Robust Optimization for Sequential Field Development Planning

Lingya Wang

UNIVERSITY OF BERGEN

Robust Optimization for Sequential Field Development Planning

Lingya Wang

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 24.09.2021

Year: 2021

Title: Robust Optimization for Sequential Field Development Planning

Name: Lingya Wang

Print: Skipnes Kommunikasjon / University of Bergen

Scientific environment

This thesis is submitted as partial fulfillment of the requirements for the degree of Philosophiae Doctor in Applied Mathematics at the University of Bergen. The supervising committee consists of Guttorm Alendal (University of Bergen), Dean S. Oliver (NORCE), Geir Evensen (NORCE), and Andreas S. Stordal (NORCE).

Acknowledgements

This thesis has been a wonderful journey. It would not have been possible without the help and support of many people. I am very grateful for all the people I have met during my Ph.D.

First, I would like to thank my supervisors Guttorm Alendal, Dean S. Oliver, Geir Evensen, Andreas S. Stordal for the opportunity to pursue my Ph.D. A big thanks to Dean for his excellent supervision and guidance throughout the research. It was a great learning experience for me to work closely with him. I highly appreciate his incredible support, guidance, patience, and encouragement throughout my Ph.D. study.

Last but not least, many thanks to my family and friends for giving me unconditional support whenever I need it.

Lingya Wang
Bergen, May 2021

Abstract

To achieve high profitability from an oil field, optimizing the field development strategy (e.g., well type, well placement, drilling schedule) before committing to a decision is critically important. The profitability at a given control setting is predicted by running a reservoir simulation model, while determining a robust optimal strategy generally requires many expensive simulations. In this work, we focus on developing practical and efficient methodologies to solving reservoir optimization problems for which the actions that can be controlled are discrete and sequential (e.g., drilling sequence of wells).

The type of optimization problems I address must take into account both geological uncertainty and the reduction in uncertainty resulting from observations. As the actions are discrete and sequential, the process can be characterized as sequential decision-making under uncertainty, where past decisions may affect both the possibility of the future choices of actions and the possibility of future uncertainty reduction. This thesis tackles the challenges in sequential optimization by considering three main issues: 1) optimizing discrete-control variables, 2) dealing with geological uncertainty in robust optimization, and 3) accounting for future learning when making optimal decisions.

As the first contribution of this work, we develop a practical online-learning methodology derived from A* search for solving reservoir optimization problems with discrete sets of actions. Sequential decision making can be formulated as finding the path with the maximum reward in a decision tree. To efficiently compute an optimal or near-optimal path, heuristics from relaxed problems are first used to estimate the maximum value constrained to past decisions, and then online-learning techniques are applied to improve the estimation accuracy by learning the errors of the initial approximations obtained from previous decision steps. In this way, an accurate estimate of the maximized value can be inexpensively obtained, thereby guiding the search toward the optimal solution efficiently. This approach allows for optimization of either a complete strategy with all available actions taken sequentially or only the first few actions at a reduced cost by limiting the search depth.

The second contribution is related to robust optimization when an ensemble of reservoir models is used to characterize geological uncertainty. Instead of computing the expectation of an objective function using ensemble-based average value, we develop various bias-correction methods applied to the reservoir mean model to estimate the expected value efficiently without sacrificing accuracy. The key point of this approach is that the bias between the objective-function value obtained from the mean model and the average objective-function value over an ensemble can be corrected by only using information from distinct controls and model realizations. During the optimization process, we only require simulations of the mean model to estimate the expected value

using the bias-corrected mean model. This methodology can significantly improve the efficiency of robust optimization and allows for fairly general optimization methods.

In the last contribution of this thesis, we address the problem of making optimal decisions while considering the possibility of learning through future actions, i.e., opportunities to improve the optimal strategy resulting from future uncertainty reduction. To efficiently account for the impact of future information on optimal decisions, we simplify the value of information analysis through key information that would help make better future decisions and the key actions that would result in obtaining that information. In other words, we focus on the use of key observations to reduce the uncertainty in key reservoir features for optimization problems, rather than using all observations to reduce all uncertainties. Moreover, by using supervised-learning algorithms, we can identify the optimal observation subset for key uncertainty reduction automatically and evaluate the information's reliability simultaneously. This allows direct computation of the posterior probability distribution of key uncertainty based on Bayes' rule, avoiding the necessity of expensive data assimilation algorithms to update the entire reservoir model.

List of papers

A Wang, L. and Oliver, D.S., 2019, *Efficient Optimization of Well-Drilling Sequence with Learned Heuristics*, SPE Journal, 24(5), 2111-2134.

B Wang, L. and Oliver, D.S., 2021, *Fast Robust Optimization Using Bias Correction Applied to the Mean Model*, Computational Geosciences, 25, 475–501.

C Wang, L. and Oliver, D.S., 2021, *Improving Sequential Decisions – Efficiently Accounting for Future Learning*, Journal of Petroleum Science and Engineering, 205, 108770.

Contents

Part I

Scientific background

Chapter 1

Introduction

In petroleum field development planning, optimization is crucially important to substantially improve the profitability of reservoirs (e.g., net present value (NPV), total oil production). Decisions such as the number of wells, well location, well type, drilling schedule of wells and well control settings can have a significant impact on the revenues and costs. However, in general, many time-consuming reservoir simulations are needed to compute an optimal strategy. Therefore, to obtain an optimal or near-optimal solution efficiently, optimization methods that are appropriate to the optimization problem are called for. In this work, we develop robust and efficient methodologies to search for an optimal sequential solution for reservoir optimization problems with discrete sets of actions, deal with uncertainty in reservoir characteristics and account for future learning possibilities through actions. An overall goal of the development of these methodologies is to solve sequential field development tasks efficiently while considering the effects of geological uncertainty and future uncertainty reduction on the optimal strategy (i.e., dynamic sequential decision-making problems under uncertainty). In the first section of this introductory chapter, we provide the motivation for this research, identifying problems and objectives. The next section lists the main contributions of this work. Finally, we outline of the remainder of the thesis.

1.1 Motivation

In the past few decades, a significant amount of work has been done in reservoir-development planning with a focus on optimization techniques. Various gradient-based and derivative-free optimization methods have been investigated, but most of the work has been aimed at optimizing well placement or well control. The optimization of other factors (e.g., drilling sequence, well type) can also play important roles in increasing the reservoir's profitability. However, there are few methods for how to solve optimization problems efficiently with these discrete-control variables. As field development is performed over time, uncertainty in reservoir properties can be reduced by considering past observations before making the next decision to proceed. Accounting for this sequential nature of the problem, the optimization of complete development strategies is an observation-based sequential decision-making process under uncertainty, whereby past decisions will affect future choices of actions and future possibilities of uncertainty reduction. In this work we aim to develop robust and efficient methods for the following three optimization problems:

Optimization with discrete-control variables

Determination of well-drilling schedule is an important element of field development plan. For some petroleum fields, the profitability could be improved as much as more than 20% by optimizing the drilling sequence of wells. In theory, the true optimal solution can be found by evaluating all possible drilling sequences. However, because the number of possible sequences grows rapidly with the number of wells, this naïve approach is infeasible when there are more than about six wells that need to be sequenced. For example, when optimizing the drilling schedule of eight wells, there would be 40320 (8!) possible combinations of drilling sequences; hence, there is a need for an efficient and practical method for reservoir optimization problems in which the actions that can be controlled are discrete and sequential.

In one earlier study, Beckner and Song [9] applied simulated annealing to optimize both the locations and drilling sequence of wells. This approach can reduce the probability of ending up trapped in a local minimum (i.e., there are almost certainly multiple minima for the drilling-order problem), but requiring an infeasible number of simulations. In recent times, gradient-based optimization with an approximate gradient (i.e., one computed using an ensemble of control perturbations) [67, 23, 94, 35] has become increasingly popular and provided encouraging results in production optimization under uncertainty. For discrete-control (not differentiable) variables, new continuous-control variables, such as the drilling priorities that map to discrete-drilling sequences [43, 42, 65], have been proposed to determine the optimal drilling order of wells by the use of approximate-gradient methodologies. Gradient-based optimization schemes converge to local optimal solutions, however, and the ordering of wells problem might have a local optimum that is far away from the true optimum. When a greedy optimizer is used, it is generally unclear if the optimized strategy is close to the true optimal solution. Although repeating the optimization with different initial guesses can improve the solution, it requires additional computation effort, making this approach impractical for large problems.

To avoid the dependency of the solution on the gradient (or approximation gradient) for optimization, Lamas et al. [63] described two methodologies based on best-first search and evaluation algorithm, respectively, for drilling-sequence optimization in deterministic models. In the best-first search algorithm, the optimal drilling order is computed by ranking the wells based on an approximation of the maximum NPV, which is obtained by assuming that multiple wells are drilled simultaneously. During the optimization process, the search extends in a direction determined by the highest estimated value. A complete optimized strategy can then be obtained efficiently, but the quality of the solution depends heavily on the estimated optimal NPV. In their test example of a heavy-oil reservoir model, drilling all remaining wells simultaneously resulted in an approximate value that was much lower than the actual NPV when all wells were drilled sequentially. In that case, a better solution might be ignored due to the underestimated values (i.e., poor estimates will misguide the search direction). The second methodology was based on a normal NPV distribution generated from a large number of random drilling schedules, which were used to identify the ordering of wells with high values. This optimization procedure, however, does not seem to be efficient enough for application to problems with large numbers of wells or with many high-values combinations. Hence, one of the main goals in this work is to develop

more efficient and practical methods to solve general reservoir optimization problems with discrete sets of actions (i.e., sequential decision-making problems).

Robust optimization under uncertainty

Due to the complexity of subspace and limited observations of the reservoir, reservoir properties such as porosity, permeability, and fault transmissibilities may be highly uncertain. In this case, the general goal of reservoir optimization problems is to maximize the *expected* value of the objective-function (e.g., expected NPV), namely robust optimization (RO) under uncertainty. The ensemble-based method has been widely used as a powerful tool for quantifying geological uncertainty. In general, the expectation of an objective function is estimated by averaging over all realizations, i.e., sample average approximation (SAA). Van Essen et al. [99] have demonstrated that RO performed with such an approximation of expected NPV is superior to deterministic optimization from a single realization. The straightforward computation of the expected value requires simulations of each individual realization, however. Consequently, RO for which uncertainty is quantified by an ensemble of reservoir models can be computationally demanding.

To reduce the cost of reservoir simulation-based RO, many researchers have studied how to accelerate reservoir simulations or reduce the number of simulations needed for optimization. The former can be achieved by using a simplified model from reduced order modeling or upscaling model parameters [52, 14, 15, 98, 31], but such a proxy model is generally less accurate and may result in a suboptimal solution in RO [28]. Alternatively, the number of simulations required in RO can be reduced by using more efficient methods of optimization. For general RO problems with continuous variables, ensemble-based optimization [23, 35] can be very efficient because the gradient of the expected value with respect to control variable settings is computed using the same ensemble that is used to represent uncertainty in model parameters. When the control variables to be optimized are discrete, optimization algorithms that do not rely on using the gradient are generally preferable. Nevertheless, even if the optimization is improved such that many fewer iterations are needed for obtaining the maximum, the amount of computation required for RO performed with SAA over a large ensemble could still be very expensive.

As mentioned previously, the main issue facing RO over an ensemble of model realizations is that the expected value computed using SAA requires many expensive simulations; hence, one of the most effective ways to improve the efficiency of RO is to reduce the cost of evaluating the expected value. In many studies, instead of using the full ensemble, a subset of ensemble members is employed to represent uncertainty for RO [54, 89, 77]. However, the representation of uncertainty in reservoir properties may not be modeled well with only a few realizations. To achieve the optimal balance between the uncertainty representation and the reduction in the cost of RO, the subset of model realization must be selected carefully. Barros et al. [7] presented an automated scenario reduction approach for obtaining an optimal subset of model realizations that are able to represent the full ensemble. Although the representation of uncertainty is reduced through model selection, the cost of RO performed over a subset of ensemble members still increases linearly with the number of representative realizations and the number of iterations required for optimization. To address this issue, a fast approach is to directly use the mean reservoir model for the estimation of

the expected objective-function value [19]. The dependence of an objective function on the model parameters is generally highly nonlinear, however, so that the mean model may offer a poor estimate of the expected value and result in loss of the solution quality. Therefore, in addition to developing efficient optimization methods, we also need to design a fast and effective approach to estimate the expectation of an objective function when a number of model realizations characterize uncertainty.

Decision-making accounting for future learning

Geological uncertainty creates challenges in optimization, but the uncertainty can be reduced through history matching or data assimilation based on past observations. In traditional closed-loop reservoir management (CLRM)[51, 83, 101, 23], production optimization and history matching are combined to improve the profitability of reservoir development by updating the reservoir model continuously with new observations and then re-optimizing the production strategy in the updated reservoir model. In an application of the drilling-order problem, Hanea et al. [42] investigated the impact of history matching well data on creating value (named the value of learning), after adjusting the drilling schedule for the remaining wells over the current uncertainty assessment. They demonstrated that information obtained at the early decision stages has a larger potential for improving the optimal strategy. However, as in most applications of CLRM optimization, the consequences of a current decision on the future uncertainty state are not taken into account, i.e., the possibility of reducing uncertainty through future observations is neglected in optimization. To act optimally, when choosing actions, we should also take into account the opportunities to improve the optimal strategy resulting from future uncertainty reduction (i.e., future learning possibilities), rather than solely accounting for the possibility of achieving the maximum NPV over the current uncertainty state.

Robust optimization that accounts for the possibility of future learning through the future information from remaining actions is closely related to the concept of the value of information (VOI) [86, 40, 12]. Barros et al. [5, 6, 8] presented the procedure for evaluating the VOI obtained from ensemble-based history matching and reservoir optimization (i.e., an optimal CLRM strategy with traditional production observations). Hong et al. [50] carefully articulated the concept of VOI from the perspective of decision analysis and illustrated the general workflow of VOI analysis using a simple application with saturation information in a 2D waterflooded reservoir for the design of a polymer flood. Barros et al. [4] used the VOI framework to optimize bottom-hole pressure controls on wells in a single inverted 5-spot pattern. Although these studies demonstrated that additional value is created when accounting for the potential future uncertainty reduction before committing to a decision, the use of VOI decision tree is impractical for most reservoir optimization problems because it would be computationally intractable to compute the VOI in a rigorous way that considers all possible data values that might be obtained from an action. That is, one must update the reservoir model and re-optimize the strategy for each possible outcome of the data.

Several approaches have been proposed to estimate the VOI [39, 18, 34, 48, 4] for problems in which production flow data must be assimilated. However, the workflow of the VOI analysis that combines history matching and optimization is still computationally prohibitive for realistic problems. Torrado et al. [97] formulated

the sequential drilling problem as a partially observable Markov decision process (POMDP), and then applied partially observable Monte-Carlo planning (POMCP) to compute an approximate optimal solution considering the future learning possibilities through all remaining actions. The optimization procedure in POMCP is similar to the standard VOI analysis, in which the effects of all possible combinations of future observations are considered, while only the strategies with potentially high expected values would be evaluated. In this approach, both the posterior probabilities of uncertainty and the expected NPV are estimated by sampling deterministic realizations at given previous observations and control settings. Therefore, there is no need to update the reservoir model through history matching. Even so, many expensive simulations were still required to optimize the drilling sequence of wells in the case with only two possible observations from each well. When all possible future observations from all remaining actions are taken into account, the computational effort required in making decisions will increase exponentially with the decision and the observation spaces. Hence, most applications that have considered future learning have had very few decision options (e.g., drill or not to drill a well) or few possible data [25, 6, 50]. In a realistic problem, however, each step includes many possible decision alternatives and many possible data that could be obtained from decisions. Therefore, it is desirable to make the computation of VOI manageable and design a more practical way to account for the future learning possibilities through actions.

1.2 Main contributions

The overall aim of this dissertation is to develop efficient and robust methods for the optimization of reservoirs' profitability through discrete sets of actions, taking into account the effects of geological uncertainty and future uncertainty reduction, i.e., observation-based dynamic sequential decision-making under uncertainty. To achieve this goal, the following research questions arise based on the above discussion,

- How can an optimal strategy be computed efficiently when the goal (e.g., maximize NPV) is not a continuous function of the control variables (e.g., drilling order)?

- How can field development be optimized efficiently when an ensemble of model realizations characterizes the uncertainty?

- How can future learning possibilities be accounted for efficiently through actions when optimizing field development plan under uncertainty?

Figure 1.1 shows the research problems, the objectives, and the proposed methods for addressing each of the above research questions. Below we summarize the main contributions of this work:

Efficient sequential optimization with learned heuristics Paper A presents an optimization methodology based on heuristic search and online-learning techniques for efficiently solving optimization problems with discrete actions. We formulate the sequential decision-making problem as finding the best path in a directed tree search, and use information from previous and future actions to estimate the maximum NPV constrained to past decisions. To obtain accurate approximations, online-learning techniques are designed to improve accuracy by learning the observations obtained

from previous decision steps. In this way, the search direction can be effectively guided toward the optimal solution. This approach can be used to either optimize a complete strategy or optimize only the first few actions at a reduced cost by limiting the search depth.

Fast robust optimization using bias-corrected mean model Paper B describes a fast and effective approach that requires only simulation of the mean reservoir model with a bias correction factor for estimating the expectation of the objective function. First the mean reservoir model provides an initial approximation of the expected value, which is generally poor compared to SAA. Then information from distinct controls and model realizations is used to correct the bias between the objective-function value obtained from the mean model and SAA, obtaining an estimate close to SAA while requiring many fewer simulations. By using such a bias-corrected mean model to account for model uncertainty, we can improve the efficiency of RO significantly without sacrificing the estimation accuracy of the expected value. This approach allows for the application of fairly general optimization methods.

Improving decisions - efficiently accounting for future learning Paper C proposes a flexible workflow built on a key-feature-based VOI analysis to make optimal decisions efficiently while considering the future learning possibilities. Instead of taking into account all possible future observations, we consider only the important information from key actions to reduce the uncertainties that have large influences on optimal decisions, i.e., information that will be most helpful in making better future decisions. To identify key information efficiently, we built supervised-learning algorithms that can automatically detect the optimal combination of observations to reduce key uncertainties and simultaneously estimate the information's reliability. The posterior probabilities of key uncertainties then can be computed directly based on Bayes' rule, avoiding the need for history matching or data assimilation. This workflow is practical for sequential decision-making under uncertainty, considering the opportunities to improve the optimal strategy resulting from future uncertainty reduction

1.3 Thesis outline

The thesis is structured in two parts. Part I introduces the scientific background, and Part II presents the main scientific contributions in the form of three peer reviewed papers. The remainder of the thesis is organized as follows:

- Chapter 2 - This chapter gives an introduction to general reservoir optimization problems and robust sequential decision-making under uncertainty.

- Chapter 3 - In this chapter, we present the development of a learned heuristic search method for efficiently solving optimization problems with discrete sets of actions.

- Chapter 4 - We describe various bias-corrected methods to estimate accurately and efficiently the expected objective-function value for RO under uncertainty.

- Chapter 5 - This chapter presents the VOI analysis through key actions and key information for improving decisions accounting for future learning.

- Chapter 6 - The final chapter summarizes each of the papers in Part II and provides an outlook for future research.
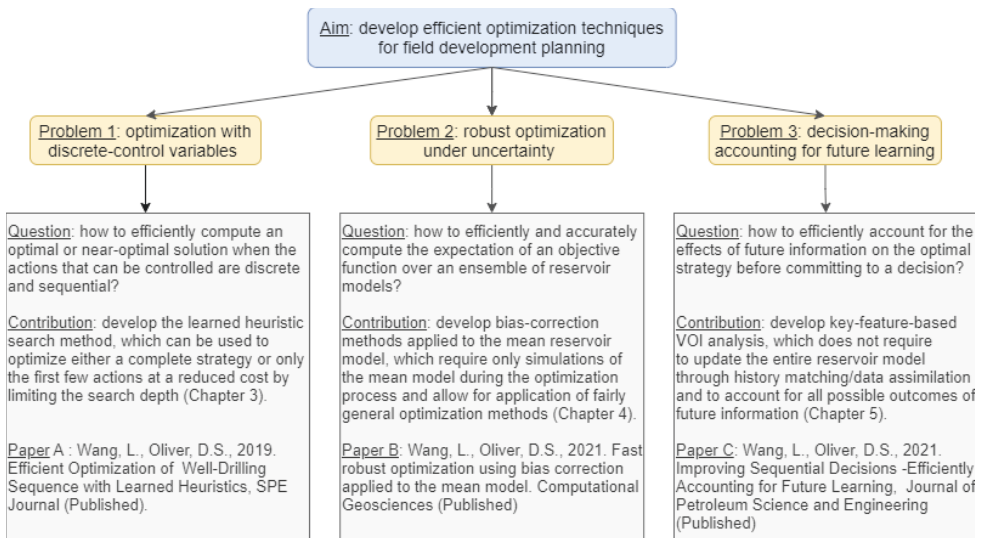
Figure 1.1: A summary of research problems, objectives and main contributions of thesis work.

Chapter 2

Optimization problem

The main focus of this work is on reservoir optimization problems with discrete and sequential actions, for example, maximizing the expected NPV by optimizing the drilling sequence of wells. In this chapter, we first give a brief introduction to reservoir development optimization under geological uncertainty, i.e., robust optimization under uncertainty. We then illustrate sequential decision-making using the example of the drilling-order problem, and we discuss the optimal strategies obtained from the traditional CLRM that only considers the current assessment of uncertainty in optimization and the standard VOI decision analysis considering the effects of all possible future information, respectively.

2.1 Robust optimization under uncertainty

The general purpose of field development optimization is to identify an optimal strategy $\mathbf{x}^*$ that maximizes a given objective function J,

$$\mathbf{x}^* = \arg\max J(\mathbf{x}), \tag{2.1}$$

where $\mathbf{x}$ is a vector of control inputs (e.g., well rates, well type), and J generally is an economic measure of oil reservoirs (e.g., NPV, oil recovery factor).

NPV is the most widely used economic objective function. It measures the difference between discounted cash inflows and outflows over a period of time. Assuming that all revenues are from oil production, and all costs are caused by water production, water injection and drilling wells, the NPV for a single model realization at control $\mathbf{x}$ can be computed as

$$J(\mathbf{x}) = \mathrm{NPV}(\mathbf{x}) = \sum_{k=1}^{T} \frac{\left(q_{o,k}(\mathbf{x}) \cdot r_o - q_{w,k}(\mathbf{x}) \cdot r_w - q_{wi,k}(\mathbf{x}) \cdot r_{wi}\right) \cdot \Delta t_k}{(1+b)^{t_k/\tau}} - \sum_{n=1}^{N_w} \frac{W_n}{(1+b)^{t_n/\tau}}, \tag{2.2}$$

where $q_{o,k}$, $q_{w,k}$ and $q_{wi,k}$ are the rates of oil production, water production and water injection at time t_k, respectively; r_o, r_w and r_{wi} are the oil price, water production cost, and water injection cost, respectively; T is the number of time steps; Δt_k is the time interval in days; b is the discount rate for a certain reference time τ (typically one year); N_w is the total number of wells to be drilled; W_n denotes the cost of drilling the nth well; t_n is the cumulative time in days up to the open time for each well.

Due to limited observations of the reservoir, the reservoir characterization is incomplete with considerable uncertainty in properties. Geological uncertainty could

influence the optimal decisions significantly, so that optimization based on a single model realization may lead to a suboptimal solution deviating severely from the actual optimality. To obtain a robust optimal strategy, it is usually necessary to consider the effect of geological uncertainty in optimization. In general, when taking into account geological uncertainty, the objective of optimization is to maximize the expected objective-function value,

$$\mathbf{x}^* = \arg\max E[J(\mathbf{x}, \mathbf{m})], \tag{2.3}$$

where $\mathbf{m}$ is the vector of model parameters, and $E[J(\mathbf{x}, \mathbf{m})]$ is the expectation of the objective function J (e.g., expected NPV) at control $\mathbf{x}$ over uncertain model parameters $\mathbf{m}$.

As mentioned previously, an appropriate approach to account for uncertainty is to use an ensemble of reservoir models that have been sampled from the probability distribution for model parameters. The expected value can then be estimated by averaging all objective-function values over the ensemble. However, this approach requires many expensive simulations and will result in the cost of RO to increase linearly with the ensemble size and with the number of iterations required for optimization. To deal efficiently with uncertainty, we developed bias-correction methods applied to the mean reservoir model to estimate the expected value over an ensemble of model realizations. The use of bias-corrected mean model allows for RO performing with an approximation of expected value close to the ensemble average value, while requiring only the simulations of the mean reservoir model during the optimization process. We will elaborate on the bias-correction methodology later in Chapter 4.

2.2   Robust sequential decision-making

Although geological uncertainty is unavoidable and results in large uncertainty in reservoir performance, the uncertainty can be reduced based on the past observations through history matching or data assimilation. In field development optimization, wells are usually drilled and completed sequentially in time rather than simultaneously. After drilling a new well, we can update the uncertainty in reservoir properties by assimilating the observations that are collected from all drilled wells, including the new well. In traditional closed-loop field development (CLFD) [88], which is similar to CLRM optimization, the next production strategy is obtained by performing a re-optimization over the current assessment of uncertainty. In other words, the optimal decision is typically determined by maximizing the expected NPV over the current uncertainty state. This approach adds new observations to reduce geological uncertainty before making the next decision and thereby potentially improves the production strategy performance.

Figure 2.1 shows the process flow of CLFD, which involves four steps, as follows: 1) collecting observations after drilling a new well, 2) updating the geological description with all acquired information, 3) performing optimization in the updated reservoir model with re-estimated geological uncertainty, and 4) drilling the optimal next well based on the current uncertainty state. This procedure is repeated until all available wells have been drilled or until drilling additional wells will not increase the expected NPV over the reservoir life. In the optimization step, the control variables may include the number, type, location, control, and drilling schedule of wells. In this section, we will use the drilling-sequence optimization problem as an example to demonstrate robust sequential decision-making under uncertainty.
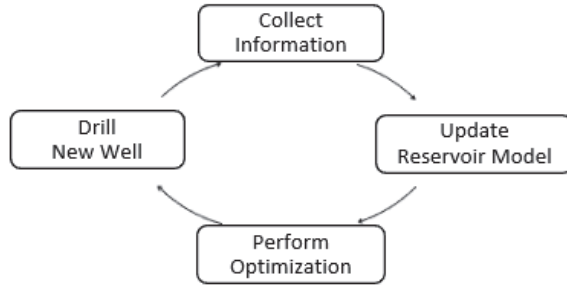
Figure 2.1: Schematic of closed-loop field development [88]

.

Suppose that we need to determine the optimal drilling schedule of N_w wells. Each drilled well provides observations that can be used to re-estimate uncertainty before choosing the next well to be drilled. Figure 2.2 shows the observation-based dynamic sequential decision process with N_w drilling wells. The set of actions $a_1, a_2, \ldots, a_{N_w}$ represents an ordered sequence of N_w wells drilled at times $t_0, t_1, \ldots, t_{N_w-1}$. The observations $o_1, o_2, \ldots, o_{N_w}$ denote the data collected from the drilling of each well, where o_j obtained from each past action a_j might be a single datum (e.g., types of facies), a collection of measured values of porosity/permeability, or a sequence of time-dependent data (e.g., production/injection data of various types over a time interval). The state $s_0, s_1, \ldots, s_{N_w}$ represents the specific environments at a particular time for each decision step constrained to the past decisions. We assume that observations from all drilled wells are immediately available for updating the reservoir model. Hence, uncertainty in each environmental state s_j is re-estimated based on $o_1, o_2, \ldots, o_j$ from all previously j drilled wells, before optimizing the next decision. The objective is to maximize the expected NPV, which is the sum of rewards $R_1, R_2, \cdots, R_{N_w}$ over the time periods $\Delta t_1, \Delta t_2, \cdots, \Delta t_{N_w}$.
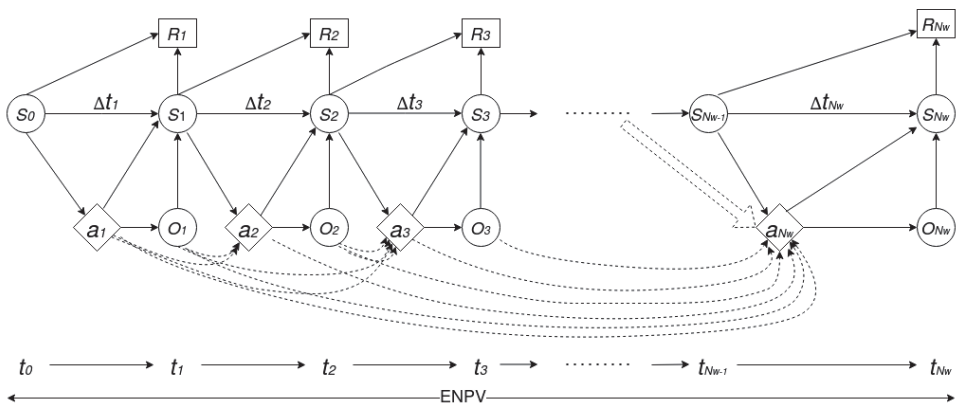
Figure 2.2: Example of an observation-based dynamic drilling-sequence planning with N_w wells

As shown in Fig. 2.1, the optimal action at each step in the traditional CLRM is determined based on current geological knowledge, i.e., performing a re-optimization in the reservoir model updated using all currently available past observations. In that

case, after completing the drilling of j wells, the optimal next well is the one that will achieve the maximum expected NPV over the current assessment of uncertainty u_j based on the observations in history h_j, i.e.,

$$a_{j+1}^* = \underset{a_{j+1} \in A_{j+1}(h_j)}{\operatorname{argmax}} \ \mathrm{EV}^* \left(h_j, a_{j+1}, u_{h_j} \right), \qquad (2.4)$$

where h_j is an observable history consisting of a sequence of past actions (i.e., j drilled wells) and observation pairs, $h_j = (a_1, o_1, \ldots, a_j, o_j)$; $A_{j+1}(h_j)$ is the current action space at a given history h_j, which consists of the $(N_w - j)$ remaining wells; u_j is the current uncertainty state evaluated based on the past observations $o_1, o_2, \ldots, o_j$ from j drilled wells in history h_j; $\mathrm{EV}^* (h_j, a_{j+1}, u_j)$ is the maximum achievable expected NPV for complete drilling sequences over the uncertainty state u_j constrained to history h_j, followed by taking a_{j+1} as the next decision. Note that here a_{j+1} is the current decision alternative instead of the previous decision. Thus, $\mathrm{EV}^* (h_j, a_{j+1}, u_j)$ is evaluated over the current uncertainty state u_j instead of u_{j+1}. To compute the optimal next decision a_{j+1}^*, we generally need to evaluate the entire optimal sequence because the expected NPV over the reservoir life-cycle is related to complete sequences with all available actions taken sequentially. In this work, we developed a general methodology for efficiently computing a_{j+1}^* without finding the entire optimal sequence. During the optimization process, the maximum expected value EV^* constrained to the past decisions can be accurately estimated by using the information obtained from both previous and future actions. The methodology will be elaborated later in Chapter 2.

In an observation-based sequential decision-making process under uncertainty (Fig. 2.2), the past decisions not only directly affect the maximum achievable expected NPV but also influence the future observations that can be used to reduce uncertainty. In other words, the decision we make at the current time will affect the possibilities of future learning through actions (i.e., the opportunities to improve the optimal strategy resulting from future uncertainty reduction). Therefore, the true optimal action for each decision step depends on both the past decisions and the consequences on future uncertainty reduction. To act optimally, we must also account for the effects of future information before committing to a decision. However, this future learning possibility is ignored in most applications of CLRM optimization because it would be computationally intractable to update the reservoir model and re-optimize to account for all possible outcomes of future observations. Consequently, a_{j+1}^* that is obtained by solving Eq.2.4 may lead to a suboptimal solution to the final uncertainty state updated based on all observations from the complete strategy.

In theory, the optimal decision accounting for the future learning possibilities through all remaining actions could be computed using the standard decision analytic approach [36]. In this approach, after sequentially drilling j wells, the optimal choice for the next well is based on the expected values over all possible observations from all remaining wells,

$$a_{j+1}^{*^{\mathrm{fs}}} = \underset{a_{j+1} \in A_{j+1}(h_j)}{\operatorname{argmax}} \sum_{o \in O_{a_{j+1}}} p(o|h_j, a_{j+1}) Q_{N_w-(j+1)}^*(h_{j+1}), \qquad (2.5)$$

where $O_{a_{j+1}}$ is the observation space obtained from a_{j+1}; $p(o|h_j, a_{j+1})$ are the marginal probabilities of distinct observations o from a_{j+1} following history h_j; and $Q_{N_w-(j+1)}^*(h_{j+1})$ is the optimal expected value considering all possible future observations constrained to history h_{j+1} including the observations from a_{j+1}. This optimal

expected value is calculated in a backward induction procedure,

$$Q^*_{N_w-(j+k)}(h_{j+k}) = \max_{\substack{a_{j+k+1}\in A_{j+k+1}(h_{j+k}) \\ o\in O_{a_{j+k+1}}}} \sum p(o|h_{j+k}, a_{j+k+1})Q^*_{N_w-(j+k+1)}(h_{j+k+1}),$$

$$\text{for} \quad k = 1, 2, \ldots, N_w - (j+1),$$

$$(2.6)$$

where $Q^*_0(h_{N_w})$ is the expected NPV over the final uncertainty state updated using all sequential observations from a complete drilling sequence in history h_{N_w} (i.e., all N_w wells have been drilled sequentially). This optimization procedure is also known as the standard VOI decision analysis process with extensive form [100, 78], which is a fully structured decision tree that considers all possible combinations of the sequences of remaining actions with distinct observations [50].
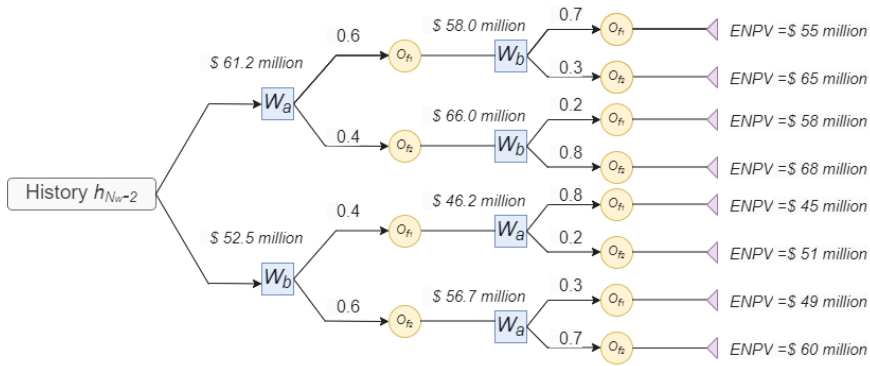
Figure 2.3: A fully structured decision tree for determining the order of two remaining wells in consideration of all possible future observations [104]

Figure 2.3 shows a simple illustration of determining the optimal next well from two remaining wells W_a, W_b through the backward induction procedure (Eqs. 2.5-2.6). We assume that N_w -2 wells have been drilled sequentially resulting in history h_{N_w-2}, and each remaining well can provide two possible distinct observations o_{f_1}, o_{f_2} about the type of facies. In this case, determining the optimal choice of next well while considering the effect of all possible future observations, requires information about the expected NPV from all 8 possible combinations of sequences with distinct observations and the information about the marginal probabilities of all possible observations from two remaining wells. Because the number of options is small, the optimal action $a^{*\text{fs}}_{j+1}$ (i.e., to drill W_a as the next well) can be easily determined using the backward induction. However, as this decision procedure requires the consideration of all possible combinations of sequences with distinct observations, the size of the decision tree is exponential in the number of distinct states related to both the action space and the observation space obtained from each action. If there are 8 possible remaining wells and each well provides only two distinct observations, then there would be $8! \times 2^8 \approx 1 \times 10^7$ possible combinations of drilling sequences with distinct sequential observations. The use of such a fully structured decision tree will be computationally intractable even before accounting for the cost of updating the reservoir model. Consequently, in

practice, it is only applicable to problems with small numbers of distinct actions and distinct observations.

