rock stress to cause fracturing. The main disadvantage of the cyclic steam stimulation process is that it typically only allows about 15% of the oil to be recovered before the oil-steam ratio becomes too low.

In steam flooding, steam is forced continuously into specific injection wells and oil have driven to separate production wells. The area around the injection well is heated to the saturation temperature of the steam and these areas extend towards the production well. Steam flooding process needs a lot more steam than steam stimulation. It is often economical to switch the steam flooding after initial operation of the field by steam stimulation. The oil recovery rate of steam flooding can reach 50% or even higher.

Hot water flooding is usually less effective than steam flooding since the temperature of the hot water flooding is lower than steam flooding. In-situ combustion is the method injecting air or oxygen to the combustion zone through a well drilled from the surface. In this case, it is much cheaper to generate steam than other methods. In-situ combustion tends to be less stable than steam processes. It is common that premature arrival of the combustion front at the production wells [6].

1.1.1 The Definition of SAGD

In the late 1970s, Butler et al. introduced the definition of the Steam Assisted Gravity Drainage (SAGD) process. A cross-section of the SAGD process is displayed in Figure 1.1. There is a pair of horizontal wells drilled into the reservoir. Steam is injected into the formation through the top well, which is the injection well. The production well typically is placed 5 to 10 m below and parallel to the injection well, and the production well usually is located several meters above the base of pay. A steam chamber grows around and above the injection well. At the edge of the steam chamber, heated bitumen and steam condensate flow to the production well under the influence of gravity. Generally, the lengths of the injection

well and the production well are between 500 and 1000 m (Figure 1.2).

At the beginning of the SAGD process, a pre-heating period is needed, which may take several weeks to months. One method of pre-heating is injecting the high-temperature steam into both horizontal wells and the steam can be circulated back through the annulus of the two wells. Because of the high-temperature steam, the bitumen around the wells is heated up and may become mobile oil. Generally, a hydraulic connection can be generated between two wells. Once the connection is created, the top horizontal well becomes the injector and the bottom one becomes the producer. Because of the effect of gravity, the injected high-temperature steam rises up, touches the cold reservoir rocks and bitumen at the boundary of the steam chamber, the steam trans its heat to the surrounding reservoir rocks and bitumen, and the steam becomes condensed water. Since the bitumen becomes mobile oil, the viscosity of the bitumen is also decreased [4]. Because of the gravity force, the mobile oil and the condensed water drain down to the bottom of the steam chamber where we produce oil from the producer. As the oil flows away and is produced, the steam chamber expands in both upwards and sideways direction. The growth rate precedes rather irregular but fairly fast way until it is limited by the top of the reservoir. The upwardly moving interface tends to be in the form of steam fingers in which the oil flows between them in a rather unstable manner. Instead, the side and down interfaces move in a very stable manner; it is stabilized by gravity.

The nature of this process is that the cold oil is heated and replaced by a steam system and maintains the temperature of the oil as it flows to the lower production well. The next step in this process when the chamber has reached the top of the reservoir and is expending the sideways beneath the overburden, which is shown in Figure 1.3 [7].

Generally, there are three ways of looking at SAGD. In the early development of SAGD, it is focused on the flow that heated oil around the steam chamber and the theories about developing the expected drainage rates calculation. These theories show that the oil adjacent to the chamber is heated by conduction and discharged downward; the viscous resistance of the oil is balanced by gravity force. Based on these theories, oil rates can be

predicted successfully. The second view of SAGD looks on the steam chamber as a large condenser. The steam condenses in the presence of a large heat transfer zone, which is produced as the oil flows to the production well. A large number of reservoirs are heated and as the heated oil slides down, a cooler reservoir is uncovered. In a mature SAGD process, the surface of the room above the production well maybe as much as 10 hectares. When the oil is discharged, the steam fills the pores and maintains the reservoir pressure and drive. SAGD produces a higher recovery than cyclic steam stimulation. The last view of SAGD is that horizontal wells allow for economical production without the need for steam coning. Even if steam is injected above the production well, the steam chamber can be pumped to the production well for about 1 meter without significant steam generation [3]. Compared to traditional California-style steam drives, SAGD has the distinct advantage of achieving high recovery without the need for taper and 4 hectares (10 acres) mode spacing. Therefore, for comparable recoveries, a well spacing of the order of 0.16 hectares (5/8 acre) is required. However, coning of steam is still a problem since reservoir contact is not as good as that provided by horizontal wells [5].

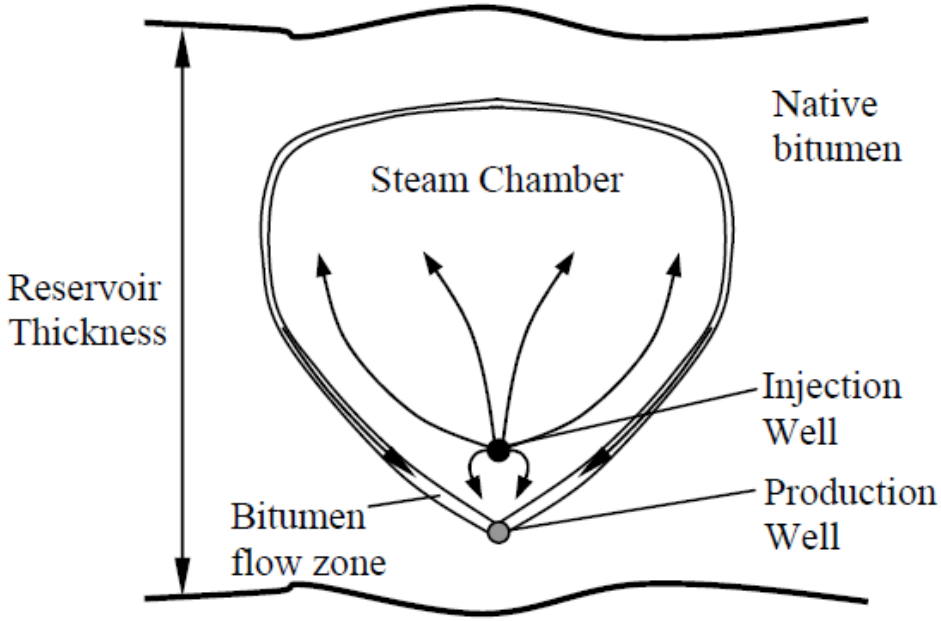

Figure 1.1: Cross-section of the Steam Assisted Gravity Drainage (SAGD) process [21].

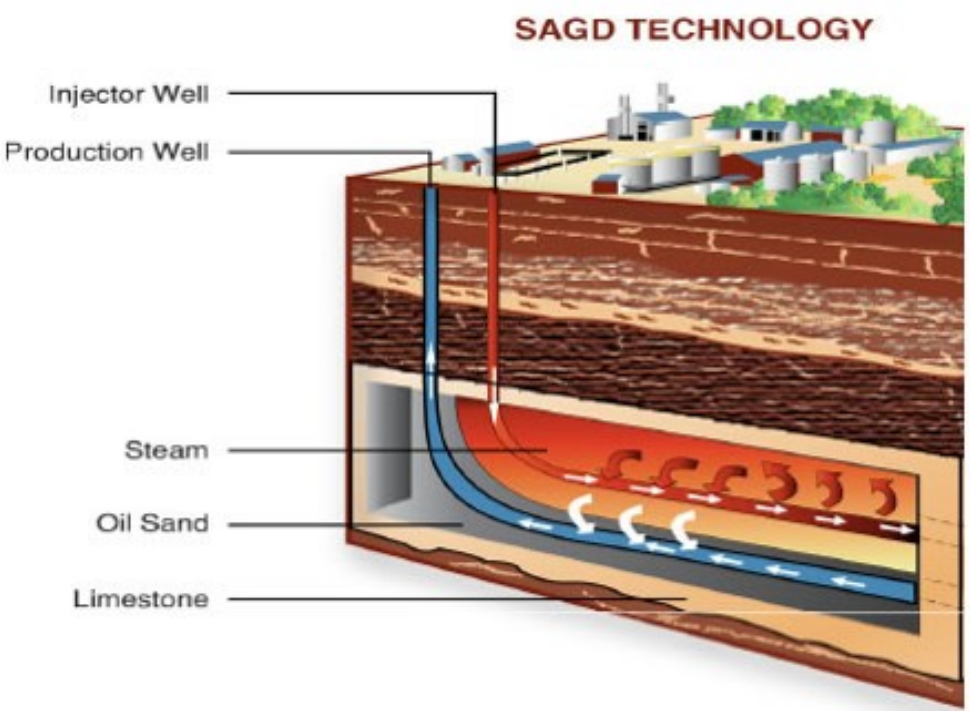

Figure 1.2: The Steam Assisted Gravity Drainage (SAGD) process [13].

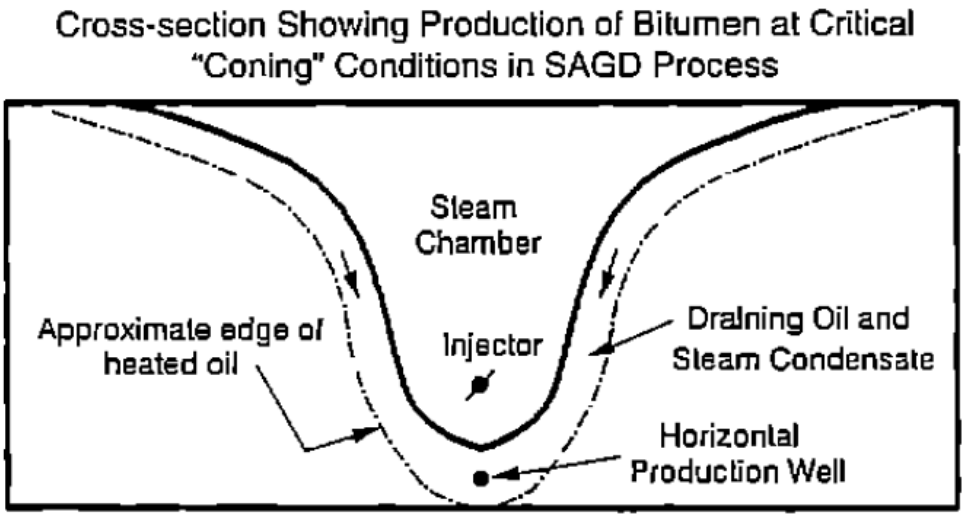

Figure 1.3: Cross-sectional diagram showing the steam-assisted gravity drainage process with the chamber in contact with the overburden. In a commercial project, the lateral spreading of the steam chamber becomes limited when it joins similar chambers on either side. After this, the rate gradually declines as the head available to create drainage decreases [7].

From Figure 1.1 and Figure 1.3, we see that the higher speed of growing the steam chamber, the faster the heavy crude oil produced. To have a high speed growing steam chamber, it is desirable to prevent the direct steam generation and to combine the steam

in the chamber so that when the steam latent heat chamber condenses at the edge of the chamber, and the latent heat of steam can be transferred to the asphalt surrounding the chamber [35]. The key point is to maintain a liquid pool around the producer, in this case, injected steam is produced from the production well [20]. From Figure 1.1, we can say that the heated mobile oil and condensate water flow to the bottom of the steam chamber along the edge of the chamber. The oil and the water generate a liquid pool, and then they are produced to the surface (see Figure 1.4). When the oil production rates go too high, the liquid pool cannot be generated, the injected steam will be produced through the production well directly, without touching the cold rocks and the bitumen. This can reduce the energy efficiency of the SAGD process. On the other hand, when the oil production rates go too low, the liquid pool level may even above the injection well, the area where the hot steam contact with the cold rocks and bitumen is decreased, in that case, the steam chamber grow up rate is also decreased.

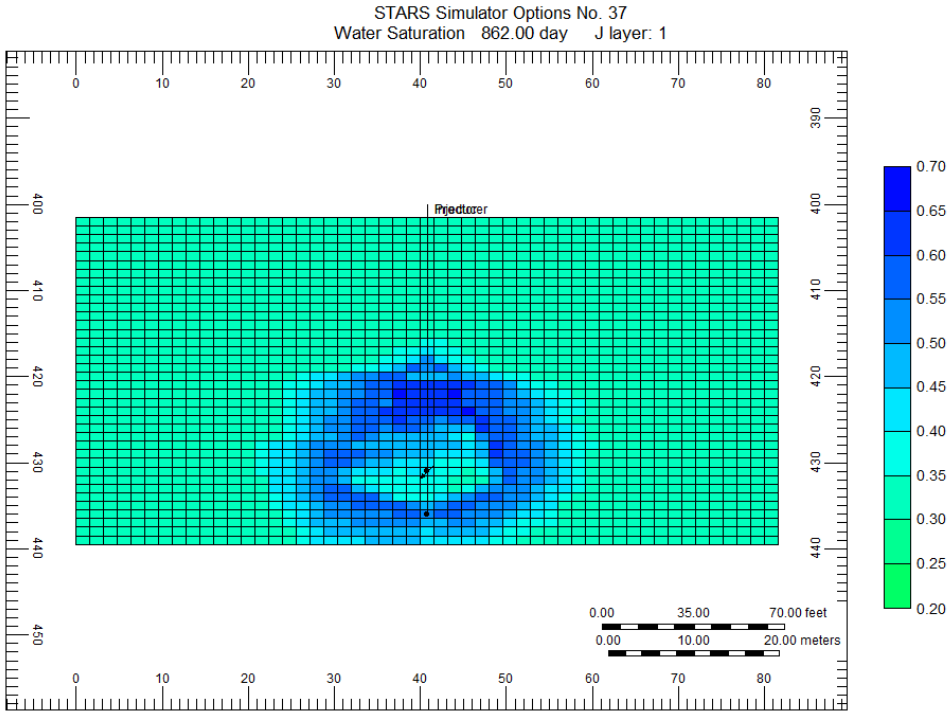

Figure 1.4: Water saturation after SAGD process 862 days. The bottom of the steam chamber, there is a high water saturation place, which is the liquid pool.

1.1.2 SAGD Optimization

Life-cycle production is a subsequent step of assisted history matching in reservoir management [25]. Life cycle production optimization is an optimal well control problem with the goal of determining the conditions for maximizing NPV or cumulative oil production over the assumed reservoir life. [8].

In Figure 1.5, the top part of the figure represents the physical system, the output of this physical system is also the responses from the reservoir simulator. With different geology, fluid properties, many system models can be generated. However, in this research, only one reservoir model which may be considered as a model determined after history matching of the historical production data [27].

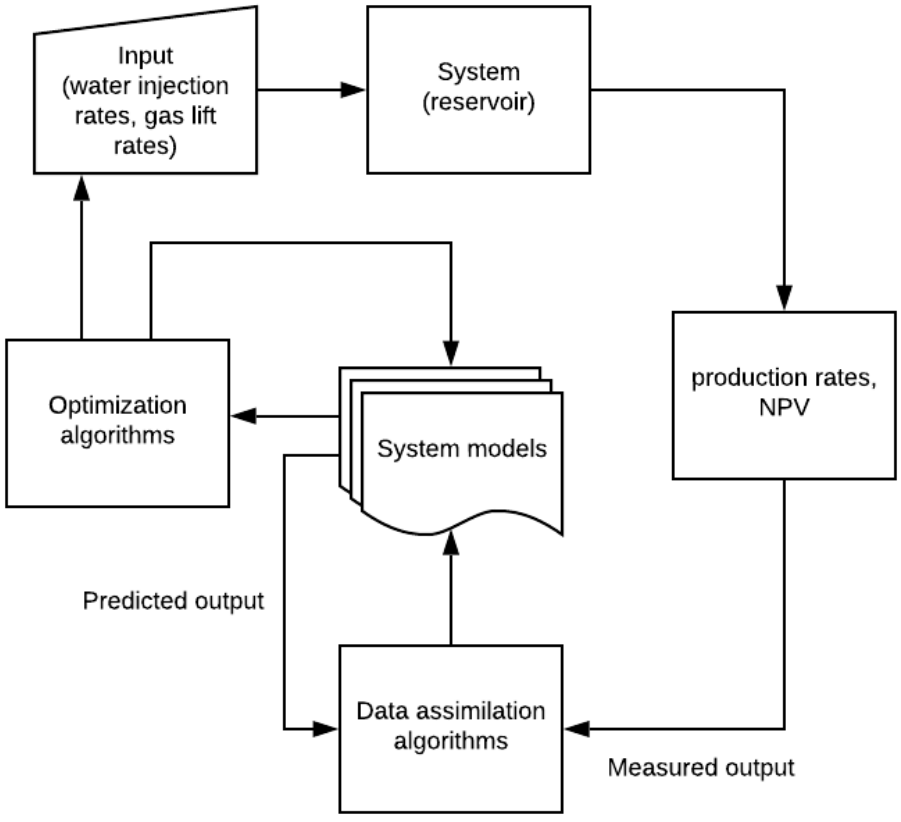

Figure 1.5: Key elements of the closed-loop reservoir management process [27].

During SAGD process a lot of high-temperature steam needs to be injected. This

requires a lot of energy. To find out an economic way to enhance oil recovery, the best way is to manage the heat efficiency, i.e., recover more oil with less steam injection. Egermann et al. developed a general methodology based on numerical investigations to obtain and maintain an optimized development of the steam chamber in SAGD process during the whole production life-cycle for better oil recovery and heat management [15]. In this method, the optimization of the steam chamber development is obtained by adjusting the steam injection rate which can also be affected by the fluid and geology in the reservoir. To keep the steam chamber as large as possible by monitoring productivity during the SAGD process, but away from the production well to prevent the steam breakthrough. The oil production rate can be adjusted by the temperature difference between the injected steam and the production fluid, which is called subcool temperature. The most expensive part in SAGD process is the generation of the high-temperature steam. The thermal efficiency is measured by the cumulative steam-oil ratio (cSOR). The higher the cSOR, the higher the steam usage. The lower the cSOR, the lower the steam used per unit volume of produced bitumen, which implies a more economic process [22]. Kisman et al. [29] evaluated SAGD process performance and determine the sensitivity of the performance of the base case to selected parameters to see whether have a significant impact on performance. Their results showed that carbonate flow barriers, wettability change in the steam zone and the choice of three-phase relative permeability show only a small effect on performance. Permeability, oil viscosity, thermal conductivity, steam gas ratio and volatility, oil relative permeability endpoint value and wettability change in the temperature transition region has a moderate effect on overall performance. The operation pressure in the steam zone has a large effect on the calendar day oil rate (CDOR), and the lower the pressure, the poor the performance. Ito and Suzuki [26] show that optimum subcool temperature is between 30 to 40 °C, which is lower than the steam saturation temperature. Queipo et al. [37] present a technique to use a surrogate model based on Neural Networks, Design and Analysis of Computer Experiments (DACE) modeling and adaptive sampling to optimize SAGD process. They figure out that the neural network-based efficient global optimization (NEGO) algorithm can be used effectively and

improve oil recovery efficiently. On the other hand, this optimization method is useful in the optimization of the objective functions, including performing computationally expensive mathematical models (reservoir numerical simulators), which is found not only in oil recovery processes but also in another petroleum engineering. Edmunds and Chhina [14] used a reservoir simulator to find out the optimum economical SAGD process. They show that SAGD economies are more sensitive to resource quality than other operating parameters. Gates et al. [22] also found that the optimized injection pressure starting with high pressure, then gradually decrease to a lower level. That is because at early times the steam chamber is not in contact with the overburden and overlying water, the heat from the steam directly trans into the bitumen. With the high injection pressure before the steam chamber reaches the top of the rich zone and contacts the overburden, which means high saturation temperature, which can decrease the bitumen viscosity and increase oil productivity. After the steam chamber reaches the overburden, the injection pressure is reduced, which means the saturation temperature is reduced thus reducing the overburden heat loss. As the steam chamber expands laterally across the area, the injection pressure is further reduced. Bao et al. [2] also found that SAGD performance is more sensitive to injection pressure. The higher the injection pressure the earlier steam breakthrough into the overlying water zone. Optimization through varying the injection pressure and steam trap control leads to the optimal steam chamber operation strategies, which has similar results as Gates et al. [22].