Equation 2.5 can also be expressed using Bellman equation [10] and solved within the optimization framework of POMDP [108, 91, 55]. For small problems, a number of algorithms (e.g., enumeration, incremental pruning, linear support) have been proposed to solve POMDP exactly [24, 92, 16]. For large problems, POMDP can be solved approximately using approaches such as point-based value iteration [72], heuristic search value iteration [90], and Monte-Carlo sampling [96]. However, because the evaluation of an objective function in reservoir applications require expensive simulation and the number of the states that need to be evaluated in a POMDP usually is large, the cost of solving POMDP can be prohibitive, especially in cases with many combinations that are likely to generate high expected values. Therefore, computing an optimal decision while considering the effects of all possible future information (Eq. 2.5) is generally intractable for reservoir simulation-based problems. In this work, we developed a key-feature-based VOI analysis obtained by identifying key actions that will provide important information for improving future decisions and key information that will be most helpful in exploring key reservoir features of the optimization problem. This approach does not require an exhaustive history matching and optimization procedure accounting for all possible outcomes of future observations. It is therefore practical to obtain an optimal or near-optimal decision efficiently while considering the possibilities of future learning. We will present this methodology in detail in Chapter 5.

Chapter 3

Sequential planning with online learned heuristics

Field development planning with discrete actions is a sequential decision-making task that involves choosing an entire sequence of actions to maximize the reservoir's profitability. This optimization problem can be cast as a search problem in a directed graph, i.e., finding the path with the maximum reward from a start node (beginning of simulation) to a goal node (end of simulation). In Chapter 3, we present an efficient search method derived from learned heuristics for solving optimization problems with a discrete set of actions (i.e., sequential decisions-making problems), which allows for optimizing either a complete strategy or only the first few actions by limiting the search depth. During the optimization process, an accurate approximation of the maximum achievable value constrained to past decisions can be inexpensively obtained using only the information of previous and future actions, thereby guiding the search toward an optimal or near-optimal solution efficiently and effectively. Section 3.1 introduces one of the most important heuristic search algorithms, i.e., A*, which is regarded as the gold standard for search methods and has been widely applied to solve problems in various areas. Section 3.2 then describes how to make the heuristic search method efficient and practical for reservoir applications using online-learning techniques. Finally, Section 3.3 shows the space reduction and restoration techniques for finding an approximate solution quickly for large problems.

3.1 Heuristic search

Search algorithms can naturally be divided into informed search (also called heuristic search) and uninformed search (also called blind search). Informed search approaches (e.g., best-first search, greedy best-first search, A* search) use additional knowledge about the problem to provide hints about the best solution. By contrast, uninformed search approaches (e.g., breadth-first search, uniform-cost search, depth-first search) search blindly toward the goal without using any domain specific knowledge. Hence, informed search is generally more efficient and acceptable than uninformed search [82, 68].

The best-first search method is a generic informed search strategy that uses an evaluation function $f(n_s)$ with problem-specific knowledge to describe the desirability of expanding a particular state node n_s. The basic idea of the best-first search is to repeatedly expand the most promising node until a goal node is reached. For problems

that aim to find a high-reward path, the most promising search direction is determined by the node with the highest evaluation-function value obtained thus far. Two important variants of best-first search are greedy best-first search [29] and A* search [47]. The difference between these two search algorithms is that the greedy best-first search (also called pure heuristic search) completely ignores the reward to get to the current state node when evaluating the promise of a node, while A* search does not. Although A* search uses more memory than greedy best-first search, it finds the optimal solution with optimal efficiency. Because of its effectiveness, optimality and completeness, the A* algorithm is currently the most popular heuristic search method and many of the current search algorithms are variants of A* search [73, 74, 62, 93, 59, 20]. These algorithms have been applied to solve problems in various fields such as routing of telephone traffic, the alignment of DNA sequences, navigation through a maze, and real-time path planning in robotics and digital games. In the following, we give a general description of A* search and describe its search progress and properties with a simple example of the drilling-sequence planning problem.

3.1.1   A* algorithm

The evaluation function $f(n_s)$ in A* search consists of two elements,

$$f(n_s) = g(n_s) + h(n_s), \tag{3.1}$$

where $g(n_s)$ is the true reward from a start node to node n_s, depending only on the past decisions, $h(n_s)$ is the estimated reward from node n_s to a goal node. This estimated value is obtained using a heuristic function h that incorporates additional knowledge about the problem. When solving maximization problems over a set of actions, $f(n_s)$ is an estimate of $f^*(n_s)$ that is the maximum objective-function value of complete paths constrained to go through the state node n_s, and the search terminates when a complete strategy (i.e., all actions have been performed sequentially) for which the objective-function value is larger than the evaluation-function values of all visited nodes is found.

Suppose that we need to maximize the NPV of a reservoir model by optimizing the drilling schedule of wells (i.e., all wells have to be drilled sequentially). Figure 3.1 shows the procedure for using A* search to determine the optimal drilling sequence of four wells ($N_w = 4$). The sequence consists of 13 visited nodes, including the start node S_0, which indicates the beginning of the simulation at time t_0, and four expanded nodes, A, C, F, K, which were selected as the most promising search directions based on the evaluation-function value. Each node represents a specific environment state associated with past decisions at a particular time. The reward g_i along a partial path is the true contribution to NPV over a certain time period from a given sequence of N_s selected wells. The heuristic value h is an estimated maximum future reward resulting from previously drilled wells and all remaining wells over the remaining time period. Then, the evaluation function value $f = \sum_{i=1}^{N_w} g_i + h$ is the estimated maximum achievable NPV of complete drilling sequences constrained to previously drilled wells.

As illustrated in Fig. 3.1, four successor nodes A, B, C, D are generated from the start node S_0, each corresponding to a particular search direction starting with a specific well. At the first decision stage, the path through node A is the most promising direction because it has the highest evaluation-function value $f(A)$, which indicates that
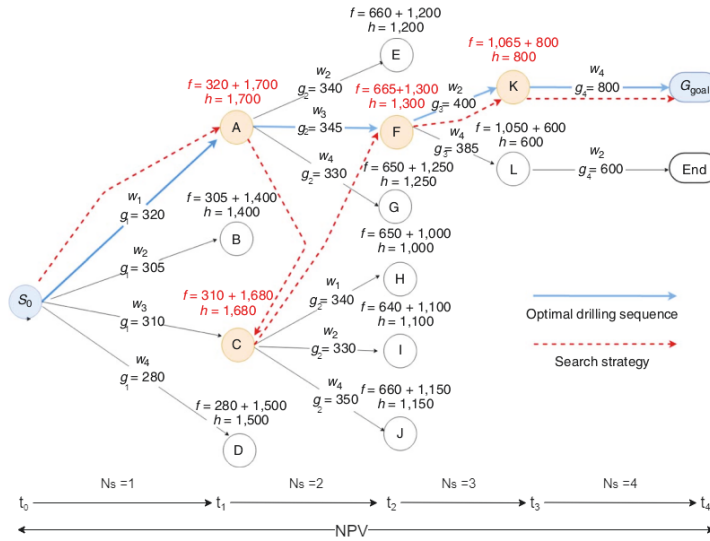
Figure 3.1: Search progress in A* algorithm for optimizing the drilling sequence of four wells. Nodes are labeled with their f-values and h-values, and paths are labeled with their g-values [102]

a drilling sequence starting with W_1 is more likely to generate higher economic value. Consequently, node A is extended, and its successor nodes E, F, G are evaluated for choosing the next search direction. However, the evaluation-function values obtained from all these three successor nodes are lower than $f(C)$. Hence, the search moves to the path through node C that is selected as the most promising node at the second decision stage. After evaluating all the successor nodes H, I, J, the search returns to node F with the currently highest f-value. In this case, the search terminates with an optimal drilling sequence $[W_1 \rightarrow W_3 \rightarrow W_2 \rightarrow W_4]$ along the full path $[S_0 \rightarrow A \rightarrow F \rightarrow K]$ for which the NPV is higher than the estimated maximum NPV of all evaluated paths. Note that the f-values through nodes K and L with complete drilling sequences are the actual NPVs rather than estimates. If any visited node has a higher evaluation-function value, the search will continue to explore other directions until the stopping criterion is reached.

3.1.2   Behavior of A*

A deeper search generally results in a more accurate evaluation and A* algorithm enables the search direction to be adjusted in consideration of all investigated paths; hence, this approach can guide the search toward an optimal solution effectively. For a sequential decision-making problem with a set of N_a actions (i.e., $N_a!$ possible action sequences), it may be necessary to evaluate only $\frac{N_a \times (N_a+1)}{2}$ nodes to find an optimal complete strategy using A* search with a strong heuristic function. If the heuristic function is admissible, namely, $h(n)$ never underestimates the true maximum reward $h^*(n)$ from the current state node n to a goal state, A* search is guaranteed to find the true optimal solution.

Suppose the true optimal solution has the maximum objective-function value J^*; then, all nodes n^* along the true optimal path would have evaluation-function value

$f(n^*) = g(n^*) + h(n^*) \geq J^*$ based on an admissible heuristic function $h(n^*) \geq h^*(n^*)$. During the search process, there must exist at least one visited node n^*, so that a sub-optimal solution G does not terminate the search since $f(G) = J(G) < J^* < f(n^*)$. $f(n^*)$ will guide the search toward the true optimal solution. However, an admissible heuristic function that results in many nodes along suboptimal solutions with $f(n) > J^*$ has a severe drawback: the search does not terminate immediately when the true optimal solution is obtained, and it will continue to expand all visited nodes with $f(n) > J^*$. In the worst-case scenario, this approach might lead to an exhaustive search (i.e., evaluating all $N_a! \times N_a$ possibles state nodes), which is worse than using the most naïve method to compute the true optimal solution (i.e., evaluating N_a all possibles sequences). Compared to the use of a weak admissible heuristic function, it is preferable to compute an optimal or near-optimal solution quickly using an effective inadmissible heuristic function.

As discussed above, the important component of the evaluation function is the heuristic function, which is the main driving force of heuristic search. Heuristic functions incorporate problem-specific knowledge and are usually different for various problem domains. There are several ways to derive heuristic functions such as through relaxed problems with fewer restrictions on actions [106], abstract problems [76], or subproblems [75]. For the drilling-order problem, one possible heuristic function is obtained by drilling all remaining wells simultaneously at the next step to estimate the maximum NPV of complete drilling sequences constrained to past decisions [63]. Such a heuristic can be obtained inexpensively but might lead to an exhaustive search due to large estimated values in some cases. In many applications, it is more important to find an approximate solution quickly because of time limitations. A weighted evaluation function $f(n) = g(n) + \omega h(n)$ (e.g., weighted A* [73], dynamic weighted A* [74, 33]) that has a weighting parameter ω on the heuristic function $h(n)$ often performs well in rapidly finding solutions by giving up the guarantee of optimality. However, an appropriate ω is difficult to find in terms of both running time and solution quality, and an appropriate value might not exist for some applications or be extremely challenging to obtain without resorting to trial and error. Strong and admissible heuristic functions are usually learned from experience within the problem domain and empirical studies are usually needed. For reservoir optimization problems, the cost of finding an accurate heuristic function can be prohibitive because the evaluation of heuristic functions requires simulations. To address this issue, we developed online-learning techniques for heuristic search , i.e., learned heuristic search, that considers learning the initial heuristic during the search process to improve the accuracy of the evaluation function.

3.2   Online learning techniques

To efficiently find a solution that is optimal or near optimal, the evaluation function $f(n_s)$ should be close to the true maximum reward $f^*(n_s)$, such that the search would not frequently change direction due to inaccurately estimated values. In this section, we describe how to efficiently obtain an "adequate" evaluation function when it is difficult to design a heuristic function $h(n_s)$ that is accurate for all state nodes. The main idea is to improve a crude heuristic using online-learning techniques that aim to adjust the initial approximation $f(n_s)$ toward $f^*(n_s)$ by learning the observations obtained from previous decision steps.

One possible way to improve the evaluation function is to use the set of estimated additive errors along all remaining steps, i.e.,

$$\hat{f}(n_s) = f(n_s) + \sum_{n=n_s}^{n_{\text{goal}}} \hat{\epsilon}_{fn},$$ (3.2)

where $\hat{f}(n_s)$ and $f(n_s)$ are the improved and the initial evaluation-function values through node n_s, respectively, and $\hat{\epsilon}_{fn}$ is an estimate of the error across a single step, which can be obtained by observing the initial evaluation-function values $f(S_0)$, $f(n_1)$, $f(n_2)$, ..., $f(n_s)$ from the expanded nodes along the current optimal path.

Thayer et al. [95] presented a technique with mean single-step error $\bar{\epsilon}_{fn_s}$ to compute the improved $\hat{f}(n_s)$ in Eq. 3.2,

$$\hat{f}(n_s) = f(n_s) + d(n_s)\bar{\epsilon}_{fn_s},$$ (3.3)

where $d(n_s)$ is the number of remaining actions at node n_s, and the mean additive error $\bar{\epsilon}_{fn_s}$ can be calculated using only the information about the number of selected actions and the initial evaluation-function values at the start node S_0 and the current state node n_s, i.e.,

$$\bar{\epsilon}_f(n_s) = \frac{f(n_s) - f(S_0)}{d_{\text{goal}} - d(n_s)},$$ (3.4)

where d_{goal} denotes the number of all possible actions at the start node S_0. Compared to the use of the initial $f(n_s)$, the corrected evaluation-function value $\hat{f}(n_s)$ (Eq. 3.3), which is dependent on the temporal difference learning, often provides better guidance for search [95].

For reservoir optimization problems, the absolute difference in the initial estimated maximum NPV at various decision steps is often highly variable, and the additive single-step error ϵ_{fn} decreases with the addition of more observations when searching deeper. Consequently, instead of using the additive error, we developed online-learning techniques associated with the ratio between two neighboring nodes to improve $f(n_s)$. Based on the learning period, we divide the techniques into two classes: single-step adjustment and multiple-time-period learning. The main difference between them is that the former directly adjusts $f(n_s)$ over the entire production period, while the latter simultaneously adjusts the contributions to $f(n_s)$ over different time periods.

3.2.1 Single-step adjustment

The improved $\hat{f}(n_s)$ corrected by using an estimated single-step ratio of the evaluation function f across all remaining actions can be described as

$$\hat{f}(n_s) = f(n_s) \prod_{n=n_s}^{n_{\text{goal}}} \hat{\gamma}_{fn},$$ (3.5)

where $\hat{\gamma}_{fn}$ is the estimate of the single-step ratio of $f(n)$ along the current optimal path to the goal node n_{goal}.

There are two possible ways to obtain $\hat{\gamma}_{fn}$, depending on the trend of the observed ratio γ_{fn_1}, γ_{fn_2}, ..., γ_{fn_s} from past decisions. If the observed ratio is approximately

constant, the improved $\hat{f}(n_s)$ can be computed with an estimate of the mean single-step ratio $\bar{\gamma}_{fn_s}$, i.e.,

$$\hat{f}(n_s) = f(n_s)\bar{\gamma}_{fn_s}^{d(n_s)}, \tag{3.6}$$

where the mean single-step ratio $\bar{\gamma}_{fn_s}$ can be calculated from the observed initial evaluation-function values $f(S_0), f(n_1), f(n_2), \ldots, f(n_s)$ along the current optimal path using the following formula,

$$\bar{\gamma}_{fn_s} = \frac{\sum_{n_i=n_1}^{n_s} \gamma_{fn_i}}{d_{\text{goal}} - d(n_s)} = \frac{\sum_{n_i=n_1}^{n_s} \frac{f(n_i)}{f(n_{i-1})}}{d_{\text{goal}} - d(n_s)}. \tag{3.7}$$

If the observed ratio γ_{fn} through $S_0, n_1, n_2, \ldots, n_s$ is not stationary, instead of using a fixed single-step ratio $\bar{\gamma}_{fn_s}$ to improve $f(n_s)$, we use a dynamic γ_{fn} that considers the trend of γ_{fn} for all remaining steps. Let $\bar{\mu}_{ns}$ be the trend indicator of γ_{fn} that is evaluated from the observed ratios $\gamma_{fn_1}, \gamma_{fn_2}, \ldots, \gamma_{fn_s}$ along the current optimal path, i.e.,

$$\bar{\mu}_{n_s} = \frac{\sum_{n_i=n_1}^{n_s} \mu_{n_i}}{d_{\text{goal}} - d(n_s)} = \frac{\sum_{n_i=n_1}^{n_s} \frac{\gamma_{fn_i}}{\gamma_{fn_{i-1}}}}{d_{\text{goal}} - d(n_s)}. \tag{3.8}$$

The estimated $\hat{\gamma}_{fn}$ for each remaining step can then be represented as,

$$\hat{\gamma}_{fn} = \gamma_{fn_s}\bar{\mu}_{n_s}^{d(n_s)-d(n)}. \tag{3.9}$$

In consideration of the trend of observed γ_{fn} (Eq. 3.8), the improved evaluation-function value $\hat{f}(n_s)$ (Eq. 3.5) with a dynamic estimated $\hat{\gamma}_{fn}$ (Eq. 3.9) can be formulated as

$$\hat{f}(n_s) = f(n_s) \prod_{n=n_s}^{n_{\text{goal}}} \hat{\gamma}_{fn} = f(n_s) \prod_{i=1}^{d(n_s)} \gamma_{fn_s}\bar{\mu}_{n_s}^i = f(n_s)\gamma_{fn_s}^{d(n_s)}\bar{\mu}_{n_s}^{\frac{(d(n_s)+1)d(n_s)}{2}}. \tag{3.10}$$

3.2.2   Multiple-time-period learning

Despite abandoning the guarantee of optimality, we strive to improve the accuracy of $\hat{f}(n_s)$ and efficiently compute a solution that is optimal or nearly optimal. A single learning technique to minimize the error of $\hat{f}(n_s)$ at all levels is difficult to design, especially at the early decision stages with limited past observations. The techniques presented in the single-step adjustment (Eqs. 3.6 and 3.10) are derived from the estimated ratio $\hat{\gamma}_{fn}$ associated with the entire production period, which means that all contributions to $f(n_s)$ at node n_s over different time periods are adjusted with a fixed $\hat{\gamma}_{fn}$. However, the rewards corresponding to various time periods might have different trends. To reduce the instability of the observations of $\gamma_{\Delta t}$, we can improve $f(n_s)$ through a set of $\hat{\gamma}_{\Delta t}$ obtained from different learning periods Δt (i.e., multiple-time-period learning). In that case, the improved evaluation-function value would be more likely to capture the characteristics of the changes in the initial estimates along the current optimal path.

Depending on whether the operating strategy is complete, the entire production period $T = \sum_{i=1}^{N_a} \Delta t_i$ through a set of N_a discrete actions can simply be divided into two major learning periods $T = \Delta t_1^L + \Delta t_2^L$, where Δt_1^L is the time period $\sum_{i=1}^{N_a-1} \Delta t_i$ during which all N_a actions are supposed to be performed sequentially, and Δt_2^L corresponds to the last time period Δt_{N_a} from the time the last action has been taken to the end of simulations. In general, past decisions have a considerable impact on the economic

value of the near future. To potentially reduce the impact on the observations of the nodes for learning, we could use the first N_a-1 time periods with a short-term period Δt_d as the first learning period, i.e., $\Delta t_1^L = \sum_{i=1}^{N_a-1} \Delta t_i + \Delta t_d$, where Δt_d is the time period when all actions have already been taken, while the second learning period $\Delta t_2^L = \Delta t_{N_a} - \Delta t_d$ begins at the time when the last action would be performed for time Δt_d.

Given these two periods, the initial evaluation function $f(n_s)$ constrained to pass node n_s can be expressed as

$$f(n_s) = h_{\Delta t_1^L}(n_s) + h_{\Delta t_2^L}(n_s). \tag{3.11}$$

where $h_{\Delta t_1^L}(n_s)$ is the estimated reward over the time period $\Delta t_1^L = \sum_{i=1}^{N_a-1} \Delta t_i + \Delta t_d$ and $h_{\Delta t_2^L}(n)$ is the estimated reward corresponding to the time period $\Delta t_2^L = \Delta t_{N_a} - \Delta t_d$.

The second period Δt_2^L can be further divided into multiple shorter periods for learning, i.e., $\Delta t_2^L = \Delta t_{N_a} - \Delta t_d = (\Delta t_2^L, \Delta t_3^L, \Delta t_4^L, \dots, \Delta t_{N_L}^L)$. We can then describe $f(n_s)$ as

$$f(n_s) = h_{\Delta t_1^L}(n_s) + \sum_{i=2}^{N_L} h_{\Delta t_i^L}(n_s), \tag{3.12}$$

where $h_{\Delta t_i^L}(n_s)$ is the estimated reward over each learning Δt_i^L, which can be improved by observing the initial estimated values corresponding to the same time period along the current optimal path.

Based on the online-learning techniques presented in Eqs. 3.6-3.10, the improved evaluation-function value obtained via multiple-time-period learning can be computed by simultaneously adjusting the heuristic values to various time periods,

$$\hat{f}(n_s) = \hat{h}_{\Delta t_1^L}(n_s) + \sum_{i=2}^{N_L} \hat{h}_{\Delta t_i^L}(n_s) = h_{\Delta t_1^L}(n_s) \prod_{n=n_s}^{n_{\text{goal}}} \hat{\gamma}_{h_{\Delta t_1^L}}(n) + \sum_{i=2}^{N_L} \left(h_{\Delta t_i^L}(n_s) \prod_{n=n_s}^{n_{\text{goal}}} \hat{\gamma}_{h_{\Delta t_i^L}}(n) \right). \tag{3.13}$$

Note that the estimated $\hat{\gamma}_h$ in the above equation is obtained from the estimated rewards over the corresponding learning period, differing from the estimated $\hat{\gamma}_{fn}$ of the cumulative reward over the entire period.

3.2.3 Multilearned heuristics

In this work, the techniques designed for improving the initial evaluation function are based on the ratio of the rewards between two neighboring nodes. If more online-learning mechanisms are available, a useful evaluation-function value is more likely to be obtained. Suppose that there is a set of online-learning mechanisms, $\Phi_1, \Phi_2, \Phi_3, \dots, \Phi_L$. Each can be used independently to improve the initial evaluation function $f(n)$. The best-improved evaluation function $\hat{f}_\Phi(n)$ after multiple online-learning techniques can be defined as

$$\hat{f}_\Phi(n) = \max\left(\hat{f}_{\Phi_1}(n), \hat{f}_{\Phi_2}(n), \hat{f}_{\Phi_3}(n), \dots, \hat{f}_{\Phi_L}(n) \right). \tag{3.14}$$

Here we choose the maximum estimated value as the best-improved $\hat{f}_\Phi(n)$ because it is more likely to overestimate the true maximum reward $f^*(n)$ and guide the search close to the optimal solution, although this approximation might not be the one with the best accuracy.

Performing heuristic search with multiple online-learning techniques (i.e., multi-learned heuristic search) is similar to the idea of simultaneously using multiple heuristics to search [17, 49, 80, 84, 2], but the use of online-learning techniques is much more efficient for simulation-based optimization problems. Although it might be possible to generate an "adequate" evaluation function in consideration of multiple heuristics, the cost of optimization increases linearly with the number of heuristics because the evaluation of each heuristic requires simulation runs. Moreover, due to the complexity of optimization problems and the expensive cost of designing an accurate initial heuristic, most of the potential initial heuristics $f(n)$ are probably not "helpful". In contrast, online-learning techniques, whose designs are based on available information, can be used to create a set of latent heuristics, for which only one initial $f(n)$ requires simulation runs while all the information-available techniques do not. Because the error in the approximation is reduced after each single online-learning technique, the potential heuristics $\hat{f}_{\Phi_1}(n), \hat{f}_{\Phi_2}(n), \ldots, \hat{f}_{\Phi_k}(n)$ are generally more accurate than the multiple initial heuristics obtained without any learning. Hence, multilearned heuristic search (MLHS) is preferred over general multiheuristic search for reservoir applications.

3.3   Space reduction and restoration

As shown in Fig.3.1, all successor nodes from each selected direction are evaluated in the general heuristic search method, i.e., all the remaining actions are considered as the next possible action. At each decision step, all visited nodes are considered for choosing the next search direction, i.e., the search allows for changing direction among all visited paths). This search method can be used to obtain an optimal solution considering the possibility of achieving high values from the sequences starting with various past decisions, but the number of remaining actions limits optimization efficiency. In some cases, there may be a large number of strategies with an objective-function value very close to the true maximum, resulting in frequent changes in the search direction. To address these issues, we can accelerate the search progress using space reduction techniques, which allow computation of a valid solution quickly.

3.3.1   Space reduction

The basic idea of space reduction is to prune unimportant nodes that are likely to generate low values. Ranking the node importance has theoretical and practical significance in decision making [70, 13] and has been noted previously in the heuristic search method [95, 105]. Depending on whether the evaluation of a heuristic is required for pruning the nodes, search space can be reduced in two ways: prior space reduction and posterior space reduction.

Prior space reduction

Prior space reduction is achieved by avoiding the evaluation of all successor nodes, which means that only some of the remaining actions are evaluated at the next possible action. An appropriate heuristic might be able to provide useful information to identify the decision alternatives that are more likely to generate high objective-function values without requiring additional simulations. For example, using a heuristic in which all remaining wells are drilled simultaneously at the next step, we can also obtain all the

information needed to compute the production well economic indicator (PWEI) and the injection well economic indicator (IWEI) of individual wells [37],

$$PWEI = \sum_{k=1}^{T} \frac{\left(q_{o,k}^p \cdot r_o - q_{w,k}^p \cdot r_w\right) \cdot \Delta t_k}{(1+b)^{t_k/365}} - \frac{W_n^p}{(1+b)^{t_n^p/365}}, \tag{3.15}$$

$$IWEI = \sum_{j=k}^{T} \frac{q_{wi,k}^i \cdot r_{wi} \cdot \Delta t_k}{(1+b)^{t_k/365}} + \frac{W_n^i}{(1+b)^{t_n^i/365}}, \tag{3.16}$$

where $q_{o,k}^p$, $q_{w,k}^p$ and $q_{wi,k}^i$ are the rates of oil production, water production and water injection for an individual well. W_n^p and W_n^i are the costs of drilling a producer and an injector, respectively. Note that uncertainty in reservoir properties is neglected in Eqs. 3.15-3.16. Taking into account geological uncertainty, PWEI and IWEI are the expected values over uncertainty.

In general, the wells with higher PWEI and IWEI starting operating in the early stages are likely to achieve more profitability of a field. Producers and injectors therefore can be approximately ranked based on their contribution to the reservoir's profitability. The economic indicator is representative of an individual well, however, and might not accurately rank the possibility of achieving high NPV from complete drilling sequences. When such an economic indicator is used to prune the search space, a better solution might be ignored after prior space reduction. To reduce the risk of solution quality loss, the search space should not be pruned excessively, e.g., at least half of the remaining actions at each decision step should be considered as the next possible actions.

Posterior space reduction

In contrast to prior space reduction that reduces the number of successor nodes before evaluation, posterior space reduction limits the number of expanded nodes based on evaluation-function values, i.e., only the successor nodes with high estimated values are considered during the search process, but those with low estimated values are pruned. This approach is an efficient way to control frequent changes in search direction. Because the search space is pruned by ranking the node based on the evaluation-function value, which is more accurate than the economic indicator of individual actions for prior space reduction, posterior space reduction is less likely to miss the optimal solution or reduce the solution quality. Although posterior space reduction attempts to control the selected direction along the optimal sequential solution, it will not accelerate the search progress if the pruned nodes would not have been expanded during MLHS without space reduction. In contrast, prior space reduction is guaranteed to accelerate the search process, but it is more likely to lose the solution quality if the search space is reduced excessively.

3.3.2 Space restoration

To find a solution as quickly as possible, we can perform MLHS with both prior and posterior space reduction (MLHS-SR); however, this approach may reduce the solution quality if the space is pruned excessively. The solution could be improved by adding the removed space back gradually and utilizing MLHS iteratively to continue the search. Algorithm 1 presents pseudo-code of MLHS with space reduction and restoration

(MLHS-SRR), in which MLHS is executed twice. That is, MLHS with space reduction is first utilized to rapidly find a valid solution, and then MLHS is performed with space restoration to improve the solution.

Efficiency of MLHS-SRR

MLHS with gradual space restoration is similar to the anytime algorithm [46, 107, 45], but it outperforms that algorithm in terms of both running time and accuracy of the estimated value. For the anytime algorithm, a solution is found quickly by sacrificing the accuracy of the estimated value. Then, the heuristic search is executed iteratively by improving the accuracy gradually. In contrast, MLHS-SRR is executed without harming the accuracy of the estimated value, since a solution is rapidly found by pruning the paths that are likely to generate low values rather than reducing the accuracy of the estimated value. Although MLHS is performed at least twice in this approach (space reduction and restoration), it is not necessary to evaluate all of the visited nodes for each search process. Most have already been evaluated and their evaluation-function values have been obtained during the first search process with space reduction. Thus, MLHS with space restoration requires simulations only for the nodes that were not previously evaluated. Moreover, since a valid solution is already obtained by MLHS-SR, performing MLHS-SRR is based on the true objective-function value of the current solution. In other words, only the nodes with evaluation-function values that are higher than that of the current solution would be considered during the search process. Hence, MLHS-SRR can find an approximate solution faster than MLHS. When the improved evaluation function overestimates the maximum objective-function value, MLHS-SRR is guaranteed to find the same solution as the MLHS without space reduction. The performance of various heuristic search methods, MLHS, MLHS-SR, and MLHS-SRR is studied in Paper A via the application of drilling-order problem in a synthetic field model.

Partial optimization

In addition, solving optimization problems with discrete actions with online learned heuristics has an advantage in that it allows for optimizing only the first few decisions at a reduced cost without finding the entire optimal solution. The reservoir model will almost certainly be updated based on the information from past decisions. The optimal strategy of the remaining actions obtained from the updated reservoir model generally differs from the initial optimal strategy. It is, therefore, more important to identify the first few actions rather than finding an entire optimal sequence for all remaining actions. The most straightforward approach to optimizing the first few actions is to terminate the search at a certain depth, namely depth-limited search [32]. For example, to compute an optimized sequence of the first N_s actions, we can terminate the search at the first expanded node with $N_s + 1$ selected actions. Note that we prefer to cut off the search with more selected actions because online learning techniques with the addition of more observations along a longer path can further improve the approximations and potentially generate a better solution. In addition, we can also perform depth-limited search iteratively to find the partial solution faster. The basic idea behind this approach is to iteratively optimize the next action determined by the first expanded node with more selected actions (e.g., two more actions) until a solution of $N_s + 1$ selected actions is

---

Algorithm 1 Multilearned heuristic search with space reduction and restoration (MLHS-SRR)

Input: Geological model, Initial heuristic function, A start node n_0
Output: Optimal or near-optimal control sequence

1: $G_{\text{goal}} \leftarrow 0$
2: $OPEN \leftarrow \{n_0\}$
3: $CLOSED \leftarrow \emptyset$
4: Reduction $\leftarrow$ True
5: **while** $OPEN \neq \emptyset$ or Reduction is False **do**
6:     $OPEN \leftarrow \{n_0\}$"
7:     $CLOSED \leftarrow \emptyset$
8:     **while** $OPEN \neq \emptyset$ **do**
9:         $n \leftarrow \text{argmax}_{n \in OPEN} \hat{f}_L(n)$
10:         $OPEN \leftarrow$ Remove node n with the highest $\hat{f}_L(n)$ from $OPEN$
11:         $CLOSED \leftarrow$ Add node n to $CLOSED$
12:         **if** Reduction is True **then**
13:             Economic indicator $\leftarrow$ Calculate the economic indicator for each successor node n' of n
14:             Prior space reduction $\leftarrow$ Prune the successor nodes with lower value of economic indicator
15:             $successor(n) \leftarrow$ Expand the remaining successor nodes n' of n after prior space reduction
16:         **else**
17:             $successor(n) \leftarrow$ Expand all remaining successor nodes n' of n
18:         **end if**
19:         $OPEN \leftarrow$ Add $successor(n)$ to $OPEN$
20:         **for each** $n' \in successor(n)$ **do**
21:             **if** all actions are taken sequentially at node n' **then**
22:                 $\hat{f}_L(n') \leftarrow g(n')$
23:                 **if** $\hat{f}_L(n') > G_{\text{goal}}$ **then**
24:                     $G_{\text{goal}} \leftarrow g(n')$
25:                 **end if**
26:             **else**
27:                 **if** n' has already been evaluated **then**
28:                     Obtain $\hat{f}_L(n')$ directly without running additional simulations
29:                 **else**
30:                     Evaluate the initial heuristic function value $h(n')$ at each time period
31:                     Evaluate the initial estimated maximum objective-function value $f(n')$ through node n'
32:                     Utilize multiple online learning mechanisms to improve $f(n)$ simultaneously
33:                     $\hat{f}_L(n') \leftarrow \max_{L_1 \leq L_i \leq L_n} \hat{f}_{L_i}(n) \leftarrow f(n') \leftarrow h(n') \leftarrow n'$
34:                 **end if**
35:             **end if**
36:             **if** Reduction is True **then**
37:                 Posterior space reduction $\leftarrow$ Prune the nodes n' with lower $\hat{f}_L(n')$
38:             **end if**
39:         **end for**
40:         Prune all nodes $n \in OPEN$ if $\hat{f}_L(n) < G_{\text{goal}}$
41:     **end while**
42:     **for each** node n along the control sequence of G_{goal} **do**
43:         **if** $\hat{f}_L(n) < G_{\text{goal}}$ **then**
44:             $\hat{f}_L(n) \leftarrow G_{\text{goal}}$
45:         **end if**
46:     **end for**
47:     **if** Reduction is True **then**
48:         Reduction $\leftarrow$ False
49:     **else**
50:         **Break**
51:     **end if**
52: **end while**
53: **return** a sequence of actions with objective-function value $\leftarrow G_{\text{goal}}$

---

found. Because only the paths extended from past decisions are considered, this method can avoid evaluating unnecessary paths along other directions due to underestimated values. Learned heuristics with accurately estimated values generally will not change direction frequently, so the optimized sequence of the first few actions obtained by limiting the search depth is likely to be near the final optimized complete sequence. In Paper B, we investigated the possibility of optimizing only the first few wells using learned heuristic search with limited search depth.