Shin and Polikar tested the sensitivity of the control parameters which include the length of preheating time period, injector to producer spacing, operating pressure, maximum steam injection rates, and reservoir thickness [39]. Yang et al. [41] using experimental design and Designed Exploration and Controlled Evolution (DECE) optimized SAGD process, which is faster and more accurate than traditional history matching. They controlled the steam injected rates and the production rates, their optimized well controls are more energy-efficient and environmentally friendly since with these well controls less natural gas and clean water are used. They used Monte Carlo simulation to quantify the uncertainty of the NPV. They added geological uncertainty to set up robust optimization. They generated more than

100 realizations and rank the realizations by applying a reliable ranking method. After that, they picked 9 realizations to reduce the computational cost. Then, they found the optimal well location and operating strategy and applied optimal well location and operating strategy to 100 realizations. Finally, they performed risk analysis and robustness validation for the optimal strategy [42]. Nurmammadov [35] used the simultaneous perturbation stochastic approximation (SPSA) and EnOpt algorithms to optimize the SAGD process. He presented that cSOR values obtained using the SPSA algorithm lower than EnOpt algorithm. Comparing with NPV, the cSOR is not a good choice as the objection function for the production optimization. When the steamchamber reaches the cap layer, the injection rate decreases at a later stage. The bottom-hole pressure control of the final production well increases at an early stage and remains stable for a period of time and then decreases.

Ghasemi et al. [23] found that co-injection liquid solvent with steam is highly efficient for increasing the oil recovery from the SAGD process. The optimum solvent has a molecular weight in the range of 100-120g/mol. The optimum solvent amount (C6 Volume Fraction) is 0.15 when the solvent amount is the only optimization variable. NPV is affected by both the solvent amount and operating condition. Kaiser et al. [28] presented that the injector designs can be validated to accommodate uncertainties in reservoir characteristics, which can optimize SAGD process. A similar optimization can be made for the producer or a combination of modifications to the two wells, depending on whether the change can be reliably determined before the well is completed and then refined through operational experience. Nascimento [32] showed that dynamic flow simulations can be used to test different operational procedures during the pre-circulation phase. With field data for the pre-circulation phase, dynamic flow simulations can estimate conductivity and the amount the injected steam that migrates to the formation. Dynamic flow simulator can also optimize the steam splitter distribution along with the tubing. For shut-ins, dynamic simulations can verify how the temperature decreases in the well or near the well-bore, evaluating the oil viscosity with time.

1.2 Research Scope

This study focuses on optimizing the Steam Assisted Gravity Drainage (SAGD) process by maximizing Net Present Value (NPV) as the objective function by use of Stochastic Simplex Approximate Gradient (StoSAG) and GPR as a ML-based method. In this research, the commercial thermal reservoir simulator CMG-STARS is used to simulate the SAGD process to generate the training data to be used in ML method, and to compute the sensitivities to bu used in StoSAG method. The control variables are the steam injection rates and the bottom-hole pressures (BHP) at the producer and the temperature of injected steam. We assume that there is only one pair of horizontal wells in this study.

1.3 Thesis Overview

This thesis is composed of four main parts. First, we introduce and discuss the NPV function to be maximized for life-cycle optimization of SAGD process (Chapter 2). Then, optimization methods; StoSAG and ML-based GPR method are introduced in Chapter 3. Following, we provide the methodology used for sampling and training to build ML-based proxy models of NPV (Chapter 4). Last, we provide applications of life-cycle optimization of SAGD process by use of the GPR proxy models in comparison with gradient based StoSAG method (Chapter 5).

CHAPTER 2

NPV FUNCTION FOR SAGD

In this chapter, the net-present value (NPV) function to be maximized by optimizing well control variables is introduced. The well control variables and their constraints used in optimization are given.

2.1 NPV Function

In production optimization problems, net-present value (NPV) is always used as the objective function. By adjusting the well controls during the production life-cycle of the reservoir to maximize NPV, the optimized well controls can be found. In this study, given a history-matched the reservoir model, the NPV for SAGD process is denoted by $J(u)$, defined by

$$J(u) = \sum_{n=1}^{N_t} \frac{\Delta t_n}{(1+b)^{\frac{t_n}{365}}} \left[\sum_{j=1}^{N_P} \left(r_o \times \overline{q}_{o,j}^n - r_w \times \overline{q}_{w,j}^n \right) - \sum_{i=1}^{N_I} \left(\hat{r}_{wi} \times \overline{q}_{wi,i}^n \right) \right] \tag{2.1}$$

where u is the vector of control variables; N_t is the total number of the control steps; Δt_n is the number of days in the nth control step; t^n is the cumulative days up to the nth control step; b is the discount rate; N_P is the total number of the producers; N_I is the total number of the injectors; r_o is the price of the produced oil ($\$/m^3$); r_w is the cost to deal with the produced water ($\$/m^3$); $\overline{q}_{o,j}^n$ and $\overline{q}_{w,j}^n$, respectively, represent the average oil and water production rates (m^3/D) at the jth producer over the nth control step; $\overline{q}_{wi,i}^n$is the average cold water equivalent steam injection rates (m^3/D) at the ith injector over the nth control step, and $\hat{r}_{wi}$ represents to the steam injection cost ($\$/m^3$). However, the steam injection cost can be separated into two parts, one for the water injection cost r_{wi} ($\$/m^3$) and the other one for the steam heat up cost r_{heat} ($\$/(m^3 \cdot {}^\circ\mathrm{F})$). Then $\hat{r}_{wi}$ can be given as the sum

of these tow costs as the steam injection cost

$$\hat{r}_{wi} = r_{wi} + r_{heat} \times T_{steam} \tag{2.2}$$

As mentioned before, in this research, only one pair of horizontal wells are considered for a SAGD process. This means there are only one producer and one injector (see Figure 1.2). In this case, Equation 2.1 can be simplified to

$$J(u) = \sum_{n=1}^{N_t} \frac{\Delta t_n}{(1+b)^{\frac{t_n}{365}}} [\big(r_o \times \overline{q}^n_{o,\mathrm{p}} - r_w \times \overline{q}^n_{w,\mathrm{p}}\big) - (\hat{r}_{wi} \times \overline{q}^n_{wi,\mathrm{i}})] \tag{2.3}$$

2.2 Control Variables

In this research, the number of the control steps is defined as N_t =36, and the length of each control step is considered to be a constant and equal to Δt_n =90 days. u is the vector of well controls which include the steam injection rates, the bottom-hole pressure (BHP) at the producer, and the temperature of the injected steam. The steam injection rates and the BHP can be different in each control step, however, the temperature of the injected steam is constant during the whole life-cycle. In this case, there are 36 steam injection rates, 36 BHPs, and 1 temperature of the injected steam. So the total number of well controls to be optimized by maximizing $J(u)$ given by Equation 2.3 is N_u =73.

2.3 Constraints

We have to consider constraints for well control variables. These constraints should depend on reality. We apply constraints to steam injection rates as

$$0 < q_{wi,\mathrm{i}} \leq 500\ m^3/d, \tag{2.4}$$

In addition, we constraints for BHP at the producer such that

$$600\ psi \leq p_{wf,\mathrm{p}} < 2500\ psi, \tag{2.5}$$

The BHP at the injector should not exceeding 2500 psi, i.e.,

$$p_{wf,\mathrm{i}} < 2500\ psi \tag{2.6}$$

Since the distance between the injector and the producer is usually small, e.g., 5 m, the BHP at the producer is significantly affected by the BHP at the injector. The temperature of the injected steam is constrained such that

$$392^{\circ}F \leq T_{steam} \leq 752^{\circ}F. \tag{2.7}$$

CHAPTER 3

OPTIMIZATION METHODS

In this chapter, we describe the optimization methods considered in this study. As we compare the performance of the ML-based GPR proxy model for maximizing NPV with the gradient based StoSAG method, we first start introducing the StoSAG method, and then the ML-based GPR proxy method.

3.1 Gradient Method

The objective of this study is to maximize NPV by optimizing well control variables discussed in Chapter 2. To maximize NPV, generally, people should find a starting point, then determine a search direction to reach a maximum point although the maximum point has a high probability to be a local maximum point. The global maximum point is generally hard to reach.

In this study, the steepest ascent algorithm is used for optimization. The steepest ascent algorithm is given by

$$u_{l+1} = u_l + a_l d_l, \tag{3.1}$$

for l = 1, 2, ..., until convergence where l is the iteration index. u_l is the control vector which is the well controls at lth iteration, a_l is the step size obtained via a line search in the search direction d_l at iteration l [17].

3.1.1 StoSAG

Currently, systematic descriptions of the basic Ensemble-Based Optimization (EnOpt) method is given by Chen (2008) [10] and Chen et al. (2009) [11].

In EnOpt, u is considered to be a random vector with a Gaussian distribution with

the mean u_l and covariance matrix C_{U_l}, which is at iteration l, $u \sim \mathcal{N}(u_l, C_{U_l})$. Throughout, it is considered that the covariance matrix is fixed, which means that $u \sim \mathcal{N}(u_l, C_U)$ at iteration l. N_p is the number of independent control perturbations and $\hat{u}_{l,j}$ is the generated N_p independent control perturbations (random samples) from the distribution $\mathcal{N}(u_l, C_U)$ at iteration l. The mean of the sample is denoted by $\overline{\hat{u}_l}$ and is given by:

$$\overline{\hat{u}_l} = \frac{1}{N_p} \sum_{j=1}^{N_p} \hat{u}_{l,j}, \tag{3.2}$$

and at each iteration l,$\overline{J(u_l)}$ is defined by:

$$\overline{J(u_l)} \equiv \frac{1}{N_p} \sum_{j=1}^{N_p} J(\hat{u}_{l,j}). \tag{3.3}$$

With the basic EnOpt method, the search direction is computed as:

$$d_{l,EnOpt} = \frac{1}{N_p - 1} \sum_{j=1}^{N_p} (\hat{u}_{l,j} - \overline{\hat{u}_l})(J(\hat{u}_{l,j}) - \overline{J(u_l)}). \tag{3.4}$$

To calculate the search direction by using the EnOpt method, there are two assumptions need to be followed [17]. The first one is:

$$\overline{\hat{u}_l} = \frac{1}{N_p} \sum_{j=1}^{N_p} \hat{u}_{l,j} \approx u_l, \tag{3.5}$$

and the second assumption is:

$$\overline{J(u_l)} = \frac{1}{N_p} \sum_{j=1}^{N_p} J(\hat{u}_{l,j}) \approx J(\overline{\hat{u}_l}) \approx J(u_l). \tag{3.6}$$

However, these two approximations are potentially unreliable under some circumstances. Hence another method to calculate the search direction which does not require to use these approximations ought to be found [9].

A modified formulation for EnOpt for deterministic optimization was proposed by

Do and Reynolds (2013) [12, 16]. They modified the control perturbations and the resulting objective function anomalies. The resulting objective function anomalies are generated relative to their current control vector distribution means. The means are given by the control vector and its corresponding objective function is used for the most recent optimization iteration. Another modification is proposed by Fonseca et al. (2014, 2015) [19, 18], with that modification, Fonseca et al. named the modified EnOpt method Stochastic Simplex Approximate Gradient (StoSAG). The search direction is given by:

$$d_{l,StoSAG} = \frac{1}{N_p} \sum_{j=1}^{N_p} (\hat{u}_{l,j} - u_l)[J(\hat{u}_{l,j}) - J(u_l)]. \tag{3.7}$$

where the $\hat{u}_{l,j}$ and u_l are vectors, however, $J(\hat{u}_{l,j})$ and $J(u_l)$ are scalars in Equation 3.7. $\hat{u}_{l,j}$ can be written as:

$$\hat{u}_{l,j} = u_l + \Delta u_{l,j}, \quad j = 1, 2, ..., N_p \tag{3.8}$$

where

$$\Delta u_{l,j} = \epsilon_j (u_j^u - u_j^l) \tag{3.9}$$

In Equation 3.9, ϵ_j is the perturbation size, the u_j^u and u_j^l are the upper bound and the lower bound of the control variables, respectively. It is worth nothing that a high-fidelity model (like CMG-STARS) directly used to perform StoSAG calculations.

3.2 Machine Learning Method

As stated in MathWorks®, "Machine Learning is a data analytics technique that teaches computers to do what comes naturally to humans and animals; learn from experience. Machine learning algorithms are computational methods to learn information directly from data without relaying on a pre-determined equation on a model. The algorithms adaptive improve their performance as the number of samples available for learning increases."

Machine learning usually refers to changes in the system that perform tasks related to artificial intelligence (AI). These tasks involve identification, diagnosis, planning, robot

control, prediction, and more. These changes may be new systems used to enhance existing systems or to synthesize systems from scratch. Machine learning can sense and simulate the actual environment and calculate the appropriate actions by predicting the impact of task [1].

There are two major types of learning. One is called supervised learning. In supervised learning, we get a data set and already know what our correct output should look like and believe that there is a relationship between input and output. Another one is called unsupervised learning, unsupervised learning allows us to approach problems with little or no idea about what the results should look like. The only thing that can do is to derive structure from data, and it is not necessarily knowing the effect of the variables.

Supervised learning problems can be divided into "classification" and "regression" problems. In the classification problem, supervised learning is going to predict the result in discrete output. In other words, it can map input variables to discrete categories. However, in the regression problem, supervised learning is going to predict the result in continuous output, which means that it can map input variables to continuous functions. In this thesis, we consider supervised learning based on regression problem.

3.2.1 GPR Method

Gaussian Processes Regression (GPR) belongs to supervised learning. GPR is a kernel-based fully Bayesian regression algorithm. Consider a simple 1-D regression problem, mapping from an input x to an output $f(x)$. In Figure 3.1, it shows that a number of sample functions drawn at random from the prior distribution over functions specified by a particular Gaussian Process which favors smooth functions. Before viewing any data, a prior belief was used to represent the various functions which are expected to observe. Assume that the average of the sample functions at each x is zero. Although the specific random functions are drawn in Figure 3.1(a) do not have a mean of zero, if we can have enough functions, the mean of $f(x)$ values for any fixed x would become zero. At any value of x, we can characterize the variability of the sample function by calculating the variance of the

points. The shaded region denotes twice the point-wise standard deviation which means that around 95% of results should lie in this region.

Suppose that the given data set is $D = (x_1, y_1), (x_2, y_2)$, by using this data set to figure out the observation function. The dashed lines in Figure 3.1(b) show that sample functions which are consistent with data set D and the solid line represents the mean value of these functions. In Figure 3.1(b), the uncertainty is reduced close to the observed data sets. After adding more samples as data sets, the mean function can be adjusted to pass through these data points and the uncertainty would also be reduced.

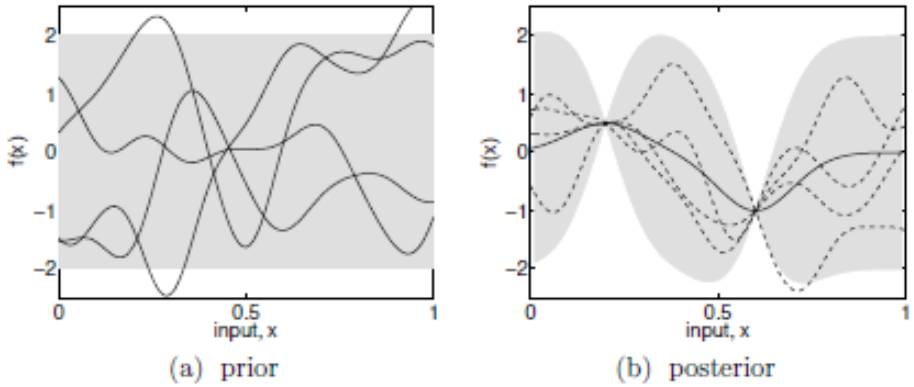

Figure 3.1: Panel (a) shows four samples drawn from the prior distribution. Panel (b) shows the situation after two data point have been observed. The mean prediction is shown as the solid line and four samples from the posterior are shown as dashed lines. In both plots the shaded region denotes twice the standard deviation at each input value x [38].

As for the Gaussian Process regression (GPR) model, one can think of a Gaussian process as a distribution of defined functions and reasoning occurs directly in the function space. Another way to describe a Gaussian process (GP) is a collection of random variables; any finite number of which have a joint Gaussian distribution. A Gaussian process is completely specified by its mean function and covariance function. It is defined that the mean function $\mu(x)$ and the covariance function $\Sigma(x, x')$ of a real process $f(x)$.

$$\mu(x) = \mathbb{E}[f(x)], \tag{3.10}$$

$$\Sigma(x, x') = \mathbb{E}[(f(x) - \mu(x))(f(x') - \mu(x'))], \tag{3.11}$$

and the Gaussian process (GP) is given by:

$$f(x) \sim \mathcal{N}(\mu(x), \Sigma(x, x')). \tag{3.12}$$

For notational simplicity, the mean function $\mu(x)$ is commonly considered to be zero, but not necessary. However, it is not a big limitation as the average of the posterior process is not limited to zero.

The specification of the covariance function implies a distribution over functions. Assume a number of input points, which play the role of test sets X_*. By using the test sets, we can generate a random Gaussian vector with the covariance matrix from

$$\mathbf{f}_* \sim \mathcal{N}(0, \Sigma(X_*, X_*)). \tag{3.13}$$

Initially, the simple special case is $(x_i, f_i) \mid i = 1, ..., n$, the joint distribution of the training outputs is considered to be $\mathbf{f}$, and the test outputs $\mathbf{f}_*$ according to the prior is:

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \Sigma(X, X) & \Sigma(X, X_*) \\ \Sigma(X_*, X) & \Sigma(X_*, X_*) \end{bmatrix}\right). \tag{3.14}$$

Considering the noisy, the observations is given by:

$$y = f(x) + \varepsilon. \tag{3.15}$$

Assuming additive independent identically distributed Gaussian noise ε with variance σ^2, the prior on the observations becomes:

$$cov(\mathbf{y}) = \Sigma(X, X) + \sigma^2 I. \tag{3.16}$$

Introducing the noisy observations part (Equation 3.15), the joint distribution of observation

is given by:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \Sigma(X,X)+\sigma^2 I & \Sigma(X,X_*) \\ \Sigma(X_*,X) & \Sigma(X_*,X_*) \end{bmatrix}\right). \tag{3.17}$$

Based on Equation 3.17, the mean function and the covariance function can be written as:

$$\mathbf{f}_* \mid X, \mathbf{y}, X_* \sim \mathcal{N}(\bar{\mathbf{f}}_*, cov(\mathbf{f}_*)), \tag{3.18}$$

where

$$\bar{\mathbf{f}}_* = \mathbb{E}[\mathbf{f}_* \mid X, \mathbf{y}, X_*] = \Sigma(X_*, X)[\Sigma(X,X)+\sigma^2 I]^{-1}\mathbf{y}, \tag{3.19}$$

and

$$cov(\mathbf{f}_*) = \Sigma(X_*,X_*) - \Sigma(X_*,X)[\Sigma(X,X)+\sigma^2 I]^{-1}\Sigma(X,X_*). \tag{3.20}$$

In covariance matrix, $\Sigma(X,X)$, each element in the matrix is called covariance function in the context of Gaussian process. The covariance function is also treated as "Kernel" function. Kernels are the key component of the Gaussian process, which determines the shape of prior and posterior of the Gaussian Process. Kernels are behaving like the inner product in some space. They code the hypothesis of the learning function by defining the "similarity" of the two data points and combining the assumption that similar data points should have similar target values. For example, kernel shows the relationship between input vectors $x^{(i)}$ and $x^{(j)}$. Kernel $k_{ij} = 1$ means that vector $x^{(i)}$ and vector $x^{(j)}$ are at the same point and kernel $k_{ij} = 0$ means that vector $x^{(i)}$ and vector $x^{(j)}$ are far from each other.

The kernel function used in this research is "Radial Quadratic kernel (RQK)." The Rational Quadratic kernel can be seen as a scale mixture (an infinite sum) of Radial Basis Function (RBF) kernels, which is a stationary kernel and is also known as the "squared exponential" kernel, with different characteristic length-scales. RQK is parameterized by a length-scale parameter $l > 0$ and a scale mixture parameter $\alpha > 0$. The Radial Quadratic kernel is given by:

$$k(x^{(i)}, x^{(j)}) = \left(1 + \frac{d(x^{(i)}, x^{(j)})^2}{2\alpha l^2}\right)^{-\alpha}, \tag{3.21}$$

where $d(x^{(i)}, x^{(j)})$ is the distance between vector $x^{(i)}$ and vector $x^{(j)}$, which is calculated by Euclidean distance. The α and l are also called hyper-parameters which can significantly affect the regression result.

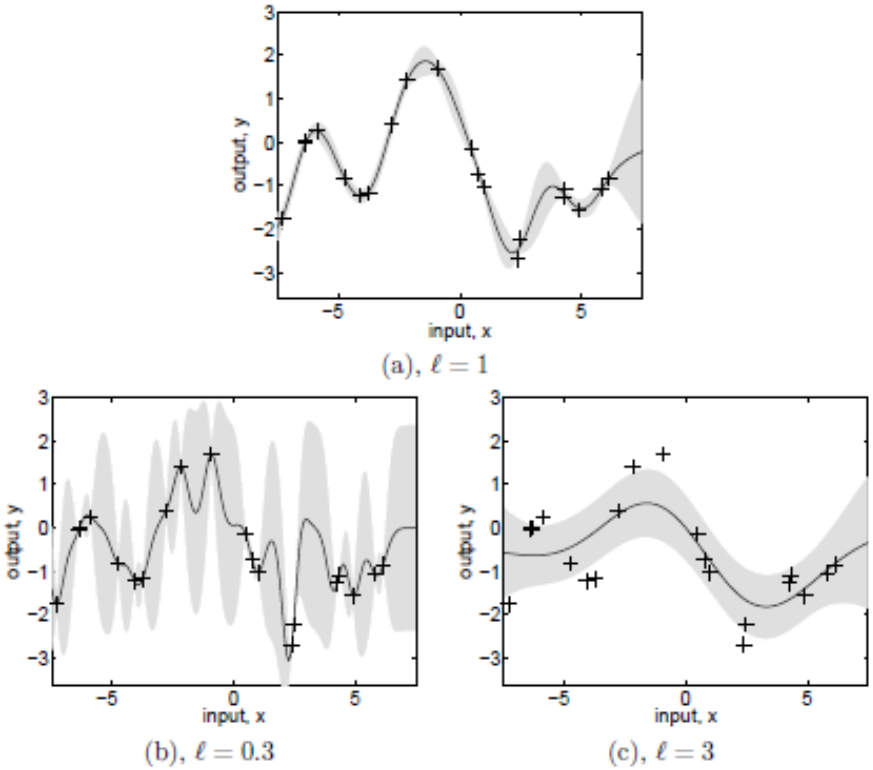

Figure 3.2: (a) Data is generated from a Gaussian Process with hyper-parameter $l = 1$, as shown by the + symbols. Using Gaussian Process prediction with these hyper-parameter to obtain a 95% confidence region for the underlying function f. Panels (b) and (c) show the hyper-parameters $l = 0.3$ and $l = 3$ respectively [38].

CHAPTER 4

METHODOLOGY FOR SAMPLING AND TRAINING

4.1 Sampling Method-Latin Hypercube Sampling (LHS)

At the beginning of the nonlinear production optimization, the first thing needs to be done is to find out the initial guess point. In this research, there are totally 73 well control variables. This means that an initial guess point is a 73 dimensional vector. We need to generate a number of its random samples for training and choosing initial guesses for optimization. The question is how to "evenly" generate these samples. In this study, we use LHS method, which may have better performance than Monte Carlo method [34].

Let S represents to the sample space, input variables $\mathbf{X} = (X_1, ..., X_K)$ and the output variable $Y = f(\mathbf{X})$. By using the computer program, for any selection of input variables $\mathbf{X}$ an output variable Y is produced by the computer program. Moreover, the computer programs are sometimes sufficiently complex so that a single set of input variables may require a quite long time even on the fastest computers in order to produce one output. A single output Y is usually a graph $Y(t)$ of output as a function of time.

From different methods of selecting the values of input variables, Latin Hypercube Sampling is used in this research. Latin Square (LHS) is an $n \times n$ array filled with n different symbols, each occurring exactly once in each row and exactly once in each column. Latin hypercube is the generalization of this concept to an arbitrary number of dimensions, whereby each sample is the only one in each axis-aligned hyperplane containing it, where hyperplane means a subspace whose dimension is one less than that of its ambient space.

Latin Hypercube Sampling (LHS) is a method of sampling random numbers that attempt to distribute samples evenly over the sample space. Ensuring each of the input variables X_k has all portions of its distribution represented by input values. Dividing the

range of each X_k into N layers of equal marginal probability $1/N$, and sample once from each layer. Let the sample be X_{ki}, $i = 1, ..., N$ and $k = 1, ..., K$, where K is the number of the dimension of the input variables and N is the number of the samples we want to sample [30].

For example, the input variable is two dimensions, and we are going to generate four samples, the LHS will generate several different samples as Figure 4.1 and Figure 4.2.

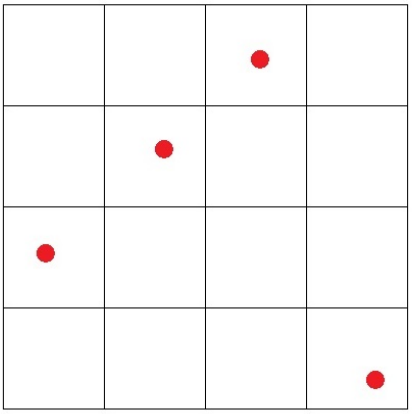

Figure 4.1: Sample 1 of LHS.

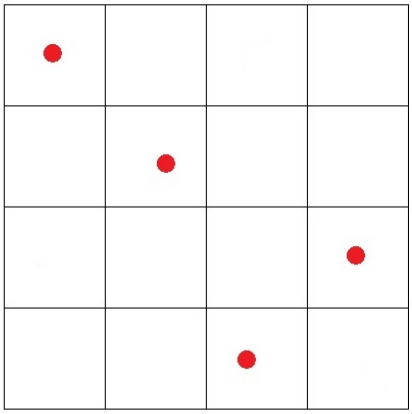

Figure 4.2: Sample 2 of LHS.

One of the advantages of LHS is that it does not require more samples for more control variables, people can get any number of samples with any number of control variables. Another advantage is that LHS is a sampling method which has memory. When doing LHS, each selected sample's location needs to be memorized, in this case, the samples can be taken

one by one.

It is necessary to note that we have used the codes publicly available in "Python" when implementing LHS.

4.2 Training Proxy Model Method

To find out the accurate proxy model that can replace computer simulator (CMG), there are two main steps: First it is to build up a proxy model by using machine-learning method; secondly it is to retrain this proxy model to improve the accuracy of results produced by this model.

4.2.1 Build Proxy Model

- First step to build is using LHS method to sample 32 different groups of initial guesses.
- Second step is using these initial guesses to write the well control that can be used by computer simulator (CMG).
- Thirdly, using the well control to run computer simulator.
- Then output useful results from CMG such as water production rate at the producer, oil production rate at the producer, water injection rate at the injector, BHP at the producer.
- Next step is using these results and Equation 2.3 to calculate the NPV.
- Finally, let the well controls be the input variables $\mathbf{X}$ and let the calculated NPV be the output variables Y. Combine these input variables and the output variables together, which can be considered as the known points. And we can use these points to build the proxy model.

4.2.2 Retrain Proxy Model

The procedures of how to retrain the proxy model are given in Figure 4.3.

- Firstly, using LHS to sample 5 initial well control vectors as 5 input variables.
- Second, find the maximum NPV by using finite difference method and this proxy model. With this maximum NPV, the optimal control variables can be found.
- Next, using the optimal well controls to run production simulator (CMG) and get the corresponding responses from CMG, which includes water and oil production rates, water injection rates and the BHP.
- Then calculate the real NPV by using the responses.
- Finally, comparing the real NPV with the NPV given by proxy model, which is given by:

$$\Delta = \frac{|NPV_{Proxy} - NPV_{CMG}|}{NPV_{CMG}} \tag{4.1}$$

If $\Delta < 0.03$, shows that the result of the proxy model is accurate, this proxy model is able to replace the production simulator. If not, the proxy model needs to be improved. To improve the proxy model, we retrain the proxy model from the very beginning step.

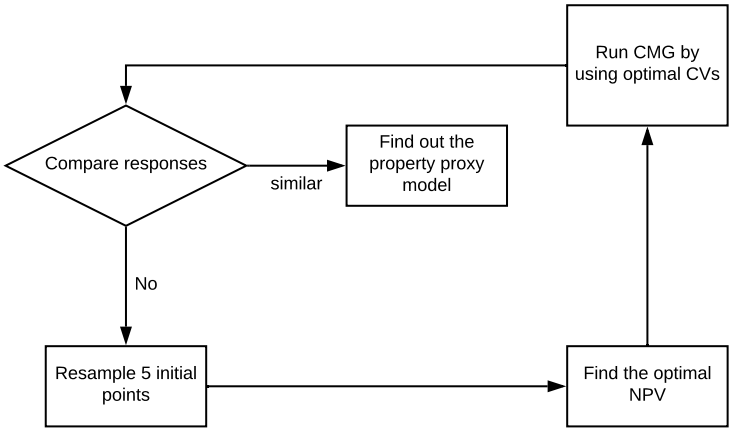

Figure 4.3: The flowchart of the proxy model optimization.

It is necessary to note that we have used the codes publicly available in "Python" when implementing GPR.

CHAPTER 5

APPLICATIONS

5.1 Model Description

The model used in this research to simulate a model of SAGD process includes only one pair of horizontal wells; one injector and one producer. The reservoir model is built by CMG-STARS, and production optimization has also been tested by CMG-STARS. The reservoir model simulates a quite small reservoir. The length of the reservoir is 900 m, the width of the reservoir is 80 m and the height of the reservoir is 38 m. To save the running simulation time cost during production optimization, a cross-section which is perpendicular to the pair of horizontal wells is simulated. That is, there is only one gridblock to be considered along the length direction of the reservoir model. The size of the cross-section is 51 $\times$ 38.That is, there are 51 gridblocks along the width direction of the model and on the vertical direction are 38 gridblocks. Totally, there are 1938 gridblocks in this reservoir model. The size of each grid-block is 900 m$\times$1.6 m$\times$1.0 m.

The permeability (k) of the model is assumed as random distribution, the lowest permeability in the reservoir is 0.0213 Darcy, and the highest permeability in the model is 67.2115 Darcy. In real applications, the reservoir model to be used for optimization would be a history matched model having a permeability distribution satisfying historical production data. So, we can consider our reservoir model here as a history matched model. The case of the permeability distribution been used in this research is given in Figure 5.1 and Figure 5.2. The reservoir model is heterogeneous and anisotropic with respect to k. The permeability in i and j direction are the same, but the permeability in k direction is equal to 40% of the permeability in i direction(i direction represents to the length, j direction represents to the width, and k direction represents to the height of the reservoir).

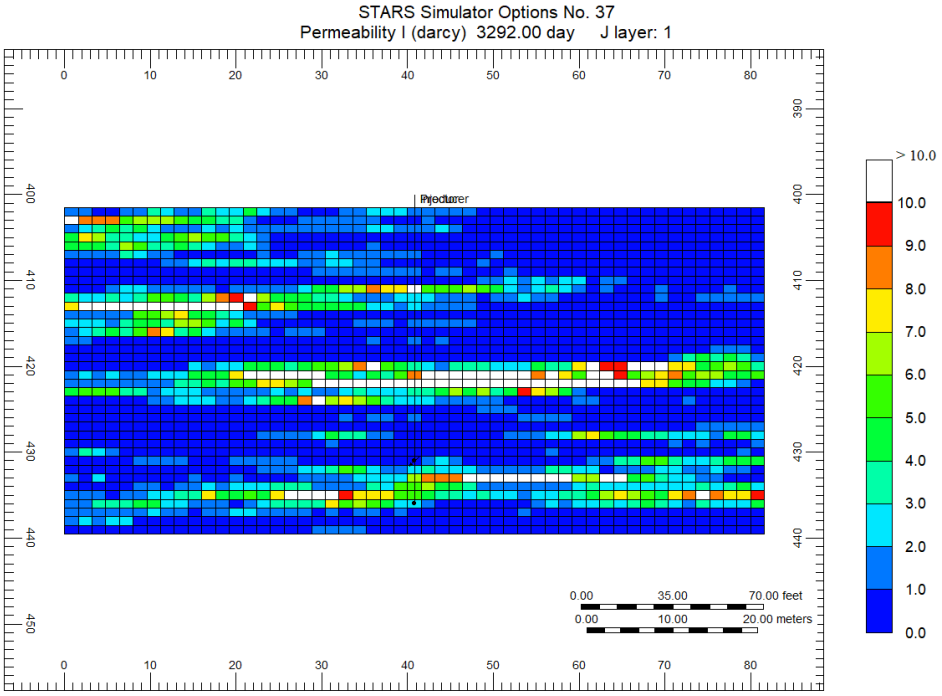

Figure 5.1: The permeability distribution in horizontal direction in this reservoir model.

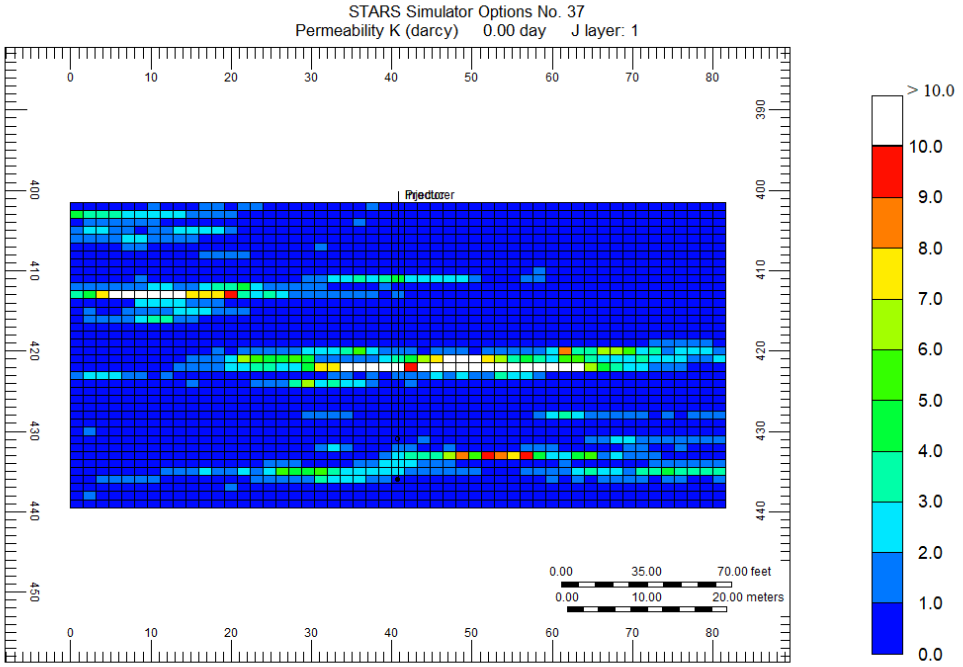

Figure 5.2: The permeability distribution in vertical direction in this reservoir model.

The top depth of the reservoir is 402 m (1318 ft), which may be considered to be

realistic for the tar sands reservoir in Canada [40]. In this study, the initial reservoir pressure is assumed as the pressure at the top of the reservoir 1300 psi and the initial reservoir temperature is assumed to be a constant as 104 °F. The injector is drilled in the center of 26 × 30 grid-block and the producer is drilled in the center of 26 × 35 grid- block, which is 3.5 m from the bottom of the reservoir. The injector is 5 m above the producer. Both wells' diameter is 8.7 inch, and both of them are fully perforated.

One of the most important ways to increase the oil recovery in heavy-oil and bitumen is to decrease the oil viscosity. In this study, the effect of pressure on oil viscosity is ignored, the oil viscosity is assumed tobe as the function of temperature. As shown in Figure 5.3, the initial oil viscosity in the reservoir is higher than 1000 cp, which is very difficult to produce. In this case, we need to inject high-temperature steam to increase the temperature in the reservoir. Figure 5.3 shows that when the temperature increases to around 392 °F (200 °C), the oil viscosity will be decreased to around 6 cp. In CMG-STARS, we will use the key word "VISCTABLE" to input the viscosity data, shown in Figure 5.3.