Chapter 4

Robust optimization using bias-corrected mean model

Chapter 3 presented an approach (i.e., learned heuristic search) to solving optimization problems with discrete-control variables for which the geologic model was not uncertain. In that case, it is possible to optimize an objective function whose evaluation requires a single model. As mentioned previously, geological uncertainty frequently results in large uncertainty in reservoir performance and ought to be taken into account in field development optimization. This chapter presents the bias-correction methodology to efficiently account for geological uncertainty in robust optimization. In Section 4.1, we introduce some possible ways to estimate the expectation of an objective function under uncertainty. Section 4.2 presents the bias-correction methods applied to the mean model to efficiently estimate the expected value. The bias-corrected mean model can be applied to fairly general problems of optimizing the expected value of an objective function for which the uncertainty is characterized by an ensemble of model realizations.

4.1   Estimation of expected value

A fast way to approximate the expected value of the objective function is to use the objective-function value based on the expected values of uncertain parameters, i.e., $J(\mathbf{x}, E[\mathbf{m}])$, which will formulate the robust optimization problem in a deterministic setting,

$$\mathbf{x}^* = \arg\max J(\mathbf{x}, E[\mathbf{m}]), \tag{4.1}$$

where $\mathbf{x}$ is a vector of control inputs and $\mathbf{m} \in \mathcal{R}^m$ is an m-dimensional vector of uncertain model parameters.

When nonlinearity in the objective function is not large, an approximation of the expected value can be obtained from the mean reservoir model $\bar{\mathbf{m}}$,

$$J(\mathbf{x}, E[\mathbf{m}]) \approx J(\mathbf{x}, \bar{\mathbf{m}}). \tag{4.2}$$

This approach requires only the simulation runs using the mean model $\bar{\mathbf{m}}$. However, $J(\mathbf{x}, \bar{\mathbf{m}})$ can be a poor estimate of the expectation of an objective function since the dependence of the objective function J (e.g., NPV) on the model parameters $\mathbf{m}$ is generally highly nonlinear,

$$J(\mathbf{x}, E[\mathbf{m}]) \neq E[J(\mathbf{x}, \mathbf{m})]. \tag{4.3}$$

If geological uncertainty has a large variance, the optimal solution obtained from the mean reservoir model $\bar{m}$ might not be near the actual optimal result. Hence, uncertainty should be treated in a better way that allows accurate estimation of the expected value. In this section, we describe some possible means of computing a good approximation of the expected value.

4.1.1   Sample average approximation

To account for the uncertainty in the RO procedure, an appropriate approach is to maximize the expectation of the objective function J over the uncertainty space, i.e.,

$$\mathbf{x}^* = \arg \max E[J(\mathbf{x}, \mathbf{m})]. \tag{4.4}$$

To quantify the uncertainty in reservoir properties, geostatistical methods can be used to generate an ensemble of equally probable realizations that describes the probability distribution for model parameters. Then, the expected value of the objective can be computed by averaging the objective-function values over all geological realizations [22, 23, 99]

$$E[J(\mathbf{x}, \mathbf{m})] \approx \bar{J}(\mathbf{x}) = \frac{1}{N_e} \sum_{j=1}^{N_e} J(\mathbf{x}, \mathbf{m_j}), \tag{4.5}$$

where $\mathbf{m_j}$ is the vector of uncertain model parameters (e.g., permeability, fault multipliers) for realization j and N_e is the number of individual realizations required to capture the uncertainty in the model parameters.

The average value $\bar{J}(\mathbf{x})$ over a sufficiently large ensemble [58, 56] can provide a good approximation of the expected objective-function value. Unfortunately, N_e simulations are required to compute this sample average approximation (SAA) since the control $\mathbf{x}$ has to be applied to every single realization $\mathbf{m_j}$ to obtain each objective-function value $J(\mathbf{x}, \mathbf{m_j})$. Due to the high computational cost of computing $\bar{J}(\mathbf{x})$ over the entire set of realizations, a subset of the reservoir model is often used to capture the overall uncertainty of the reservoir to improve the efficiency of RO. A small set of realizations, however, may not accurately characterize the uncertainty, leading to a suboptimal solution. Therefore, an appropriate subset of geological realizations must be carefully selected. Moreover, although the efficiency of RO can be significantly improved by reducing the full ensemble of models into a small subset ensemble, RO performed using the straightforward estimation of the expected value (Eq. 4.5) may still require many expensive simulations, since the cost of RO increases linearly with both the number of realizations employed and the number of iterations (i.e., the number of controls that need to be evaluated) during the optimization process. For example, if an optimization requires 500 iterations and 20 representative realizations are selected for use in RO, a total of 10^4 simulations will be required in optimization. Thus, a technique that can efficiently and effectively handle the uncertainty in robust optimization procedure is preferred.

4.1.2   Taylor series expansion

To generate a good approximation that does not require evaluation of controls applied to a large number of Monte Carlo samples, one approach is to modify the representation of

uncertainty using a Taylor series expansion of the objective function [26, 27, 11, 21, 3]. A linear approximation of the objective function is defined as

$$J^{\mathrm{lin}}(\mathbf{x}, \mathbf{m}) = J(\mathbf{x}, \bar{\mathbf{m}}) + \langle J_{\mathbf{m}}(\mathbf{x}, \bar{\mathbf{m}}), \mathbf{m} - \bar{\mathbf{m}} \rangle, \tag{4.6}$$

where $J_{\mathbf{m}}$ is the first derivative of objective function J with respect to uncertain model parameter $\mathbf{m}$.

Assuming that $\mathbf{m}$ is Gaussian distributed, the expectation over $\mathbf{m}$ of the linear approximation to the objective is simply the objective function of mean model $\bar{\mathbf{m}}$,

$$E[J^{\mathrm{lin}}(\mathbf{x}, \mathbf{m})] = J(\mathbf{x}, \bar{\mathbf{m}}). \tag{4.7}$$

As mentioned previously, $J(\mathbf{x}, \bar{\mathbf{m}})$ is generally not an accurate approximation of $E[J(\mathbf{x}, \mathbf{m})]$. However, the accuracy of the approximation can be improved by including higher-order terms in the expansion. To second order, the Taylor expansion of the objective can be described as

$$
\begin{aligned}
J^{\mathrm{quad}}(\mathbf{x}, \mathbf{m}) = {} & J(\mathbf{x}, \bar{\mathbf{m}}) + J_{\mathbf{m}}(\mathbf{x}, \bar{\mathbf{m}})(\mathbf{m} - \bar{\mathbf{m}}) \\
& + \frac{1}{2}(\mathbf{m} - \bar{\mathbf{m}})^{\mathrm{T}} J_{\mathbf{mm}}(\mathbf{x}, \bar{\mathbf{m}})(\mathbf{m} - \bar{\mathbf{m}}),
\end{aligned}
\tag{4.8}
$$

where we have neglected higher-order terms in the expansion (e.g., a cubic term); and $J_{\mathbf{mm}}$ denotes the second derivative of J with respect to $\mathbf{m}$. If $\mathbf{m}$ is multivariate Gaussian distributed with mean $\bar{\mathbf{m}}$ and covariance $\mathbf{C}$, then the expected value of the quadratic approximation of the objective can be shown [69] to be

$$E[J^{\mathrm{quad}}(\mathbf{x}, \mathbf{m})] = J(\mathbf{x}, \bar{\mathbf{m}}) + \frac{1}{2}\operatorname{tr}(\mathbf{C}^{1/2} J_{\mathbf{mm}}(\mathbf{x}, \bar{\mathbf{m}}) \mathbf{C}^{1/2}). \tag{4.9}$$

The key point of this approach is that the ensemble mean model can be used for optimization instead of performing optimization on the ensemble of realizations. The first term $J(\mathbf{x}, \bar{\mathbf{m}})$ can be obtained inexpensively from the mean model $\bar{\mathbf{m}}$, but computing $J_{\mathbf{mm}}(\mathbf{x}, \bar{\mathbf{m}})$ can be difficult, as it requires the second derivative of the objective function with respect to the model parameters. Although it might be possible in some cases to approximate higher-order derivatives, this approach is impractical for most reservoir applications. Therefore, if we want to use the mean model for RO without sacrificing the accuracy of the estimated expected objective-function value, we need a more practical method to improve the approximation of the expected value obtained from the mean model.

4.1.3   Mean-model bias correction

As discussed above, the ensemble average value $\bar{J}(\mathbf{x})$ can provide a good approximation of the expected value, but it requires many expensive simulations. Although $J(\mathbf{x}, \bar{\mathbf{m}})$ only needs the evaluation of control $\mathbf{x}$ applied to the mean model $\bar{\mathbf{m}}$, the result can be a poor estimate and misguide the optimization. This approximation can be improved using Taylor series expansion with higher-order terms; however, the computation of higher-order derivatives is impractical for most real reservoir applications.

In this work, we develop bias-correction methods that are applied to $J(x, \bar{\mathbf{m}})$ to efficiently generate an approximation of $\bar{J}(\mathbf{x})$. The main idea is to correct the bias in $J(\mathbf{x}, \bar{\mathbf{m}})$ by estimating a multiplicative correction factor $\alpha(\mathbf{x})$ between $\bar{J}(\mathbf{x})$ and $J(\mathbf{x}, \bar{\mathbf{m}})$,

$$\bar{J}(\mathbf{x}) = \alpha(\mathbf{x}) J(\mathbf{x}, \bar{\mathbf{m}}). \tag{4.10}$$

$\bar{J}(\mathbf{x})$ is computed by averaging over an ensemble (Eq. 4.5). Consequently, the correction factor $\alpha(\mathbf{x_i})$ of a fixed control $\mathbf{x_i}$ can be described as

$$\alpha(\mathbf{x_i}) = \frac{1}{N_e} \sum_{j=1}^{N_e} \frac{J(\mathbf{x_i}, \mathbf{m_j})}{J(\mathbf{x_i}, \bar{\mathbf{m}})}. \tag{4.11}$$

For control $\mathbf{x_i}$ applied to each individual model realization $\mathbf{m_j}$ and the mean model $\bar{\mathbf{m}}$, we define a partial correction factor $\beta(\mathbf{x_i}, \mathbf{m_j}, \bar{\mathbf{m}})$ between the objective-function values obtained from $\mathbf{m_j}$ and $\bar{\mathbf{m}}$,

$$\beta_{ij} = \beta(\mathbf{x_i}, \mathbf{m_j}, \bar{\mathbf{m}}) = \frac{J(\mathbf{x_i}, \mathbf{m_j})}{J(\mathbf{x_i}, \bar{\mathbf{m}})}. \tag{4.12}$$

Then the correction factor $\alpha(\mathbf{x_i})$ at control $\mathbf{x_i}$ over an ensemble of realizations (Eq. 4.11) can be written in terms of all relevant partial correction factors,

$$\alpha(\mathbf{x_i}) = \frac{1}{N_e} \sum_{j=1}^{N_e} \beta_{ij}. \tag{4.13}$$

One of the simplest ways to correct the bias in $J(\mathbf{x}, \bar{\mathbf{m}})$ is to use an approximation of $\bar{\alpha}$, which is the mean value of the correction factor. The most straightforward approach to compute $\bar{\alpha}$ is averaging the correction factors α from different controls,

$$\begin{aligned}
\bar{\alpha} &\approx \frac{1}{N_x} \sum_{i=1}^{N_x} \alpha(\mathbf{x_i}) = \frac{1}{N_x N_e} \sum_{i=1}^{N_x} \sum_{j=1}^{N_e} \frac{J(\mathbf{x_i}, \mathbf{m_j})}{J(\mathbf{x_i}, \bar{\mathbf{m}})} \\
&= \frac{1}{N_x N_e} \sum_{i=1}^{N_x} \sum_{j=1}^{N_e} \beta_{ij},
\end{aligned} \tag{4.14}$$

where N_x is the number of distinct controls. This approach is infeasible for most practical optimization problems since it requires $N_x \times (N_e + 1)$ evaluations of $J(\mathbf{x}, \mathbf{m})$.

Instead of using Eq. 4.14, a Monte Carlo estimate of $\bar{\alpha}$ can be obtained at a much lower cost by sampling control $\mathbf{x_j}$ uniformly from the space of all possible controls and sampling reservoir realizations $\mathbf{m_j}$ from the space of conditional realizations. Then an estimate of $\bar{\alpha}$ can be computed using the following formula,

$$\bar{\alpha} \approx \frac{1}{N_x} \sum_{j=1}^{N_x} \frac{J(\mathbf{x_j}, \mathbf{m_j})}{J(\mathbf{x_j}, \bar{\mathbf{m}})} = \frac{1}{N_x} \sum_{j=1}^{N_x} \beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}}), \tag{4.15}$$

where $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$ is the observed value of β at a random control $\mathbf{x_j}$, which requires two simulations, i.e., applying control $\mathbf{x_j}$ to a random individual realization $\mathbf{m_j}$ and applying control $\mathbf{x_j}$ to the mean reservoir model $\bar{\mathbf{m}}$. Therefore, in this approximation only $2 \times N_x$ simulations are needed to obtain a set of observations β from N_x distinct controls to estimate $\bar{\alpha}$, which is much lower than the cost of the straightforward application of Eq. 4.14.

Although an estimate of $\bar{\alpha}$ can be efficiently obtained from a set of observed values of β (Eq. 4.15), the accuracy of the approximation corrected using $\bar{\alpha}$ is limited by the variability in α. If the variability in $\alpha(\mathbf{x_i})$ is small as control $\mathbf{x_i}$ is varied, then $\bar{\alpha}$ can

provide a useful approximation of $\alpha(\mathbf{x_i})$ to adjust $J(\mathbf{x_i}, \bar{\mathbf{m}})$ toward the average objective-function value $\bar{J}(\mathbf{x_i})$. In the case of large variability, the correct values of α for some of the controls may be far from $\bar{\alpha}$. Moreover, for optimization algorithms based on ranking the expected values of decision alternatives, a fixed correction factor applied to all controls will not improve the optimal solution obtained by directly performing optimization in the mean model (Eq. 4.1). To ensure the quality in terms of the approximation of $\bar{J}(\mathbf{x_i})$ and the RO strategy, the bias in $J(\mathbf{x_i}, \bar{\mathbf{m}})$ should be corrected by an estimate of $\alpha(\mathbf{x_i})$ that corresponds to the specific control $\mathbf{x_i}$.

4.2 Estimation of correction factor

We developed three possible methods to estimate the correction factor α at different controls: distance-based localization, regularized localization, and covariance-based optimal weights. The latter two require additional information about the correction factor (e.g., variance of α, covariance of β), but the cost of estimating α for various controls through all these three techniques is the same, i.e., distinct controls are applied to individual realizations and to the mean reservoir model for obtaining a set of observed values of β. The key to making these approaches effective is that an appropriate weight $\omega(\mathbf{x_i}, \mathbf{x_j})$ that indicates the importance of each sampled control $\mathbf{x_j}$ for estimating the correction factor $\alpha(\mathbf{x_i})$ is assigned to the corresponding observation $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$.

4.2.1 Distance-based localization

In Eq. 4.15, the estimate of α is computed by averaging all sampled β values. This approximation is essentially an unweighted estimate, since all observed values of β are assigned equal weight $\omega = \frac{1}{N_x}$ regardless of the control that needs to be evaluated. In general, however, we expect that an estimate will be better if it is based primarily on information from similar control variables. Thus, we expect that a weighted estimate with higher weights assigned to similar controls and smaller weights for the dissimilar controls will be better than an unweighted estimate.

Suppose that N_e distinct controls (i.e., $N_x = N_e$) are applied to N_e individual realizations and the mean model for generating a set of observations β, such that each realization will provide one observed value of the partial correction factor β. The weighted linear estimate $\hat{\alpha}_l(\mathbf{x_i})$ at control $\mathbf{x_i}$ is described as

$$\hat{\alpha}_l(\mathbf{x_i}) = \frac{\sum_{j=1}^{N_e} \omega(\mathbf{x_i}, \mathbf{x_j})\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})}{\sum_{j=1}^{N_e} \omega(\mathbf{x_i}, \mathbf{x_j})}, \tag{4.16}$$

where $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$ is an observed partial correction factor from a random control $\mathbf{x_j}$ applied to an individual realization $\mathbf{m_j}$ and the mean model $\bar{\mathbf{m}}$. $\omega(\mathbf{x_i}, \mathbf{x_j})$ is the weight of $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$. The weights, $\omega(\mathbf{x_i}, \mathbf{x_j})$, depend on a measure of similarity, or distance measure, between controls $\mathbf{x_i}$ and $\mathbf{x_j}$.

With an appropriate distance measure, the sampled controls $\mathbf{x_j}$ that are similar to control $\mathbf{x_i}$ will be located at shorter distances. Because similar controls are expected to provide more useful information than dissimilar controls, weights are assigned such that $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$ for the controls at shorter distances have higher weights while $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$ for controls at larger distances have smaller weights. Lacking information about the

correlation of β with the distance between controls, we use the Gaspari-Cohn taper function [38] to compute distance-dependent weights,

$$\rho\left(\delta, L\right) = \begin{cases} -\frac{1}{4}\left(\frac{\delta}{L}\right)^5 + \frac{1}{2}\left(\frac{\delta}{L}\right)^4 + \frac{5}{8}\left(\frac{\delta}{L}\right)^3 - \frac{5}{3}\left(\frac{\delta}{L}\right)^2 + 1 \\ \qquad\qquad \text{for} \quad 0 \leq \delta \leq L \\ \frac{1}{12}\left(\frac{\delta}{L}\right)^5 - \frac{1}{2}\left(\frac{\delta}{L}\right)^4 + \frac{5}{8}\left(\frac{\delta}{L}\right)^3 + \frac{5}{3}\left(\frac{\delta}{L}\right)^2 - 5\left(\frac{\delta}{L}\right) \\ \qquad +4 - \frac{2}{3}\left(\frac{\delta}{L}\right)^{-1} \quad \text{for} \quad L < \delta \leq 2L \\ 0 \qquad\qquad\qquad\quad \text{for} \quad \delta > 2L \end{cases} \tag{4.17}$$

where $\rho\left(\delta, L\right)$ is the distance-based weight, which varies from 0 and 1; δ is the distance between controls; L is the taper length, which determines the distance at which the weighting drops to approximately 0.2, i.e., weight $\omega(\mathbf{x_i}, \mathbf{x_j}) \approx 0.2$ when distance $\delta(\mathbf{x_i}, \mathbf{x_j}) = L$; and $2L$ is the critical distance, beyond which the weighting is zero, i.e., $\omega(\mathbf{x_i}, \mathbf{x_j}) = 0$ when $\delta(\mathbf{x_i}, \mathbf{x_j}) > 2L$.

Distance measures for sequential actions

The distance measure of controls is specific to the problem of interest. For sequential planning problems with discrete actions, the control variables have no physical locations, but they are permutations of sequences of possible actions, in which case an order-based encoding is appropriate. Hence, we use the permutation encoding [81] of the sequences to design the distance measures for the controls involved with sequential actions. In this encoding, each integer value in the vector encodes the relative ordering of the execution of a specific action. Consider, for example, the drilling-order problem with four possible wells W_1, W_2, W_3, W_4 that need to be drilled sequentially. If the drilling sequences S in controls $\mathbf{x_i}$ and $\mathbf{x_j}$ are $S_{\mathbf{x_i}} = [W_1, W_2, W_3, W_4]$ and $S_{\mathbf{x_j}} = [W_3, W_1, W_4, W_2]$, respectively. Suppose the permutation encoding to the drilling sequence $S_{\mathbf{x_i}}$ is described as the vector $\mathbf{P_{x_i}} = [1, 2, 3, 4]^T$. The sequence $S_{\mathbf{x_i}}$ is then transformed to the vector $\mathbf{P_{x_j}} = [2, 4, 1, 3]^T$ based on the ordering of the drilling of each well. The distance $\delta(\mathbf{x_i}, \mathbf{x_j})$ between controls $\mathbf{x_i}$ and $\mathbf{x_j}$ is measured by the distance between the vectors $\mathbf{P_{x_i}}$ and $\mathbf{P_{x_j}}$. Note that $\mathbf{P_{x_i}}$ and $\mathbf{P_{x_j}}$ vary as the reference control sequence $\mathbf{x}_{\text{ref}}$ defined as $P_{\mathbf{x}_{\text{ref}}} = [1, 2, 3, 4]^T$ varies, but the distance $\delta(\mathbf{x_i}, \mathbf{x_j})$ between $\mathbf{P_{x_i}}$ and $\mathbf{P_{x_j}}$ is fixed.

Appropriate distance measures for ordering problems include the 'edit' distance [66], which is the minimum number of operations required to transform one sequence into another sequence, and the 'swap' or Jaro-Winkler distance [53], which counts the minimum number of swaps of two elements required to transform one sequence to another. Because the computation of swap and edit distances is relatively expensive, fitness-distance measures are commonly used as surrogates for the permutation distance [85]. Here, we introduce four of the most widely used distance metrics for measuring the similarity between sequences: Hamming distance, Manhattan distance, Euclidean distance and cosine distance.

The Hamming distance [41] is a metric for comparing sequences of equal length. It measures the similarity between sequences by the number of positions corresponding to different actions, i.e., number of actions that have different orderings in the sequences. In terms of the permutation encoded vectors, the Hamming distance $\delta(\mathbf{x_i}, \mathbf{x_j})_{\text{h}}$ between

control sequences $\mathbf{x_i}$ and $\mathbf{x_j}$ is defined as

$$\delta(\mathbf{x_i}, \mathbf{x_j})_{\mathrm{h}} = \sum_{k=1}^{N_a} u_k, \quad u_k = \begin{cases} 1 & \text{for} \quad P_{\mathbf{x_i},k} \neq P_{\mathbf{x_j},k} \\ 0 & \text{for} \quad P_{\mathbf{x_i},k} = P_{\mathbf{x_j},k} \end{cases} \tag{4.18}$$

where $P_{\mathbf{x_i},k}$ and $P_{\mathbf{x_j},k}$ denote the k-th elements in $\mathbf{P_{x_i}}$ and $\mathbf{P_{x_j}}$, respectively, which are the positions of a fixed action (e.g., drilling a specific well W_k); N_a is the number of all possible actions, which determines the length of the sequences.

The Hamming distance between two sequences is easy (Eq. 4.18) to compute, and the control sequences with small Hamming distance are highly similar. Among a set of randomly sampled controls $\mathbf{x_j}$, however, very few will have a small Hamming distance $\delta(\mathbf{x_i}, \mathbf{x_j})_{\mathrm{h}}$ to a specific control $\mathbf{x_i}$. Moreover, a sampled control $\mathbf{x_j}$ with a large $\delta(\mathbf{x_i}, \mathbf{x_j})_{\mathrm{h}}$ may provide useful information for estimating the correction factor $\alpha(\mathbf{x_i})$. To extract more similar control sequences at short distances, the position-based Manhattan distance, which measures the sum of the absolute differences between positions of the elements, is preferred over the Hamming distance. The Manhattan distance is the L1-distance between two vectors defined as

$$\delta(\mathbf{x_i}, \mathbf{x_j})_{\mathrm{L_1}} = \|\mathbf{P_{x_i}} - \mathbf{P_{x_j}}\|_1 = \sum_{k=1}^{N_a} |P_{\mathbf{x_i},k} - P_{\mathbf{x_j},k}|. \tag{4.19}$$

In addition to the Manhattan distance (L1 distance metric), the Euclidean and cosine distances are two standard distance metrics in the vector space. The Euclidean distance is also known as the L2 norm distance metric,

$$\delta(\mathbf{x_i}, \mathbf{x_j})_{\mathrm{L_2}} = \|\mathbf{P_{x_i}} - \mathbf{P_{x_j}}\|_2 = \sqrt{\sum_{k=1}^{N_a} |P_{\mathbf{x_i},k} - P_{\mathbf{x_j},k}|^2}, \tag{4.20}$$

and the cosine distance [87, 61] is a correlation-based distance measure defined as

$$\begin{aligned} \delta(\mathbf{x_i}, \mathbf{x_j})_{\mathrm{cos}} &= 1 - \cos(\mathbf{P_{x_i}}, \mathbf{P_{x_j}}) \\ &= 1 - \frac{\mathbf{P_{x_i}} \cdot \mathbf{P_{x_j}}}{\|\mathbf{P_{x_i}}\|\|\mathbf{P_{x_j}}\|} \\ &= 1 - \frac{\sum_{k=1}^{N_a} P_{\mathbf{x_i},k} \cdot P_{\mathbf{x_j},k}}{\sqrt{\sum_{k=1}^{N_a} P_{\mathbf{x_i},k}^2} \sqrt{\sum_{k=1}^{N_a} P_{\mathbf{x_j},k}^2}}. \end{aligned} \tag{4.21}$$

The choice of an appropriate distance measure to use as a measure of similarity is problem specific. In explaining the similarity of control sequences, the Manhattan, Euclidean, and cosine distance metrics are generally superior to the Hamming distance, which considers only the number of positions at which the corresponding actions are different. For optimization problems with a fixed set of actions, in which the lengths of all control sequences are identical, the cosine distance is simply a scaled version of the Euclidean distance,

$$\frac{\|\mathbf{P_{x_i}} - \mathbf{P_{x_j}}\|_2^2}{2\|\mathbf{P_{x_i}}\|_2^2} = 1 - \cos(\mathbf{P_{x_i}}, \mathbf{P_{x_j}}). \tag{4.22}$$

Consequently, there is no need to consider both the Euclidean and cosine distance measures. For large applications, the Manhattan distance is preferable to the Euclidean

distance [1]. For cases with small sets of actions, the performances of the Manhattan (L1 norm) and Euclidean (L2 norm) distances in measuring the similarity of control sequences can be similar.

Taper window selection

The distance-based weight (Eq. 4.17) is determined by the taper parameter L and the distance δ between controls. Hence, the performance of the distance-based localization for estimating the correction factor (Eq. 4.16) depends not only on the choice of distance measure, but also on the taper length L, which affects the effective sample size [57] used for computation of the correction factor. A good distance measure will effectively identify control variables with similar correction factors such that the number of realizations used for estimation is maximized and the sampling error is reduced.

Suppose that N_e is the number of observed values of β; then, a common approximation of the effective sample size [79, 30] for a weighted estimate of the correction factor $\alpha(\mathbf{x_i})$ at control $\mathbf{x_i}$ is defined as

$$n_{\text{eff}} = \frac{\left(\sum_{j=1}^{N_e} \omega(\mathbf{x_i}, \mathbf{x_j})\right)^2}{\sum_{j=1}^{N_e} \omega^2(\mathbf{x_i}, \mathbf{x_j})}, \tag{4.23}$$

where $\omega(\mathbf{x_i}, \mathbf{x_j})$ is the weight on the partial correction factor $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$ obtained from a random control $\mathbf{x_j}$ (Eq. 4.16) and n_{eff} is a measure of the equivalent number of equally weighted samples. If all weights are identical, then the effective sample size is equal to N_e, and the estimate of the correction factor will be unweighted (Eq. 4.15). If one of the normalized weights is equal to one and all others are zero, the effective sample size is 1.

The accuracy of the localized estimate $\hat{\alpha}_l(\mathbf{x_i})$ is influenced by both the effective sample size, n_{eff}, and the bias resulting from the use of partial correction factors β based on random control variables with different values of correction factor α. Reducing the taper length will decrease the bias by including only values from the control variable with very similar values of the correction factor but will also increase the sampling error by decreasing the effective sample size. Because the optimal taper length is not known a priori, we generally apply regularization to reduce the effect of a nonoptimal choice of taper length.

4.2.2 Regularized localization

As noted above, a smaller taper length will reduce the effective sample size and increase the sampling error, so that the localized estimate $\hat{\alpha}_l(\mathbf{x_i})$ obtained with a taper length that is smaller than the optimal may be far from the actual correction factor $\alpha(\mathbf{x_i})$ due to sampling error resulting from the small number of samples within a small distance to control $\mathbf{x_i}$. Although using a long taper length can avoid this situation, it may result in an unweighted estimate that is near the average value $\bar{\alpha}$.

One way to avoid inaccurate estimates caused by using small taper parameters is to improve the estimation accuracy by adding a regularization term based on the average value and the variance of the correction factor to reduce the sensitivity of the estimates to the taper length. A regularized estimate $\alpha_r(\mathbf{x_i})$ is obtained by minimizing objective

function S with both a local and a global term,

$$S\left(\alpha_r(\mathbf{x_i})\right) = \frac{n_{\text{eff}}}{\sigma_\beta^2}\left(\alpha_r(\mathbf{x_i}) - \frac{\sum_{j=1}^{N_e} \omega(\mathbf{x_i}, \mathbf{x_j})\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})}{\sum_{j=1}^{N_e} \omega(\mathbf{x_i}, \mathbf{x_j})}\right)^2$$
$$+ \frac{1}{\sigma_\alpha^2}\left(\alpha_r(\mathbf{x_i}) - \bar{\alpha}\right)^2, \tag{4.24}$$

where σ_α^2 and σ_β^2 are the variances of α and β, respectively, and n_{eff} is the effective sample size for the observations of β (Eq. 4.23). The regularized estimate is then obtained by solving $\nabla_\alpha S = 0$,

$$\hat{\alpha}_r(\mathbf{x_i}) = \left(1 + \frac{\sigma_\beta^2}{n_{\text{eff}}\sigma_\alpha^2}\right)^{-1}\left(\frac{\sum_{j=1}^{N_e}\omega(\mathbf{x_i},\mathbf{x_j})\beta(\mathbf{x_j},\mathbf{m_j},\bar{\mathbf{m}})}{\sum_{j=1}^{N_e}\omega(\mathbf{x_i},\mathbf{x_j})} + \frac{\sigma_\beta^2}{n_{\text{eff}}\sigma_\alpha^2}\bar{\alpha}\right)$$
$$= \left(1 + \frac{\sigma_\beta^2}{n_{\text{eff}}\sigma_\alpha^2}\right)^{-1}\left(\hat{\alpha}_l(\mathbf{x_i}) + \frac{\sigma_\beta^2}{n_{\text{eff}}\sigma_\alpha^2}\bar{\alpha}\right). \tag{4.25}$$

Note that the regularized objective function is a weighted average of the localized estimate $\hat{\alpha}_l(\mathbf{x_i})$ and the mean value of α. When the effective sample size n_{eff} is large compared to the ratio $\sigma_\beta^2/\sigma_\alpha^2$, the regularized estimate will be based primarily on the local samples of β, i.e., $\hat{\alpha}_r(\mathbf{x_i}) \approx \hat{\alpha}_l(\mathbf{x_i})$.

By adding a regularization term with an appropriate parameter associated with the variances σ_α^2 and σ_β^2 and the effective sample size n_{eff}, the estimation accuracy resulting from an inappropriate distance measure or taper length can be improved. Consequently, compared to the localized estimate $\hat{\alpha}_l(\mathbf{x_i})$, regularized estimate $\hat{\alpha}_r$ is potentially more accurate than $\bar{\alpha}$ for a wider range of taper lengths. When the variance of the correction factor is known, regularized localization (Eq. 4.25) is preferred over the pure distance-based localization (Eq. 4.16).

4.2.3 Covariance-based optimal weights

In this work, we define the correction factor $\alpha(\mathbf{x_i})$ as the linear combination of partial correction factors $\beta(\mathbf{x_i}, \mathbf{m_j}, \bar{\mathbf{m}})$ at the corresponding control $\mathbf{x_i}$ (Eq. 4.13). An estimate of $\alpha(\mathbf{x_i})$ is calculated based on a weighted linear combination of $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$ obtained from a set of random controls (Eq. 4.16). If the covariance of β is known, the optimal weights assigned to random observations β can be estimated.

Suppose that $\alpha(\mathbf{x_0})$ at a fixed control $\mathbf{x_0}$ must be estimated. In vector notation, $\alpha(\mathbf{x_0})$ is written as,

$$\alpha(\mathbf{x_0}) = \frac{1}{N}\sum_{j=1}^{N}\beta_{0j} = \frac{1}{N}\mathbf{1}^T\mathbf{b}_0, \tag{4.26}$$

where β_{0j} is the partial correction factor $\beta(\mathbf{x_0}, \mathbf{m_j}, \bar{\mathbf{m}})$ at the fixed control $\mathbf{x_0}$ of an individual realization $\mathbf{m_j}$ sampled from the probability distribution of model parameters (Eq. 4.12) and vector $\mathbf{b}_0$ denotes an ensemble of β_{0j} obtained at the estimation point from N individual realizations,

$$\mathbf{b}_0 = \begin{bmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0N} \end{bmatrix}^{\text{T}}.$$

The weighted linear estimate of $\alpha(\mathbf{x_0})$ is defined as

$$\hat{\alpha}(\mathbf{x_0}) = \frac{\sum_{j=1}^{N} \omega(\mathbf{x_0}, \mathbf{x_j}) \beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})}{\sum_{j=1}^{N} \omega(\mathbf{x_0}, \mathbf{x_j})} = \mathbf{w}^T \mathbf{b}, \tag{4.27}$$

where the elements in vector $\mathbf{b}$ are the observed values of β from random controls and realizations,

$$\mathbf{b} = \begin{bmatrix} \beta_{11} & \beta_{22} & \cdots & \beta_{NN} \end{bmatrix}^{\mathrm{T}}.$$

The j-th element in vector $\mathbf{b}$ is an observation of β obtained at a random control $\mathbf{x_j}$, i.e., $\beta(\mathbf{x_j}, \mathbf{m_j}, \bar{\mathbf{m}})$. Note that Eq. 4.27 is identical to Eq. 4.16, although the notation is different.

The optimal weights for estimating $\alpha(\mathbf{x_0})$ from a set of random observations β can be obtained by minimizing the expected variance of the estimate error, constrained to the unbiasedness condition (i.e., the expected error in the estimate $\hat{\alpha}(\mathbf{x_0})$ is 0),

$$\begin{aligned} \min_{\mathbf{w}} \quad & Var[\alpha(\mathbf{x_0}) - \hat{\alpha}(\mathbf{x_0})] \\ \text{s.t.} \quad & E[\alpha(\mathbf{x_0}) - \hat{\alpha}(\mathbf{x_0})] = 0, \end{aligned} \tag{4.28}$$

where the variance of the expected error can be expressed as

$$Var[\alpha(\mathbf{x_0}) - \hat{\alpha}(\mathbf{x_0})] = E\left[\left(\frac{1}{N} \mathbf{1}^T \mathbf{b}_0 - \mathbf{w}^T \mathbf{b} \right)^2 \right], \tag{4.29}$$

and an unbiased estimate requires

$$\mathbf{w}^T \mathbf{1} = 1. \tag{4.30}$$

To compute the optimal weights, we define a Lagrangian function according to Eqs. 4.28-4.30,

$$S(\mathbf{w}, \lambda) = E\left[\left(\frac{1}{N} \mathbf{1}^T \mathbf{b}_0 - \mathbf{w}^T \mathbf{b} \right)^2 \right] - 2\lambda \left(\mathbf{w}^T \mathbf{1} - 1 \right), \tag{4.31}$$

which can be minimized to solve for $\mathbf{w}$,

$$\nabla_{w,\lambda} S = 0. \tag{4.32}$$

Straightforward computation shows that

$$\nabla_{\lambda} S = \mathbf{1}^T \mathbf{w} - 1, \tag{4.33}$$

and

$$\nabla_w S = 2 \operatorname{cov}(\mathbf{b}, \mathbf{b}) \mathbf{w} - \frac{2}{N} \operatorname{cov}(\mathbf{b}, \mathbf{b}_0) \mathbf{1} - 2\lambda \mathbf{1} \tag{4.34}$$

where $\operatorname{cov}(\mathbf{b}, \mathbf{b})$ is the covariance of the variables β in $\mathbf{b}$ and $\operatorname{cov}(\mathbf{b}, \mathbf{b}_0)$ is the covariance between the observed values of β in $\mathbf{b}$ and $\mathbf{b_0}$.

Then the systems of linear equations for $\nabla_{w,\lambda} S = 0$ can be written in matrix form

$$\begin{bmatrix} \operatorname{cov}(\mathbf{b}, \mathbf{b}) & -\mathbf{1} \\ -\mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \frac{1}{N} \operatorname{cov}(\mathbf{b}, \mathbf{b}_0) \mathbf{1} \\ -1 \end{bmatrix}, \tag{4.35}$$

which we can solve for $\mathbf{w}$. However, information about the covariance of β is required to obtain $\operatorname{cov}(\mathbf{b}, \mathbf{b})$ and $\operatorname{cov}(\mathbf{b}, \mathbf{b}_0)$ (see Appendices in Paper B).

Efficiency of RO with bias-correction methods

To summarize, using the bias-corrected mean model to approximate the expected value requires two function evaluations: one evaluation using the mean reservoir model $\bar{\mathbf{m}}$ to obtain an initial approximate value, $J(\mathbf{x}, \bar{\mathbf{m}})$, and a second evaluation using information from distinct controls and individual realizations to estimate a multiplicative bias correction factor $\hat{\alpha}(\mathbf{x})$,

$$E[J(\mathbf{x}, \mathbf{m})] \approx \hat{\alpha}(\mathbf{x})J(\mathbf{x}, \bar{\mathbf{m}}). \tag{4.36}$$

$\hat{\alpha}(\mathbf{x})$ is an estimate of the correction factor between the ensemble average value $\bar{J}(\mathbf{x}, \mathbf{m})$ and the objective-function value from the mean model $J(\mathbf{x}, \bar{\mathbf{m}})$ (Eq. 4.10). This estimate can be obtained using a set of observed partial correction factors β at random controls applied to individual realizations (Eq. 4.12),

$$\hat{\alpha}(\mathbf{x}) = G(\beta_1, \beta_2, \cdots, \beta_{N_x}, \mathbf{x}). \tag{4.37}$$

where G is a function (e.g., Eqs. 4.16, 4.25 or 4.35) for estimating $\hat{\alpha}(\mathbf{x})$, depending on the available information of the correction factor (e.g., distance measure, variances of α and β); N_x is the number of distinct controls applied to obtain the observations β. For N_e reservoir model realizations, we can sample N_e distinct controls such that each realization can provide one observed value of β. In this case, the number of observed values of β is the same as the ensemble size, i.e., $N_x = N_e$. If the variability in β is small, we can use a smaller number of observations β to estimate the bias correction factor, i.e., $N_x < N_e$.

RO performed using such bias-correction methods has a major advantage over SAA that is obtained by averaging over an ensemble of reservoir models: only the simulations of the mean reservoir model are required during the optimization process. The initial sampling of control and model realizations is used to create an ensemble of partial correction factors β, which allows for the estimation of the bias correction factor α at different controls (Eq. 4.37). When evaluating specific controls during the RO phase, we only require additional simulations of the controls applied to the mean model to obtain the initial estimates of the expected values. The bias is then corrected by applying the corresponding correction factor estimated based on the observations of partial corrections.

Figure 4.1 shows a flowchart for the application of MLHS-SR with the bias-corrected mean model to solve general reservoir optimization problems with discrete-control variables (e.g., drilling-order of wells) when uncertainty is characterized by an ensemble of realizations. To estimate the bias correction factor α for different control sequences, we first sample N_e distinct controls and apply them to individual model realizations and the mean model to obtain an ensemble of observations β. The collection of N_e observed values of β requires $2 \times N_e$ simulations. After the initial ensemble of the observations is created from distinct controls and model realizations, estimation of the bias correction factor α for specific controls does not require additional simulations. Hence, at each iteration, we need to perform only one additional simulation in the mean model to evaluate the expected value using the bias-correction method, which is much less effort than that required in SAA. Suppose that the optimization process requires N_{iter} iterations (i.e., expected values of N_{iter} different controls need to be evaluated to compute the optimal solution). Then, the total number of simulations required in RO

with the bias-correction methodology is $N_{\text{tot}} = 2N_e + N_{\text{iter}}$, in which $2N_e$ simulations are performed to obtain N_e observed values of β and N_{iter} simulations are performed in the mean reservoir model to obtain the initial approximations of the expected values of N_{iter} different controls. If the expected values are evaluated using SAA, the cost in RO would be $N_{\text{tot}} = N_e \times N_{\text{iter}}$, which is much more expensive and will increase linearly with the ensemble size N_e and the number of iterations N_{iter}.

The use of the bias-corrected mean model allows for the application of fairly general optimization methods. If the variance and the covariance of the bias correction factor are unknown, distance-based localization (Eq. 4.16) is the preferred method for estimating the bias at specific controls. However, an appropriate distance metric to measure the similarity of controls, which is specific to the problem at hand, is required. For control variables with a set of discrete actions (e.g., drilling-order problem), the similarity between control sequences can be measured using the position-based distance (e.g., Manhattan distance). In Paper B, we investigated the estimation accuracy of the bias-correction methodology and the performance of RO with expected values estimated using different approaches in two applications: flow optimization in a one-dimensional model and the drilling-order problem in a synthetic field model.
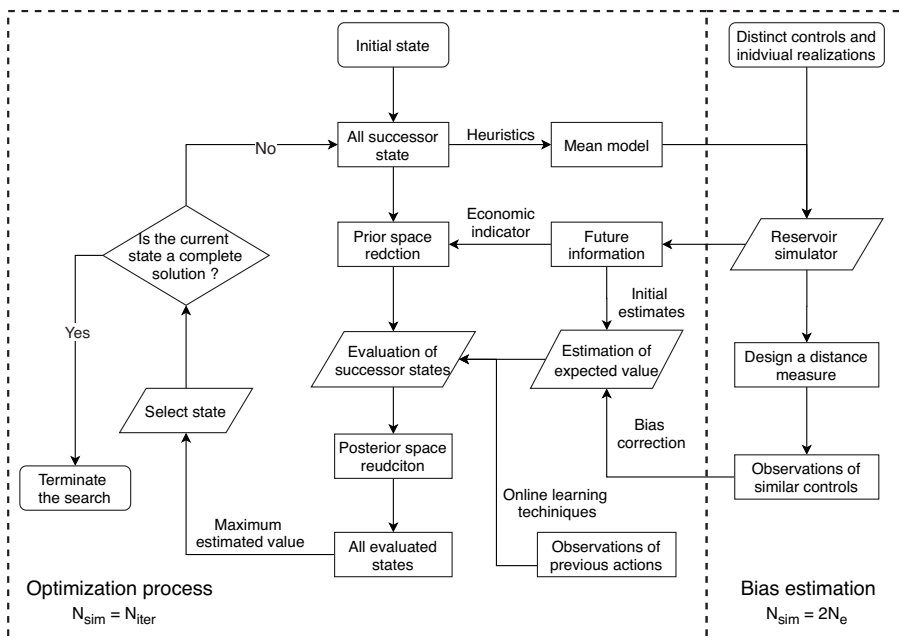

Figure 4.1: Flowchart of using MLHS-SR with bias-corrected mean model [103]

Chapter 5

Decision-making accounting for future learning

Thus far, to perform robust optimization efficiently, we have developed online learned heuristics to search for an optimal sequential solution and bias-correction methods to address uncertainty in the reservoir characteristics. However, we have not yet considered the effect of future information in making optimal decisions. As noted in Chapter 2, the previous and current decisions affect both the possibility of the future choices of actions and the possibility of future uncertainty reduction. Consequently, when choosing actions, we should also take into account the opportunities to improve the optimal strategy resulting from future uncertainty reduction (i.e., the possibility of future learning through actions), rather than that based solely on the maximization of expected NPV over the current assessment of uncertainty. In this chapter, we present a flexible workflow built on the key-feature-based VOI analysis for efficient optimal decision-making while accounting for the possibility of future learning. The first section of Chapter 5 describes the optimal strategy obtained while considering future uncertainty reduction resulting from the current decision step. In Sections 5.2 and 5.3, we discuss how to make the computation of VOI analysis manageable by identifying key actions and key observations for optimization problems. The efficiency of the method results from the focus on the use of key observations to reduce the uncertainties with large influences on the optimal decisions (i.e., key uncertainties reduction), rather than using all observations to reduce all uncertainties.

5.1 Planning for future learning

In Section 2.2, we introduced the standard decision analytic approach to computing the optimal decision $a_{j+1}^{*\text{fs}}$ (Eqs. 2.5-2.6) that considers the effects of all future possible observations from remaining actions. This method is infeasible for realistic field development problems, however, since it requires the consideration of all possible combinations of the sequences of remaining actions with distinct observations. To obtain an optimal decision while considering future possibilities for learning through actions, a more feasible way is to consider the effects of future observations resulting from only the current decision step instead of all remaining decision stages.

Using the notation in Eqs. 2.4-2.5, after taking a set of j discrete actions, the optimal next action, that considers the possibility of future learning through the current decision

step, can be expressed as

$$a_{j+1}^{*\text{fl}} = \operatorname*{argmax}_{a_{j+1}\in A_{j+1}(h_j)} \sum_{o\in O_{a_{j+1}}} p(o|h_j, a_{j+1})\, \text{EV}^*\left(h_j, a_{j+1}, u_{j+1}^o\right), \tag{5.1}$$

where $p(o|h_j, a_{j+1})$ represents the marginal probability of the observations obtained from the decision alternative a_{j+1} and $\text{EV}^*\left(h_j, a_{j+1}, u_{j+1}^o\right)$ is the maximum achievable expected NPV constrained to previous actions in history h_j and the current decision a_{j+1}. This expectation is evaluated over the updated uncertainty state u_{j+1}^o, including the future possible observation o from a_{j+1}. Note that $\text{EV}^*\left(h_j, a_{j+1}, u_{j+1}^o\right)$ is different from the expected value $Q_{N_w-(j+1)}^*(h_{j+1})$ that is computed through a backward induction procedure while considering all possible future information from all remaining actions.

Although $a_{j+1}^{*\text{fl}}$ (Eq. 5.1) might not be identical to $a_{j+1}^{*\text{fs}}$, which considers the possibilities of future learning through all remaining actions (Eq. 2.5), the cost of computing $a_{j+1}^{*\text{fl}}$ is much lower than that of computing $a_{j+1}^{*\text{fs}}$. In general, the information obtained from the later decision stages has a smaller impact on improving the optimal strategy. We expect that simplifying the fully structured decision tree by considering the effects of future information resulting from only the current decision step would not incur much performance loss, i.e., $a_{j+1}^{*\text{fl}}$ is expected to be an approximation solution near the optimal decision $a_{j+1}^{*\text{fs}}$. Instead of focusing on solving Eqs. 2.5-2.6 to compute $a_{j+1}^{*\text{fs}}$, we focus on how to efficiently improve the optimal decision a_{j+1}^* that is obtained without considering any future information (Eq. 2.4) to $a_{j+1}^{*\text{fl}}$ that considers the effect of future information on the optimal strategy.

5.1.1 Simplified VOI analysis

In the terminology of VOI, $a_{j+1}^{*\text{fl}}$ is the optimal decision determined by the expected value with additional information (EVWI) through one decision point [50],

$$a_{j+1}^{*\text{fl}} = \operatorname*{argmax}_{a_{j+1}\in A_{j+1}(h_j)} \text{EVWI}_{a_{j+1}},$$

$$\text{EVWI}_{a_{j+1}} = \sum_{o\in O_{a_{j+1}}} p(o|h_j, a_{j+1})\, \text{EV}^*\left(h_j, a_{j+1}, u_{j+1}^o\right). \tag{5.2}$$

while the optimal decision a_{j+1}^* (Eq. 2.4) ignoring the possibility of future learning is the one obtained based on the expected value without additional information (EVWOI),

$$a_{j+1}^* = \operatorname*{argmax}_{a_{j+1}\in A_{j+1}(h_j)} \text{EVWOI}_{a_{j+1}}, \quad \text{EVWOI}_{a_{j+1}} = \text{EV}^*\left(h_j, a_{j+1}, u_j\right). \tag{5.3}$$

Assuming that there is no explicit cost for obtaining information from $a_{j+1}\in A_{j+1}(h_j)$, $a_{j+1}^{*\text{fl}}$ is then generally a more robust decision than a_{j+1}^* because the consequences of the selected action on the future uncertainty state are considered before committing to a decision.

As a simple illustration, suppose that we need to determine the optimal next well from two remaining wells W_a and W_b. Each well provides only two possible observations of facies o_{f_1} and o_{f_2}. Figures 5.1 and 5.2 show the VOI decision trees for choosing the optimal next well with and without considering future learning possibilities through
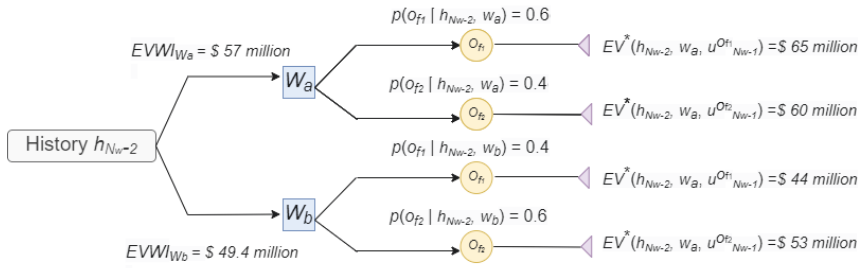
Figure 5.1: Simplified VOI analysis in consideration of the future information obtained from current decision alternatives [104]
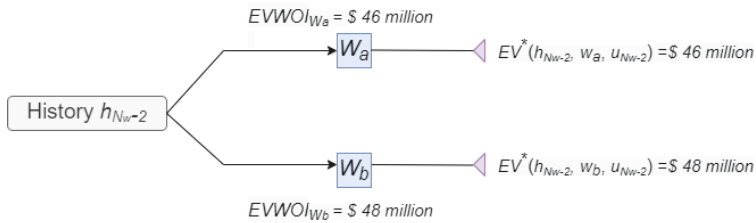

Figure 5.2: A simple example of the decision tree ignoring future learning possibilities [104]

current decision alternatives. As illustrated in Fig 5.2, drilling W_b as the next well will achieve the maximum expected NPV over current uncertainties. Taking into account the possibility of future uncertainty reduction (Fig. 5.1), however, drilling W_a as the next well turns out to be a more robust decision ($\text{EVWI}_{W_a} > \text{EVWI}_{W_b}$). As this illustration shows, optimization based on only the current assessment of uncertainty might lead to a sub-optimal solution, while taking into account the effects of future information before committing to a decision allows improvement in the optimal strategy. Compared to the use of the standard VOI analysis with extensive form (Fig. 2.3), this simplified VOI analysis does not require a fully structured decision tree with a backward induction procedure to make optimal decisions, considering the future learning possibilities. However, directly solving Eq. 5.2 is still prohibitively expensive if all possible outcomes of future observations in terms of each decision alternative have to be taken into account in the VOI analysis.

Suppose that a set of N_e model realizations is used to represent uncertainty in reservoir properties and that there are N_d possible decision alternatives at the current stage. Each ensemble member at a given decision is capable of generating a specific set of simulated observations (e.g., production/injection data over a certain period), which can be used to update the reservoir model through history matching or data assimilation before choosing the next action. Taking into account all possible future observations that can be obtained from each pair of model realization and decision, we have to update the reservoir model and perform optimization $N_e \times N_d$ times to identify the optimal decision $a_{j+1}^{*\text{fl}}$ with the highest EVWI (Eq. 5.2). This would be computationally intractable in most reservoir applications for which the costs of history matching and optimization are large. Therefore, we need to design a more practical way to obtain an optimal decision that accounts for the possibility of future learning and yet make the
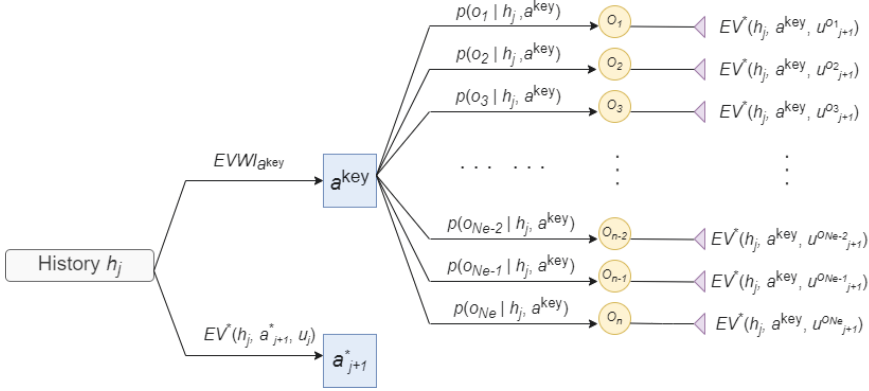
Figure 5.3: Simplified VOI decision tree with future information only from key action [104]

computation manageable.

5.2 Learning through key action

In Eq. 5.2, the effects of future information from all currently available decision alternatives are taken into account before committing to a decision. Consequently, the VOI analysis requires the evaluation of $\text{EVWI}_{a_{j+1}}$ of each possible action $a_{j+1} \in A_{j+1}(h_j)$. However, some decisions may result in little information or information that is irrelevant to the optimization of the objective, so that the optimal decision obtained with additional information may not be changed, i.e., the optimal solutions for maximizing the expected NPV over the uncertainty states u_{j+1}^o and u_j are equivalent. In this case, accounting for the future information from these actions will only increase the cost of making decisions, without improving the optimal strategy. The ability to account for the possibility of future learning through key actions is more important for evaluating decision alternatives than considering all possible decision alternatives.

5.2.1 VOI analysis through key action

In VOI analysis, the optimal decision is obtained by ranking the importance of the decision alternatives based on EVWI (i.e., assuming no explicit cost for collecting information). In this case, simplifying the VOI analysis to the decision alternatives with high expected values allows a reduction in the cost of making optimal decisions without a loss of solution quality. However, it is infeasible to identify the promising decisions by comparing the actual expected values of all possible decision alternatives, which is similar to performing a complete VOI analysis (Eq. 5.2). To address this issue, we developed a simplified VOI analysis through key actions that would provide valuable information for exploring key reservoir features of the optimization problem.

Instead of considering the possibilities of future learning through all remaining actions, we simplify the VOI decision tree to only two decision alternatives, i.e., a_{j+1}^* and a_{j+1}^{key}, and only take into account the effect of future information from a_{j+1}^{key} (Fig. 5.3). a_{j+1}^* is the optimal decision that will achieve the maximum expected NPV over the current uncertainty state, while a_{j+1}^{key} is the key action that would provide important

information for reducing key uncertainties in the optimization problem. Although information obtained from a_{j+1}^{key} is most likely to improve future decisions, taking a_{j+1}^{key} rather than a_{j+1}^* is not always worthwhile even if the information is obtained without an explicit cost. In some cases, taking a_{j+1}^{key} as the next decision could severely constrain the maximum achievable value of complete strategies, so that the maximized value may be much lower than the actual optimality. In this situation, the additional value created by improving future decisions using information from a_{j+1}^{key} may not be able to compensate for the cost caused by a sub-optimal solution constrained to a_{j+1}^{key} chosen as the next decision.

To determine whether it is preferable to take action a_{j+1}^{key} over a_{j+1}^*, we need to evaluate an implicit *net* expected value of information (EVOI) obtained from changing decision a_{j+1}^* to a^{key},

$$\mathrm{EVOI}_{a_{j+1}^* \to a_{j+1}^{\mathrm{key}}} = \mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}} - \mathrm{EV}^*(h_j, a_{j+1}^*, u_j), \tag{5.4}$$

where $\mathrm{EV}^*(h_j, a_{j+1}^*, u_j)$ is the expected value from the optimal strategy over the current uncertainty state without any future information (Eq. 5.3), and $\mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}}$ is the expected value with additional information (Eq. 5.2) from a_{j+1}^{key},

$$\mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}} = \sum_{o \in O_{a_{j+1}^{\mathrm{key}}}} p(o|h_j, a_{j+1}^{\mathrm{key}}) \times \mathrm{EV}^*(h_j, a_{j+1}^{\mathrm{key}}, u_{j+1}^o), \tag{5.5}$$

where $O_{a_{j+1}^{\mathrm{key}}}$ is the set of all possible distinct observations from a^{key}; $p(o|h_j, a_{j+1}^{\mathrm{key}})$ is the marginal probability of specific observations from a_{j+1}^{key}; and $\mathrm{EV}^*(h_j, a_{j+1}^{\mathrm{key}}, u_{j+1}^{o_i})$ is the expected value from the optimal strategy for uncertainty state u_{j+1}^o updated based on the information from a_{j+1}^{key}.

If $\mathrm{EVOI}_{a_{j+1}^* \to a^{\mathrm{key}}} > 0$, it is worth taking a^{key} to acquire the information that would help improve future decisions and vice versa. Note that $\mathrm{EVOI}_{a_{j+1}^*}$ is associated with a change in decisions. This expected value is different from the standard definition of EVOI for a^{key}, which is the difference in the expected values with and without additional information from a^{key},

$$\mathrm{EVOI}_{a_{j+1}^{\mathrm{key}}} = \mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}} - \mathrm{EVWOI}_{a_{j+1}^{\mathrm{key}}} = \mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}} - \mathrm{EV}^*(h_j, a_{j+1}^{\mathrm{key}}, u_j). \tag{5.6}$$

a^{key} is expected to provide useful information for making better future decisions, so that $\mathrm{EVOI}_{a_{j+1}^{\mathrm{key}}}$ will be a positive value. However, $\mathrm{EVOI}_{a_{j+1}^* \to a^{\mathrm{key}}}$ can be negative since it involves an indirect expected cost of information (ECOI) if taking a^{key} results in a sub-optimal solution for the current assessment of uncertainty,

$$\mathrm{ECOI}_{a_{j+1}^* \to a_{j+1}^{\mathrm{key}}} = \mathrm{EV}^*(h_j, a_{j+1}^*, u_j) - \mathrm{EV}^*(h_j, a_{j+1}^{\mathrm{key}}, u_j). \tag{5.7}$$

where $\mathrm{EV}^*(h_j, a^{\mathrm{key}}, u_j)$ is the expected value from the optimal strategy constrained to a^{key} chosen as the next action without considering information from a^{key} (i.e., uncertainty at state u_j is re-estimated only using observations in history h_j).

$\mathrm{EVOI}_{a_{j+1}^* \to a_{j+1}^{\mathrm{key}}}$ from changing the decision a_{j+1}^* to a_{j+1}^{key} (Eq. 5.4) can be rewritten as

$$\begin{aligned}
\mathrm{EVOI}_{a_{j+1}^* \to a_{j+1}^{\mathrm{key}}} = {} & \left(\mathrm{EVWI}(a_{j+1}^{\mathrm{key}}) - \mathrm{EV}^*(h_j, a_{j+1}^{\mathrm{key}}, u_j) \right) \\
& - \left(\mathrm{EV}^*(h_j, a_{j+1}^*, u_j) - \mathrm{EV}^*(h_j, a_{j+1}^{\mathrm{key}}, u_j) \right),
\end{aligned} \tag{5.8}$$

Thus, $\text{EVOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$ (Eq. 5.8) is actually an implicit *net* $\text{EVOI}_{a^{\text{key}}_{j+1}}$ (Eq. 5.6) accounting for the hidden cost $\text{ECOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$ (Eq. 5.7),

$$\text{EVOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}} = \text{EVOI}_{a^{\text{key}}_{j+1}} - \text{ECOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}. \tag{5.9}$$

If a^{key}_{j+1} is identical to a^*_{j+1}, there is no hidden cost, i.e., $\text{ECOI}_{a^{\text{key}}_{j+1} \to a^{\text{key}}_{j+1}} = 0$. In general, the key action a^{key}_{j+1} may not be the initial decision a^*_{j+1} obtained from the optimal strategy for current assessment of uncertainty, especially when the decision space is large. Due to the hidden cost, we need to assess $\text{EVOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$ before committing to a decision.

5.2.2 Key action identification

To improve the optimal decision a^*_{j+1} by accounting for future learning, a^{key}_{j+1} is the decision alternative identified from the current action space $a_{j+1} \in A_{j+1}(h_j)$ that is expected to result in a high $\text{EVOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$. As shown in Eq. 5.4, the key action is also the decision alternative with a high $\text{EVWI}_{a_{j+1}}$. Hence, performing the VOI analysis with key actions is similar to a simplified VOI analysis with the promising decision alternatives that are likely to result in high EVWI. If a^{key}_{j+1} has the maximum $\text{EVOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$, the key action a^{key}_{j+1} would be identical to the optimal decision $a^{*^{\text{fl}}}_{j+1}$ obtained accounting for the possibilities of future learning through all remaining actions (Eq. 5.2). In that case, simplifying the VOI decision tree to a^{key}_{j+1} would not incur a performance loss compared to directly solving Eq. 5.2.

As mentioned previously, it would be computationally intractable to identify the promising decision alternatives by comparing the actual $\text{EVWI}_{a_{j+1}}$ of all possible decisions $a_{j+1} \in A_{j+1}(h_j)$. However, it is possible to identify key actions without computing the actual $\text{EVOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$ of all remaining actions. According to Eq. 5.9, a^{key}_{j+1} with a high $\text{EVOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$ is the decision alternative that is expected to result in a large $\text{EVOI}_{a^{\text{key}}_{j+1}}$ but a small $\text{ECOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$. The standard $\text{EVOI}_{a^{\text{key}}_{j+1}}$ (Eq. 5.6) depends on whether a^{key}_{j+1} is able to provide useful information for making better future decisions. If taking a^{key}_{j+1} as the next action results in obtaining important information to reduce uncertainty in the key reservoir features for optimization problems, $\text{EVOI}_{a^{\text{key}}_{j+1}}$ is expected to be a high value due to the potentially significant improvement in the optimal strategy. The hidden cost $\text{ECOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$ depends on whether the optimal strategies constrained to a^{key}_{j+1} can achieve a high expected NPV for the current assessment of uncertainty. This indirect cost of information caused by taking a^{key}_{j+1} that appears to be a sub-optimal solution can be evaluated when computing the optimal decision a^*_{j+1} over the current uncertainty state, and without incurring additional costs when a^*_{j+1} is obtained using learned heuristic search, which allows for the estimation of the maximum achievable value constrained to specific past decisions without finding the actual optimal solution. Therefore, by considering the possibility of obtaining valuable information for reducing key uncertainties and the possibility of achieving high expected NPV for the current uncertainty state, we can efficiently identify the a^{key}_{j+1} that is likely to result in a high $\text{EVOI}_{a^*_{j+1} \to a^{\text{key}}_{j+1}}$, i.e., decision alternative $a_{j+1} \in A_{j+1}(h_j)$ with a high $\text{EVWI}_{a_{j+1}}$.

In the simplified VOI analysis (Fig. 5.3) based on $\mathrm{EVOI}_{a_{j+1}^* \to a_{j+1}^{\mathrm{key}}}$, the effect of future information from the initial optimal decision a_{j+1}^* is not taken into account for efficient optimal decision-making. In some cases, a_{j+1}^* may also be able to provide important observations for reducing key uncertainties. To obtain a more robust decision, we can take into account the possibilities of future learning through both a_{j+1}^{key} and a_{j+1}^*,

$$
\hat{a}_{j+1}^{*\,\mathrm{fl}} = \arg \max_{a \in [a_{j+1}^*, a_{j+1}^{\mathrm{key}}]} \mathrm{EVWI}_a = \arg \max_{a \in [a_{j+1}^*, a_{j+1}^{\mathrm{key}}]} \sum_{o \in O_a} p(o|h_j, a) \times \mathrm{EV}^*(h_j, a, u_{j+1}^o).
$$
(5.10)

However, this approach will increase the computational cost of making a decision, since the evaluation of EVWI for each action requires re-optimization and re-estimation of uncertainties multiple times to account for various outcomes of observations.

5.3   Learning through key information

Although the number of decision alternatives that need to be evaluated in the VOI analysis is reduced by identifying the key action that will provide important information for making better future decisions, directly solving Eq. 5.4 is likely to be impractical when all distinct observations are accounted for, i.e., updating the reservoir model and performing re-optimization N_e times to obtain the maximum expected values $\mathrm{EV}^*(h_j, a_{j+1}^{\mathrm{key}}, u_{j+1}^o)$ from all posterior ensembles for the computation of $\mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}}$. In this section, we discuss how to make the computation manageable by identifying key information for improving future decisions, i.e., observations that are most helpful in exploring key reservoir features for optimization problems.

5.3.1   VOI analysis through key information

As described previously, the additional value with information is achieved by the reduction in uncertainties that affect the optimal decisions. Instead of using all observations to reduce all uncertainties, we can approximately compute $\mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}}$ by using key observations to reduce key uncertainties of the optimization problem. We expect that such an approximation can be used to indicate the importance of a_{j+1}^{key} and would not incur a performance loss in the VOI framework. The VOI analysis is performed by ranking the importance of decision alternatives based on the expected values, while $\mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}}$ deals with the information content of hypothetical data. When actual observations are obtained from an action that has been executed, an actual history match is performed with all observations to update the entire reservoir model.

The action a_{j+1}^{key} generally could provide a large number of individual observations from different information sources. The reduction in key uncertainties from some observations might be very small. The computational penalty of including these nonessential observations in updating the reservoir model, however, can be large. To efficiently account for the effect of future information on the optimal strategy, we use the key observations to reduce key uncertainties in the optimization problem, rather than using all observations to reduce all uncertainties. Thus, there is no need for full history matching to update the entire reservoir model. To ensure that the computation of $\mathrm{EVWI}_{a_{j+1}^{\mathrm{key}}}$ is manageable, first, we divide the entire observation space R_b^n associated with the key observation subset into a limited number of disjoint subspaces (e.g., $R_b^n = \Omega_1^b \cup \Omega_2^b$ and

$\Omega_1^b \cap \Omega_2^b = \emptyset$). Second, we update the reservoir model for each subspace instead of each possible outcome of future observations that might be obtained from a_{j+1}^{key}.

Suppose that observations located in the same subspace have almost the same prediction precision to reduce key uncertainties. The posterior probability distributions of key uncertainties conditioned on observations in the same subspace would be similar. In this case, there is no need to compute N_e posterior ensembles considering all distinct sets of observations from individual realizations. $\text{EVWI}_{a_{j+1}^{\text{key}}}$ could be efficiently evaluated by performing the optimization process only in a few posterior ensembles obtained based on the observation subspaces Ω_k^b,

$$\text{EVWI}_{a_{j+1}^{\text{key}}} = \sum_{k=1}^{N_{\Omega^b}} p(o^b \in \Omega_k^b | h_j, a_{j+1}^{\text{key}}) \times \text{EV}^*(h_j, a_{j+1}^{\text{key}}, u_{j+1}^{\Omega_k^b}), \tag{5.11}$$

where N_{Ω^b} is the number of observation subspaces, which is much smaller than the ensemble size ($N_{\Omega^b} \ll N_e$), and $u_{j+1}^{\Omega_k^b}$ is the updated uncertainty state for observed values $o^b \in \Omega_k^b$.

To ensure the usefulness of key observations and their subspaces in reducing key uncertainties, the entire key observation space is divided such that each subspace Ω_k^b has a high probability $P(\Omega_k^b | \Theta_k^m)$ for indicating a specific subregion Θ_k^m of key uncertainties, while the probability $P(\Omega_k^b | \Theta_i^m)$ for key uncertainties located in other subregions Θ_i^m is low. Suppose that distribution of key uncertainties is divided into N_{Ω^b} disjoint subregions $\Theta^m = [\Theta_1^m, \Theta_2^m, \ldots, \Theta_{N_{\Omega^b}}^m]$. The best observation space division $\Omega^b = [\Omega_1^b, \Omega_2^b, \ldots, \Omega_{N_{\Omega^b}}^b]$ can be described as

$$\Omega^b = \underset{\Omega^b}{\text{argmax}} \sum_{k=1}^{N_{\Omega^b}} \left[P(\Omega_k^b | \Theta_k^m) - \sum_{\substack{i=1 \\ i \neq k}}^{N_{\Omega^b}} P(\Omega_k^b | \Theta_i^m) \right]. \tag{5.12}$$

The division of Θ^m can then be optimized based on the performance of the corresponding Ω^b,

$$\Theta^m = \underset{\Theta^m}{\text{argmax}} \sum_{k=1}^{N_{\Omega^b}} P(\Omega_k^b | \Theta_k^m), \tag{5.13}$$

which is a simplification of Eq. 5.12 since $\sum_{k=1}^{N_{\Omega^b}} P(\Omega_k^b | \Theta_i^m) = 1$ at a given key uncertainties subregion Θ_i^m. If key uncertainties have a limited number of categories, we can set each category to one specific subregion Θ_i^m.

When the prior probability $P(\Theta_i^m)$ of each subregion Θ_i^m and the information's reliability $P(\Omega_k^b | \Theta_i^m)$ are known, the posterior probability $P(\Theta_i^m | \Omega_k^b)$ can be computed using Bayes' theorem,

$$P(\Theta_i^m | \Omega_k^b) = \frac{P(\Omega_k^b | \Theta_i^m) \times P(\Theta_i^m)}{\sum_{i=1}^{N_{\Omega^b}} P(\Omega_k^b | \Theta_i^m) \times P(\Theta_i^m)}, \tag{5.14}$$

where $\sum_{i=1}^{N_{\Omega^b}} P(\Omega_k^b | \Theta_i^m) \times P(\Theta_i^m)$ is the marginal probability of observing $o^b \in \Omega_k^b$ in the prior ensemble.