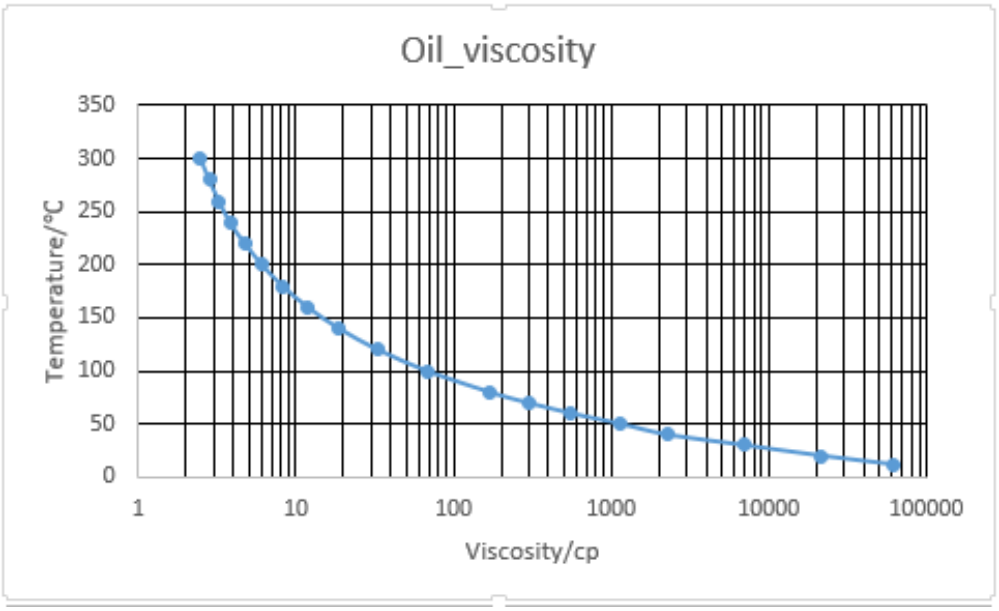

Figure 5.3: Semi-log Plot of Temperature vs Oil Viscosity.

Before any SAGD process, there is a pre-heating time period to be concerned. Actually, there are two ways to conduct a pre-heating process. A cyclic injection of high-

temperature steam into both injector and producer is a very common pre-heating method [36]. On the other hand, lack of steam injectivity is one of the most serious obstacles in the thermal recovery of tar sands and oil shale deposits. Because the huge hydrocarbon reservoirs exist in deposits where the oil viscosity and saturation are so high that sufficient steam injectivity cannot be achieved by current methods. In this case, using electric heating to heat up the reservoir temperature and increase the mobility of the heavy oil is a good choice. However, due to the high cost of the electrical heating, this method cannot be viewed as a replacement for steam injection [24].

In CMG-STARS, using the keyword "UHTR", we can add this preheating process into the model, and by using "TMPSET" we can set the highest temperature for the gridblocks which contain the wells. The pre-heating time period is set to 52 days and the specified preheating temperature is set to 465 °F. However, because of the length of the pre-heating time period and heat loses, the reservoir temperature around the wells can only reach to around 392 °F. If the length of the pre-heating time period was expended, the temperature of the gridblocks could reach to 465 °F.

The original temperature of the injected steam is 465 °F, and the steam quality is 0.9. It is assumed that the steam quality is kept constant through the whole life-cycle production optimization. The initial reservoir oil saturation is assumed to be constant in the whole reservoir, which is 0.7. As mentioned in Chapter 2.2, the total production life is 3240 days, which is divided into 36 control steps. The control step size is 90 days. Since there are two wells in this model, there are 73 well control variables need to be adjusted for production optimization during maximization of NPV function (Equation 2.3). In NPV function the price of the produced oil r_o assume to be 80 $\$/m^3$, the cost to deal with the produced water r_w assume to be 1 $\$/m^3$, the steam injection cost r_{wi} is assumed to be 5 $\$/m^3$, the steam heat up cost r_{heat} is assumed to be 0.01 $\$/(m^3 \cdot °\text{F})$, and the annual discount rate b is 0.05.

5.2 Results of StoSAG

As we compare the results of using the ML-based GPR proxy model with those of

using the gradient-based StoSAG, we first present our results obtained by using StoSAG. The results of StoSAG can be affected by two parameters; one is the size of perturbation, which is the percentage of perturbation in the control variables and another is the smoothness of the control variables as the function of time, which is affected by correlation length. The higher the value of the correlation length, the smoother the control variables as the function of time.

To investigate the effect of these parameters on StoSAG performance for SAGD process, we consider three cases. We define the perturbation sizes of 0.5% and 0.1%. The correlation lengths are set as 18 and 36. As for the number of perturbations N_p is 20.

The first case is perturbation size 0.1% and the correlation length is 18. The initial guess for the control variables is the steam injection rates at the injector is 1258 bbl/day ($\simeq$ 200 m^3/day) for the whole life-cycle, the BHP at the producer is 1200 psi for the whole life-cycle, and the temperature of the injected steam is 392 °F. The initial oil saturation in the reservoir is assumed to be 0.7. The oil saturation distribution in Case 1 after several control steps are shown in Figs. 5.4 to 5.7.

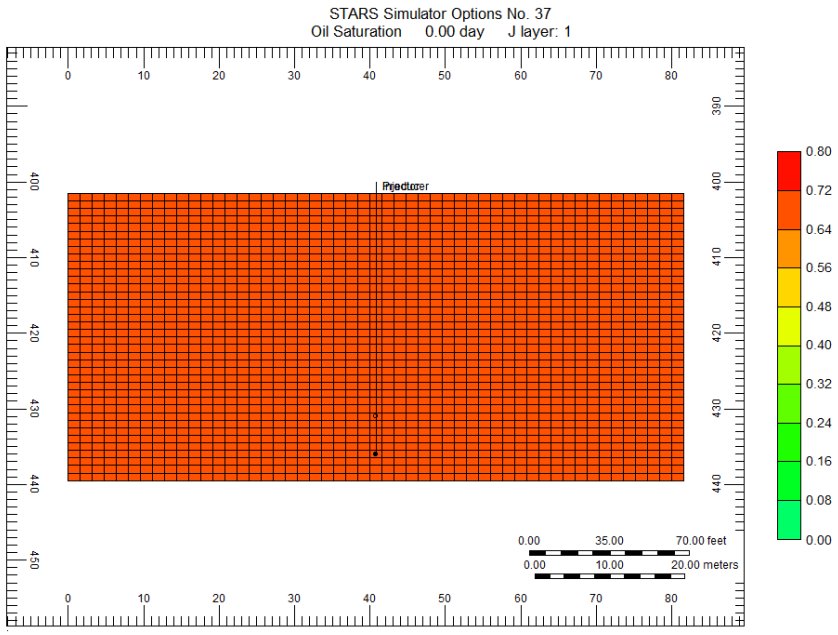

Figure 5.4: Oil Saturation distribution in the reservoir for Case 1.

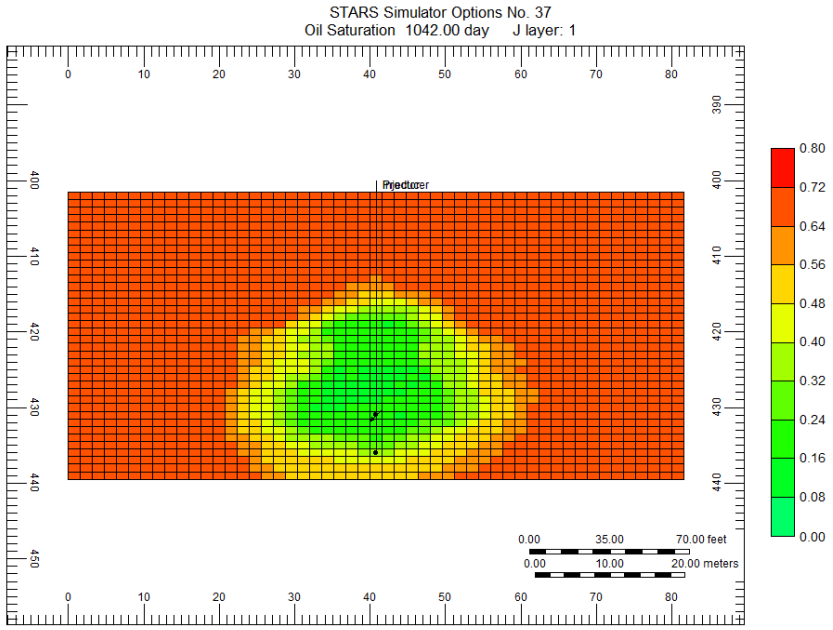

Figure 5.5: Oil Saturation distribution in the reservoir for Case 1 at 1042 days.

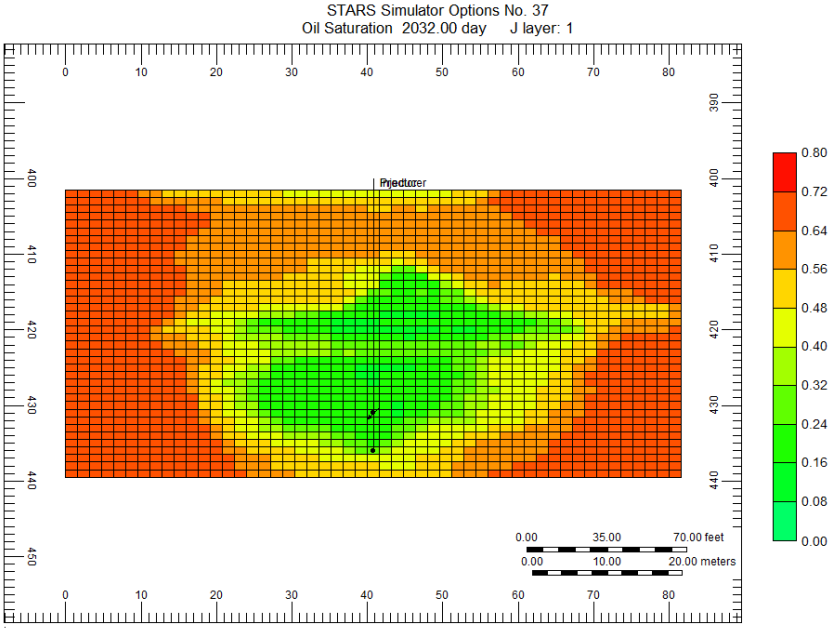

Figure 5.6: Oil Saturation distribution in the reservoir for Case 1 at 2032 days.

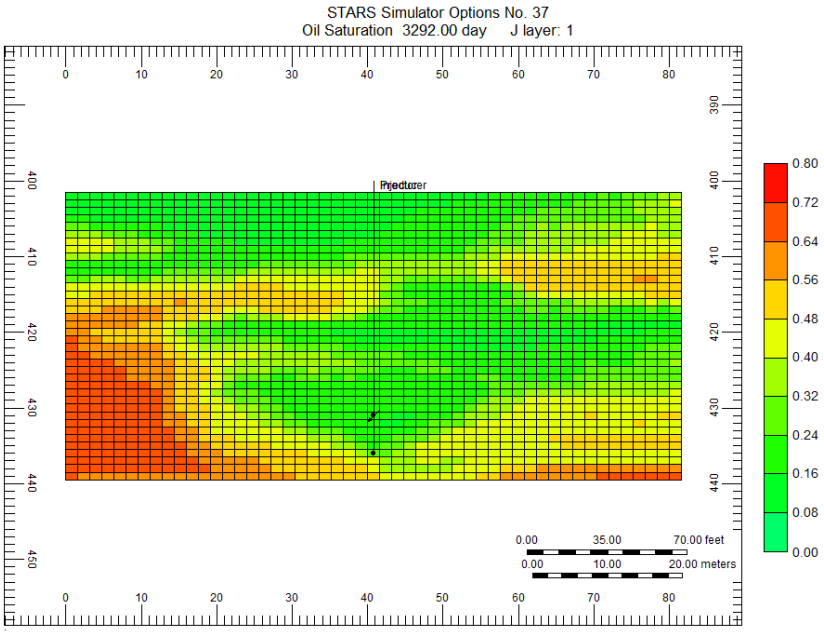

Figure 5.7: Oil Saturation distribution in the reservoir for Case 1 at 3292 days.

Comparisons of the initial values of well controls (BHP at the producer and steam injection rate) with the values of corresponding optimized well controls by using StoSAG method with the perturbation size 0.1% and correlation length 18 are shown in Figure 5.8 and Figure 5.9.

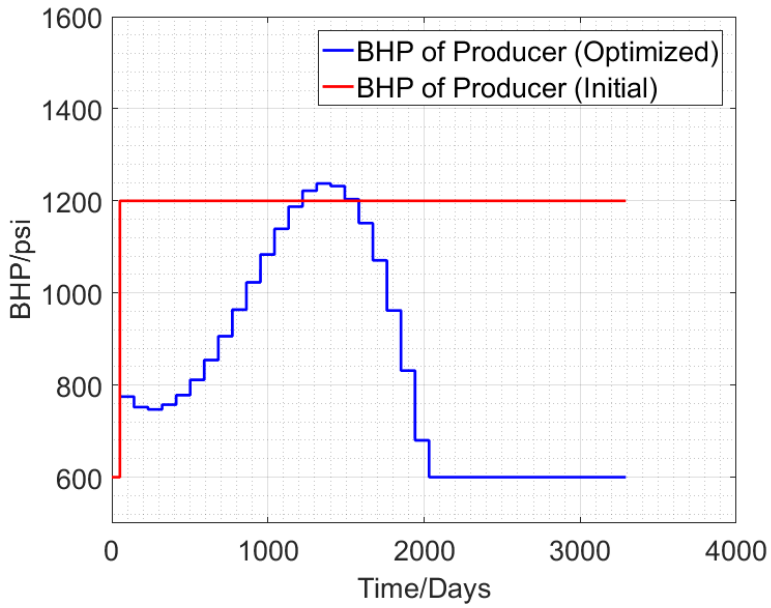

Figure 5.8: A comparison of initial and optimized BHP at the producer for Case 1.

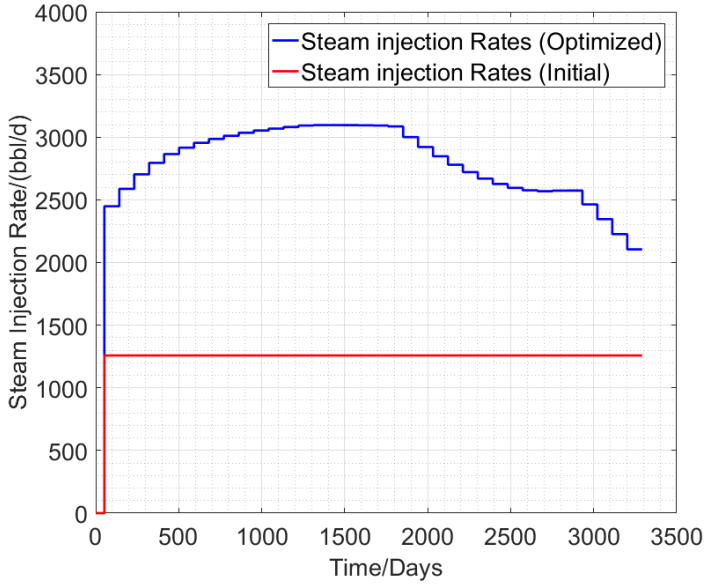

Figure 5.9: A comparison of initial and optimized steam injection rate at the injector for Case 1.

In the beginning, the optimized BHP decreases for around 200 days and it increases till 1500 days. The highest BHP achieved by optimization is 1238 psi. After that, the optimized BHP keeps decreasing to its lower bound of the BHP and is maintained at its lower bound from 2000 days to the end of the life-cycle. This maybe because there is not that much oil left in the chamber that can be produced, which is shown in the oil saturation distribution in Figure 5.6 and Figure 5.7. The steam injection rate gradually increases until it reaches its highest optimized value of 3100 bbl/day from 1400 days to 1600 days. The highest injection rates are around 3100 bbl/day ($\simeq$ 490 m^3/day), which is very close to its upper bound. After 1600 days, the steam injection rate decreases to 2600 bbl/day ($\simeq$ 410 m^3/day) and is maintained at this rate from 2800 days to 2900 days. Finally, the steam injection rate decreases to around 2100bbl/day ($\simeq$ 335 m^3/day) at the end of the life cycle (Figure 5.9). The cumulative oil production is shown in Figure 5.10. It can be seen that the optimization is really necessary to maximize NPV. The cumulative oil production at the initial values of well controls at the end of life-cycle is 0.86 million bbl, whereas the cumulative oil production at the optimized values of well the controls is 2.39 million bbl. The NPV generated by the initial values of well controls is 3.0253 million dollars ($MM),

whereas the NPV generated by the optimized values of well controls is 13.0993 million dollars ($MM). It is shown from Figure 5.11 that the NPV is converged after 20 iteration runs.

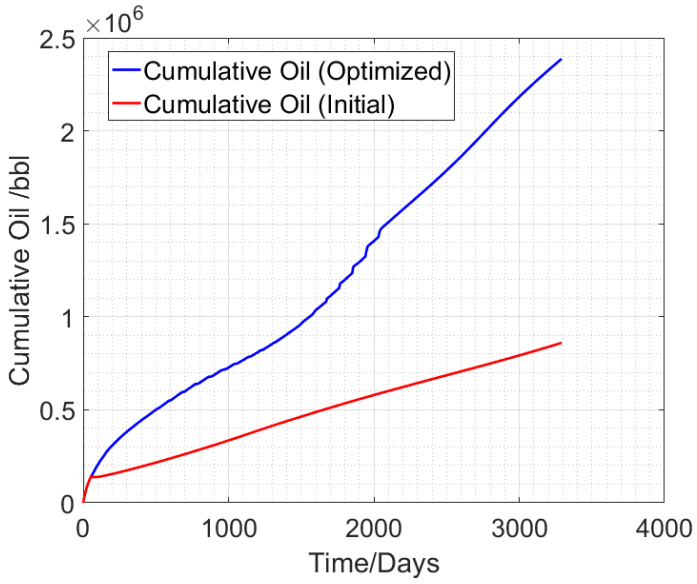

Figure 5.10: A comparison of initial and optimized cumulative oil production from the producer for Case 1.

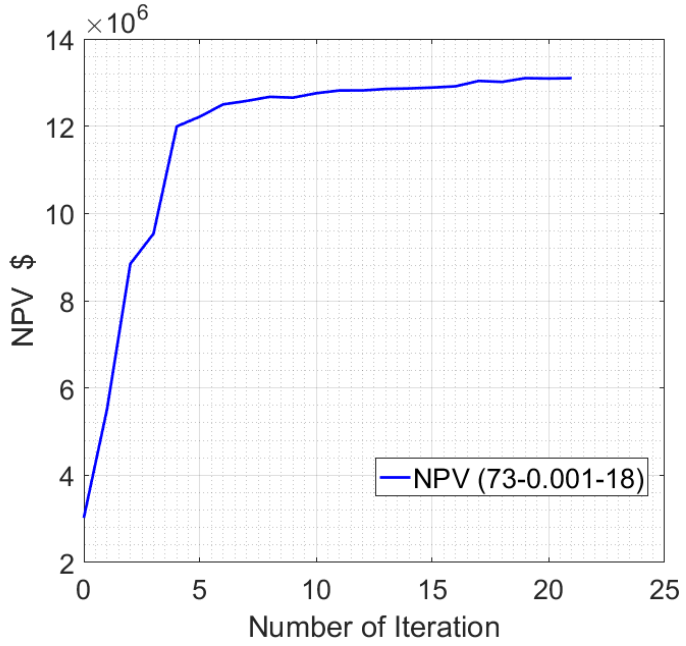

Figure 5.11: NPV for Case 1.

In the second case, the perturbation size is 0.1% and the correlation length is 36. The other conditions are the same as Case 1. The oil saturation distribution after the life-cycle

is given in Figure 5.12.

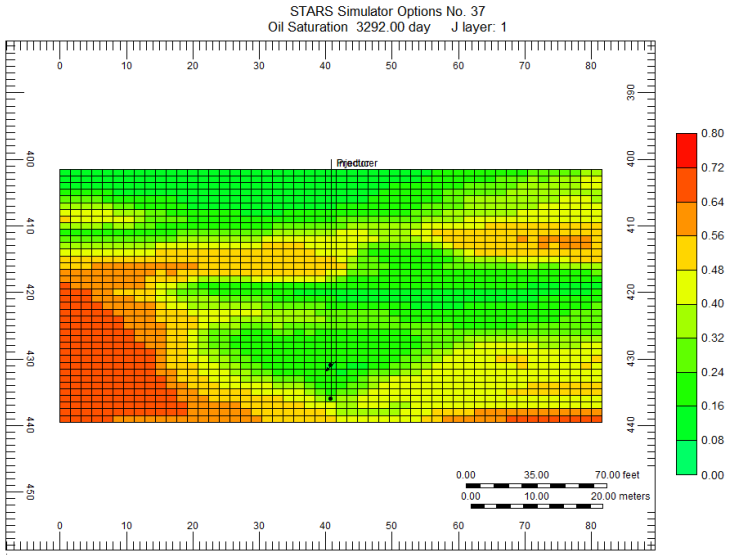

Figure 5.12: Oil Saturation distribution in the reservoir for Case 2 at 3292 days.

Comparisons of the initial values of well controls (BHP at the producer and steam injection rate) with the values of corresponding optimized well controls by using StoSAG method with the perturbation size 0.1% and correlation length 36 are shown in Figure 5.13 and Figure 5.14.

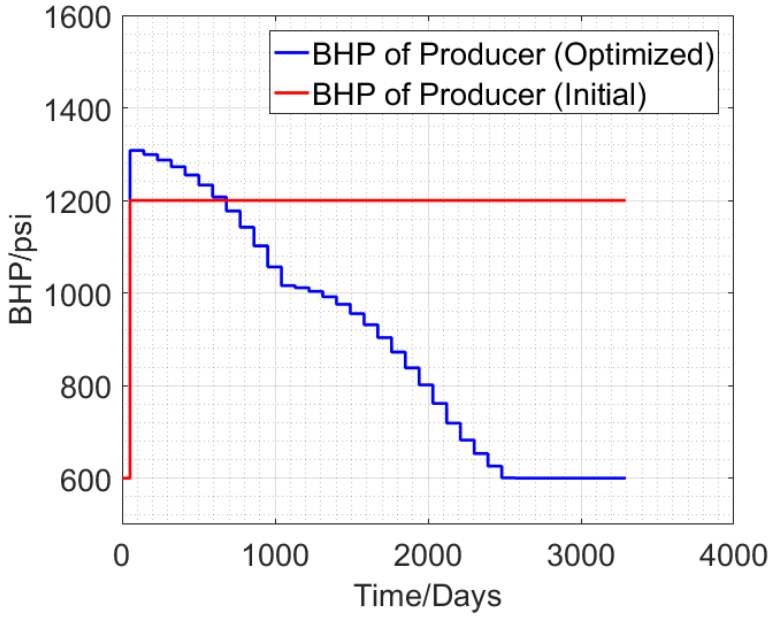

Figure 5.13: A comparison of initial and optimized BHP at the producer for Case 2.

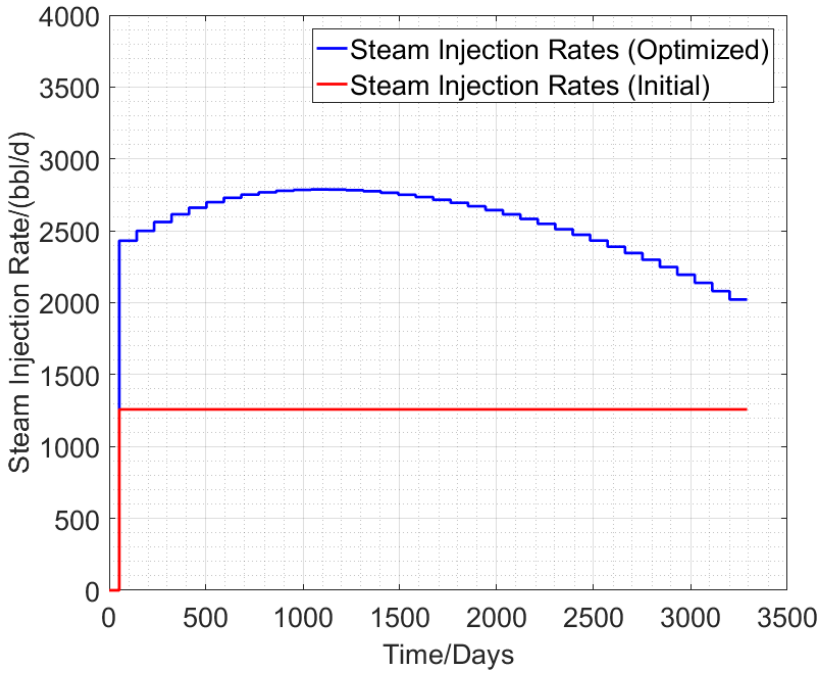

Figure 5.14: A comparison of initial and optimized steam injection rate at the injector for Case 2.

For this case, the BHP at the producer reaches its highest value of BHP 1308 psi immediately. Then, the BHP decreases to around 1004 psi at around 1000 days. The BHP is then decreases to 600 psi at around 2500 days (see Figure 5.13). The steam injection rate gradually increases until it reaches its highest value of 2800 bbl/day ($\simeq$ 443 m^3/day) at around 1000 days. After that, the steam injection rate decreases to 2000 bbl/day ($\simeq$ 320 m^3/day) till the end of the life cycle (see Figure 5.14). The cumulative oil production is shown in Figure 5.15. The cumulative oil production at the initial values of well controls at the end of life-cycle is 0.86 million bbl, whereas the cumulative oil production at the optimized values of well controls is 2.31 million bbl. The NPV generated by the initial well controls is 3.0253 million dollars ($MM), while the NPV generated by the optimized values of well controls is 12.8385 million dollars ($MM). It is shown from Figure 5.15 that the NPV is converged after 15 iteration runs.

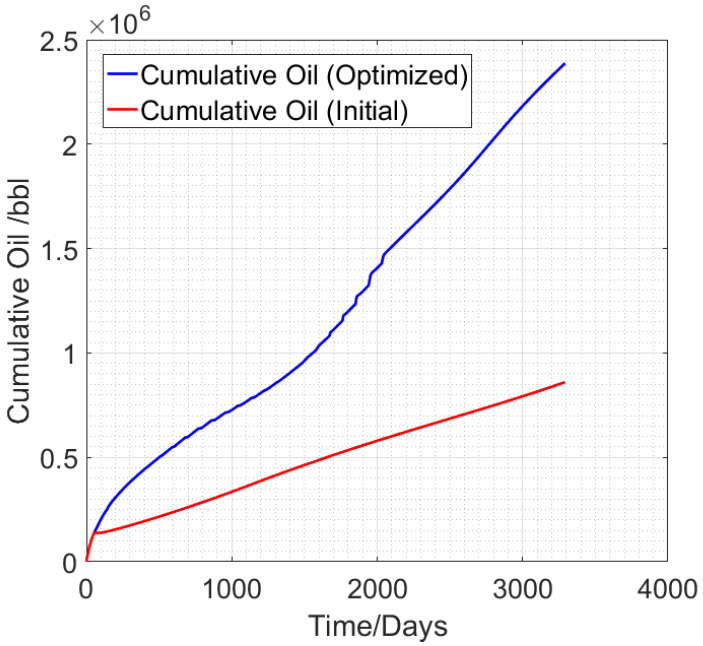

Figure 5.15: A comparison of initial and optimized cumulative oil production from the producer for Case 2.

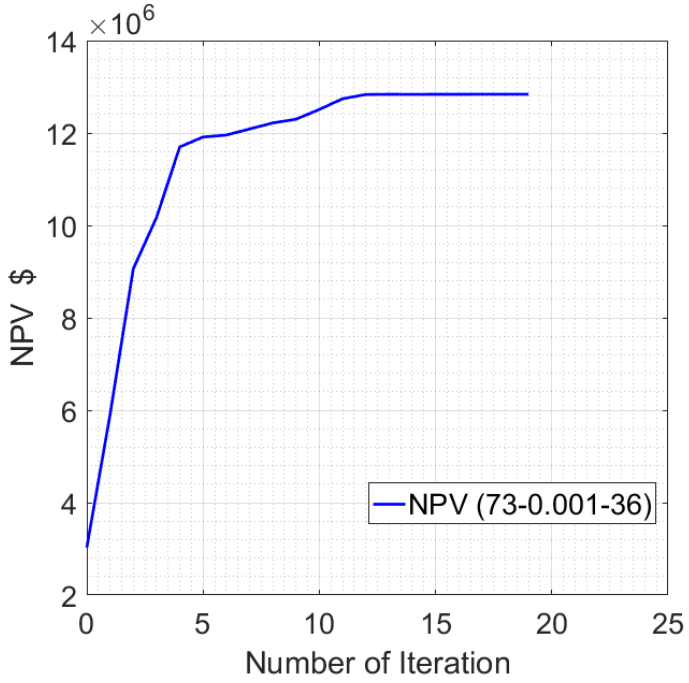

Figure 5.16: NPV for Case 2.

In the last case, the perturbation size is 0.5% and the correlation length is 18. The other conditions are the same as Case 1. The oil saturation distribution after the life-cycle

is given in Figure 5.17.

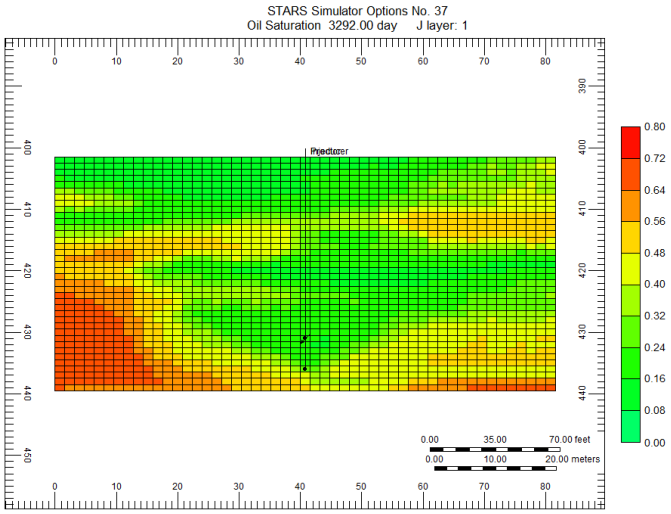

Figure 5.17: Oil Saturation distribution in the reservoir for Case 3 at 3292 days.

Comparisons of the initial values of well controls (BHP at the producer and steam injection rate) with the values of corresponding optimized well controls by using StoSAG method with the perturbation size 0.5% and correlation length 18 are shown in Figure 5.18 and Figure 5.19.

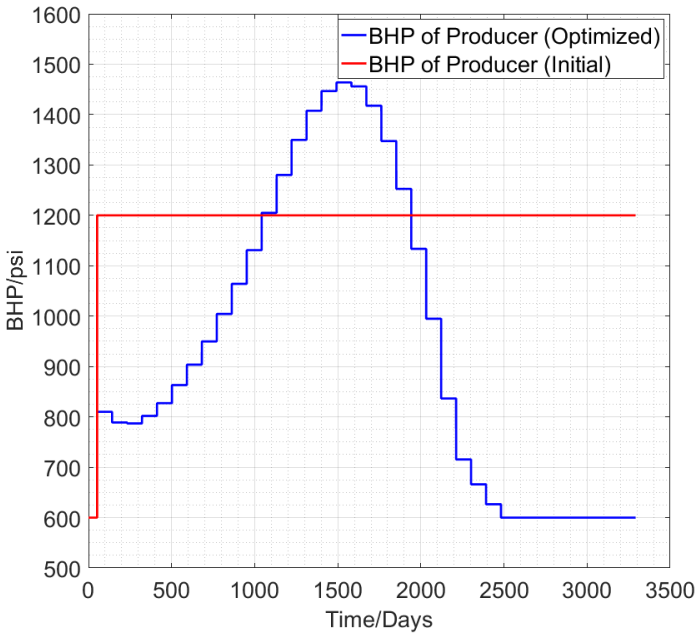

Figure 5.18: A comparison of initial and optimized BHP at the producer for Case 3.

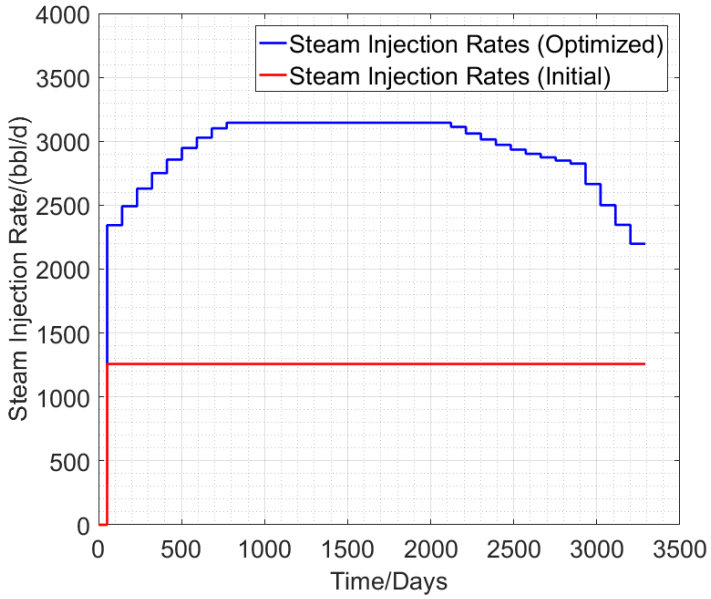

Figure 5.19: A comparison of initial and optimized steam injection rate at the injector for Case 3.

As can be seen from Figure 5.18, in the beginning, the optimized BHP decreases for around 300 days, and then it increases from 300 days to 1500 days by reaching its highest BHP value of 1464 psi. Later, the optimized BHP keeps decreasing to its lower bound of 600 psi at around 2500 days, and then is maintained at this value until the end of this life-cycle. The steam injection rate gradually increases until it reaches its highest optimized value of 3145 bbl/day ($\simeq$ 500 m^3/day) from 800 days to 2100 days. After 2100 days, the steam injection rate decreases to 2820 bbl/day ($\simeq$ 449 m^3/day) till 2900 days. Finally, the steam injection rate decreases to around 2200bbl/day ($\simeq$ 350 m^3/day) at the end of the life cycle (see Figure 5.19). The cumulative oil production is shown in Figure 5.20. The cumulative oil production at the initial values of well controls at the end of life-cycle is 0.86 million bbl, whereas the cumulative oil production at the optimized values of well the controls is 2.41 million bbl. The NPV generated by the initial values of well controls is 3.0253 million dollars ($MM), whereas the NPV generated by the optimized values of well controls is 12.7699 million dollars ($MM). It is shown from Figure 5.21 that the NPV is converged after 10 iteration runs.

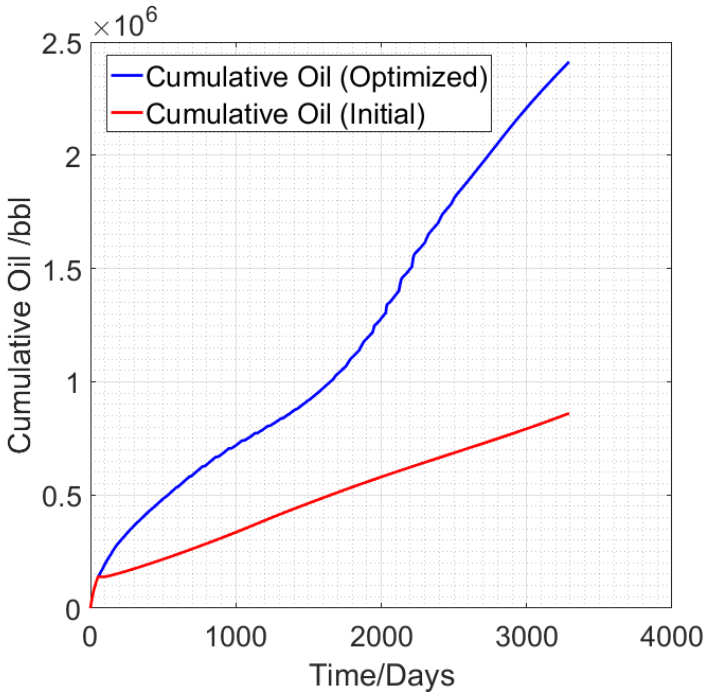

Figure 5.20: A comparison of initial and optimized cumulative oil production from the producer for Case 3.

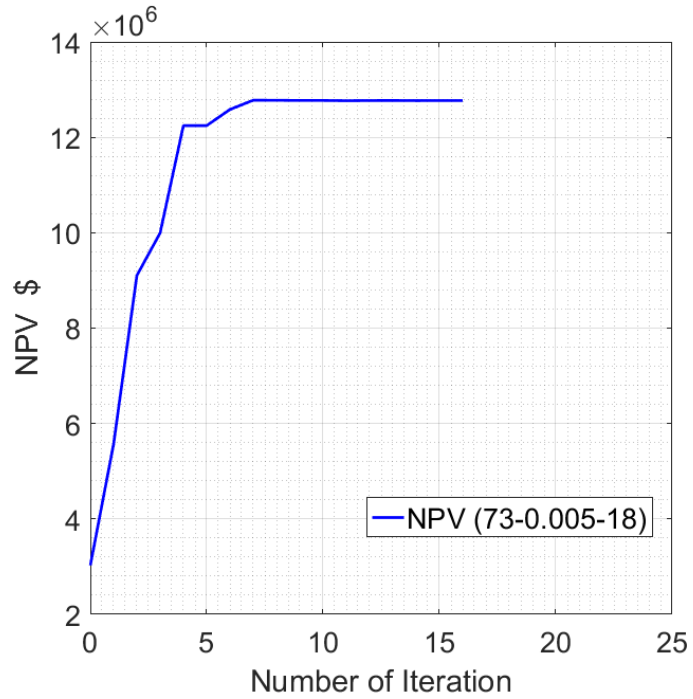

Figure 5.21: NPV for Case 3.