5.3.2 Key information identification

To efficiently identify the key observations and the best space division, we can build supervised-learning models that are able to capture the mapping between observations (inputs) and key uncertainties (outputs). This technique enables automatic detection of the optimal combination of observations and evaluation of the reliability of information $P(\Omega_k^b|\Theta_k^m)$ for each subspace. Hence, the posterior probability of key uncertainties can be computed directly based on Bayes' theorem (Eq. 5.14), avoiding the necessity of expensive data assimilation algorithms to update the reservoir model.

The use of supervised-learning algorithms requires a dataset that includes all possible observations and the corresponding distribution of key uncertainties, which can be obtained from individual realizations applied with the key action. This original dataset may contain hundreds to thousands of observations. Directly applying the learning algorithms to such a large number of input variables would make the model complex and may lead to overfitting due to the curse of dimensionality. Moreover, the original dataset may contain a large amount of nonessential observations for reducing key uncertainties. To avoid these issues, first, we can apply filter methods [64] to quickly remove redundant and irrelevant observations by ranking their importance based on some relevance measures (e.g., mRMR [71], ROC [44]), obtaining a subset of observations that are potentially useful for the reduction in key uncertainties. Second, the wrapper methods [60] involved in supervised-learning models can be applied to identify the best combination of observations that yields the optimal results for learning algorithms, i.e., key observation subset with high prediction accuracy for exploring key reservoir features.

Figure 5.4 shows the process of selecting key observations from the original dataset using filter and wrapper methods with an inductive learning model, noting that the input variables (observations) are referred as "features" in supervised-learning algorithms. To avoid overfitting in learning models, we split the original dataset into a training set to build the models and a test set to evaluate the learning models' performance. For the cases with limited data samples, resampling methods (e.g., cross-validation) can be applied to generate multiple training and test sets. If necessary, various learning models can be built simultaneously to identify the combination of observations with the best performance. This approach does not incur additional costs since all models are built on the same original dataset.

Robust decision-making through key-feature-based VOI analysis

Figure 5.5 shows an example of VOI analysis performed through the key action (Eq. 5.4) and key information (Eq. 5.11) that are identified based on the key reservoir features for the optimization problems. In this case, the effects of important future observations from a_{j+1}^{key} on the reduction in key uncertainties are considered before committing to a decision, while the entire key observation space is divided into four disjoint subspaces (i.e., $N_{\Omega^b} = 4$). It is then only necessary to update the reservoir model and perform re-optimization four times for the evaluation of $\text{EVWI}_{a_{j+1}^{\text{key}}}$. We refer to such a simplified VOI decision tree as *key-feature-based* VOI analysis.

Suppose that key observations and best space division are obtained using supervised-learning algorithms and that the learning models are built using the dataset obtained from a set of N_{sl} individual realizations, i.e., the cost of updating the reservoir model for all observation subspaces is N_{sl} simulations. Considering the cost for obtaining the
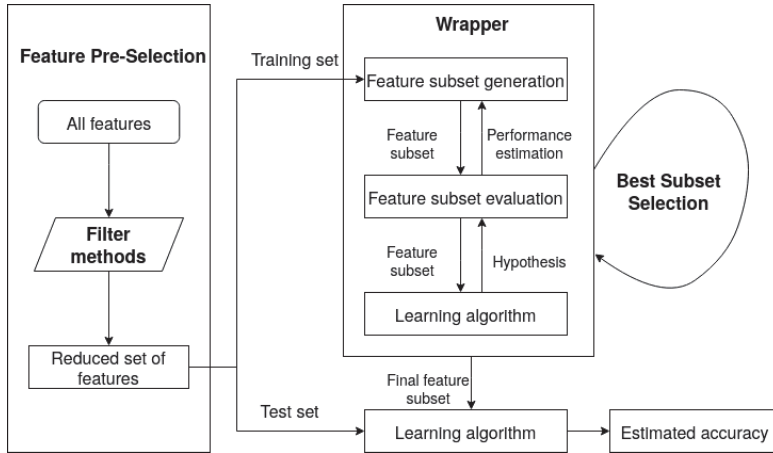
Figure 5.4: Key observation identification process based on filter and wrapper methods [104]
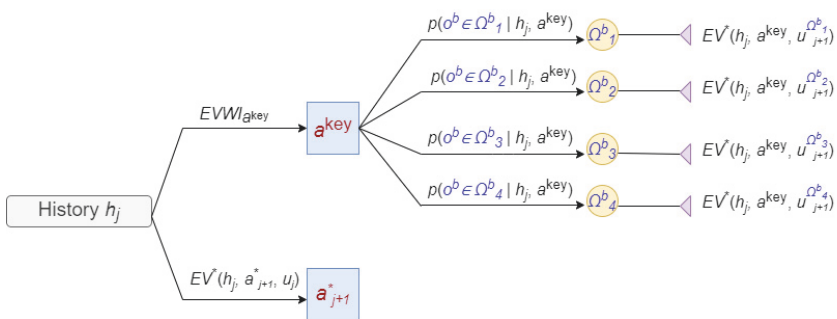

Figure 5.5: Key-feature-based VOI analysis conducted by identifying key uncertainties for optimization problems and key observations to reduce key uncertainties [104]

maximum value EV^* constrained to past decisions, only $N_{sl}+N_{opt}\times(N_\Omega^b+1)$ simulations are required to perform the key-feature-based VOI analysis, where N_{opt} is the cost of a single robust optimization. This VOI framework requires many fewer simulations than those required to directly solve Eq. 5.1 with an exhaustive history matching and optimization procedure, which requires $N_d \times N_e \times (N_{hm} + N_{opt})$ simulations to obtain the optimal decision accounting for all possible values of data that can be obtained from remaining actions, where N_{hm} is the cost of history matching and N_d is the number of possible decision alternatives. Although simplifying the VOI analysis through key action and key information might not result in the same optimal solution as the complete VOI analysis, it is a practical way to compute a near-optimal decision in consideration of the future possibilities for learning through actions. In Paper C, we investigated the performance of the key-feature-based VOI analysis applied with supervised-learning algorithms by an application of the drilling-order problem in a synthetic model, for which the drilling sequence of wells is an important contributor to the reservoir's profitability and for which the optimal solution changes significantly with key reservoir features.

The methodology can be extended to general sequential decision-making problems under uncertainty in consideration of the effect of future information. Figure 5.6 illustrates the workflow built on the key-feature-based VOI analysis to make optimal decisions efficiently while accounting for the possibilities for future learning through actions. By identifying key uncertainties for the optimization problems, we can identify the decisions that would provide the most useful information for improving future decisions. As mentioned previously, the information is obtained at the cost of making a decision to proceed, so that there is an indirect cost of information when taking action leads to a sub-optimal solution. Hence, to identify the key information-gathering actions with potentially small hidden cost of information caused by sub-optimal solutions, we need to consider both the possibility of obtaining valuable information for reducing key uncertainties and the possibility of achieving high expected value over the current uncertainty state. To avoid the cost of formal history matching, we use key observations to reduce key uncertainties, rather than using all observations to reduce all uncertainties. The key observations can be efficiently identified using supervised-learning algorithms, which allows automatic detection of the optimal observation subset for reducing key uncertainties and the estimation of the information's reliability of each observation subspace. Then, the posterior probability distortions of key uncertainties can be computed directly using Bayes' rule instead of using expensive data assimilation algorithms.

In this workflow, we only take into account the effect of future information from the key action. To determine whether it is worth taking the key action to acquire useful information for improving future decisions, we need to assess an implicit net $EVOI$ associated with changing the initial optimal action over the current uncertainty state to the key action (Eq. 5.4). In some cases, the initial optimal decision may also be able to provide useful information for reducing key uncertainties, i.e, a possible key action. To obtain a robust decision, it may be necessary to consider the possibility of future learning through both the initial optimal action and alternative key actions identified from the remaining decision alternatives. The optimal action obtained from the VOI analysis is then the one with the highest $EVWI$ (Eq. 5.10).
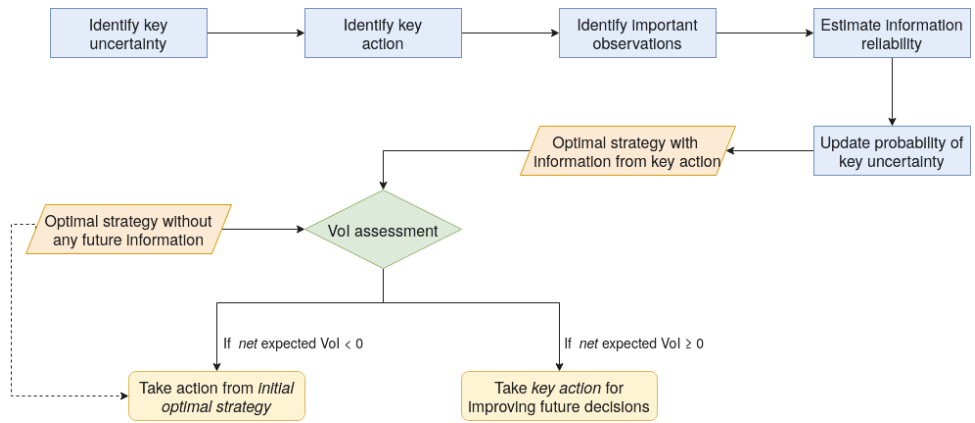
Figure 5.6: Decision-making accounting for the future learning possibilities through key action and key information for optimization problems [104]

Chapter 6

Summary of papers and outlook

In the previous chapters, we introduced new methodologies to address three main issues in field development optimization, which have been discussed in Sections 1.1 and 1.2. Our solution methods included learned heuristic search, bias-correction of predictions from the mean model, and key-feature-based VOI analysis for robust decision-making accounting for future learning. Combining these methods, we can efficiently solve sequential field development tasks while considering both the effect of geological uncertainty and the opportunity to improve the optimal strategy resulting from future uncertainty reduction. In last Chapter 6, we provide a brief summary of the articles included in Part II, and present the deficiencies and potential improvements for further research.

6.1   Summary of the papers

Paper A

Title: Efficient Optimization of Well-Drilling Sequence with Learned Heuristics
Authors: Lingya Wang and Dean S. Oliver
Journal: SPE Journal (Published)
DOI: 10.2118/195640-PA

Paper A presented a non-parametric online learning methodology for efficiently searching for an optimal or near-optimal sequential solution. We described the optimization problem with a discrete set of actions as finding the best path in a direct tree search and investigated the performance of various heuristic search methods via an application of the drilling-order problem for fixed well locations and deterministic geology. A crude heuristic in which all remaining wells are assumed to be drilled simultaneously at the next step is used to initially estimate the maximum NPV constrained to previous wells. Results show that directly using this approximation in the A* algorithm will lead to an exhaustive search caused by a large estimated value, although it is guaranteed to find the true optimal solution when the actual maximum NPV is always overestimated. By adding a small weighting parameter ($\omega < 1$) to the heuristic value, a sub-optimal solution can be found quickly. However, experiments of weighted A* search show that an appropriate weighing parameter that can speed up the search

progress without sacrificing the estimation accuracy or the solution quality might not exist for some applications.

In contrast, the online learning techniques designed based on the observation from previous decision steps have proved to be effective in improving the accuracy of the estimated value and accelerating the search process, thereby efficiently obtaining a solution with high value. Moreover, the use of space reduction techniques allows to find a solution with the same quality faster. For the case with 8 wells, which has 40320 (8!) possible drilling schedules, it was only necessary to perform 22 simulations to obtain an optimized drilling sequence in a deterministic setting. The average error in the estimated values is reduced from 5.57% to 0.19% after using online-learning mechanisms. The obtained solution is then very likely to have a high NPV and be near the true optimal solution.

Paper B

Title: Fast Robust Optimization Using Bias Correction Applied to the Mean Model
Authors: Lingya Wang and Dean S. Oliver
Journal: Computational Geosciences (Published)
DOI: 10.1007/s10596-020-10017-y

In Paper B, the effect of geological uncertainty is taken into account for computing a robust optimal solution, with a goal to maximizing the expectation of an objective function. Instead of using sample averaging to approximate the expected value, which requires many expensive simulations, we developed bias-correction methods applied to the mean model. The key point of this approach is that information from distinct controls and model realizations can be used to correct the bias in the value obtained from the mean model, so that only simulations of the mean model are required during the optimization process. The effectiveness of the bias-corrected mean model is illustrated by two applications: flow optimization in a one-dimensional model and the drilling-order problem in a synthetic field model.

The problem in the first example is small enough that we can evaluate the Taylor series approximation of the expected value with a higher-order term. The results show that there is clearly a large bias in both linear and quadratic Taylor approximations, while the bias-correction approach gives better results, and the improvement is more apparent at a larger ensemble size. The second example extends the study from Paper A by considering geological uncertainty in the drilling-order problem. The estimation error of the expected NPV from the mean model is reduced from -9% to 0.56% after using the bias-correction methods. In the case with an ensemble of 100 realizations, learned heuristic search applied with the bias-corrected mean model requires only 246 simulations to compute the RO drilling sequence of 8 wells. Optimization results from both examples show that the RO solutions obtained using the bias-corrected mean model and the sample average approximation are of similar quality and are superior to the deterministic solutions for individual realizations and the mean reservoir model (i.e., linear Taylor approximation).

Paper C

Using the techniques presented in Papers A and B, we establish a framework for efficient optimal decision-making over the *current assessment of uncertainty*. To act optimally, we should also take into account the effect of future information before committing to a decision, i.e., the opportunity to improve the optimal strategy resulting from future uncertainty reduction. In Paper C, we developed a key-feature-based VOI analysis to efficiently account for the possibility of future learning through actions. This approach is a simplified VOI analysis obtained by identifying key actions and key observations for the optimization problem, i.e., focusing on the use of important future information to reduce the uncertainties with large influences on the optimal decisions, rather than using all observations to reduce all uncertainties.

The methodology is illustrated by the application of the drilling-order problem in a synthetic field model, for which the drilling sequence of wells is an important contributor to the reservoir's profitability and for which the optimal solution changes significantly with key reservoir features. The key observations are identified using supervised-learning algorithms, which allow automatic detection of the optimal observation combination and direct computation of the posterior probability distributions of key uncertainties based on Bayes' rule, so that there is no need for full history matching to update the entire reservoir model. Results show that accounting for the future learning through the key information obtained from key actions can efficiently improve the expected outcome of an optimal strategy. Taking the key action to acquire useful information for reducing key uncertainties is not always worthwhile, however, even if there is no explicit cost in obtaining that information. The cause of that is there may be a large indirect cost of information when taking the key action leads to a sub-optimal solution. Thus, it is necessary to consider both the possibility of key uncertainty reduction and the possibility of high expected NPV to determine whether it is worth taking the key action earlier in sequence to improve future decisions.

6.2   Outlook

The methodologies presented in this work can be extended to general sequential field development tasks, although they are illustrated by the application of drilling-sequence optimization problem. A few aspects require further consideration or research to ensure the methods are effective and efficient:

Learned heuristic search This approach is efficient for solving optimization problems with discrete actions. It can be used either to optimize a complete control sequence or only the first few actions at a reduced cost by limiting the search depth. The key

to making this method efficient is developing online learning mechanisms that can be used to improve a crude heuristic. The techniques presented in Section 3.2 work well when the initial evaluation-function value has an exponential trend and the ratio changes at a stationary rate. Thus, they might not be suitable for other reservoir applications or heuristics that have distinct properties. Since past observations are obtained sequentially, it might generally be more useful to build a distinct dynamic sequential model based on observed data to improve the evaluation function. For large optimization problems, we might need to design a practical space-reduction approach to find a valid solution quickly.

Bias-corrected mean model The major advantage of using the bias-corrected mean model for RO is that it requires only simulations in the mean model to obtain initial estimates of expected value, for which the bias can be corrected using the information from distinct controls and model realizations. To ensure the effectiveness of distance-based localization for bias correction, we need to design appropriate distance measures for the identification of sets of similar controls. When the actions that can be controlled are discrete and sequential, we can use a position-based distance metric (e.g., Manhattan distance) to measure the similarity between control sequences. The bias-corrected methods can be modified for use in multiple objective optimizations under uncertainty (e.g., standard deviation, percentiles, expected value). For example, it would be possible to compute an estimate of the variance of the objective-function value, by squaring the value of the objective function from the mean model and multiplying it by the variance of partial corrections.

Key-feature-based VOI analysis By properly accounting for the possibility of future uncertainty reduction, improvement of the optimal strategy could be achieved. A simplified VOI analysis, which considers only the important future information from key actions to improve future decisions, can be applied to consider future learning possibilities efficiently when making optimal decisions. The key point is to effectively identify key actions and key information associated with key uncertainties in the optimization problem. For small problems, key actions can be identified by estimating the reliability of information for reducing key uncertainties and the hidden cost of obtaining information from key actions. For large problems, it may be necessary to explore more generalized and efficient approaches for identifying key actions. Our method considers the effect of future information resulting from only the current decision step. If the key information can only be obtained by taking at least two actions, we can extend the VOI analysis by considering the possibility of future learning through the following two decision steps. This approach, however, will increase the computational cost of making optimal decisions and the hidden cost of obtaining key information caused by sub-optimal solutions, which will be constrained to more past decisions.

Bibliography

[1] C. C. Aggarwal, A. Hinneburg, and D. A. Keim, *On the surprising behavior of distance metrics in high dimensional space*, in Database Theory — ICDT 2001, vol. 1973 of Lecture Notes in Computer Science, Berlin, Heidelberg, 2001, Springer Berlin Heidelberg, pp. 420–434.

[2] S. Aine, S. Swaminathan, V. Narayanan, V. Hwang, and M. Likhachev, *Multi-Heuristic A**, International Journal of Robotics Research, 35 (2016), pp. 224–243.

[3] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, *Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations*, SIAM/ASA Journal on Uncertainty Quantification, 5 (2017), pp. 1166–1192.

[4] E. G. D. Barros, P. M. J. V. den Hof, and J. D. Jansen, *Informed production optimization in hydrocarbon reservoirs*, Optimization and Engineering, 21 (2020), pp. 25–48.

[5] E. G. D. Barros, J. D. Jansen, and P. M. J. Van den Hof, *Value of information in parameter identification and optimization of hydrocarbon reservoirs*, IFAC-PapersOnLine, 48 (2015), pp. 229–235.

[6] E. G. D. Barros, O. Leeuwenburgh, P. M. J. Van den Hof, and J. D. Jansen, *Value of multiple production measurements and water front tracking in closed-loop reservoir management*, in SPE Reservoir Characterisation and Simulation Conference and Exhibition, Abu Dhabi, UAE, 2015, Society of Petroleum Engineers.

[7] E. G. D. Barros, S. Maciel, R. J. Moraes, and R. M. Fonseca, *Automated clustering based scenario reduction to accelerate robust life-cycle optimization*, in ECMOR XVI-16th European Conference on the Mathematics of Oil Recovery, Barcelona, Spain, 2018, European Association of Geoscientists & Engineers.

[8] E. G. D. Barros, P. M. J. Van den Hof, and J. D. Jansen, *Value of information in closed-loop reservoir management*, Computational Geosciences, 20 (2016), pp. 737–749.

[9] B. L. Beckner and X. Song, *Field development planning using simulated annealing-optimal economic well scheduling and placement*, in SPE Annual Technical Conference and Exhibition, Dallas, Texas, 1995, Society of Petroleum Engineers.

[10] R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, NJ, USA, 1 ed., 1957.

[11] H. BEYER AND B. SENDHOFF, *Robust optimization – a comprehensive survey*, Computer Methods in Applied Mechanics and Engineering, 196 (2007), pp. 3190–3218.

[12] R. B. BRATVOLD, J. E. BICKEL, AND H. P. LOHNE, *Value of information in the oil and gas industry: Past, present, and future*, SPE Reservoir Evaluation & Engineering, 12 (2009), pp. 630–638.

[13] C. BUDAK, D. AGRAWAL, AND A. EL ABBADI, *Limiting the spread of misinformation in social networks*, in Proceedings of the 20th International Conference on World Wide Web, New York, NY, USA, 2011, Association for Computing Machinery, p. 665–674.

[14] M. A. CARDOSO AND L. J. DURLOFSKY, *Linearized reduced-order models for subsurface flow simulation*, Journal of Computational Physics, 229 (2010), pp. 681–700.

[15] M. A. CARDOSO AND L. J. DURLOFSKY, *Use of reduced-order modeling procedures for production optimization*, SPE Journal, 15 (2010), pp. 426–435.

[16] A. CASSANDRA, M. L. LITTMAN, AND N. L. ZHANG, *Incremental pruning: A simple, fast, exact method for partially observable markov decision processes*, in Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence, San Francisco, CA, USA, 1997, Morgan Kaufmann Publishers Inc., pp. 54–61.

[17] P. CHAKRABARTI, S. GHOSE, A. PANDEY, AND S. DE SARKAR, *Increasing search efficiency using multiple heuristics*, Information Processing Letters, 30 (1989), pp. 33–36.

[18] B. CHEN, J. HE, X.-H. WEN, W. CHEN, AND A. C. REYNOLDS, *Uncertainty quantification and value of information assessment using proxies and Markov chain Monte Carlo method for a pilot project*, Journal of Petroleum Science and Engineering, 157 (2017), pp. 328–339.

[19] C. CHEN, Y. WANG, G. LI, AND A. C. REYNOLDS, *Closed-loop reservoir management on the Brugge test case*, Computational Geosciences, 14 (2010), pp. 691–703.

[20] J. CHEN, R. C. HOLTE, S. ZILLES, AND N. R. STURTEVANT, *Front-to-end bidirectional heuristic search with near-optimal node expansions*, in 26th International Joint Conference on Artificial Intelligence, 2017, pp. 489–495.

[21] P. CHEN, U. VILLA, AND O. GHATTAS, *Taylor approximation and variance reduction for PDE-constrained optimal control under uncertainty*, Journal of Computational Physics, 385 (2019), pp. 163–186.

[22] Y. CHEN, *Efficient ensemble-based closed-loop production optimization*, PhD thesis, University of Oklahoma, Oklahoma, USA, 2008.

[23] Y. CHEN, D. S. OLIVER, AND D. ZHANG, *Efficient ensemble-based closed-loop production optimization*, SPE Journal, 14 (2009), pp. 634–645.

[24] H.-T. CHENG, *Algorithms for Partially Observable Markov Decision Processes*, PhD thesis, University of British Columbia, Vancouver, 1988.

[25] P. Cunningham and S. Begg, *Using the value of information to determine optimal well order in a sequential drilling program*, The American Association of Petroleum Geologists, 92 (2008), pp. 1393–1402.

[26] J. Darlington, C. C. Pantelides, B. Rustem, and B. A. Tanyi, *An algorithm for constrained nonlinear optimization under uncertainty*, Automatica, 35 (1999), pp. 217–228.

[27] J. Darlington, C. C. Pantelides, B. Rustem, and B. A. Tanyi, *Decreasing the sensitivity of open-loop optimal solutions in decision making under uncertainty*, European Journal of Operational Research, 121 (2000), pp. 343–362.

[28] D. Denney, *Pros and cons of applying a proxy model as a substitute for full reservoir simulations*, Journal of Petroleum Technology, 62 (2010), pp. 634–645.

[29] J. E. Doran and D. Michie, *Experiments with the graph traverser program*, Proceedings of the Royal Society (A), 294 (1966), pp. 235–259.

[30] A. Doucet, S. Godsill, and C. Andrieu, *On sequential Monte Carlo sampling methods for Bayesian filtering*, Statistics and Computing, 10 (2000), pp. 197–208.

[31] L. J. Durlofsky, *Upscaling and gridding of fine scale geological models for flow simulation*, in 8th International Forum on Reservoir Simulation Iles Borromees, Stresa, Italy, 2005.

[32] R. E. Korf, *Depth-limited search for real-time problem solving*, Real-Time Systems, 2 (1990), pp. 7–24.

[33] R. Ebend and R. Drechsler, *Weighted A* search – unifying view and application*, Artificial Intelligence, 173 (2007), pp. 1310–1342.

[34] J. Eidsvik, G. Dutta, T. Mukerji, and D. Bhattacharjya, *Simulation-regression approximations for value of information analysis of geophysical data*, Mathematical Geosciences, 49 (2017), pp. 467–491.

[35] R. M. Fonseca, B. Chen, J. D. Jansen, and A. C. Reynolds, *A stochastic simplex approximate gradient (StoSAG) for optimization under uncertainty*, International Journal for Numerical Methods in Engineering, 109 (2017), pp. 1756–1776.

[36] S. French and D. R. Insua, *Statistical Decision Theory*, Edward Arnold, 2000.

[37] A. T. F. S. Gaspar, E. Muñoz Mazo, and D. Schiozer, *Case study of the structure of the process for production strategy selection*, International Journal of Modeling and Simulation for the Petroleum Industry, 4 (2011), pp. 9–15.

[38] G. Gaspari and S. E. Cohn, *Construction of correlation functions in two and three dimensions*, Q. J. R. Meteorol. Soc., 125 (1999), pp. 723–757.

[39] T. Goda, K. Ohno, and K. Sato, *Value of information in optimizing reservoir development under geological uncertainty*, Journal of the Japan Petroleum Institute, 60 (2017), pp. 41–52.

[40] C. J. Grayson, *Decisions under uncertainty: Drilling decisions by oil and gas operators*, Harvard University Press, Boston, Massachusetts, 1960.

[41] R. Hamming, *Coding and Information Theory*, Prentice-Hall, 1980.

[42] R. Hanea, P. Casanova, L. Hustoft, R. Bratvold, R. Nair, C. W. Hewson, O. Leeuwenburgh, and R.-M. Fonseca, *Drill and learn: A decision-making work flow to quantify value of learning*, SPE Reservoir Evaluation & Engineering, 22 (2019), pp. 1131–1143.

[43] R. G. Hanea, R. M. Fonseca, C. Pettan, M. O. Iwajomo, K. Skjerve, L. Hustoft, A. G. Chitu, and F. Wilschut, *Decision maturation using ensemble based robust optimization for field development planning*, in ECMOR XV-15th European Conference on the Mathematics of Oil Recovery, Amsterdam, Netherlands, 2016, European Association of Geoscientists & Engineers.

[44] J. Hanley and B. McNeil, *The meaning and use of the area under a receiver operating characteristic (roc) curve*, Radiology, 143 (1982), pp. 29–36.

[45] E. A. Hansen and R. Zhou, *Anytime heuristic search*, Journal of Artificial Intelligence Research, 28 (2007), pp. 267–297.

[46] E. A. Hansen, S. Zilberstein, and V. A. Danilchenko, *Anytime heuristic search: First results*, Tech. Rep. 97–50, Computer Science Department, University of Massachussetts Amherst, 1997.

[47] P. E. Hart, N. J. Nilsson, and B. Raphael, *A formal basis for the heuristic determination of minimum cost paths*, IEEE Transactions on Systems Science and Cybernetics, 4(2) (1968), pp. 100–107.

[48] J. He, P. Sarma, E. Bhark, S. Tanaka, B. Chen, X.-H. Wen, and J. Kamath, *Quantifying expected uncertainty reduction and value of information using ensemble-variance analysis*, SPE Journal, 23 (2018), pp. 428–448.

[49] M. Helmert, *The fast downward planning system*, Journal of Artificial Intelligence Research, 26 (2006), pp. 191–246.

[50] A. Hong, R. Bratvold, P. Thomas, and R. Hanea, *Value-of-information for model parameter updating through history matching*, Journal of Petroleum Science and Engineering, 165 (2018), pp. 253–268.

[51] J. D. Jansen, D. R. Brouwer, G. Nævdal, and C. P. J. W. Van Kruijsdijk, *Closed-loop reservoir management*, First Break, 23 (2005), pp. 43–48.

[52] J. D. Jansen and L. J. Durlofsky, *Use of reduced-order models in well control optimization*, Optimization and Engineering, (2016), pp. 105–132.

[53] M. A. Jaro, *Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida*, Journal of the American Statistical Association, 84 (1989), pp. 414–420.

[54] M. Jesmani, B. Jafarpour, M. C. Bellout, and B. Foss, *A reduced random sampling strategy for fast robust well placement optimization*, Journal of Petroleum Science and Engineering, 184 (2020), p. 106414.

[55] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, *Planning and acting in partially observable stochastic domains*, Artificial Intelligence, 101 (1998), pp. 99–134.

[56] S. Kim, R. Pasupathy, and S. G. Henderson, *A guide to sample average approximation*, in Handbook of Simulation Optimization, Springer, 2015, pp. 207–243.

[57] L. Kish, *Survey sampling*, John Wiley and Sons, Inc, New York, 1965.

[58] A. J. Kleywegt, A. Shapiro, and T. Homem-de Mello, *The sample average approximation method for stochastic discrete optimization*, SIAM Journal on Optimization, 12 (2002), pp. 479–502.

[59] S. Koeniga, M. Likhachev, and D. Furcy, *Lifelong Planning A**, Artificial Intelligence Journal, 155 (2004), pp. 93–146.

[60] R. Kohavi and G. H. John, *Wrapper for feature subset selection*, Artificial Intelligence, 97 (1997), pp. 273–324.

[61] T. Korenius, J. Laurikkala, and M. Juhola, *On principal component analysis, cosine and euclidean measures in information retrieval*, Information Sciences, 177 (2007), pp. 4893–4905.

[62] R. E. Korf, *Depth-first iterative-deepening: An optimal admissible tree search*, Artificial Intelligence, 27 (1985), pp. 97–109.

[63] L. F. Lamas, V. E. Botechia, D. J. Schiozer, and M. Delshad, *Optimization for drilling schedule of wells in the development of heavy oil reservoirs*, Brazilian Journal of Petroleum and Gas, 11 (2017).

[64] C. Lazar, J. Taminau, S. Meganck, D. Steenhoff, A. Coletta, C. Molter, V. De Schaetzen, R. Duque, H. Bersini, and A. Nowé, *A survey on filter techniques for feature selection in gene expression microarray analysis*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 9 (2012), pp. 1106–1119.

[65] O. Leeuwenburgh, A. G. Chitu, R. Nair, P. J. P. Egberts, L. Ghazaryan, T. Feng, and L. Hustoft, *Ensemble-based methods for well drilling sequence and time optimization under uncertainty*, in ECMOR XV-15th European Conference on the Mathematics of Oil Recovery, Amsterdam, Netherlands, 2016, European Association of Geoscientists & Engineers.

[66] V. I. Levenshtein, *Binary codes capable of correcting deletions, insertions and reversals.*, Soviet Physics Doklady, 10 (1966), pp. 707–710.

[67] R. J. Lorentzen, A. Berg, G. Naevdal, and E. H. Vefring, *A new approach for dynamic optimization of water flooding problems*, in SPE Intelligent Energy Conference and Exhibition, Amsterdam, Netherlands, 2006, pp. 11–13.

[68] J. Maharshi, L. Ronit, and P. Sonal, *Comparative analysis of search algorithms*, International Journal of Computer Applications, 179 (2018), pp. 40–43.

[69] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker, Inc., New York, 1992.

[70] A. E. MOTTER, *Cascade control and defense in complex networks*, Physical Review Letters, 93 (2004), p. 098701.

[71] H. PENG, F. LONG, AND C. DING, *Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy*, IEEE transactions on pattern analysis and machine intelligence, 27 (2005), pp. 1226–1238.

[72] J. PINEAU, G. GORDON, AND S. THRUN, *Point-based value iteration: An anytime algorithm for POMDPs*, in Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI'03, San Francisco, CA, USA, 2003, Morgan Kaufmann Publishers Inc., pp. 1025–1030.

[73] I. POHL, *Heuristic search viewed as path finding in a graph*, Artificial Intelligence, 1 (1970), pp. 193–204.

[74] I. POHL, *The avoidance of (relative) catastrophe, heuristic competence, genuine dynamic weighting and computational issues in heuristic problem solving*, in Proceedings of the 3rd International Joint Conference on Artificial Intelligence, Stanford, USA, 1973, Morgan Kaufmann Publishers Inc., pp. 12–17.

[75] A. E. PRIEDITIS, *Machine discovery of effective admissible heuristics*, Machine Learning, 12 (1993), pp. 117–141.

[76] A. E. PRIEDITIS AND B. JANAKIRAMAN, *Generating effective admissible heuristic by abstraction and reconstitution*, in Proceedings of the 11th National Conference on Artificial Intelligence, 1993, pp. 743–748.

[77] S. RAHIM AND Z. LI, *Well placement optimization with geological uncertainty reduction*, IFAC-PapersOnLine, 48 (2015), pp. 57–62. 9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2015.

[78] H. RAIFFA AND R. SCHLAIFER, *Applied statistical decision theory*, Harward University Press, Boston, USA, 1961.

[79] C. P. ROBERT AND G. CASELLA, *Introducing Monte Carlo Methods with R*, Springer, 2010.

[80] G. RÖGER AND M. HELMERT, *The more, the merrier: Combining heuristic estimators for satisficing planning*, in Proceedings of the 20th international conference on automated planning and scheduling, Toronto, Ontario, Canada, 2010, AAAI, pp. 246–249.

[81] S. RONALD, *More distance functions for order-based encodings*, in 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence, 1998, pp. 558–563.

[82] S. RUSSELL AND P. NORVIG, *Artificial Intelligence: A Modern Approach*, Pearson Education Limited, Harlow, United Kingdom, 3 ed., 2016.

[83] P. SARMA, L. J. DURLOFSKY, K. AZIZ, AND W. H. CHEN, *Efficient real-time reservoir management using adjoint-based optimal control and model updating*, Computational Geosciences, 10 (2006), pp. 3–36.

[84] T. SAYAG, S. FINE, AND Y. MANSOUR, *Combining multiple heuristics*, in STACS 2006, Proceedings, B. Durand and W. Thomas, eds., vol. 3884 of Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2006, pp. 242–253.

[85] T. SCHIAVINOTTO AND T. STÜTZLE, *A review of metrics on permutations for search landscape analysis*, Computers & Operations Research, 34 (2007), pp. 3143–3153.

[86] R. SCHLAIFER, *Probability and statistics for business decisions : an introduction to managerial economics under uncertainty*, McGraw-Hill Book Company, New York, 1959.

[87] S. SHARMA, *Applied Multivariate Techniques*, John Wiley & Sons, Inc., New York, USA, 1996.

[88] M. G. SHIRANGI AND L. J. DURLOFSKY, *Closed-loop field development under uncertainty by use of optimization with sample validation*, SPE Journal, 20 (2015), pp. 908–922.

[89] V. L. S. SILVA, A. A. EMERICK, P. COUTO, AND J. L. D. ALVES, *History matching and production optimization under uncertainties–application of closed-loop reservoir management*, Journal of Petroleum Science and Engineering, 157 (2017), pp. 860–874.