Comparing the results of Case 1, Case 2 and Case 3, we find that the optimized well

control variables in the three cases are quite different, depending on the chosen perturbation size and correlation length in the StoSAG method.

In Figure 5.22 to Figure 5.27, the red curve represents the well controls and well responses generated by the perturbation size of 0.1% and correlation length of 18 (Case 1), the blue curve represents the results for the perturbation size of 0.1% and correlation length of 36 (Case 2), and the green curve represents the results for the perturbation size of 0.5% and correlation length of 18 (Case 3). It is clear from the results presented in Figure 5.22 to Figure 5.27 that the different perturbation sizes and different correlation lengths yield quite different values of optimized well controls. However, inspection of the results of Figure 5.26 and Figure 5.27, we can say that although the oil and water production rates are quite different depending on the cases, the cumulative oil production for all three cases are similar and there is not a significant difference in NPV for three cases.

Based on these three cases, it can be concluded that the perturbation sizes and the correlation lengths do not have significant effect on NPV. In this study, we found that the Case 1 yields the maximum NPV value, though it is only slightly higher than those for Case 2 and Case 3. The perturbation size set as 0.1% and the correlation length set as 18 in this study.

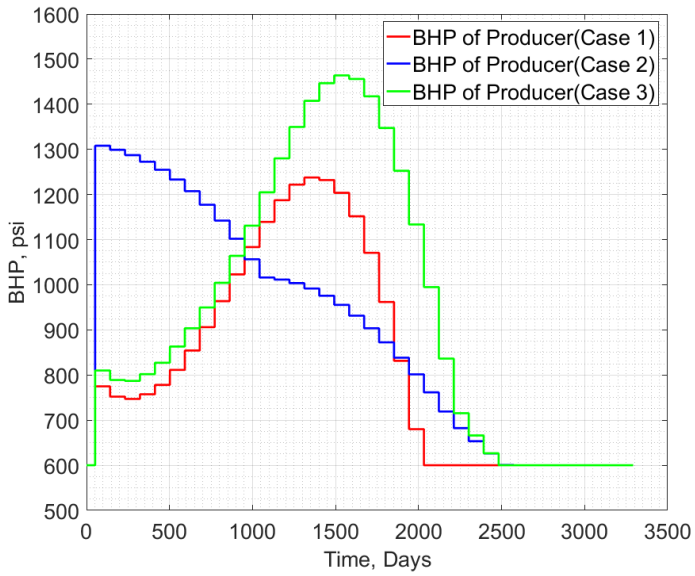

Figure 5.22: A comparison of the optimized BHP for three different cases.

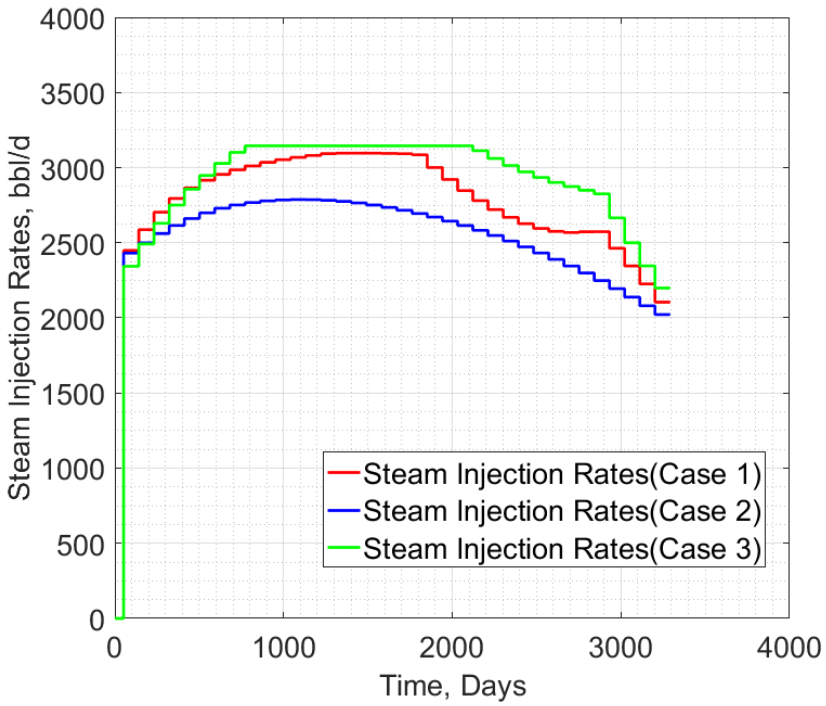

Figure 5.23: A comparison of the optimized steam injection rate for three different cases.

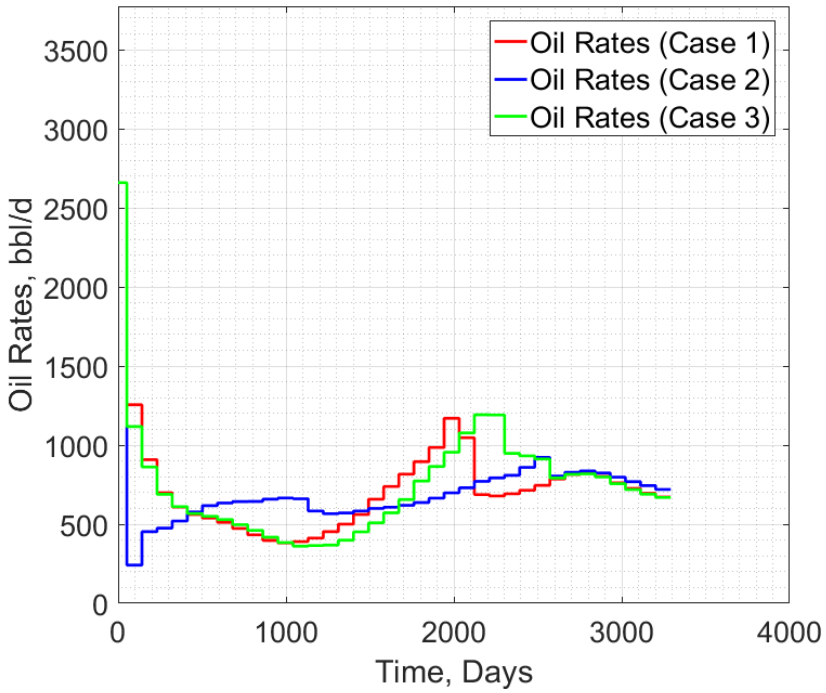

Figure 5.24: A comparison of the oil rate responses for three different cases.

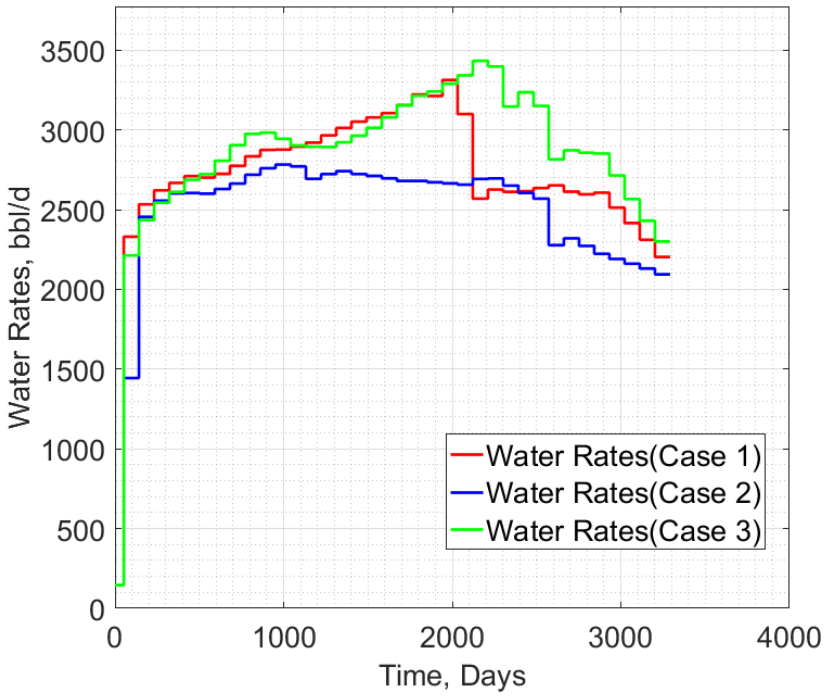

Figure 5.25: A comparison of the water rate responses for three different cases.

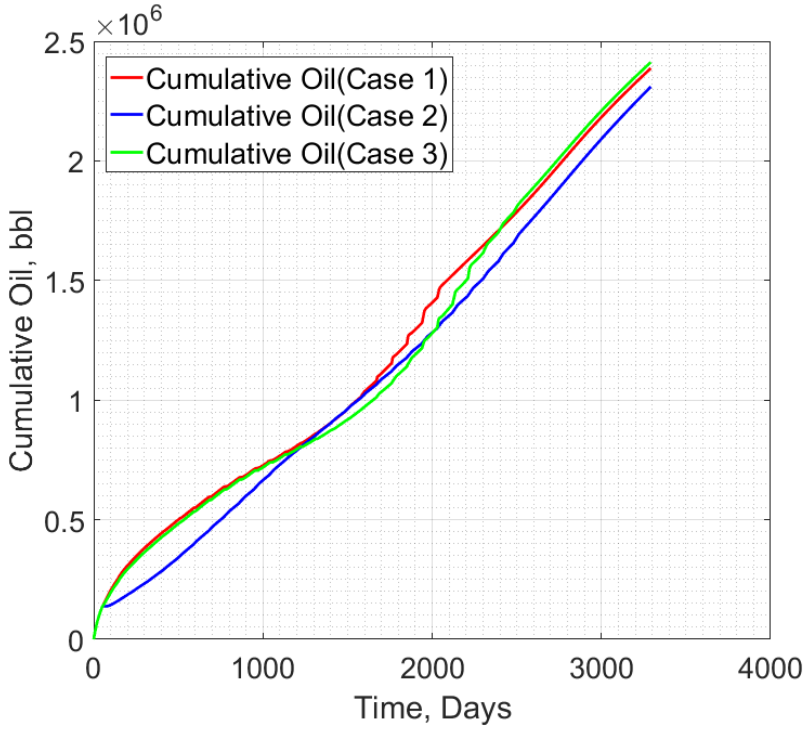

Figure 5.26: A comparison of the cumulative oil production for three different cases.

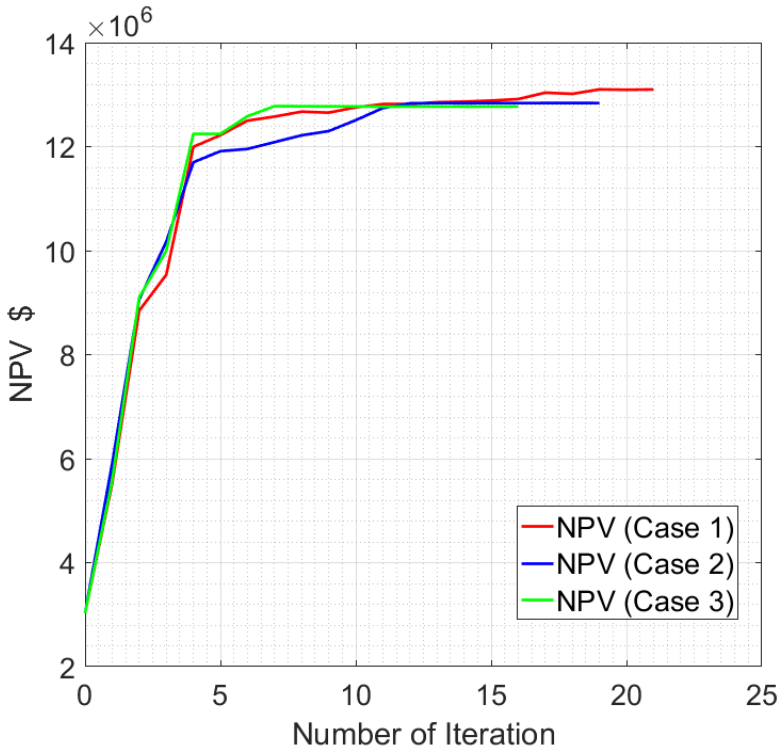

Figure 5.27: A comparison of the NPV for three different cases.

5.3 Results of GPR

Before we present our results with the GPR method, it is important to not that a learning algorithm that fits a training set well does not mean it is a good hypothesis. It could overfit. This means that the hypothesis can only fit the training set well but the predictions on the test set would be poor. As mentioned in Chapter 3.2.1, the results of machine learning regression can be affected by the hyper-parameters in the kernel functions. For this reason, finding appropriate values of hyper-parameters is very important. To achieve this, we break down the data set into three sets:

- Training set: 60%
- Cross validation set: 20%
- Test set: 20%

Calculate three separate error values for the three different sets using the following method:

1. Train the proxy model using the training set.

2. Find the property hyper-parameters with the least error by adjusting the range of the hyper-parameters using the cross validation set.

3. Estimate the generalization error using the test sets.

The error can be represented by a correlation coefficient. The correlation coefficient formulas tell us how strong a relationship is between two sets of data. The formula returns a value somewhere between negative 1 and positive 1. The positive correlation indicates a positive relationship. Positive 1 indicates that for every positive increase in one variable, there is a positive increase in a fixed proportion in the other. A correlation of 0 means that for every increase, there is not a positive or negative increase, the two sets of data is absolutely no correlation whatsoever. And a negative correlation means there is a correlation but it is in a negative direction. The correlation of negative 1 means that for every positive increase in one variable, there is a negative decrease in a fixed proportion in the other.

First, we construct an GPR proxy model by using the initial well controls without any normalization, however, the correlation coefficient for this proxy model is only 0.5058 (see Figure 5.28). Since there are 36 BHP, 36 steam injection rates, and 1 temperature of the injected steam as the input well control variables, the ranges of these input variables are very uneven, which can cause the calculation of the gradient very inefficiently and may incorrectly. The way to prevent or to improve this situation is to modify the range of the input variables so that they all have the same range: $0 \leq x_{(i)} \leq 1$. The technique to help with this is scaling. Scaling involves dividing the input values by the range (i.e. the maximum value minus the minimum value) of the input variable, resulting in a new range of just 1. The formula of scaling is given by:

$$x_{scaled} = \frac{x_{initial} - min(x_{initial})}{max(x_{initial}) - min(x_{initial})}, \tag{5.1}$$

where $min(x_{initial})$ and $max(x_{initial})$ are the minimum and maximum values in the initial input variables, respectively.

Using the scaled input well control variables, we generated a GPR proxy model for NPV, the correlation coefficient now increases to 0.7140 (see Figure 5.29). By adjusting the hyper-parameters in the kernel function, the correlation coefficient can be further improved to 0.8382 (see Figure 5.30).

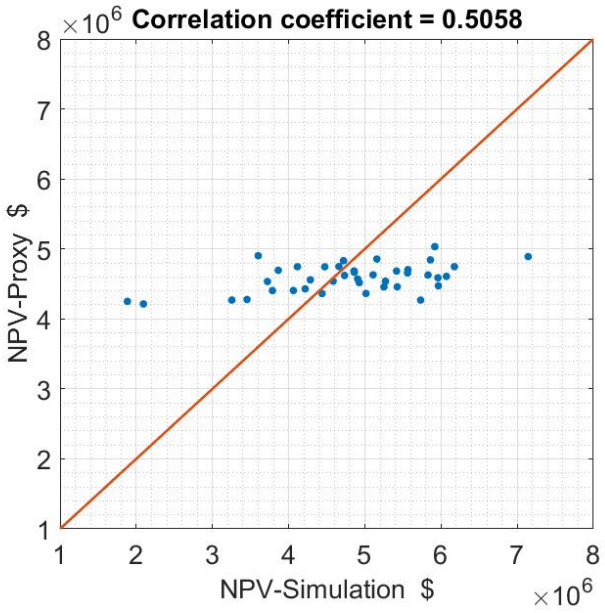

Figure 5.28: A cross-plot of NPV GPR proxy vs. NPV-simulation (CMG) based on unscaled well control variables.

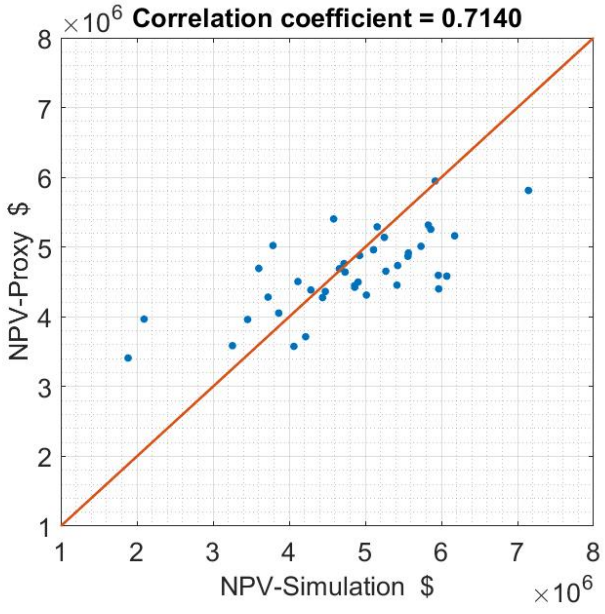

Figure 5.29: A cross-plot of NPV GPR proxy vs. NPV-simulation (CMG) based on scaled well control variables.

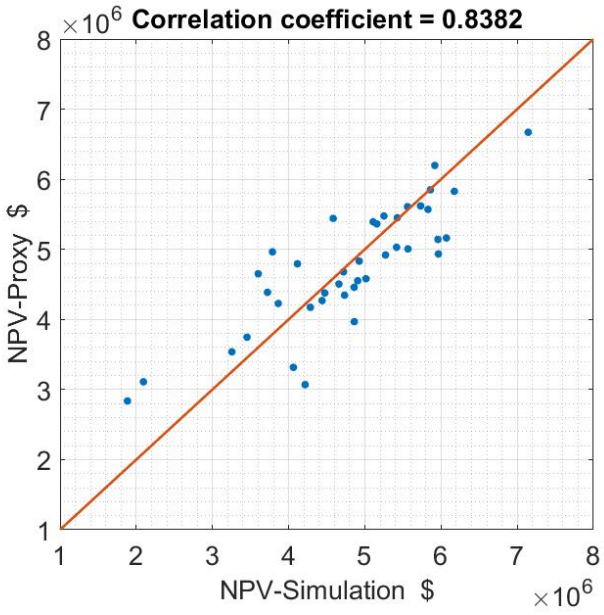

Figure 5.30: A cross-plot of NPV GPR proxy vs. NPV-simulation (CMG) based on scaled well control variables and adjusted hyper-parameters.