[90] T. SMITH AND R. SIMMONS, *Heuristic search value iteration for POMDPs*, in Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence, UAI '04, Arlington, Virginia, USA, 2004, AUAI Press, pp. 520–527.

[91] E. J. SONDIK, *The optimal control of partially observable Markov processes*, PhD thesis, Stanford University, Stanford, CA, USA, 1971.

[92] E. J. SONDIK, *The optimal control of partially observable markov processes over the infinite horizon: Discounted costs*, Operations Research, 26 (1978), pp. 282–304.

[93] A. STENTZ, *The focussed D* algorithm for real-time replanning*, in Proceedings of the International Joint Conference on Artificial Intelligence, 1995.

[94] A. S. STORDAL, S. P. SZKLARZ, AND O. LEEUWENBURGH, *A theoretical look at ensemble-based optimization in reservoir management*, Mathematical Geosciences, 48 (2016), pp. 399–417.

[95] J. T. THAYER, A. J. DIONNE, AND W. RUML, *Learning inadmissible heuristics during search*, in Proceedings of the 21st International Conference on Automated Planning and Scheduling, Freiburg, Germany, 2011, AAAI Press, pp. 250–257.

[96] S. THRUN, *Monte carlo POMDPs*, in Proceedings of the 12th International Conference on Neural Information Processing Systems, NIPS'99, Cambridge, MA, USA, 1999, MIT Press, pp. 1064–1070.

[97] R. R. TORRADO, J. RIOS, AND G. TESAURO, *Optimal sequential drilling for hydrocarbon field development planning*, in Proceedings of the 29th AAAI Conference on Innovative Applications (IAAI-17), 2017, pp. 4734–4739.

[98] R. VAN DOREN, JORN F. M.AND MARKOVINOVIĆ AND J. D. JANSEN, *Reduced-order optimal control of water flooding using proper orthogonal decomposition*, Computational Geosciences, 10 (2006), pp. 137–158.

[99] G. VAN ESSEN, M. ZANDVLIET, P. VAN DEN HOF, O. BOSGRA, AND J. D. JANSEN, *Robust waterflooding optimization of multiple geological scenarios*, SPE Journal, 14 (2009), pp. 202–210.

[100] J. VON NEUMANN AND O. MORGENSTERN, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, USA, 1944.

[101] C. WANG, G. LI, AND A. C. REYNOLDS, *Production optimization in closed-loop reservoir management*, SPE Journal, 14 (2009), pp. 506–523.

[102] L. WANG AND D. S. OLIVER, *Efficient optimization of well drilling sequence with learned heuristics*, SPE Journal, 24 (2019), pp. 2111–2134.

[103] L. WANG AND D. S. OLIVER, *Fast robust optimization using bias correction applied to the mean model*, Computational Geosciences, 25 (2021), pp. 475–501.

[104] L. WANG AND D. S. OLIVER, *Improving sequential decisions – efficiently accounting for future learning*, Journal of Petroleum Science and Engineering, 205 (2021), p. 108770.

[105] C. WILT AND W. RUML, *Effective heuristics for suboptimal best-first search*, Journal Of Artificial Intelligence Research, 57 (2016), pp. 273–306.

[106] S. YOON, A. FERN, AND R. GIVEN, *Learning heuristic functions from relaxed plans*, in Proceedings of the Sixteenth International Conference on International Conference on Automated Planning and Scheduling, 2006, pp. 162–170.

[107] R. ZHOU AND E. A. HANSEN, *Multiple sequence alignment using anytime A**, in Eighteenth National Conference on Artificial Intelligence, USA, 2002, American Association for Artificial Intelligence, pp. 975–976.

[108] K. ÅSTRÖM, *Optimal control of markov processes with incomplete state information*, Journal of Mathematical Analysis and Applications, 10 (1965), pp. 174–205.

Part II

Scientific results

Paper A

Efficient Optimization of Well Drilling Sequence with Learned Heuristics

Lingya Wang, Dean S. Oliver

Paper B

Fast robust optimization using bias correction applied to the mean model

Lingya Wang, Dean S. Oliver

Fast robust optimization using bias correction applied to the mean model

Lingya Wang[a,b], Dean S. Oliver[a]

[a] - NORCE Norwegian Research Centre AS, Norway
[b] - University of Bergen, Norway

Abstract

Ensemble methods are remarkably powerful for quantifying geological uncertainty. However, the use of the ensemble of reservoir models for robust optimization (RO) can be computationally demanding. The straightforward computation of the expected net present value (NPV) requires many expensive simulations. To reduce the computational burden without sacrificing accuracy, we present a fast and effective approach that requires only simulation of the mean reservoir model with a bias correction factor. Information from distinct controls and model realizations can be used to estimate bias for different controls. The effectiveness of various bias-correction methods and a linear or quadratic approximation is illustrated by two applications: flow optimization in a one-dimensional model and the drilling-order problem in a synthetic field model. The results show that the approximation of the expected NPV from the mean model is significantly improved by estimating the bias correction factor, and that RO with mean model bias correction is superior to both RO performed using a Taylor series representation of uncertainty and deterministic optimization from a single realization. Use of the bias-corrected mean model to account for model uncertainty allows the application of fairly general optimization methods. In this paper, we apply a nonparametric online learning methodology (learned heuristic search) for efficiently computing an optimal or near-optimal robust drilling sequence on the REEK Field example. This methodology can be used either to optimize a complete drilling sequence or to optimize only the first few wells at a reduced cost by limiting the search depths.

Keywords: Robust optimization; Bias correction; Geological uncertainty; Ensemble-based methods; Drilling sequence optimization; Sequential-decision making

Introduction

In the development of a hydrocarbon field, optimization is an important process that can substantially improve the profitability through reduction in production and development costs or by increasing or accelerating the production of hydrocarbons. The increase in profitability of a field for a given change in controls is predicted using a reservoir simulation model, in which the geological uncertainties such as porosity, permeability, and fault transmissibility have critical effects on production forecasts. Ensemble methods are remarkably powerful for quantifying the geological uncertainty. Van Essen et.al [39] demonstrated that robust optimization (RO) performed using the expectation of NPV

by averaging over an ensemble of reservoir models outperforms the nominal optimization (NO) from a single realization. However, the straightforward computation of the expected value requires many expensive simulations for each control variable setting.

Many approaches have been proposed to reduce the cost of simulation-based RO. These methods can be mainly classified into two categories: accelerating the reservoir model simulations and reducing the number of simulations required to obtain the maximum. A simplified model such as one obtained from reduced-order modeling and upscaling model parameters [21, 4, 5, 38, 13] is a computationally inexpensive substitute for full reservoir simulation. However, such a proxy model is generally less accurate due to the challenges posed by the high nonlinearity in the reservoir model; consequently, RO performed in a simplified model is likely to yield a suboptimal solution [11]. Techniques for reducing the number of simulations needed for optimization of the expected value of an objective function can be further classified into two main categories: improving the optimization algorithms for certain problems and reducing the cost of evaluating the expected value.

Ensemble-based optimization [8, 15] is an efficient gradient-based approach for fairly general RO problems with continuous variables. The cost of evaluating the gradient of the expected value of the objective function is reduced by using the same ensemble to represent the reservoir model uncertainty and to compute the gradient of the expectation of the objective function with respect to control variable settings. One downside to a gradient-based approach is that it is easily possible to get stuck in a local minimum (or maximum). For optimization problems with discrete control variables, Wang and Dean [40] proposed a nonparametric online learning methodology with heuristic controls to efficiently solve sequential decision-making problems. This approach can reduce the number of iterations required during the optimization process. Nevertheless, the amount of computation required for RO performed using sample average approximation (SAA) over a large ensemble could still be expensive since the cost increases linearly with the number of realizations required to represent the uncertainty in model properties.

There are several possible ways to compute the expected value at a lower cost than SAA. One is to reduce the representation of uncertainty through model selection, in which a subset of the ensemble members are used for optimization [23, 37, 32]. A small subset of model realizations may not span the uncertainties accurately, leading to a suboptimal solution. To address this issue, Barros et al. [2] proposed an automated scenario reduction approach for selecting a subset that contains an optimal number of realizations that are able to capture the range of the uncertainties of the full ensemble. A much faster approach is to evaluate the objective function in the mean reservoir model. Chen et al. [6] obtained relatively good results by using the ensemble mean model updated from the ensemble Kalman filter (EnKF) data assimilation for production optimization. Their optimized design increased the expected NPV compared to the reference case, but was not as good as the optimized results obtained using SAA. One major drawback to direct use of the mean model is the large bias in predicted reservoir behavior that may result from the use of a model with reduced heterogeneity, in which case the mean model may offer a poor estimate of the expected NPV. To improve this approximation, one possible approach is to modify the representation of uncertainty using a Taylor series expansion of the objective function [9, 10, 3, 1, 7]. The first term of Taylor expansion can be inexpensively obtained using the mean of reservoir model realizations. To accurately estimate the expected value, the Taylor series expansion generally includes higher-order

terms (e.g., a quadratic or cubic term). The computation of higher-order derivatives is impractical for most real reservoir applications, however.

Instead of using higher-order terms to improve the estimation accuracy, we develop a fast and efficient approach to correcting the bias in the NPV obtained from the mean reservoir model by estimating a multiplicative bias correction factor based on the information from distinct controls and individual model realizations. To accurately estimate the expected NPV, we apply distance-based localization to estimate the bias for specific controls, considering the similarity between samples and control variables in terms of the bias correction factor. To avoid poor estimates caused by sampling error due to small sample sizes, a regularization term based on the average value and the variance of the bias correction factor is used to reduce the sensitivity of the estimates to the taper length, thereby allowing more accurate estimates to be generated for a wider range of taper lengths. The initial sampling of control and model realizations is used to create an ensemble of partial corrections factors. During the RO phase, when it is necessary to estimate the expected value of NPV for a control, we only require simulation of the control applied to the mean model. The bias correction factor is estimated from partial bias correction factors on similar controls. Hence, after creation of the initial ensemble, RO performed using bias-correction methods requires only one additional simulation of the mean reservoir model at each iteration, which is much less than the effort required in SAA. This bias-correction methodology can be applied to fairly general problems of optimizing the expected value of an objective function. But an appropriate distance metric to measure the similarity of controls in terms of the bias correction factor is required, which is specific to the problem at hand.

The performance of various bias-correction methods and a linear or quadratic approximation is investigated in two applications. The objective in the first example is to locate an injection well such that flow rate for a fixed pressure in a one-dimensional model is minimized; this problem is small enough that we can evaluate both linear and quadratic Taylor approximations of the objective function. The second example is to maximize the expected NPV in a synthetic field model by optimizing the drilling sequence of wells at fixed locations. Here we use well-position based distance to measure the similarity of drilling sequences for the bias correction factor. We formulate the sequential drilling optimization problem as one of finding a path with the maximum reward in a decision tree, and apply learned heuristic search [40] with mean model bias correction (MMBC) to compute the RO drilling sequence under geological uncertainty and to optimize only the first few wells at a reduced cost by limiting the search depths.

This paper is organized as follows. Section 2 describes the bias-correction methods for estimating the expected NPV and the learned heuristic search method for optimizing either a complete drilling sequence or only the first few wells. Sections 3 and 4 describe two numerical case studies (i.e., flow optimization in a one-dimensional model and drilling-order problem in a synthetic model). The conclusions are presented in Section 5.

Methodology

Estimation of expected value

Correction factor

In this work, we use maximization of the NPV as the objective for an optimal reservoir management problem. Because reservoir characterization is always incomplete, the optimization problem is based on a reservoir model with uncertainty in parameter values. An appropriate approach to account for the uncertainty is to use the expected NPV over an ensemble of reservoir models that have been sampled from the probability distribution for model parameters. The expected value of the objective can be written as

$$E[f(x,m)] \approx \bar{f}(x) = \frac{1}{N_e} \sum_{j=1}^{N_e} f(x, m_j), \tag{1}$$

where x is a vector of control inputs; $m \in \mathcal{R}^m$ is an m-dimensional vector of uncertain model parameters; j is the index of individual realizations; and N_e indicates the number of reservoir models.

This ensemble-based average value can provide a good approximation of the expected NPV if the ensemble of model realizations is sufficiently large [26, 24]. RO performed using such straightforward estimation of the expected NPV, however, requires many expensive simulations when the ensemble size N_e or the number of iterations needed for the optimization process is large. Instead of computing the expectation of the objective function by using SAA, one could consider optimizing the expectation for a linear or quadratic approximation of the objective function [9, 10, 3, 1]. To second order, the Taylor expansion of the objective function is

$$\begin{aligned} f^{\text{quad}}(x,m) = f(x, \bar{m}) + f_m(x, \bar{m})(m - \bar{m}) \\ + \frac{1}{2}(m - \bar{m})^{\text{T}} f_{mm}(x, \bar{m})(m - \bar{m}), \end{aligned} \tag{2}$$

where we have neglected higher-order terms in the expansion. f_m and f_{mm} are first and second derivatives of f with respect to uncertain model parameter m, respectively.

If m is distributed as multivariate Gaussian with mean $\bar{m}$ and covariance C, then the expected value of the quadratic approximation of the objective can be shown [31] to be

$$E[f^{\text{quad}}(x,m)] = f(x, \bar{m}) + \frac{1}{2} \text{tr}(C^{1/2} f_{mm}(x, \bar{m}) C^{1/2}). \tag{3}$$

A possible advantage of this approach to approximating the expected value of the objective function is that optimization of the expectation does not require evaluation of controls applied to a large number of Monte Carlo samples [1]. Computing $f(x, \bar{m})$ in Eq. 3 will be easy as it only requires the mean of the realizations. Computing $f_{mm}(x, \bar{m})$ is more difficult as it requires the second derivative of the objective function with respect to model parameters. Although it might be possible in some cases to approximate the second derivative, computing second derivatives will be impractical for most reservoir applications. The linear approximation seems more likely to be feasible in practice,

$$f^{\text{lin}}(x,m) = f(x, \bar{m}) + \langle f_m(x, \bar{m}), m - \bar{m} \rangle. \tag{4}$$

Assuming that m is distributed as Gaussian, the expectation over m of the linear approximation to the objective is simply a function of $\bar{m}$ and x,

$$E[f^{\text{lin}}(x, m)] = f(x, \bar{m}). \tag{5}$$

The key point of this is that the mean of the ensemble can be used for the optimization, instead of performing the optimization on the ensemble of realizations.

Such an approach has a major advantage over SAA in that it significantly reduces the number of simulations required for evaluating the expected value. However, since the NPV is generally not a linear function of the uncertain model parameters m, NPV from the mean model is generally not the same as the expected value over uncertainty space, $f(x, E[m]) \neq E[f(x, m)]$. If we want to use the mean model for optimization problems without sacrificing the accuracy of the estimated expected NPV, we need a method for improving the approximation of expected NPV from the mean model.

In this paper, we compute a multiplicative correction factor between the ensemble average of values $\bar{f}(x_i)$ and the value obtained from the mean model $f(x_i, \bar{m})$ at a fixed control x_i,

$$\bar{f}(x_i) = \alpha(x_i) f(x_i, \bar{m}), \tag{6}$$

where $\alpha(x_i)$ is the correction factor for a fixed control x_i. Instead of directly computing the ensemble average of values of the objective function, we develop an approximation of expected value $\bar{f}(x)$ by estimating the correction factor $\alpha(x_i)$ of control x_i. If it were feasible to compute the value of the NPV at control x_i for all samples of model parameters, then $\alpha(x_i)$ could be computed in the following way,

$$\alpha(x_i) = \frac{1}{N_e} \sum_{j=1}^{N_e} \frac{f(x_i, m_j)}{f(x_i, \bar{m})}, \tag{7}$$

where $f(x_i, m_j)$ is the economic value at control x_i of an individual realization m_j.

For each individual model realization m_j and control x_i, we define a partial correction factor $\beta(x_i, m_j, \bar{m})$,

$$\beta_{ij} = \beta(x_i, m_j, \bar{m}) = \frac{f(x_i, m_j)}{f(x_i, \bar{m})}. \tag{8}$$

The correction factor $\alpha(x_i)$ at control x_i of all ensemble realizations (Eq. 7) can be written in terms of the partial correction factors,

$$\alpha(x_i) = \frac{1}{N_e} \sum_{j=1}^{N_e} \beta_{ij}. \tag{9}$$

We also define the mean value of the correction factor $\alpha(x_i)$,

$$\bar{\alpha} \approx \frac{1}{N_x} \sum_{i=1}^{N_x} \alpha(x_i) = \frac{1}{N_x N_e} \sum_{i=1}^{N_x} \sum_{j=1}^{N_e} \frac{f(x_i, m_j)}{f(x_i, \bar{m})}$$
$$= \frac{1}{N_x N_e} \sum_{i=1}^{N_x} \sum_{j=1}^{N_e} \beta_{ij}, \tag{10}$$

where N_x is the number of relevant controls. Straightforward application of Eq. 10 requires $N_x \times (N_e + 1)$ evaluations of $f(x, m)$ to compute the average correction factor $\bar{\alpha}$ from N_x different controls on an ensemble of N_e model realizations.

A Monte Carlo estimate of $\bar{\alpha}$ can be obtained at a much lower cost by sampling control x_j uniformly from the space of all possible controls, and sampling reservoir realizations m_k from the space of conditional realizations. Then $\bar{\alpha}$ can be estimated using the following formula,

$$\bar{\alpha} \approx \frac{1}{N_x} \sum_{j=1}^{N_x} \frac{f(x_j, m_j)}{f(x_j, \bar{m})} = \frac{1}{N_x} \sum_{j=1}^{N_x} \beta(x_j, m_j, \bar{m}), \tag{11}$$

where $\beta(x_j, m_j, \bar{m})$ is the observed value at control x_j, which requires two simulations, i.e., apply control x_j to a random individual realization m_j and apply control x_j to the mean reservoir model $\bar{m}$. Therefore, it would require $2 \times N_x$ simulations to obtain a set of observations β from N_x distinct controls.

For N_e reservoir model realizations, we can sample N_e distinct controls to obtain observations of β so that each realization will provide one observed value of β. In that case, the number of observed values of β is the same as the ensemble size (i.e., $N_x = N_e$). In some cases, more observations might be needed to obtain a good approximate of $\bar{\alpha}$ (e.g., obtain two observations of β from each realization (i.e, $N_x = 2N_e$) or use a larger ensemble size). If the value of β is similar for most realizations, we can use information from a subset of the ensemble members and a smaller number of distinct controls (i.e., $N_x < N_e$) to estimate the bias.

Although an estimate of $\bar{\alpha}$ can be efficiently obtained from the observed values of β, the accuracy of an estimate of the expected NPV obtained using $\bar{\alpha}$ for correction is limited by the variability in α. Use of the bias correction factor obtained by averaging samples of β (Eq. 11) will result in the same correction factor being used for all controls, even though the correct values of α for some of the controls may be far from $\bar{\alpha}$. In such a case, the accuracy level of the estimates from different controls is limited by the actual $\bar{\alpha}$ value. Nevertheless, if the variability in $\alpha(x_i)$ is small as control x_i is varied, then it is possible that $\bar{\alpha}$ can provide a useful approximation to $\alpha(x_i)$ for estimation of the value of $\bar{f}(x_i)$ from the value of $f(x_i, \bar{m})$.

Distance-based localization

In general, however, we expect that an estimate of the correction factor will be better if it is primarily based on information from similar control variables. Thus, we expect that a weighted estimate will be better than an unweighted estimate. In our work, we use weighted linear estimation. Suppose that N_e distinct controls are applied to N_e individual realizations and the mean model for generating a set of observations β. The weighted linear estimate $\hat{\alpha}(x_i)$ at control x_i is defined as

$$\hat{\alpha}(x_i) = \frac{\sum_{j=1}^{N_e} \omega(x_i, x_j) \beta(x_j, m_j, \bar{m})}{\sum_{j=1}^{N_e} \omega(x_i, x_j)}, \tag{12}$$

where $\beta(x_j, m_j, \bar{m})$ is the observed correction factor at a random control x_j applied to an individual realization m_j and the mean model $\bar{m}$, and $\omega(x_i, x_j)$ is the weight for $\beta(x_j, m_j, \bar{m})$. The weights, $\omega(x_i, x_j)$, should depend on a measure of similarity, or distance measure, between controls x_i and x_j.

With an appropriate measure of distance between control sequences, weights are assigned such that $\beta(x_j, m_j, \bar{m})$ at shorter distances will have higher weights while

partial correction factors for controls that are more dissimilar will have smaller weights. Lacking information about the correlation of β with distance between control sequences, we use the Gaspari-Cohn taper function [17] to compute distance-dependent weights,

$$\rho(\delta, L) = \begin{cases} -\frac{1}{4}\left(\frac{\delta}{L}\right)^5 + \frac{1}{2}\left(\frac{\delta}{L}\right)^4 + \frac{5}{8}\left(\frac{\delta}{L}\right)^3 - \frac{5}{3}\left(\frac{\delta}{L}\right)^2 + 1 \\ \qquad \text{for} \quad 0 \leq \delta \leq L \\ \frac{1}{12}\left(\frac{\delta}{L}\right)^5 - \frac{1}{2}\left(\frac{\delta}{L}\right)^4 + \frac{5}{8}\left(\frac{\delta}{L}\right)^3 + \frac{5}{3}\left(\frac{\delta}{L}\right)^2 - 5\left(\frac{\delta}{L}\right) \\ \qquad +4 - \frac{2}{3}\left(\frac{\delta}{L}\right)^{-1} \qquad \text{for} \quad L < \delta \leq 2L \\ 0 \qquad\qquad\qquad \text{for} \quad \delta > 2L \end{cases} \tag{13}$$

where $\rho(\delta, L)$ is the distance-based weight and varies from 0 and 1; δ is the distance between controls; L is the taper length determining the distance at which the weighting drops to approximately 0.2, and $2L$ is the critical distance, beyond which the weighting is zero.

For the drilling-order problem, the control variables x have no physical locations but are permutations of sequences of possible actions, in which case an order-based encoding is appropriate. We have chosen to use the permutation encoding [34] of the drilling sequence. In this encoding, each integer value in the vector encodes the relative ordering of the drilling of a specific well. Consider, for example, two possible control sequences x_i and x_j in which four wells are drilled, i.e., $S_{x_i} = [W_1, W_2, W_3, W_4]$ and $S_{x_j} = [W_3, W_1, W_4, W_2]$. The permutation encodings for these two sequences are $\mathbf{P}_{x_i} = [1, 2, 3, 4]^T$ and $\mathbf{P}_{x_j} = [2, 4, 1, 3]^T$, respectively.

Distance between two control sequences is then measured by the distance between the vectors $\mathbf{P}_{x_i}$ and $\mathbf{P}_{x_j}$. Appropriate distance measures for ordering problems include the 'edit' distance [30], which is the minimum number of operations required to transform one sequence to another sequence, and the 'swap' or Jaro-Winkler distance [22], which counts the minimum number of swaps of two elements required to transform one sequence to another. Because computation of swap and edit distances are relatively expensive, it is common to use fitness-distance measures as surrogates for the permutation distances [35]. The Hamming distance [18] between two sequences of equal length is the number of positions at which the corresponding actions are different, i.e., the number of wells that have different positions in the drilling sequence. The Manhattan distance (also known as the 'position-based distance') measures the sum of the absolute differences between positions of the elements. In terms of the permutation encoded vectors, the Manhattan distance is

$$\delta(x_i, x_j)_{\mathrm{L}_1} = \|\mathbf{P}_{x_i} - \mathbf{P}_{x_j}\|_1 = \sum_{k=1}^{N_w} |P_{x_i,k} - P_{x_j,k}|, \tag{14}$$

where the k-th elements in $\mathbf{P}_{x_i}$ and $\mathbf{P}_{x_j}$ are $P_{x_i,k}$ and $P_{x_j,k}$, respectively, which are the positions of a fixed well W_k. N_w is the number of wells.

In addition to the Hamming and Manhattan distance measures for permutation encodings, we initially considered the use of two standard distance metrics on vector spaces. The Euclidean distance is defined as

$$\delta(x_i, x_j)_{\mathrm{L}_2} = \|\mathbf{P}_{x_i} - \mathbf{P}_{x_j}\|_2 = \sqrt{\sum_{k=1}^{N_w} |P_{x_i,k} - P_{x_j,k}|^2} \tag{15}$$

and the cosine distance [36, 27] is a correlation-based distance measure defined as

$$
\begin{aligned}
\delta(x_i, x_j)_{\cos} &= 1 - \cos(\mathbf{P}_{x_i}, \mathbf{P}_{x_j}) \\
&= 1 - \frac{\mathbf{P}_{x_i} \cdot \mathbf{P}_{x_j}}{\|\mathbf{P}_{x_i}\|\|\mathbf{P}_{x_j}\|} \\
&= 1 - \frac{\sum_{k=1}^{N_w} P_{x_i,k} \cdot P_{x_j,k}}{\sqrt{\sum_{k=1}^{N_w} P_{x_i,k}^2}\sqrt{\sum_{k=1}^{N_w} P_{x_j,k}^2}}.
\end{aligned}
\tag{16}
$$

When the lengths of all sequences are identical, as they are when the sequences are all perturbations of a base sequence, then the cosine distance is simply a scaled version of the Euclidean distance,

$$
\frac{\|\mathbf{P}_{x_i} - \mathbf{P}_{x_j}\|_2^2}{2\|\mathbf{P}_{x_i}\|_2^2} = 1 - \cos(\mathbf{P}_{\mathbf{x_i}}, \mathbf{P}_{\mathbf{x_j}})
\tag{17}
$$

consequently, there is no need to consider both the Euclidean and cosine distance measures.

The choice of an appropriate distance measure to use as a measure of similarity is problem specific. In the application section, we show that the Manhattan, Euclidean, and cosine distance metrics are all very similar when applied to the well ordering problem. They are all superior to the Hamming distance in explaining similarity of drilling-order sequences in terms of the bias correction factor.

Taper window selection

The performance of localization for estimation of correction factor depends not only on the choice of distance measure, but also on the taper parameter L which affects the weights and the effective sample size [25] used for computation of correction factor. A good distance measure will effectively identify control variables with similar correction factors so that the number of realizations used for estimation is maximized and sampling error is reduced.

Suppose that N_e is the number of observed values of β, then

$$
n_{\mathrm{eff}} = \frac{\left(\sum_{j=1}^{N_e} \omega_j\right)^2}{\sum_{j=1}^{N_e} \omega_j^2},
\tag{18}
$$

is a common approximation of the effective sample size [33, 12]. In this equation, ω_j is the weight on the jth partial correction factor. If all weights are identical, then the effective sample size is equal to N_e, while if one of the normalized weights is equal to one and all others are zero, the effective sample size is 1. n_{eff} is a measure of the equivalent number of equally weighted samples. The weights are determined by the distance between the estimation location and the jth drilling sequence.

The accuracy of the estimate of the correction factor for control variable sequence, x_i, is influenced both by the effective sample size, n_{eff}, and the bias resulting from the use of partial correction factors based on random control variables with different values of α. Reducing the taper length will decrease the bias by only including values from control variable sequences with very similar values of the correction factor, but will also increase the sampling error by decreasing the effective sample size. Because the optimal taper length is not known a priori, we generally apply regularization to reduce the effect of a non-optimal choice of taper length.

Regularization

The major disadvantage of pure distance-based localization is that use of a taper length that is smaller than optimal one may result in an estimate of α that is far from the correct value due to sampling error resulting from the small number of samples within a small distance of the estimation point. Instead of using a long taper length to avoid such a situation, it is generally possible to improve the accuracy of estimated α value by adding a regularization term based on the average value and the variance of correction factor to reduce the sensitivity of the estimate to taper length while still generating accurate estimates.

A regularized estimate of $\alpha(x_i)$ is obtained by minimizing an objective function with both a local and a global term,

$$
\begin{aligned}
S\left(\alpha_r(x_i)\right) = & \frac{n_{\text{eff}}}{\sigma_\beta^2} \left(\alpha_r(x_i) - \frac{\sum_{j=1}^{N_e} \omega(x_i, x_j)\beta(x_j, m_j, \bar{m})}{\sum_{j=1}^{N_e} \omega(x_i, x_j)} \right)^2 \\
& + \frac{1}{\sigma_\alpha^2} \left(\alpha_r(x_i) - \bar{\alpha} \right)^2 .
\end{aligned}
\tag{19}
$$

In Eq. 19, σ_α^2 is an estimate of the variance of α over the domain of interest, σ_β^2 is the variance of β, and n_{eff} is the effective sample size for the observations of β (Eq. 18). The regularized estimate, $\hat{\alpha}_r(x_i)$, is obtained by solving $\nabla_\alpha S = 0$, obtaining,

$$
\begin{aligned}
\hat{\alpha}_r(x_i) &= \left(1 + \frac{\sigma_\beta^2}{n_{\text{eff}}\sigma_\alpha^2} \right)^{-1} \left(\frac{\sum_{j=1}^{N_e} \omega(x_i, x_j)\beta(x_j, m_j, \bar{m})}{\sum_{j=1}^{N_e} \omega(x_i, x_j)} + \frac{\sigma_\beta^2}{n_{\text{eff}}\sigma_\alpha^2}\bar{\alpha} \right) \\
&= \left(1 + \frac{\sigma_\beta^2}{n_{\text{eff}}\sigma_\alpha^2} \right)^{-1} \left(\hat{\alpha}_{\text{loc}}(x_i) + \frac{\sigma_\beta^2}{n_{\text{eff}}\sigma_\alpha^2}\bar{\alpha} \right) .
\end{aligned}
\tag{20}
$$

Note that the regularized objective function is a weighted average of the localized estimate $\hat{\alpha}_{\text{loc}}(x_i)$ from Eq. 12 and the mean value of α. When the effective sample size is large compared to the ratio $\sigma_\beta^2/\sigma_\alpha^2$, the regularized estimate will be based primarily on the local samples of β.

Estimation with regularized localization has a major advantage over an approach that relies only on localized estimation: by improving the accuracy of the estimated values that are obtained with an inappropriate distance measure or taper length, regularized estimate $\hat{\alpha}_r$ is potentially more accurate than $\bar{\alpha}$ for a wider range of taper lengths. When the variance of α is unknown, it might be difficult to select the optimal value, but as shown in experiments, results are not strongly sensitive to the exact choice.

Optimal weights

Here we show how the optimal weights can be estimated if the covariance of partial correction factors is known. In that case, an estimate of α at a fixed control x_0 is calculated based on a linear combination of observations β_{ij} with weights w_i. In vector notation, the estimate is written as

$$
\hat{\alpha}(x_0) = \frac{\sum_{j=1}^{N} \omega(x_0, x_j)\beta(x_j, m_j, \bar{m})}{\sum_{j=1}^{N} \omega(x_0, x_j)} = \mathbf{w}^T \mathbf{b},
\tag{21}
$$

where elements in vector $\mathbf{b}$ are the observed values of β_{jk} from random controls and realizations. The collection of observations will be denoted by the vector $\mathbf{b}$, i.e.

$$\mathbf{b} = \begin{bmatrix} \beta_{11} & \beta_{22} & \cdots & \beta_{NN} \end{bmatrix}^{\mathrm{T}}.$$

Although the notation is different, Eq. 21 is identical to Eq. 12.

The quantity $\alpha(x_0)$, that is to be estimated, is defined to be the linear combination of β_{0j} at the estimation location,

$$\alpha(x_0) = \frac{1}{N}\mathbf{1}^T\mathbf{b}_0, \tag{22}$$

where the j-th element of $\mathbf{b}_0$ is β_{0j} at a fixed control x_0 of an individual realization m_j that is sampled from the probability distribution for model parameters, i.e.,

$$\mathbf{b}_0 = \begin{bmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0N} \end{bmatrix}^{\mathrm{T}}.$$

Imposing the constraint $\mathbf{w}^T\mathbf{1} = 1$ provides an unbiased estimate for which the expected error is 0. The optimal weights for estimating $\alpha(x_i)$ from a set of random observations β are obtained by minimizing the expected variance of the estimate, constrained to the unbiasedness condition. For estimation of $\alpha(x_0)$, the variance of the expected error is

$$S_w(\mathbf{w}) = E\left[\left(\frac{1}{N}\mathbf{1}^T\mathbf{b}_0 - \mathbf{w}^T\mathbf{b}\right)^2\right] \tag{23}$$

To minimize the variance in estimation error, subject to the constraint that $\mathbf{w}^T\mathbf{1} = 1$, we define a Lagrangian function

$$S(\mathbf{w}, \lambda) = S_w(\mathbf{w}) - 2\lambda\left(\mathbf{w}^T\mathbf{1} - 1\right),$$

where λ is the Lagrangian parameter and $S_w(\mathbf{w})$ is the variance of the estimator error (Eq. 23). The optimal weights are then obtained by solving for $\mathbf{w}$ that minimizes $S(\mathbf{w}, \lambda)$,

$$\nabla_{w,\lambda}S = 0.$$

Straightforward computation shows that

$$\nabla_w S = 2\,\mathrm{cov}(\mathbf{b}, \mathbf{b})\mathbf{w} - \frac{2}{N}\,\mathrm{cov}(\mathbf{b}, \mathbf{b}_0)\mathbf{1} - 2\lambda\mathbf{1} \tag{24}$$

and

$$\nabla_\lambda S = \mathbf{1}^T\mathbf{w} - 1. \tag{25}$$

Then the weights $\mathbf{w}$ can be found from the following systems of linear equations for $\nabla_{w,\lambda}S = 0$, which can be written in matrix form

$$\begin{bmatrix} \mathrm{cov}(\mathbf{b}, \mathbf{b}) & -\mathbf{1} \\ -\mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \frac{1}{N}\,\mathrm{cov}(\mathbf{b}, \mathbf{b}_0)\mathbf{1} \\ -1 \end{bmatrix}, \tag{26}$$

where $\mathrm{cov}(\mathbf{b}, \mathbf{b})$ denotes the covariance of the variables β in $\mathbf{b}$, and each element in $\mathrm{cov}(\mathbf{b}, \mathbf{b}_0)$ is the covariance function between the corresponding observed values of β in $\mathbf{b}$ and $\mathbf{b_0}$.

Learned heuristic search

Heuristic function

Heuristic search [20] is an efficient approach for solving sequential decision-making problems by repeatedly expanding the partial path with the largest estimated value until a complete path for which the true objective-function value is higher than the estimated values of all evaluated partial paths is found. The estimated value of a partial path could be obtained using an evaluation function $f(n_s)$, which estimates the value of objective-function for the optimal complete path constrained to the previous actions. This estimated objective-function value consists of two elements,

$$f(n_s) = g(n_s) + h(n_s), \tag{27}$$

where $g(n_s)$ is the true reward from the initial state to a specific state n_s through a set of selected actions, and $h(n_s)$ is a heuristic function that estimates the maximum reward from current state n_s to a goal state.

For drilling-order problem, the objective of robust optimization is maximization of expected NPV over uncertainty by optimizing the drilling sequence of wells. The expected NPV computed using mean model bias correction (MMBC) can be mathematically represented as

$$J = \alpha(x_i) \left\{ \sum_{j=1}^{T} \frac{\left(q_{o,j}^{\bar{m}} r_o - q_{w,j}^{\bar{m}} r_w - q_{wi,j}^{\bar{m}} r_{wi} \right) \Delta t_j}{(1+b)^{t_j/\tau}} - \sum_{n=1}^{N_w} \frac{W_n}{(1+b)^{t_n/\tau}} \right\}, \tag{28}$$

where $q_{o,j}^{\bar{m}}$, $q_{w,j}^{\bar{m}}$ and $q_{wi,j}^{\bar{m}}$ denote the rates of produced oil, produced water and injected water, respectively, from the mean model in m^3/day; r_o, r_w, and r_{wi} are the oil price, water production cost and water injection cost, respectively; T represents the number of time steps; t_j is the cumulative time in days up to time step j; Δt_j is the time interval in days; b is the discount rate for a certain reference time τ (365 days); N_w is the total number of drilling wells; W_n denotes the cost of drilling the nth well; t_n is the cumulative time in days up to the open time for each well; $\alpha(x_i)$ is the bias correction factor for the corresponding control x_i.

The cost of finding a strong and admissible heuristic for drilling-problem can be prohibitive since the evaluation of heuristic function requires simulations. In this paper, we use a heuristic function in which all remaining wells are drilled simultaneously at the next step [40, 28], as this estimate can be obtained inexpensively and generally provides an overestimate of the NPV. Such a heuristic is guaranteed to find the true optimal drilling order. However, it might lead to an exhaustive search due to large estimated values.

Online learning techniques

To efficiently find a solution that is optimal or near optimal, the evaluation function $f(n)$ in Eq. 27 should be close to the true maximum value $f^*(n)$. It is difficult to design a heuristic function that is accurate in all situations, but a crude heuristic function can be

improved by online-learning techniques, i.e., estimate the error of the initial approximate value by learning the observations from previous decision steps. From a set of available online-learning mechanisms $\Phi_1, \Phi_2, \Phi_3, \ldots, \Phi_n$, the best-improved evaluation function $\hat{f}_\Phi(n)$ with multiple online-learning techniques is defines as

$$\hat{f}_\Phi(n) = \max\left(\hat{f}_{\Phi_1}(n), \hat{f}_{\Phi_2}(n), \hat{f}_{\Phi_3}(n), \ldots, \hat{f}_{\Phi_n}(n)\right), \tag{29}$$

which might not be the most accurate value, but it is more likely to overestimate the actual maximum value and guide the heuristic search close to the optimal solution.

Wang and Dean [40] have proposed two possible online-learning techniques (i.e, single-step adjustment and multiple-time-periods learning) for improving the initial approximate values obtained from heuristic sequences (i.e., all the remaining wells are drilled simultaneously and opened at the next time step) by estimating a set of forecast errors of $f(n)$ and $h(n)$ for the remaining decision steps.

The single-step adjustment is defined as

$$\begin{aligned}
\hat{f}(n_s) &= f(n_s) \prod_{n=n_s}^{n_{\text{goal}}} \hat{\gamma}_{fn} = f(n_s) \prod_{i=1}^{d(n_s)} \gamma_{fn_s} \bar{\mu}_{n_s}^i \\
&= f(n_s) \gamma_{fn_s}^{d(n_s)} \bar{\mu}_{n_s}^{\frac{(d(n_s)+1)d(n_s)}{2}},
\end{aligned} \tag{30}$$

where $d(n_s)$ is the number of remaining actions at n_s, and $\hat{\gamma}_{fn}$ is the forecast error of $f(n)$ in future decision step n, which is estimated by using the ratio γ_{fn_s} associated with $f(n_s)$ and $f(n_{s-1})$ and the mean single-step ratio $\bar{\mu}_{n_s}$ of $\gamma_{fn_1}, \gamma_{fn_2}, \ldots, \gamma_{fn_s}$ along the current optimal path.

Multiple-time-periods learning is calculated by correcting the heuristic values of various time periods simultaneously

$$\hat{f}(n_s) = h_{\Delta t_1^\Phi}(n_s) \prod_{n=n_s}^{n_{\text{goal}}} \hat{\gamma}_{h_{\Delta t_1^\Phi}}(n) + \sum_{i=2}^{N_L} \left(h_{\Delta t_i^\Phi}(n_s) \prod_{n=n_s}^{n_{\text{goal}}} \hat{\gamma}_{h_{\Delta t_i^\Phi}}(n) \right). \tag{31}$$

Note that estimated value $h_{\Delta t_1^\Phi}(n_s)$ of the first learning period Δt_1^Φ contains both true contribution to the expected NPV at time periods $t_s = \sum_{i=1}^{N_s} \Delta t_i$ from N_s sequentially drilled wells and heuristic values at a certain time period $\Delta t_1^\Phi - t_s$ when all wells are open.

To summarize, multi-learned heuristic search with space reduction (MLHS-SR) based on the economic indicator [16] and improved evaluation function $\hat{f}_\Phi(n)$ (Eq. 29) can be used to find a solution to an optimization problem faster without losing quality [40]. Figure 1 shows the flowchart of using MLHS-SR with the bias-corrected mean model for the drilling-order optimization problem under geological uncertainty. To estimate the bias correction factor α for different drilling sequences, we first sample N_x distinct controls and apply them to individual model realizations to obtain the initial observations of β. This step requires $2 \times N_x$ simulations. Both bias-correction methods (i.e., improve the estimates of expected NPV) based on the observations from distinct controls and model realizations, and online learning techniques (i.e., improve the estimates of maximum expected NPV) based on the observations from previous drilling steps do not require any simulations. Hence, we only need to perform one additional simulation in the mean model at each iteration for evaluating the expected NPV of

one specific control. The use of bias correction applied to the mean model allows the application of fairly general optimization methods. During the optimization process, it requires only simulations in the mean model for obtaining initial estimates of expected NPV. For an ensemble with N_e realizations, we assume that the information from N_e distinct controls is used to estimate the bias correction factor, and the optimization process requires N_{iter} iterations (i.e., N_{iter} different controls have to be evaluated to obtain the optimal solution). Then, the total number of simulations required in RO is $N_{\text{tot}} = 2N_e + N_{\text{iter}}$, in which $2N_e$ simulations are performed to obtain N_e observed values of β and N_{iter} simulations are performed in the mean reservoir model to obtain the initial approximations of expected NP for N_{iter} different controls. If N_{iter} different controls were to be evaluated using SAA, the cost in RO would be $N_{\text{tot}} = N_e \times N_{\text{iter}}$, which is much more expensive and will increase linearly with the ensemble size N_e and the number of iterations N_{iter}.
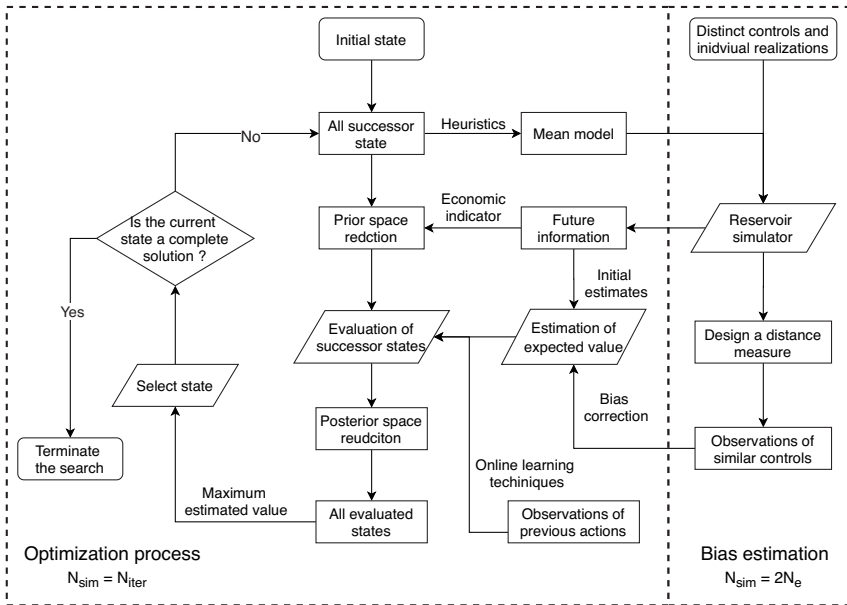
Figure 1: Flowchart of using MLHS-SR with mean model bias correction

Depth-limited search

The reservoir model will almost certainly be updated based on information obtained from drilling the first few wells. After model updating, the optimal order of the remaining wells will differ from the initial estimate of optimal order. It is therefore more important to correctly identify the first few wells to be drilled rather than provide an entire drilling sequence. One advantage of using learned heuristic search for the drilling-order problem is that a partial solution of the first wells can be obtained at a reduced cost by cutting off the search at a specified depth.

The most straightforward approach to partial sequence optimization is through the depth-limited search (DLS) [14], in which the learned heuristic search is terminated at a certain depth. To find the solution of the first N_s wells, we prefer to terminate the

search at the first-visited best partial path with $N_s + 1$ selected wells, because online learning techniques with more observations along a longer path can further improve the approximations and potentially generate a better solution. Note that the last well along the first selected partial path of $N_s + 1$ wells might not be the optimal well because the search might change direction after evaluating its extended paths.

A faster approach to finding the partial solution is to use iterative depth-limited search (IDLS). It works by iteratively optimizing the next well based on the first selected partial path with two more wells until a solution of $N_s + 1$ selected wells is found. Because only the partial paths extended from previous decisions are considered, this approach can avoid evaluating unnecessary paths along other directions caused by underestimated values. Learned heuristic search with accurately estimated values generally will not change direction frequently, so that the optimized sequence of the first few wells with a limited depth is likely to be near the final optimized drilling sequence.

Case study 1: flow optimization in one-dimensional model

The purpose of this simple example is to thoroughly investigate methods of efficient optimization on a flow problem for which the dependence of the objective function on the model parameters is highly nonlinear. The problem is chosen to be small enough that we can evaluate both linear and quadratic Taylor series approximations of the objective function, and can evaluate the correct optimal solution.

The objective in this example is to locate an injection well, operating at fixed pressure in a one-dimensional flow domain with uncertain permeability, such that the total flow rate out of the reservoir through fixed pressure boundaries is minimized. The reservoir is discretized with 150 grid cells (i.e., 150 possible injector locations), each of length Δx, and cross-sectional area A. The permeability k_i is distributed as log-gaussian. Pressures at both ends of the grid are fixed at 0. An injector is located at the interface between cells $i = i_w - 1$ and $i = i_w$, where the pressure is fixed at $P = P_w$. Instead of permeability, we use the log-permeability θ_i as a parameter in the problem, $k_i = \exp(\theta_i)$.

(a) True log-permeability field and best approximation after observations   (b) Samples of the log-permeability distribution after conditioning.
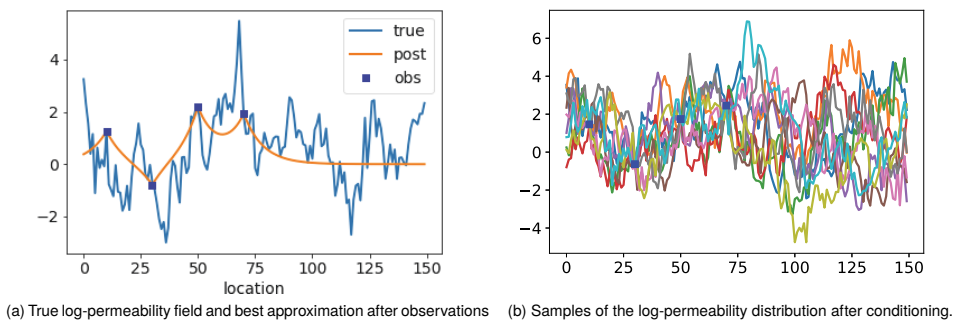
Figure 2: Light blue curve shows the true θ that generated observations. Blue squares show observations of θ with noise added. Correlation range of prior distribution for log-permeability is 26.

In this notation, the total flow rate is the sum of the flow to the left and flow to the

right:

$$q(i_w, \theta) = \frac{AP_w}{\mu \Delta x} \left(\sum_{i=0}^{i_w - 1} \exp(-\theta_i) \right)^{-1} + \frac{AP_w}{\mu \Delta x} \left(\sum_{i=i_w}^{N_x} \exp(-\theta_i) \right)^{-1} \tag{32}$$

where all variables are in consistent units. Because the permeability is uncertain, we minimize the *expected value* of flow rate q, by adjusting the location of the injection well, i_w. To avoid a problem in which the optimal injector location is at the center of the reservoir because of symmetry, we modify the prior probability for θ by assuming observations of θ at four locations ($x = 10, 30, 50, 70$). The mean of the posteriori distribution for θ is shown as the orange curve in Fig. 2a, along with the true log-permeability (blue). Ten realizations of the log-permeability field from the posteriori distribution are shown in Fig. 2b.
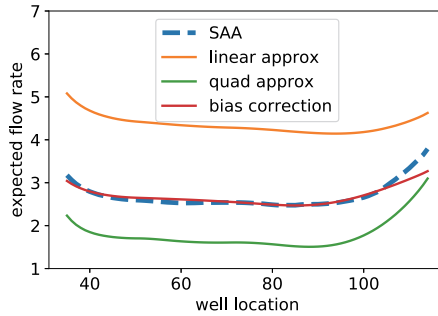

Figure 3: *The expected value of the total injection rate, $E[q(x, \theta)]$, computed using the sample average approximation, and two levels of Taylor series approximation.*

The most straightforward approach to approximating the expected value of the objective function is through the use of SAA or averaging of Monte Carlo samples from the posterior distribution for θ. We use SAA with 400 samples for each control location as a benchmark for other methods. Hence the SAA method uses 150×400 function evaluations to generate the expected value of $q(i_w, \theta)$ for optimization.

Taylor series expansions of the objective function provide much less expensive approximations of the expected value of $q(i_w, \theta)$. For the linear approximation, the expected value of $q(i_w, \theta)$ is approximated using the mean of the log-permeability field (Eq. 5). The quadratic approximation is considerably more expensive as it requires computation of the second derivative of the objective function with respect to well location. For this 1D steady flow problem, that is still a manageable computation. Figure 3 compares the linear and quadratic Taylor series approximations of the expected value of flow rate to the sample average approximation. Although there is clearly a large bias in the values from both the linear and quadratic approximations, the shapes are quite similar to the SAA, and the location of the minimum for each curve is also approximately the same. For this particular problem (minimizing total flow rate), a uniform bias does not affect the optimization result.

To correct the nonuniform bias in α, we use a regularized localization approach (Eq. 20) in which higher weights are given to samples that are closer to the control variable for which the bias correction is being estimated. In this optimization problem,
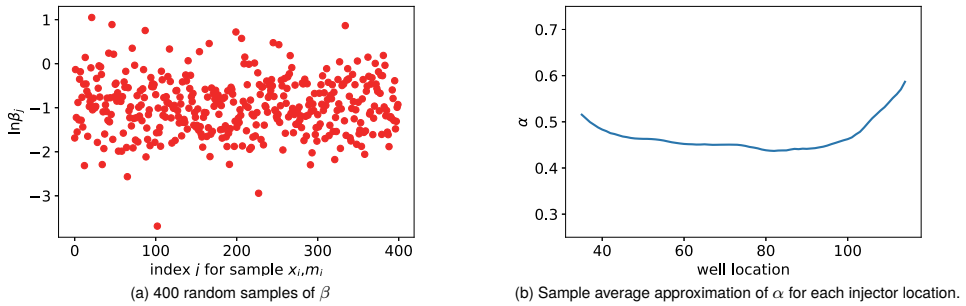
(a) 400 random samples of β

(b) Sample average approximation of α for each injector location.

Figure 4: Variability of the partial bias correction β and the bias correction factor α for the 1D flow problem.

the actual distance between well locations is an appropriate measure of similarity. The variability of values of β is quite large (Fig. 4a) because the rate is strongly affected by the occurrence of low permeability values between the injector and the boundary. The variability in α is much smaller than the variability in β (Fig. 4b). Unfortunately, while the variability in β can be estimated from samples, the variability in α will generally be unknown. Because it is computed by averaging over samples of β, however, we should expect it to generally be smaller than the variability in β, so that the ratio $\sigma_\alpha^2/\sigma_\beta^2$ which appears in Eq. 20 will generally be substantially smaller than 1.

In Figure 5 we compare the optimal well locations obtained using three different methods. In the SAA approach, we use N_e samples of the permeability field for each well location between 35 and 115 to compute an approximation of the expected value of flow rate. In the linear approximation, we simply compute the flow rate using the mean of the log-permeability field, and for the bias-correction approach, we used regularized localization with $\gamma = n_{\text{eff}}\sigma_\alpha^2/\sigma_\beta^2$ and a taper length of 20 to obtain a bias-correction to the linear approximation of expected value of flow rate. Each experiment was repeated 100 times to reduce the effect of sampling error. For both ensemble sizes, the bias-correction approach give better results than the linear approximation, although the difference is more apparent at larger ensemble size because the spread in the results is reduced in that case.
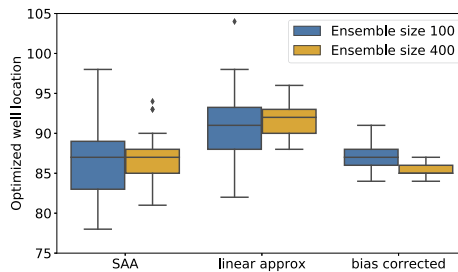

Figure 5: Optimized well locations for three different methods and two ensemble sizes.

The most important criterion for judging success of the methodology is the ability to actually minimize the flow rate. Figure 6 compares the distribution of total flow rates obtained from the optimal well locations applied to the truth case for each of the

methods. The test was repeated 100 times so there are 100 different truth cases and 100 "optimal" locations for each method. Note that the bias-correction method is now clearly superior to the linear approximation, even in this example in which the optimization depends only on relative differences in the value of the objective function. Optimization of the well location based on single realizations from the posterior distribution for model parameters gives very poor results (Fig. 6b).
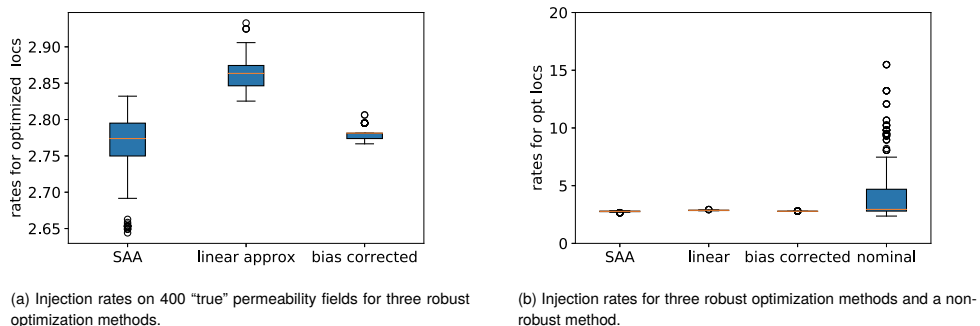

(a) Injection rates on 400 "true" permeability fields for three robust optimization methods.

(b) Injection rates for three robust optimization methods and a non-robust method.

Figure 6: Values of the objective function from optimization of well location in the 1D flow problem. (Small values are better.)

Case study 2: drilling-order problem in synthetic model

Reservoir model

In this example, we design a set of experiments to study the bias properties of the drilling-order problem applied to the synthetic REEK Field model [40, 19, 29], and the performances of various methods for estimating the bias for both complete and heuristic sequences by using the information from distinct controls and individual realizations. The objective of the robust optimization problem is maximization of the 10-year expected NPV with respect to the drilling schedule of wells. Learned heuristic search applied to the mean model with bias correction is used to optimize the drilling sequence under geological uncertainty. To illustrate the quality of the robust optimal solution, we compute both optimal solution based on SAA and deterministic solutions of individual realizations. Moreover, we also investigate the possibility of optimizing the first few wells by using learned heuristic search with limited search depths.

The REEK model is a three-phase black-oil reservoir model with $40 \times 64 \times 14$ grid cells, of which 34,770 are active cells. An ensemble of 100 geologically consistent model realizations is used to empirically represent uncertainty in the porosity field, permeability field, and fault transmissibility multipliers. Table 1 shows the reservoir properties and control variables in REEK field. For the test problem, we assume that eight vertical fully penetrating wells (5 producers and 3 injectors) with fixed locations need to be drilled sequentially, and the first well is drilled at the beginning of simulation. The assumed drilling period is six months for all wells and wells begin operating immediately after drilling. Figure 7 shows the well locations and initial oil saturation of one randomly chosen model realization. The injectors are positioned around the oil-water contact and controlled by a maximum injection rate of 10,000 m³/day and a

maximum BHP of 320 bars. The producers are distributed throughout the oil-containing area and controlled by a maximum production rate of 6,000 m³/day and a minimum BHP of 250 bars.

Field	REEK model
Number of grid blocks	$40 \times 64 \times 14$
Number of active cells	34770
Permeability (md)	0 to 3500 (average 733)
Porosity	0 to 0.45 (average 0.159)
Fault transmissibility multiplier	0 to 1 (average 0.105)
Number of geological realizations	100
Number of wells (all vertical wells)	8 (5 producers and 3 injectors)
Number of all possible drilling sequences	8! = 40320
Maximum production rate (m³/day)	6000
Minimum BHP of producers (bars)	250
Maximum injection rate (m³/day)	10000
Maximum BHP of injectors (bars)	320

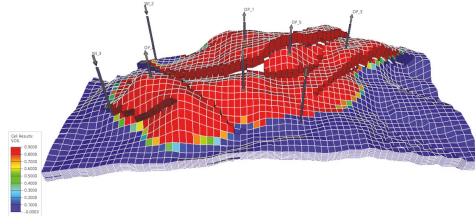Table 1: Reservoir properties and control variables in REEK model

Figure 7: Well locations and initial oil saturation in one realization of the REEK model ensemble

Estimation of correction factor

Properties of α and β

We randomly select 5 different complete drilling sequences and apply them to the entire ensemble of realizations and the reservoir mean model to obtain the actual values of the multiplicative bias correction factor α (Eq. 7) and their partial correction factor β of individual realizations (Eq. 8). Figure 8 shows the observed values in α and β from 5 random complete drilling sequences. Symbols that are in the same color indicate the partial correction factor $\beta(x_i, m_j, \bar{m})$ at a fixed drilling sequence x_i applied to different reservoir models and the mean model. The x-axis displays the index j of a single reservoir model m_j. We observe that $\beta(x_i, m_j, \bar{m})$ at a fixed control changes significantly with geological uncertainty, but $\beta(x_i, m_j, \bar{m})$ at a fixed reservoir model changes slightly with different controls. The horizontal dashed lines represent the bias correction factors α_i for the 5 drilling sequences, computed by averaging over samples of β. The variability in α_i for these 5 drilling sequences is very small and the mean for each is close to 1.1.

To obtain more statistically reliable results, we sample 560 different drilling sequences for studying the bias properties. For 5 producers and 3 injectors, there are 56 possible combinations of drilling sequences based on the types of wells (e.g., [P, P, P, P, W, P, W, W], [P, P, W, P, P, P, W, W]). We randomly select 10 distinct drilling sequences from each combination and compute their actual values of α and β. Figure 9 shows the distributions of α and β values obtained from these 560 random complete drilling
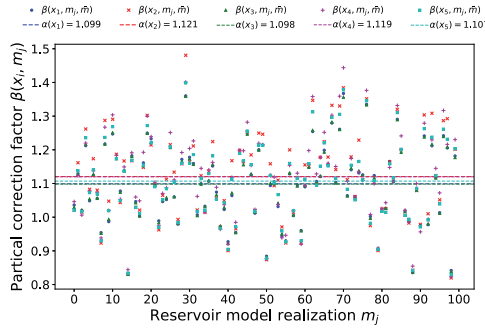
Figure 8: Comparison of bias correction factor α and partial correction factor β of 5 different complete drilling sequences

sequences. α has a value near 1.108 for most of the drilling sequences (Fig. 9a), indicating that SAA of NPV is almost 10% higher than the NPV computed using the mean model. To accurately estimate the expected NPV, it is necessary to correct the bias in the initial approximation that is obtained in the reservoir mean model. Out of the 560 drilling sequences, only 22 (approximately 4%) sequences have α values lower than 1.09, whereas β changes between 0.6 and 1.5 with a larger variability (Fig. 9b). Interestingly, almost all of the drilling sequences with atypical bias correction factor α are from two extreme drilling sequences, i.e., either all producers or all injectors are drilled first. The results of these two combinations are plotted in yellow. This finding was unexpected and suggests that to more accurately estimate the bias for general drilling sequences, we should avoid sampling controls from these two extreme combinations since they seem to provide less useful information and vice versa.

(a) Distribution of bias correction factor α

(b) Distribution of partial correction factor β
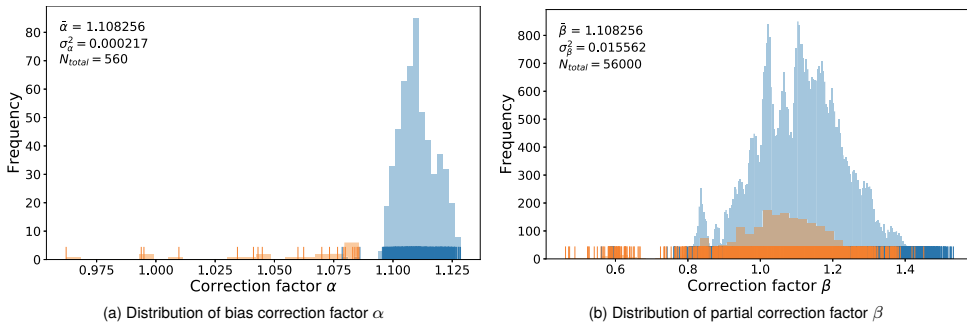
Figure 9: Comparison of the distributions of α and β of 560 different complete general drilling sequences

Figure 10 compares the distributions of variance in β at a fixed reservoir model and a fixed control after eliminating two extreme combinations (i.e., [P, P, P, P, P, W, W, W] and [W, W, W, P, P, P, P, P]). As previously observed in Fig. 8, the variability of β among different drilling sequences at a fixed reservoir model is smaller than the variability of β among different reservoir models at a fixed drilling sequence. The average value of α can be estimated by averaging all observed values in β from distinct controls. In this example, the variance in α is only 0.000058 for general drilling sequences. The estimates of expected NPV with a bias correction factor $\bar{\alpha}$ could be accurate for most of the drilling sequences, but this will not always be the case in other problems. For the

case of large variability in α, we should consider the similarity between the samples and control variables for estimating the bias for specific controls. In this case, the variability in β is much larger than the variability in α for general complete drilling sequences, and the ratio $\sigma_\beta^2/\sigma_\alpha^2$ is almost 267. For a small number of samples, local estimates $\hat{\alpha}_{\text{loc}}$ based on β (Eq. 12) might be away from the actual bias correction factor. The error of local estimates caused by sampling error due to small sample size could be as large as 40%. Hence, it is critical for this problem to use a taper length with a relatively large effective sample size for pure distance-based localization or apply regularized localization to avoid generating estimates that are far away from $\bar{\alpha}$.
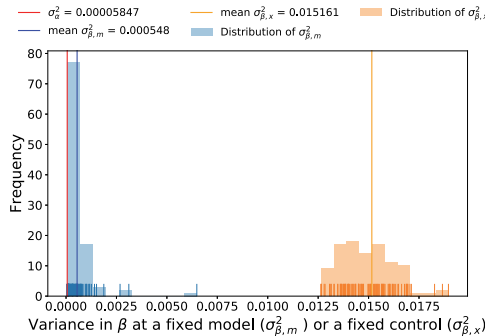

Figure 10: Distribution of variance in β at a fixed model or a fixed control

To measure the similarity of drilling sequences for the bias correction factor, we investigate the use of four different distance metrics (i.e., the Manhattan, Euclidean, cosine, and Hamming distances). Figure 11a shows the empirical variograms of α for these four distance metrics. We observe that all four variograms show an approximately linear increase with distance over most of their ranges. Overall, each of these four distance metrics can measure the similarity of drilling sequences for expected NPV. For the purpose of identifying similar drilling sequences, a measure of distance is better than another if it correlates better over larger numbers of samples.

Figure 11b shows the cumulative fraction of drilling sequences within an upper bound on the variogram of α. Among a set of random controls, very few of them will have small Hamming distances. We note that, compared with the other three distance metrics, the Hamming distance yields fewer similar drilling sequences at a fixed threshold variance of α. In other words, when a fixed fraction of samples is considered for local estimation, drilling sequences extracted by the Hamming distance will potentially have more significant variability in α. By contrast, the Manhattan, Euclidean and cosine distance metrics have similar behaviors and all perform better than the Hamming distance for local estimation based on the observations from randomly selected drilling sequences. In our subsequent experiments, we use the Manhattan distance to measure the distance between two drilling sequences for estimating α values for specific controls.

Accuracy of estimates of α

In this experiment, we compare various methods for estimating the bias correction factor α of different drilling sequences. We first sample a number (N_x) of distinct

(a) Empirical variogram of the bias correction factor α as obtained using different distance metrics

(b) Cumulative fraction of sequences within a certain upper bound on the variogram of α
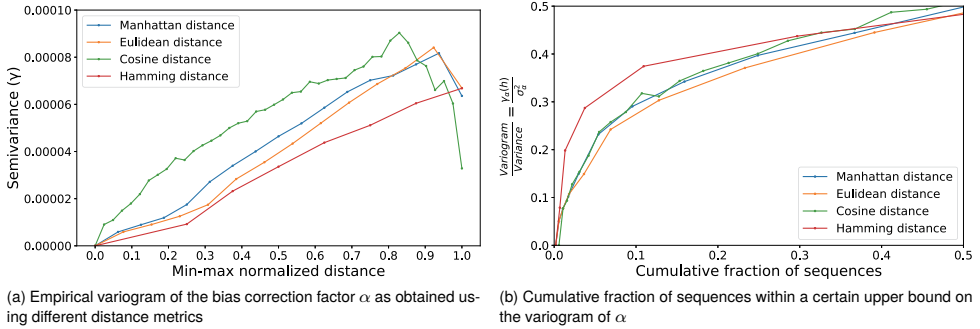
Figure 11: Performance comparison of four distance metrics applied in the drilling-order problem

general drilling sequences to obtain a set of observations of β, which requires $2 \times N_x$ simulations. To evaluate the quality of each of the three approaches (i.e., average bias correction factor, pure-distance localization and regularized localization), we use each method to estimate α for 540 different drilling sequences. Estimation of α for various drilling sequences using these approaches does not require additional samples of β, but to estimate the expected NPV of a specific drilling sequence, we need to perform one additional simulation using the mean model to obtain the initial approximate value of expected NPV. In other words, the amount of computation required for computing the expected NPV of 540 sequences by using the information from 100 distinct controls and individual realizations is $2 \times 100 + 540 = 740$ simulations, which is much less than the cost for SAA (54,000 simulations). In this experiment, we use the accuracy of the estimated α as a criterion for judging the quality of the expected NPV based on the reservoir mean model with bias correction.

Figure 12 shows RMSE of the estimates of the 540 values of α based on a set of 100 observed values in β. The black horizontal line indicates the result from a fixed estimated α obtained by averaging samples of β (Eq. 11). Pure distance-based localization and regularized localization are applied with different taper lengths between 1 to 200 (for reference, the maximum Manhattan distance between complete sequences with eight wells is 32). We observe that purely local estimates $\hat{\alpha}_{\text{loc}}$ are potentially more accurate than the estimates computed by averaging over samples of β when a taper length of $L > 26$ is used. In this case, the effective sample size is relatively large ($n_{\text{eff}} > 80$), and the best value of the taper length in this test is approximately 34. When the taper length extends over 170, all samples are assigned high weights and the performance is close to the approximation of $\bar{\alpha}$. On the contrary, the quality of local estimates is very sensitive to the taper length because of sampling error when the taper length is small (e.g., $n_{\text{eff}} \approx 10$ at $L = 15$). The error of estimated values resulting from the small number of samples is significant.

Since our estimate of optimal weighting in regularized estimation of α depends on the value of σ_α^2, which we are unlikely to know accurately, we investigated the sensitivity of results to variation in the magnitude of the weighting term. Three different levels of the ratio of the variance of α and β (i.e., $\lambda = 0.1 \cdot \sigma_\beta^2/\sigma_\alpha^2 n_{\text{eff}}$, $\sigma_\beta^2/\sigma_\alpha^2 n_{\text{eff}}$, and $10 \cdot \sigma_\beta^2/\sigma_\alpha^2 n_{\text{eff}}$) were evaluated. For $\lambda = 10 \cdot \sigma_\beta^2/\sigma_\alpha^2 n_{\text{eff}}$, the regularized estimates $\hat{\alpha}_r$ are close to $\bar{\alpha}$ and the RMSE is similar. For $\lambda = 0.1 \cdot \sigma_\beta^2/\sigma_\alpha^2 n_{\text{eff}}$, the effect of regularization is small;

consequently, the estimates at a small taper length are inaccurate due to sampling errors. However, the results show that $\lambda = \sigma_\beta^2/\sigma_\alpha^2 n_{\text{eff}}$ reduces the sensitivity of the estimates to the taper length, while still generating more accurate values even for cases with a small number of samples. As a result, the corresponding regularized estimates are potentially more accurate for a wider range of taper lengths.
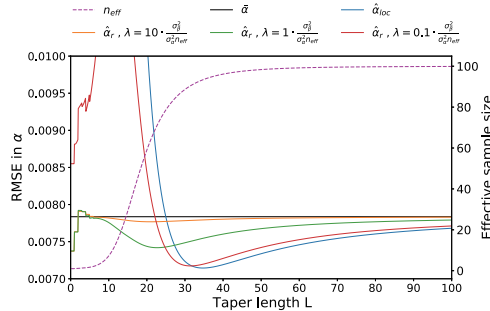

Figure 12: Comparison of the RMSEs of estimates of α obtained using different approaches based on 100 observed values of β

To reduce the influence of sampling error on the conclusions, we repeat each experiment 100 times, each time using a fresh sample of controls to obtain observations of β. For every single test, we apply three different sample sizes ($N_x = 100$, 200, and 400) to investigate the effect of the sample size on the accuracy of the estimated α and the best value of the taper length for distance-based localization. Figure 13 shows the RMSEs of α estimates obtained using different approaches with the three different sample sizes after repeating each test 100 times. In general, the estimates are improved by using more samples. The improvement, however, is small for estimates obtained without localization (horizontal dashed lines) as the estimate of the mean for β is already quite accurate for $N_x = 200$.