In the beginning, we tried to use 73 scaled well control variables to train the proxy model, however, the results from the GPR proxy model shows that the optimized well control variables are trapped in the boundaries. From Figure 5.31, we can say that the optimized BHPs are trapped in the lower bound. In Figure 5.32, the steam injection rates are trapped in the upper bound. This shows that with too much control variables, the optimized result from GPR model may be trapped in a local maximum.

Following the procedures in Chapter 4.2.1, we built a proxy model by GPR method. Re-sample 5 new initial well controls, they finally converge to 3 different maximum NPV. Once GPR proxy model of NPV is built, we preformed optimization of well control variables to maximize NPV by using the GPR proxy model with five different initial guesses of well control vectors, sampled from the LHS method. The results obtained for the optimal values of NPV from CMG-STARS and GPR proxy model with five different initial guesses of well control vectors are presented in Table 5.1. As can be seem from the results of Table 5.1, we basically have three different maximum values for the NPV predicted by the GPR proxy model.

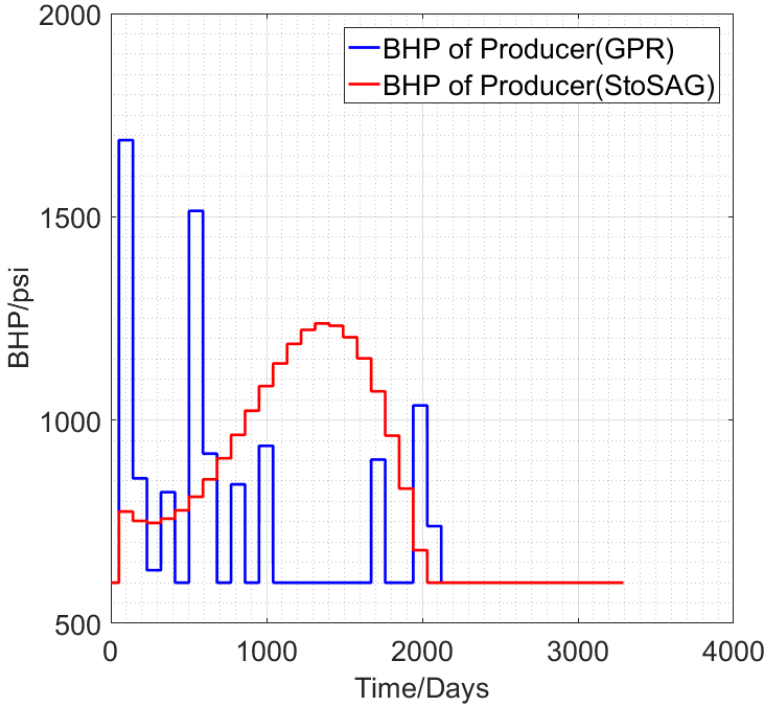

Figure 5.31: A comparison of the optimized BHP for GPR proxy and StoSAG methods.

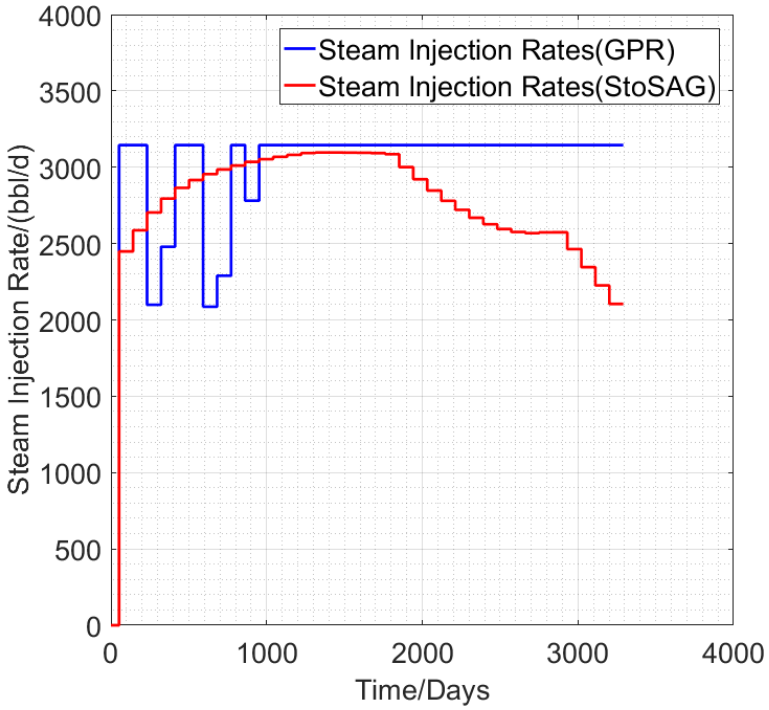

Figure 5.32: A comparison of the optimized steam injection rate for GPR proxy and StoSAG methods.

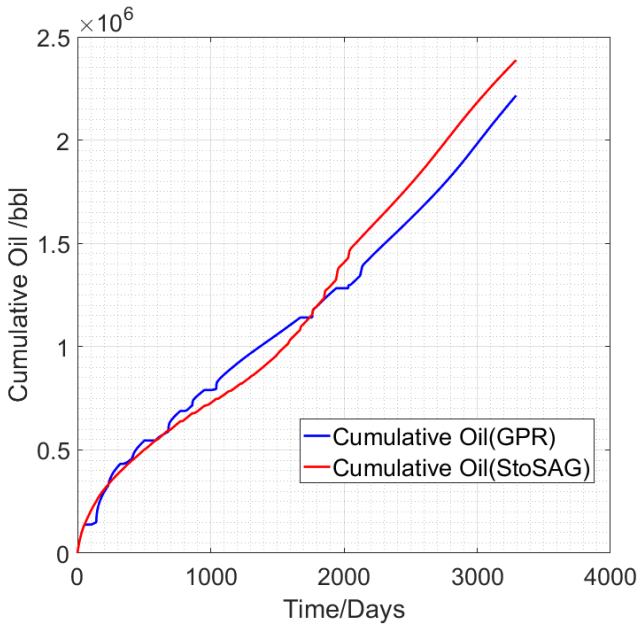

Figure 5.33: A comparison of the cumulative oil production for GPR proxy and StoSAG methods.

Since the number of the control variables is too large, the GPR model can be trapped in a local maximum. So, the number of control variables is reduced to 9. The method to decrease the control variables is shown in Figure 5.34 to Figure 5.36. Assuming that the value of every 9 consecutive variables is the same, which means that the steam injection rates and the BHP at the producer from 52 days to 862 days, 862 days to 1672 days, 1672 days to 2482 days and 2482 days to 3292 days are constants, in Figure 5.34, using BHP as the example. After normalization, the BHP is in the range of 0 to 1. To improve the accuracy of the proxy model, the method we use in this study is to add the training samples around the optimal point. In this case, we define a interval which is 0.2 far from the optimized BHP. By using this method, a trust region can be generated as shown in Figure 5.35 (Black area). Using the LHS method sample 5 initial guesses in this trust region, and increase the number of the control variable to 25, which means that the value of every 3 consecutive variables is the same. Using these initial guesses and follow the retraining procedures (Chapter 4.2.2), after optimization, the optimized BHP is shown in Figure 5.36.

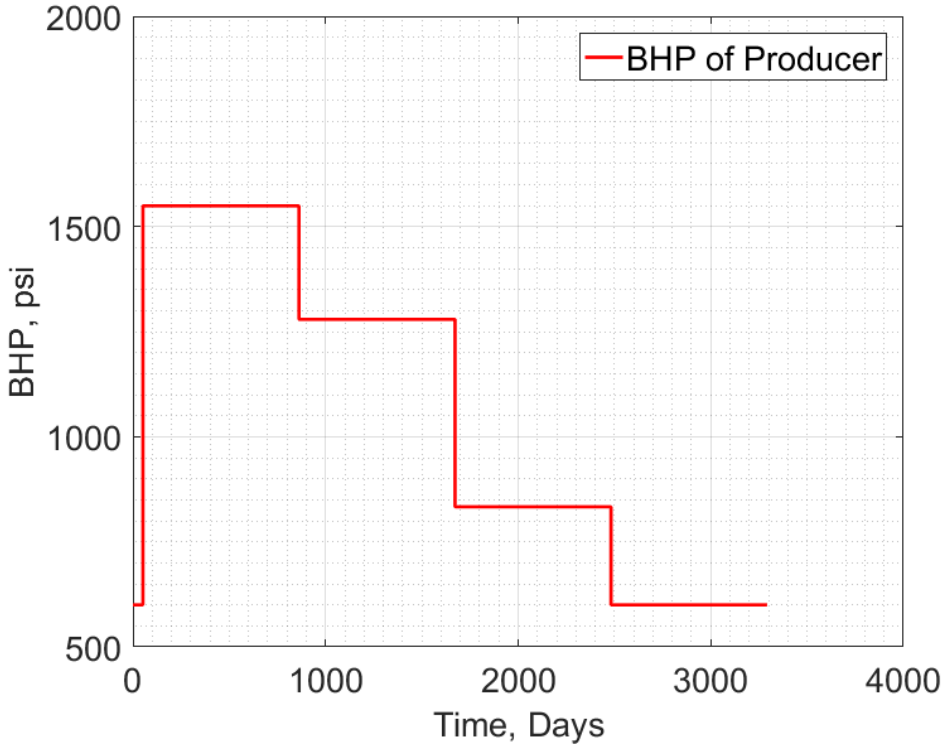

Figure 5.34: The initial trained proxy model result (optimized BHP) by 9 control variables.

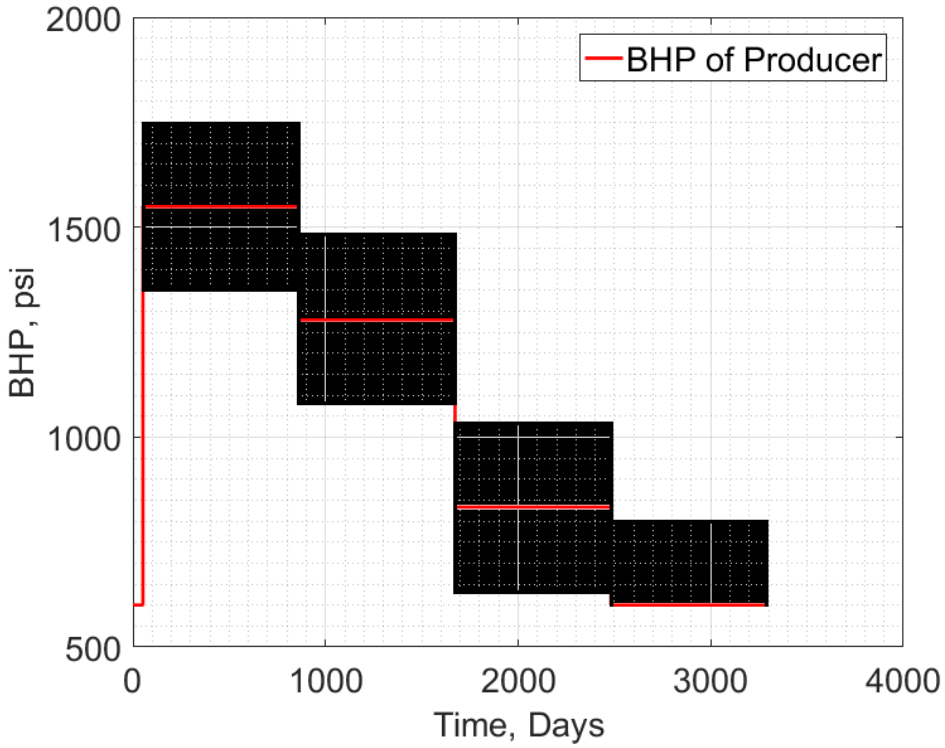

Figure 5.35: The trust region.

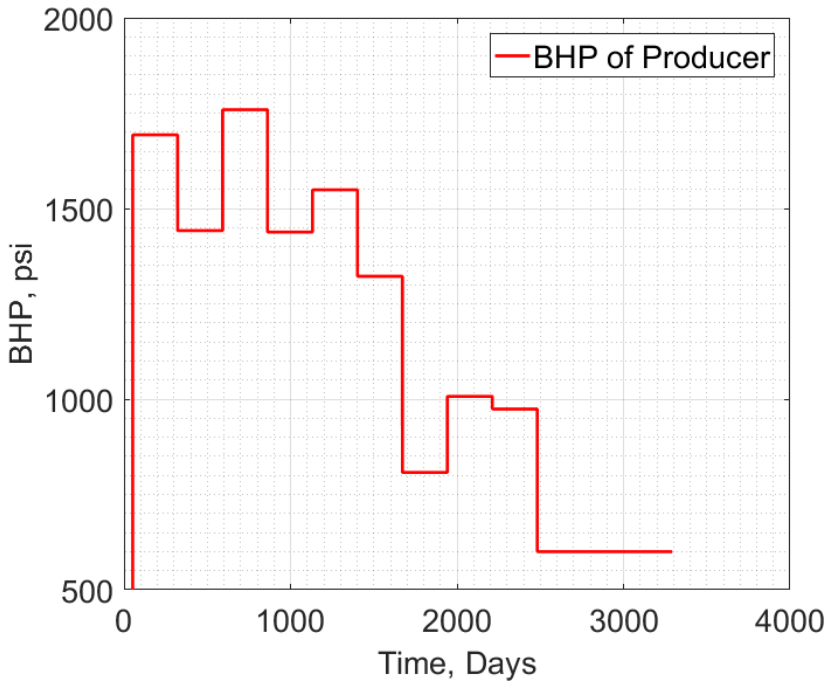

Figure 5.36: The retrained proxy model result (optimized BHP) by 25 control variables.

Table 5.1: NPV of CMG and GPR proxy model for three optimal control variables

Selected Optimal CV	NPV (CMG), $MM	NPV (GPR),$MM	Difference
Optimal CV1	11.234	11.800	5.04%
Optimal CV2	11.966	11.537	3.59%
Optimal CV3	11.307	11.695	3.43%
Optimal CV4	10.581	11.817	11.68%
Optimal CV5	10.581	11.842	11.92%

From Table 5.1, the difference between the results obtained from the GPR proxy model and the simulator (CMG) is more than 3%. Following the procedure in Chapter 4.2.2, we retrain the proxy model. After 2 retraining and optimizing steps, the results given in Table 5.2 are obtained:

After retraining the proxy model, the maximum NPV increased a little bit compared with the maximum NPV generated by the initial proxy model. The difference between the results from the retrained proxy model and the simulator (CMG) is around 1%, in this case, this GPR proxy model can be considered sufficient to replace the reservoir simulator (CMG).

Table 5.2: NPV of CMG and GPR proxy model for three optimal control variables after two iterations

Selected Optimal CV	NPV (CMG), $MM	NPV (GPR),$MM	Difference
Optimal CV1	11.928	12.058	1.09%
Optimal CV2	11.859	11.984	1.05%
Optimal CV3	12.002	12.026	0.20%
Optimal CV4	11.937	12.045	0.90%
Optimal CV5	11.734	12.007	2.33%

From Figure 5.34, it shows similar trends like the optimized BHP controls in Figure 5.22, the BHP increased to the highest point and generally decrease to the lower bound, which is 600 psi and maintains at the lower bound until the end of the life-cycle. From Figure 5.35, the optimized steam injection rates are trapped at the upper bound, which is quite different from the optimized well controls in StoSAG Method.

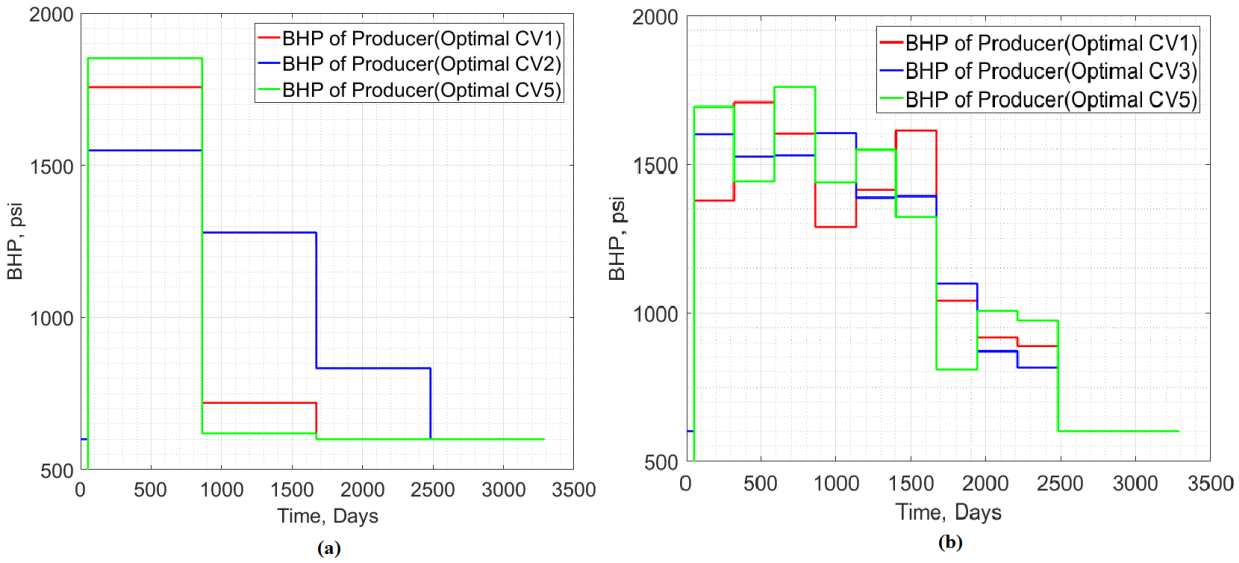

Figure 5.37: A comparison of the optimized BHP of producer from the initial proxy model (a) and from the retrained proxy model (b).

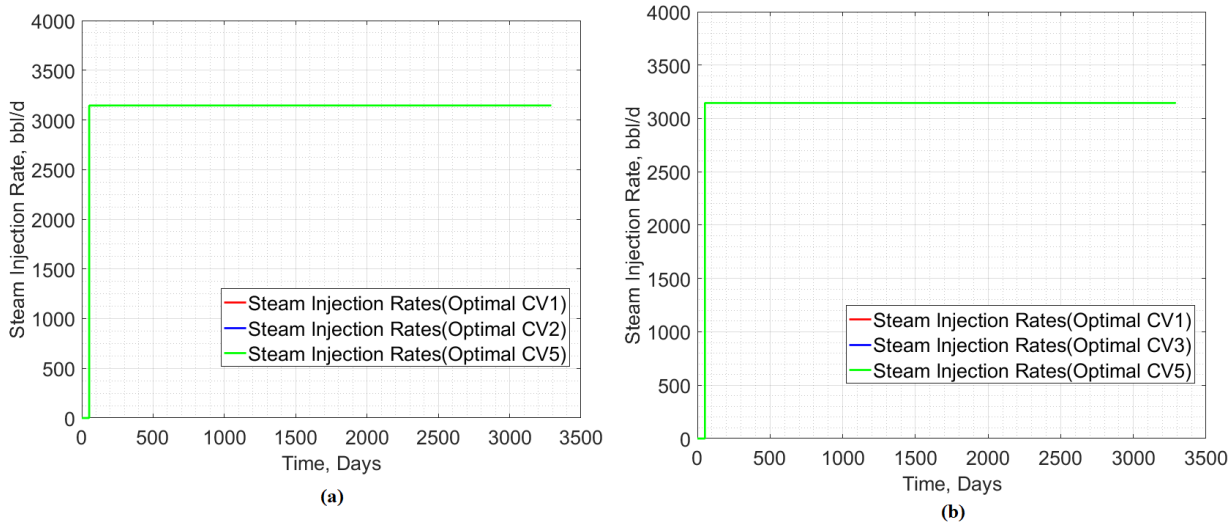

Figure 5.38: A comparison of the optimized steam injection rates from the initial proxy model (a) and from the retrained proxy model (b).

5.4 Comparison of Results

Comparing the results from StoSAG method and GPR method (Figure 5.39 to Figure 5.41), in Figure 5.41, we can state that, StoSAG method can get higher NPV than the GPR method. Using StoSAG method, the maximum NPV is 13.099 $\$MM$, which is 8.63% higher than GPR method, however, StoSAG method needs more than 400 simulator runs, GPR method only needs 47 simulator runs. In this case, the GPR method can save a lot of computational cost compare with StoSAG method (Table 5.3).

Table 5.3: NPV from StoSAG and GPR proxy model and the number of simulation runs

Selected Optimal CV	NPV, \$MM	No. Simulation runs
Optimal CV1	12.058	47
Optimal CV3	12.026	47
Optimal CV5	12.007	47
StoSAG CV	13.099	> 400

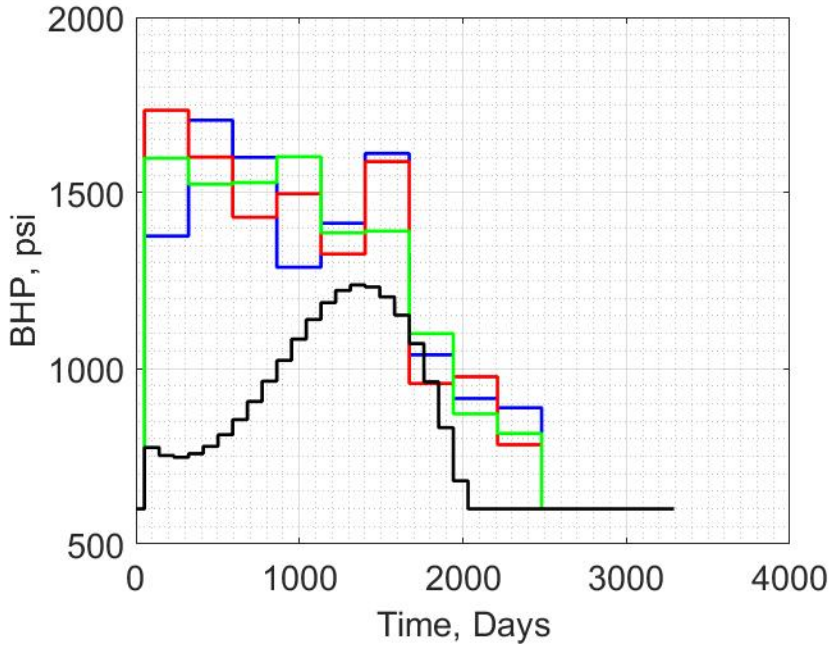

Figure 5.39: A comparison of the optimized BHP for StoSAG and GPR proxy model.

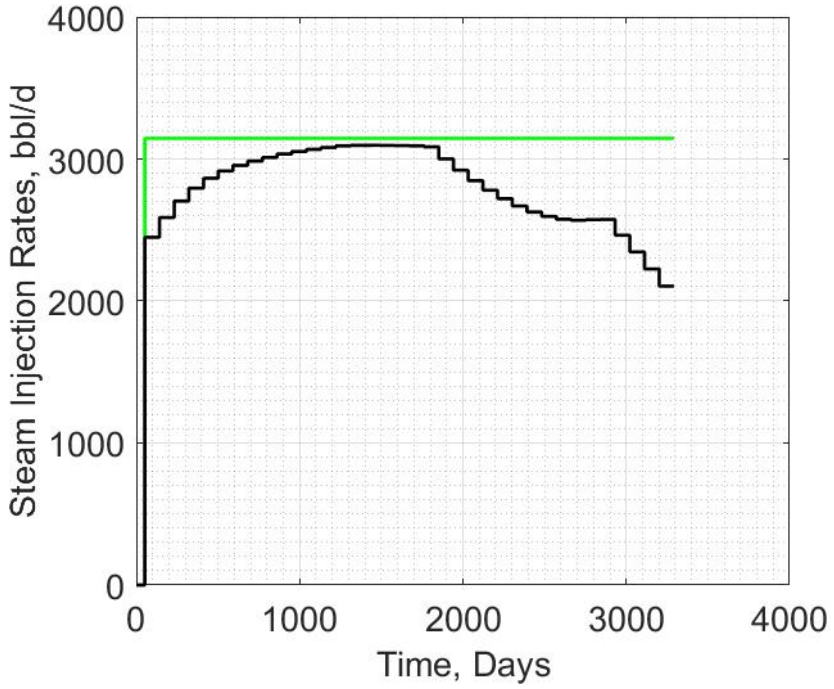

Figure 5.40: A comparison of the optimized steam injection rate for StoSAG and GPR proxy model.

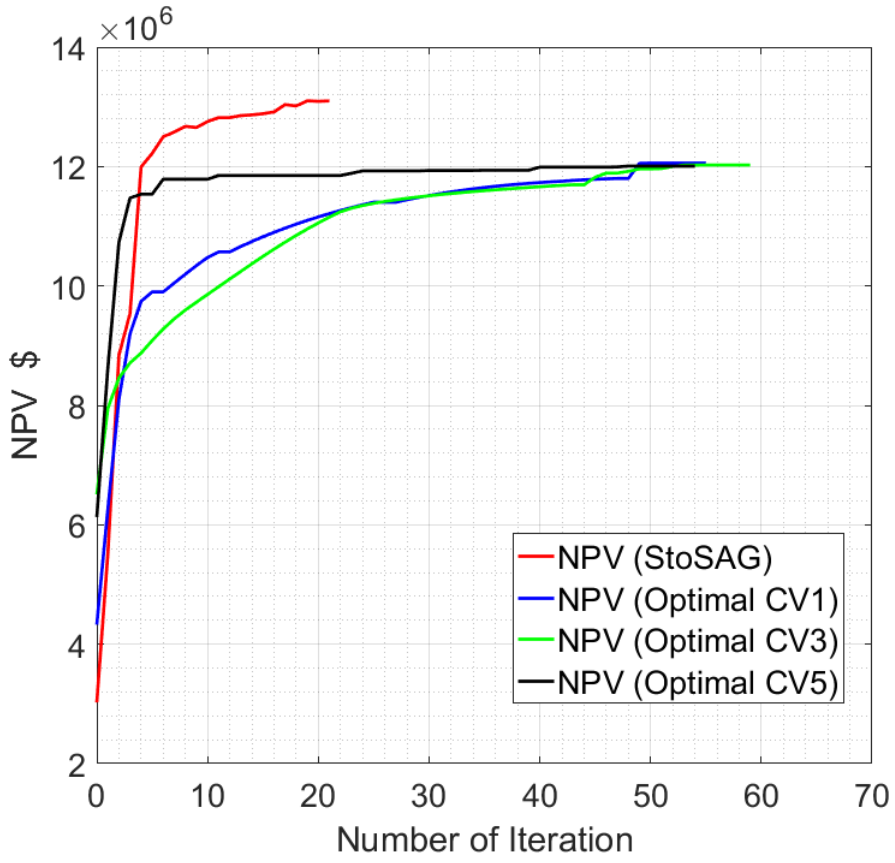

Figure 5.41: A comparison of the NPV for StoSAG and GPR proxy model.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

From this study, the following conclusions are warranted:

- In this study, for the gradient based StoSAG method, two parameters which include the perturbation size and the correlation length showed not significant effect on the maximized NPV, although they can affect the optimized well control variables.

- The optimized well control variables shows the similar results as Gates et al. and Bao et al. [2, 22]. In the early time, the BHP at the injector and the producer is at a high level, after the steam chamber reaches the top layer, the BHP generally decreases to a lower level.

- The GPR method results can be affected by the hyper-parameters in the kernel function. Compare with using non-scaled well control variables, using scaled well control variables to train the proxy model can improve the accuracy of the model significantly.

- For the GPR method, too many well control variables can make the optimized result trapped in local maximums, which may not yield a desired result for NPV. If we decrease the number of well control variables, the optimized result obtained from the GPR model can be improved.

- The steam injection rates optimized by GPR method have always been trapped in the upper bound considered in this study.

- In this study, the maximum NPV generated by gradient-based StoSAG method is higher than the maximum NPV generated by ML-based GPR method. However, the

number of simulator runs required by the StoSAG method is ten times that of the GPR method, which means a high computational cost.

6.2 Recommendations for Future Work

In future, we can use different kinds of machine learning methods to built the proxy model. For instance, we can use Least-Squares Support Vector Regression (LS-SVR) which have been found useful for other life-cycle optimization process, e.g., CO_2 huff-n-puff and water flooding. Comparing the results generated by the GPR method can help to find the more appropriate ML-based method to built a proxy model to be used in the optimization of SAGD process. We can use different permeability distribution map to assess the uncertainty in the life-cycle optimization of a SAGD process.

NOMENCLATURE

Greek Symbols

α	Scale mixture parameter
Δ	Tolerance of proxy model and CMG
ϵ	Perturbation size
μ	Mean function
Σ	covariance function
σ	Standard deviation
ε	Noise

Other Symbols

$\mathbb{E}$	Expectation
$\mathcal{N}$	Gaussian distribution
$\sim$	Distributed according to
a	Step length
b	Discount rate
$cov()$	Gaussian process posterior covariance
I	Identity matrix
$J(u)$	Objective function

l Length-scale parameter

Subscripts and Superscripts

i Injector

p Producer

i Samples index

j Perturbation index

k Input variables index

l Iteration index

n Control step index

o Oil

w Water

wf Bottom-hole

wi Cold water equivalent injected steam

BIBLIOGRAPHY

[1] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2009.

[2] Xia Bao, Zhangxin John Chen, Yizheng Wei, Chao Charlie Dong, Jian Sun, Hui Deng, Song Yu, et al. Numerical simulation and optimization of sagd process in surmont oil sands lease. In *Abu Dhabi International Petroleum Exhibition and Conference*. Society of Petroleum Engineers, 2010.

[3] R Butler et al. The steam and gas push (sagp). *Journal of Canadian Petroleum Technology*, 38(03), 1999.

[4] RM Butler, GS McNab, and HY Lo. Theoretical studies on the gravity drainage of heavy oil during in-situ steam heating. *The Canadian journal of chemical engineering*, 59(4):455–460, 1981.

[5] Roger Butler et al. Sagd comes of age! *Journal of Canadian Petroleum Technology*, 37(07), 1998.

[6] Roger M Butler. *Thermal recovery of oil and bitumen*, volume 46. Prentice Hall Englewood Cliffs, NJ, 1991.

[7] Roger M Butler et al. Steam-assisted gravity drainage: concept, development, performance and future. *Journal of Canadian Petroleum Technology*, 33(02):44–50, 1994.

[8] Bailian Chen, Albert C Reynolds, et al. Ensemble-based optimization of the water-alternating-gas-injection process. *SPE Journal*, 21(03):786–798, 2016.

[9] Bailian Chen and Jianchun Xu. Stochastic simplex approximate gradient for robust life-cycle production optimization: Applied to brugge field. *Journal of Energy Resources Technology*, pages 1–17, 2019.

[10] Y Chen. *Efficient ensemble based reservoir management.* PhD thesis, PhD Thesis, University of Oklahoma, USA, 2008.

[11] Yan Chen, Dean S Oliver, Dongxiao Zhang, et al. Efficient ensemble-based closed-loop production optimization. *SPE Journal*, 14(04):634–645, 2009.

[12] Sy T Do and Albert C Reynolds. Theoretical connections between optimization algorithms based on an approximate gradient. *Computational Geosciences*, 17(6):959–973, 2013.

[13] Jerome Downey. Downey: Understanding sagd technology.

[14] Neil Edmunds, Harbir Chhina, et al. Economic optimum operating pressure for sagd projects in alberta. *Journal of Canadian Petroleum Technology*, 40(12), 2001.

[15] P Egermann, G Renard, E Delamaide, et al. Sagd performance optimization through numerical simulations: Methodology and field case example. In *SPE International Thermal Operations and Heavy Oil Symposium*. Society of Petroleum Engineers, 2001.

[16] Rahul Fonseca, Olwijn Leeuwenburgh, Ernesto Della Rossa, Paul MJ Van den Hof, Jan-Dirk Jansen, et al. Ensemble-based multiobjective optimization of on/off control devices under geological uncertainty. *SPE Reservoir Evaluation & Engineering*, 18(04):554–563, 2015.

[17] Rahul Rahul-Mark Fonseca, Bailian Chen, Jan Dirk Jansen, and Albert Reynolds. A stochastic simplex approximate gradient (stosag) for optimization under uncertainty. *International Journal for Numerical Methods in Engineering*, 109(13):1756–1776, 2017.

[18] RM Fonseca, SS Kahrobaei, LJT Van Gastel, O Leeuwenburgh, JD Jansen, et al. Quantification of the impact of ensemble size on the quality of an ensemble gradient using principles of hypothesis testing. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2015.

[19] RM Fonseca, AS Stordal, O Leeuwenburgh, PMJ Van den Hof, and JD Jansen. Robust ensemble-based multi-objective optimization. In *ECMOR XIV-14th European conference on the mathematics of oil recovery*, 2014.

[20] Ian D Gates and Christopher Leskiw. Impact of steam trap control on performance of steam-assisted gravity drainage. *Journal of Petroleum Science and Engineering*, 75(1-2):215–222, 2010.

[21] Ian Donald Gates, Joseph Kenny, Ivan Lazaro Hernandez-Hdez, Gary L Bunio, et al. Steam injection strategy and energetics of steam-assisted gravity drainage. In *SPE International Thermal Operations and Heavy Oil Symposium*. Society of Petroleum Engineers, 2005.

[22] ID Gates, N Chakrabarty, et al. Optimization of steam-assisted gravity drainage in mcmurray reservoir. In *Canadian International Petroleum Conference*. Petroleum Society of Canada, 2005.

[23] Mohammad Ghasemi, Curtis Hays Whitson, et al. Numerical investigation and integrated optimization of solvent-sagd process. In *International Petroleum Technology Conference*. International Petroleum Technology Conference, 2014.

[24] CA Glandt, Hsu Chia-Fu, et al. Electric preheating in low-injectivity tar sand deposits. In *SPE/DOE Enhanced Oil Recovery Symposium*. Society of Petroleum Engineers, 1992.

[25] Zhenyu Guo, Albert C Reynolds, et al. Robust life-cycle production optimization with a support-vector-regression proxy. *SPE Journal*, 2018.

[26] Y Ito, S Suzuki, et al. Numerical simulation of the sagd process in the hangingstone oil sands reservoir. In *Annual Technical Meeting*. Petroleum Society of Canada, 1996.

[27] J-D Jansen, DR Brouwer, G Naevdal, and CPJW Van Kruijsdijk. Closed-loop reservoir management. *First Break*, 23(1):43–48, 2005.

[28] TMV Kaiser, SP Taubner, et al. Injector design optimisation for sagd conformance and cost. In *SPE Thermal Well Integrity and Design Symposium*. Society of Petroleum Engineers, 2016.

[29] KE Kisman, KC Yeung, et al. Numerical study of the sagd process in the burnt lake oil sands lease. In *SPE international heavy oil symposium*. Society of Petroleum Engineers, 1995.

[30] Michael D McKay, Richard J Beckman, and William J Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.

[31] Richard F Meyer, Emil D Attanasi, and Philip A Freeman. Heavy oil and natural bitumen resources in geological basins of the world. Technical report, US Geological Survey, 2007.

[32] CM Nascimento et al. Design, optimization and operation of sagd wells using dynamic flow simulations. In *SPE Western Regional Meeting*. Society of Petroleum Engineers, 2016.

[33] Tawfik Noaman Nasr, Oluropo Rufus Ayodele, et al. Thermal techniques for the recovery of heavy oil and bitumen. In *SPE international improved oil recovery conference in Asia Pacific*. Society of Petroleum Engineers, 2005.

[34] CP Nathanail, MS Rosenbaum, et al. Probabilistic slope stability assessment using latin hypercube sampling. In *7th ISRM Congress*. International Society for Rock Mechanics and Rock Engineering, 1991.

[35] Emil Nurmammadov. *Optimization of SAGD Operations Using Derivative Free Algorithms*. PhD thesis, University of Tulsa, 2011.

[36] Songliang Peng, Aiwu Yuan, Hua Jiang, Shoujun Zhang, et al. Cyclic preheating test

in sagd horizontal wells of liaohe oilfield. In *International Oil and Gas Conference and Exhibition in China*. Society of Petroleum Engineers, 2010.

[37] Nestor V Queipo, Javier V Goicochea, and Salvador Pintos. Surrogate modeling-based optimization of sagd processes. *Journal of Petroleum Science and Engineering*, 35(1-2):83–93, 2002.

[38] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71. Springer, 2003.

[39] Hyundon Shin, Marcel Polikar, et al. Optimizing the sagd process in three major canadian oil-sands areas. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2005.

[40] HY Wang et al. Application of temperature observation wells during sagd operations in a medium deep bitumen reservoir. *Journal of Canadian Petroleum Technology*, 48(11):11–15, 2009.

[41] Card Yang, C Card, L Nghiem, et al. Economic optimization and uncertainty assessment of commercial sagd operations. In *Canadian International Petroleum Conference*. Petroleum Society of Canada, 2007.

[42] Chaodong Yang, Colin Card, Long X Nghiem, Eugene Fedutenko, et al. Robust optimization of sagd operations under geological uncertainties. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2011.