Distance-based localization offers a further increase in accuracy, and the improvement is more significant with a larger sample size. In this example, we observe that the RMSE of the local estimates at the optimal taper length is close to that of the estimates computed from the optimal weights based on the covariance of the partial correction factors (Eq. 26). For the drilling-order problem, the well-position based distance distribution obtained from a fixed set of random controls changes slightly when the control variable for estimating the bias correction factor is varied. In such a case, the distributions of elements in covariance matrix $\text{cov}(\mathbf{b}, \mathbf{b}_0)$ (Eq. 42) in terms of the distance is stable. The variance of the correction factor increases almost linearly with distance, making the optimal weight based on the covariance of partial correction factors decreases with the distance of the samples. As a result, the optimal weights would be close to the local weights at the optimal taper length for the drilling-order problem.

As observed previously, regularization with $\lambda = \sigma_\beta^2/\sigma_\alpha^2 n_{\text{eff}}$ appears to always produce more accurate estimates of α than those obtained by averaging of all values of β. When the taper length is greater than a certain threshold, it seems that pure distance-based localization slightly outperforms regularized localization. The reason for this is that local estimates are potentially more accurate than $\bar{\alpha}$ with large effective sample sizes; however, the regularization term reduces the effect of localization and results in estimates that are closer to $\bar{\alpha}$. Both the variance of α and the best value of the taper length are

generally unknown. Therefore, instead of using regularized localization based on an uncertain value of $\sigma_\alpha^2/\sigma_\beta^2$, for this problem we use pure distance-based localization with a relatively large effective sample size to estimate α values for different controls.
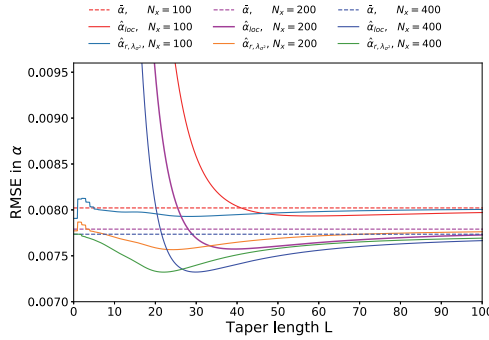

Figure 13: RMSEs of estimates of α with three different sample sizes

We investigated the effect of taper length on the error in the estimate of bias correction factor, α, for three different sample sizes. For each sample size, we repeated the experiment 100 times to obtain reliable estimates. The results show that the optimal taper length for local estimates of α decreases as the sample size increases (Fig. 13). As the optimal taper length is also subject to sampling error, we investigated the variability. Figure 14 compares the frequency distributions of the optimal taper lengths for each ensemble size. When only 100 samples are used to estimate α, the best value of taper length is shown to be highly sensitive to the initial samples (histogram in blue). In approximately 1/3 of the cases, the optimal taper length is larger than 200, which means that the average value of all β will provide a better estimate than a purely distance based weighting with a small taper length. For some controls, the actual α values might be very close to $\bar{\alpha}$, such that more samples will be required for local estimation to provide any further improvement in accuracy. When N_x is increased from 100 to 400, the optimal taper length is less variable, varying between 25 and 30 in most cases (histogram in green). It seems that for cases with sample sizes larger than the ratio $\sigma_\beta^2/\sigma_\alpha^2$, the best value of the taper length is likely to lie near a point where the value of $\sigma_\beta^2/\sigma_\alpha^2 n_{\text{eff}}$ is approximately 1. In that case, the variability in the local estimates will be
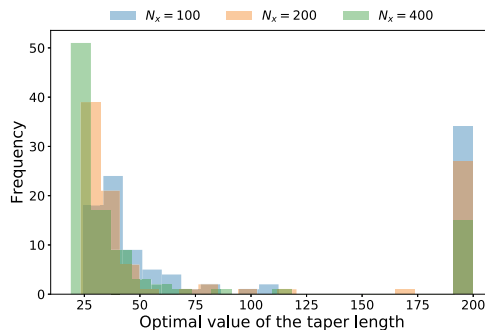

Figure 14: Distribution of the optimal taper length for the pure distance-based localization
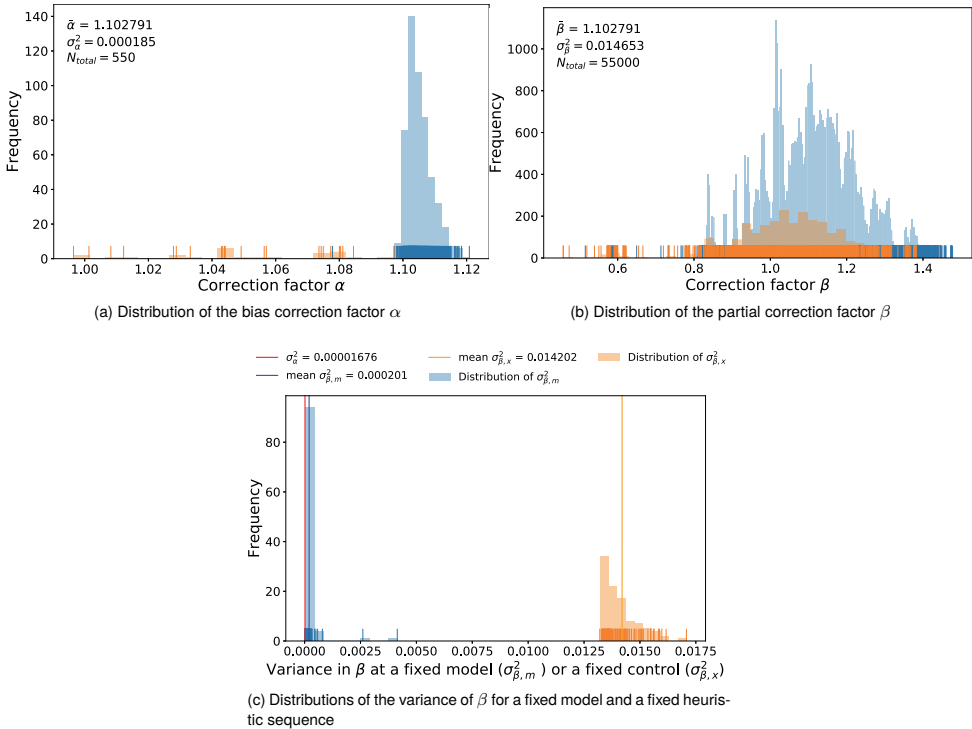
(a) Distribution of the bias correction factor α

(b) Distribution of the partial correction factor β

(c) Distributions of the variance of β for a fixed model and a fixed heuristic sequence

Figure 15: Distributions of α, β, and σ_β^2 of heuristic drilling sequences

close to the variability in the actual α values. When the variance of α is unknown, it might be difficult to select the best value of the taper length. However the results show that the local estimates are not highly sensitive to the taper length when the effective sample size is relatively large, in which case local estimates are still more accurate than estimates obtained using the average value of β.

Correction factors for heuristic sequences

During the heuristic search process, a set of heuristic sequences with different numbers of selected wells is used in combination with online learning techniques to estimate the maximum value among the possible complete drilling sequences constrained by previous wells. Hence, we need to compute the expected NPVs of some specific heuristic sequences to optimize either an entire drilling sequence or only the first few wells when applying a learned heuristic search. In this experiment, we study the bias properties of heuristic sequences and the possibility of using mean model bias correction to accurately estimate their expected NPVs.

Figures 15a and 15b show the distributions of α and β from 550 random heuristic sequences, 400 of which have fewer than four selected wells (i.e., all possible heuristic sequences with $N_s \leq 3$ are considered). The histogram in orange represents the bias from the heuristic sequences in which either all producers or all injectors are drilled sequentially first. Figure 15c compares the distributions of the variance of β for a fixed reservoir model and a fixed heuristic sequence after the elimination of controls from

those two extreme combinations. Relative to the observations for complete drilling sequences, the bias for heuristic sequences has some similar properties: (1) the $\bar{\alpha}$ for heuristic sequences is close to that for complete sequences, but the variability in α is smaller; (2) heuristic sequences in which either all producers or all injectors are drilled first have atypical and relatively smaller values of α; (3) β values for heuristic and complete sequences have similar distributions, and the variability in β is much larger than the variability in α; and (4) the variance of β among different heuristic sequences for a fixed model is much smaller than that among different realizations for a fixed heuristic sequence. We observe that the variance of α for heuristic sequences is only 0.0000168. In this case, the expected NPV that is estimated from the mean model with $\bar{\alpha}$ will have a high level of accuracy. Distance-based localization might require a large number of samples to improve the estimates, and even then, the improvement will be small.

Figure 16 shows the empirical variogram of α for heuristic sequences based on the Manhattan distance. The number of selected wells constrains the distance between heuristic sequences. For eight wells, the maximum Manhattan distance between heuristic sequences with fewer than four selected wells is only 12. For the sequences in which at least half of the wells are drilled simultaneously ($N_s \leq 4$), while the variance of α increases with distance, the correlation of α is still strong even for the sequences with the largest distance. For $N_s \leq 5$ and $N_s \leq 6$, however, the variance of α decreases beyond a certain distance. The main reason for this is that heuristic sequences with a significant gap in N_s generally have a large distance, but they might also have similar values of α. If we sample heuristic sequences with a large N_s (i.e., in which most of the wells are drilled sequentially) to estimate the bias for sequences with a small N_s (i.e., in which most of the wells are drilled simultaneously), distance-based localization might not significantly improve the estimates even with a large number of samples, since most of the samples will be at long distances, but may still have similar α values. Therefore, to ensure the effectiveness of distance-based localization, we should use samples with a small gap in N_s to estimate the bias for specific heuristic sequences with N_s selected wells.
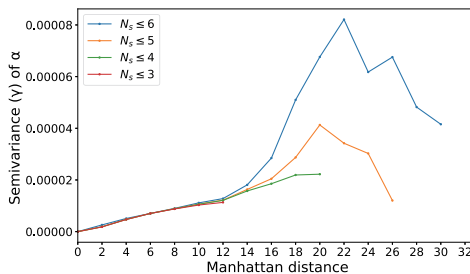
Figure 16: Empirical variogram of α for heuristic sequences based on the Manhattan distance

Figure 17 shows the RMSEs of the estimated α values for heuristic sequences with different numbers of selected wells obtained by using the information of distinct heuristic controls. As in the previous section, to reduce the sampling error, we repeat each experiment 100 times. As expected, the results show that the average value of β is a good approximation of α for most heuristic sequences due to the small variability in α. The error in the estimates can be only slightly reduced from 0.0034 to 0.0032 by
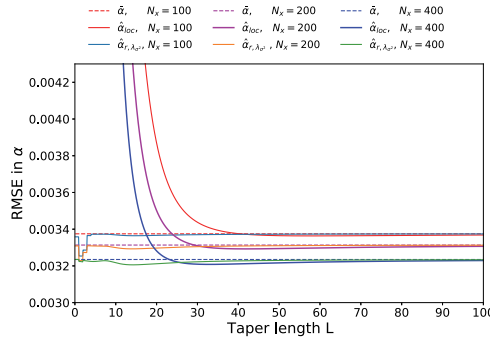
Figure 17: RMSEs of estimates of α for heuristic sequences obtained using different approaches

increasing the sample size from 100 to 400, but at the cost of 600 additional simulations. In this case, it is not necessary to use a larger sample size to improve the estimates. The estimated α based on 100 observed values of β is already very close to the actual α values, although this might not always be the case in other problems. The effect of distance-based localization is small even if we use $N_x = 400$, which means that more samples would be needed for distance-based localization to improve the estimates further. As observed in the case of complete drilling sequences, the regularized estimates obtained from heuristic sequences also prove that a regularization term based on $\bar{\alpha}$ with a parameter of $\sigma_\beta^2 / \sigma_\alpha^2 n_{\text{eff}}$ reduces the sensitivity of the estimates to the taper length, such that the estimates are potentially more accurate for a wider range of taper lengths. The results for both complete and heuristic sequences illustrate the effectiveness of using MMBC to estimate the expected NPV. Although the effect of localization on improving the estimates is relatively small in this example, this might not always be the case in other problems. For cases with considerable variability in α, the improvement in accuracy achieved by using distance-based localization with an appropriate taper length or regularized localization will be more significant.

Robust optimization under geological uncertainty

Optimization of complete sequences

In this experiment, we apply learned heuristic search with MMBC to compute the RO drilling sequence under geological uncertainty. A set of 100 random heuristic drilling sequences with different numbers of selected wells is used to generate 100 observations of the partial correction factor β. This requires 200 simulations, as the NPV must be computed for the drilling sequence run using the mean model, and for a model realization. Because the variance of α is unknown, we apply pure distance-based localization to estimate the bias for different drilling sequences. To avoid inaccuracy in the local estimates caused by sampling error due to a small sample size, we choose a taper length near 50, which is large enough to ensure that the effective sample size is relatively large. To evaluate the effectiveness of the various approaches, we compute the "gold standard" RO strategy based on SAA and the NO strategies of individual realizations. The performance of RO/NO optimal strategies is investigated by comparing their NPV distributions in the entire ensemble of 100 realizations with that of 560 random complete drilling sequences.

Figure 18 shows MLHS-SR search strategy applied to the mean model with bias correction. To speed up the search process while maintaining the highest possible solution quality after space reduction, we consider at least half of the remaining wells with high economic indicator values as the next possible actions (i.e., prior space reduction) and preserve three partial paths with the highest estimated expected NPVs for the remaining search process (i.e., posterior space reduction). For nodes marked in orange, the paths through them have the highest estimated expected NPV for complete sequences and have been expanded in the most promising directions. For each selected direction, only the partial paths go through blue nodes (or orange nodes if the paths have been extended), and black nodes are evaluated after prior space reduction. However, the paths with black nodes are pruned in the process of posterior space reduction due to their lower estimated values and will not be considered during the search process. In this case, the search evaluates 46 paths through 13 decision steps to optimize an entire drilling sequence: [OP_3, WI_1, WI_2, OP_1, OP_4, WI_3, OP_2, OP_5]. Considering the cost of obtaining the partial correction factor β (200 simulations), only 246 simulations are required to compute the RO drilling sequence for eight wells. For comparison, when MLHS-SR is used with SAA, it is necessary to perform more than 3,000 simulations to obtain a solution with the same quality. By more extensively reducing the search space, we can find the same solution faster (235 simulations). However, excessive space reduction (e.g., searching along only one direction) will lead to a drilling sequence with a less-than-optimal value (-0.71%). Because a large fraction of the cost is a result of estimation of the bias correction factor, the cost reduction achieved by evaluating fewer controls might not be very significant, especially for small fields with large ensemble sizes. In such cases, it is not advisable to prune the space significantly to reduce the cost, because the computational cost might be reduced only slightly while yielding a solution with a suboptimal value.

The percentage presented near each node in Fig. 18 represents the error of the estimated expected NPV of each partial path (i.e., heuristic sequences) compared with SAA. The average error of the estimated expected NPV from the mean model is reduced from -9.20% to 0.21% by estimation of the bias correction factor. The maximum error on the estimated expected NPV is only 0.54%. In this example, by using the information from only 100 samples of the drilling sequence and individual model realizations, the expected NPVs of different drilling sequences can be accurately estimated. The optimal drilling order from the learned heuristic search is computed by repeatedly expanding the search in the directions with the highest estimated values. Because the bias correction factor for the drilling-order problem in the REEK Field is relatively stable, adding such a multiplicative factor does not significantly change the initial ranking of the heuristic sequences in the mean model. Hence, we observe that MLHS-SR performed directly in the mean model (i.e., linear approximation) finds the same solution obtained using the bias-corrected mean model. However, for the cases with considerable variability in α, a learned heuristic search with MMBC can potentially find a better drilling sequence than that obtained using the linear approximation.

Figure 19 shows the error on the maximum expected NPV of the complete drilling sequence along the optimal path that is estimated by using heuristic sequences with MMBC and online learning techniques. We first remove the bias in the initial approximations of the expected NPV obtained from the mean model by estimating a bias correction factor α (green curve $\rightarrow$ blue curve), and we then correct the estimates of
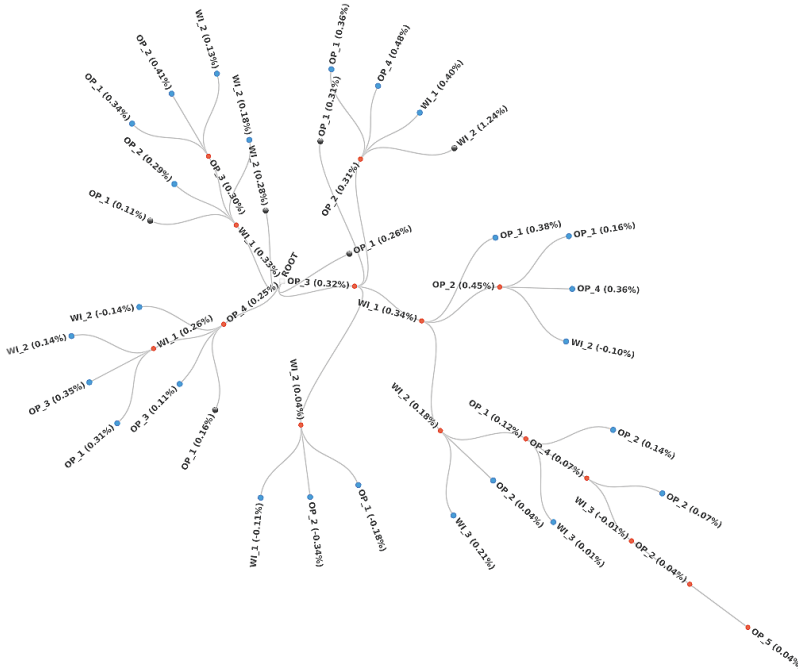
Figure 18: Search strategy of MLHS-SR

the maximum expected NPV obtained from heuristic sequences through online learning mechanisms (blue curve → orange curve). We notice that the initial estimates of the maximum expected NPV from heuristic sequences (blue curve) always overestimate the actual maximum value for the complete sequence and show an exponentially decaying trend with an increasing decay factor throughout the search process. However, the accuracy of the estimated maximum expected NPV is significantly improved by learning the errors on approximate values obtained from previous drilling steps (orange curve). In this case, the average error along the optimal path is only 0.22%. With such accurate estimates, the final optimized sequence is very likely to have a high expected NPV and to be near the true optimal sequence. When MLHS is applied without any space reduction, the average NPV of the optimized sequence is 0.60% lower than that of the solution obtained with space reduction. The main reason for this is that some estimated values underestimate the actual maximum value (e.g., the error for the partial path with three selected wells is -1.03%), meaning that if the expected NPV of the current solution is higher than all of these underestimated values, the search will terminate with the current solution. In this experiment, an appreciable amount of space reduction (e.g., at least half of the remaining wells are considered as the next possible action) is shown to effectively prune paths that are less likely to correspond to the optimal drilling order, leading to a better solution with a lower cost.

Figure 20 compares the NPV distributions for RO/NO sequences and a set of 560 random drilling sequences. The results show that the average NPV is significantly increased by optimizing the drilling sequence with a learned heuristic search. RO sequence has a higher average NPV than either the NO solutions or the set of random
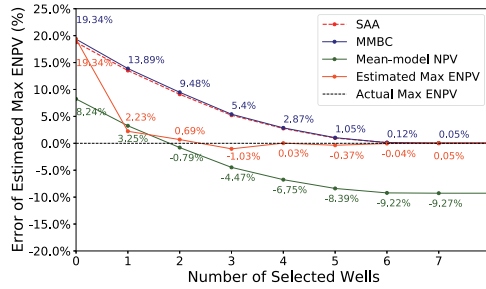
Figure 19: Error of estimates of maximum NPV along the optimal path

	Average NPV (E09 USD)	Simulations	% change
Robust Optimization	8.15	246	Ref
Nominal Optimization	7.66 to 8.15	22 to 49	-6.04% to 0
Random Sequences	6.04 to 8.12	-	-25.86% to -0.36%

Table 2: Comparison of robust and nominal optimization results

drilling sequences, being as much as 25.86% higher than the average NPVs of some of the random drilling sequences. Table 2 summarizes the optimization results for each of the approaches. At the cost of approximately 200 additional simulations, the RO strategy obtained from the mean model with bias correction is shown to be significantly better than the deterministic solutions for individual realizations. We observe many of the NO strategies have similar NPV distributions over a set of 100 realizations. It seems that in the REEK model, the optimal drilling sequence does not vary significantly with the geological uncertainty.

(a) Average NPVs of RO/NO solutions and random drilling sequences

(b) PDFs based on a set of 100 realizations of RO/NO optimal solutions and random drilling sequences
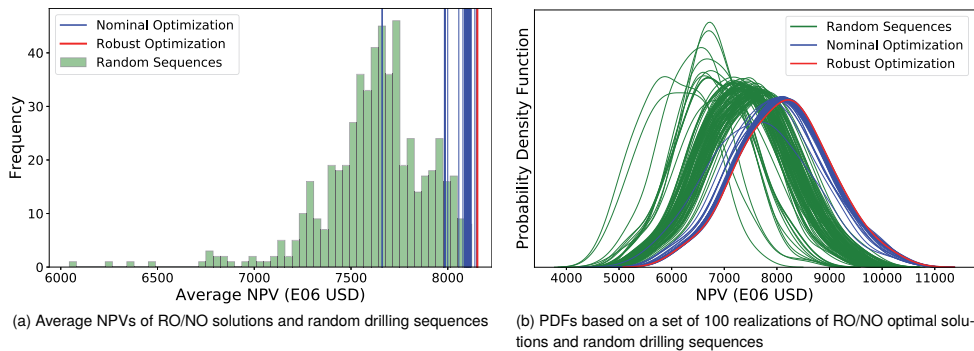
Figure 20: Results obtained through learned heuristic search strategies compared with random drilling sequences

Although the geological properties vary significantly from realization to realization, the optimized deterministic well sequences for 100 model realizations have common characteristics. In particular, the optimized sequence for each single realization always starts with one producer and one injector, in that order. For 92 of the 100 individual realizations, the last well drilled in the optimal sequence is a producer. In most cases, Producer OP_3, and Injector WI_1 are drilled first, and Producer OP_5 is the last well drilled. In the REEK Field case, with five producers and three injectors, the optimal

wells for these three drilling positions appear to be nearly independent of the geological properties. As a result, the RO solution is likely to be near several deterministic optimal solutions for this field. It is clear that this will not always be the case for other fields.

Partial optimization

Wang and Dean [40] have shown that learned heuristic search is a viable means of solving the sequential drilling-order problem for complete sequences. In many cases, however, only the first few wells in the drilling sequence may be needed. As shown in Fig. 18, the optimal partial path will not always extend in only one direction during the search. In such cases, it is not guaranteed that the first few wells can always be optimized along the final optimal complete sequence by limiting the depth of the search path. In this experiment, we investigate the possibility of optimizing the first few wells without finding the entire optimal sequence by performing learned heuristic search with a limited depth (i.e., a depth-limited search and an iterative depth-limited search).

(a) Depth-limited search strategy of the first two wells   (b) Depth-limited search strategy of the first three wells

Figure 21: Optimization of the first two and three wells using MLHS-SR with DLS

Figure 21 shows the search strategies for the first two and three wells using MLHS-SR with depth-limited search (DLS). To optimize the first few wells, we terminate the search with the first selected partial path that includes one additional well. Figure 21a shows that the first optimal partial path with three selected wells is [OP_3, WI_1, OP_2], which indicates that the optimized sequence of the first two wells is [OP_3, WI_1]. The path through Producer OP_3 and Injector WI_1 has been previously expanded as the most promising direction in the second decision step; however, the search changes direction after the evaluation of the extended partial paths due to their underestimated values. Seven additional decision steps are required to guide the search back to the path along the final optimized complete sequence. In this case, if the search were to be stopped immediately with the first selected partial path with two wells (i.e.,[OP_3, WI_1]), then the same solution for the first two wells could be obtained faster (i.e., 26 simulations would be eliminated). However, this will not always be the case for other search depths. Figure 21b shows that although the first expanded direction with three selected wells

is [OP_3, WI_1, OP_2], the path with four wells [OP_3, WI_1, WI_2, OP_1] provides a better solution for the first three wells (i.e., [OP_3, WI_1, WI_2]) since with more observations from longer paths, online learning mechanisms can improve the estimated values and guide the search closer to the true optimal solution. A total of 234 and 240 simulations are required to optimize the first two and three wells, respectively, under geological uncertainty, of which 34 and 40 simulations, respectively, are performed during the search process.

(a) IDLS strategy for the first two wells
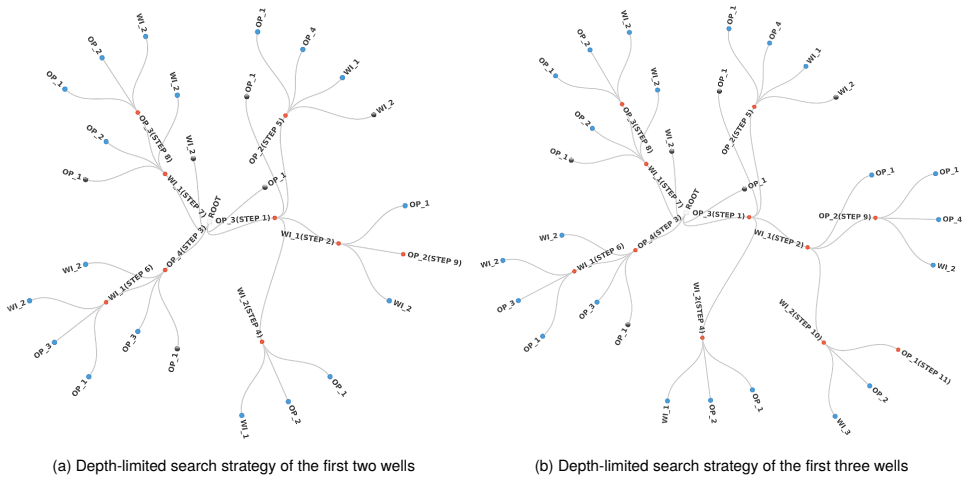(b) IDLS strategy for the first three wells
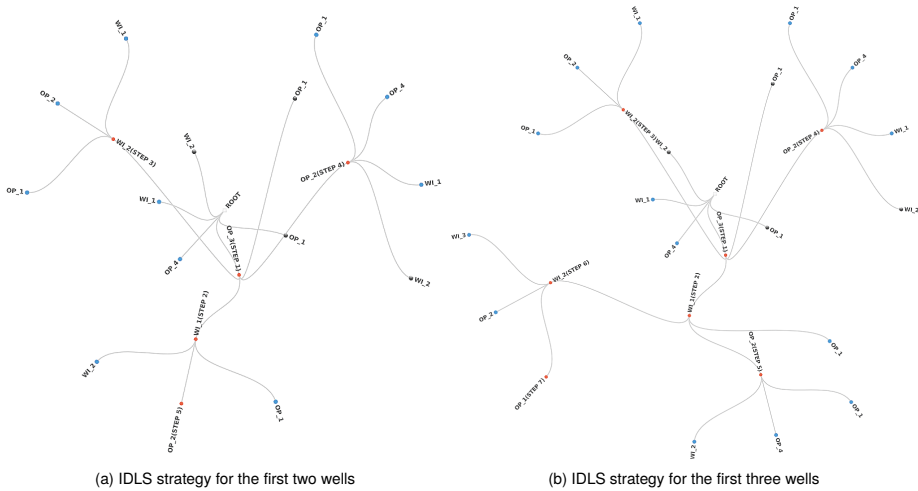
Figure 22: Optimization of the first two and three wells using MLHS-SR with IDLS

Figure 22 shows the search strategies for the first two and three wells using MLHS-SR with iterative depth-limited search (IDLS). In each decision step, we consider only the paths extended from the previous decisions and optimize the next well based on the first selected partial path with two more wells. We optimize the first well based on the first-visited best partial path with two selected wells (i.e., [OP_3, WI_1]), then we compute the second well based on the paths starting with OP_3. In contrast to the search strategies obtained using DLS, as shown in Fig. 21, the paths starting with Injector WI_1 and Producer OP_4 are pruned, thus avoiding 14 simulations. Figure 23 shows the numbers of simulations needed to optimize the first few wells by using MLHS-SR with DLS and IDLS. The results for this example show that both methods can be used to optimize the first few wells along the final optimized complete sequence by controlling the search depth, but IDLS finds a solution faster by eliminating some unnecessary node evaluations caused by underestimated values.

The performance of DLS and IDLS depends on the online learning techniques applied. When the estimated values are sufficiently accurate such that the search process will expand the path in only one direction, both DLS and IDLS can optimize the first few wells along the final optimal complete drilling sequence at the same cost. However, the estimated values are generally less accurate earlier in the search process due to the lack of previous sequential observations, meaning that the search might later change direction. Compared with DLS, IDLS is more likely to optimize the first few wells faster without any loss of the solution quality since paths with low NPV generally

will not continue to be extended when appropriate online learning mechanisms are used. In this case, because only 246 simulations are needed to optimize an entire sequence for eight wells via MLHS-SR, the cost reduction achieved by optimizing only the first few wells is not very significant. For larger fields, however, MLHS-SR with IDLS can be used to efficiently optimize only the first few wells in the sequence.
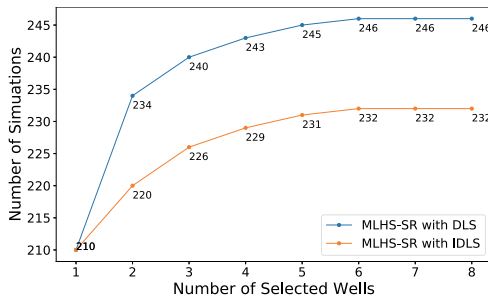
Figure 23: Comparison of the costs of optimizing the first few wells using MLHS-SR with DLS and IDLS

Conclusions

In this paper, we presented a methodology for robust optimization of the expected value of an expensive-to-compute objective function for which the uncertainty is character-ized by an ensemble of model realizations. A straightforward approach using sample averaging to approximate the expected value is prohibitively expensive in most real reservoir applications. In our approach, we compute a multiplicative bias correction to the value of the objective function computed from the first term of the Taylor series expansion. Computation of the bias correction requires two function evaluations for each random sample of the control variables: one evaluation using the mean reservoir model and a second evaluation using a random model realization. Our results show that the information from distinct controls and individual realizations can be used to accu-rately estimate the bias in the approximation of expected NPV obtained from the mean model and that robust optimization can be performed with good approximations of the expected NPV while requiring many fewer simulation runs than SAA. Based on our experiments, we arrive at the following conclusions:

- Approximations of the expected NPV obtained using the mean reservoir model are generally poor compared with estimates obtained using SAA, but approximations from the mean model can be significantly improved by estimating a multiplica-tive bias correction factor, which is estimated by simulating a modest number of randomly selected controls on both reservoir realizations and the mean reservoir model.

- Distance-based localization can improve the estimated values of the correction factor. The performance of this approach the estimation of the bias-correction factor depends on the distance measure and the taper length, both of which must be estimated.

- Regularized localization with an appropriate parameter based on the variance of the bias correction factor reduces the sensitivity of the estimates to the taper length;

consequently, regularized estimates are accurate for a wider range of taper lengths. When the variance of the bias correction factor is known, regularized localization is the preferred method.

- The RO solutions obtained using MMBC and SAA are of similar quality and are both superior to the deterministic solutions for individual realizations and the mean reservoir model (i.e., linear approximation).

- A learned heuristic search can be performed either to optimize complete drilling sequences or to optimize only the first few wells in the sequence at a reduced cost by limiting the search depth. Compared with DLS approach, a solution of the same quality can be found faster by performing an IDLS with effective online learning techniques.

The major advantage of using mean model bias correction is that we only need to perform simulation using the mean model to obtain an initial approximate value of expected NPV, for which the bias will be corrected by a multiplicative correction factor estimated based on the observations of partial corrections from similar controls. Although the methodology was applied to the problem of determining optimal drilling order, the methodology is fairly general and does not require that the objective function be general or that the control variables be continuous. The effect of localization of the estimate could be significant in the case of large variability in the bias correction factor, especially when the ensemble size is large such that more observations from distinct controls and individual realizations could be used for bias estimation. However, to ensure the effectiveness of distance-based localization for bias correction, we need to design appropriate distance measures for identification of sets of similar controls. For the drilling-order problem, a distance metric based on the well position can effectively measure the similarity between drilling sequences. In this work, we presented a method for efficiently estimating the expected value of an objective function by applying a multiplicative bias correction to the value obtained from the mean model. It would not be difficult to modify the bias-correction method for use in multiple objective optimizations under uncertainty (e.g., standard deviation, percentiles, expected value). For example, it would be possible to compute an estimate of the variance of the objective-function value, by squaring the value of the objective function from the mean model and multiplying it by the variance of partial corrections, which can also be estimated by using information from distinct controls and individual model realizations.

Nomenclature

f^*	actual expected objective function value of optimal control
f^*	maximum net present value of optimal control
α	bias correction factor
α	bias correction factor
$\bar{\alpha}$	average value of bias correction factor
$\bar{\varepsilon}$	mean observed single-step error of estimated value
$\bar{m}$	expected value of model parameter
$\bar{m}$	expected value of model parameter
$\bar{\gamma}$	mean observed single-step ratio of estimated value
$\bar{\mu}$	mean observed single-step ratio of variability in estimated value
$\bar{\mu}_{n_s}$	mean single-step ratio of variability in estimated value
β	partial correction factor for individual realizations
β	partial correction factor
Δt	time interval
Δt	time interval
Δx	length interval
δ	distance between controls
γ	single-step ratio of estimated value
γ	single-step ratio of estimated value
$\hat{\alpha}_{\mathrm{loc}}$	local estimate
$\hat{\alpha}_r$	regularized estimate
$\hat{\gamma}$	forecast error of initial evaluation function value
$\hat{f}$	learned evaluation function
$\hat{f}$	learned evaluation function
$\hat{h}$	learned heuristic function

$\hat{\gamma}$ estimated single-step ratio of estimated value

$\hat{\mu}$ estimated single-step ratio of variability in estimated value

λ regularization parameter/Lagrangian parameter (depending on context)

$\mathbf{b_0}$ vector of β at a fixed control x_0

$\mathbf{b}$ vector of β from random controls and realizations

$\mathbf{P}$ vector of drilling sequence

$\mathbf{w}$ weight vector of β

$\text{ECOI}_{a^* \to a^{\text{key}}}$ hidden cost of information caused by sub-optimal solutions constrained to decision a^{key}

EVOI expected value of information

$\text{EVOI}_{a^* \to a^{\text{key}}}$ expected value of information changing decision a^* to a^{key}

EVWI expected value with information

EVWOI expected value without information

μ single-step ratio of variability in estimated value

μ viscosity of fluid

ω weight

Ω^b key observation subspace

ρ distance-based weight

σ_α^2 variance of bias correction factor α

σ_β^2 variance of partial correction factor β

θ permeability

Θ^m subregion of key uncertainty

ε single-step error of estimated value

A cross-sectional area

A decision space

a decision alternative

a^* optimal decision over the current assessment of uncertainty

$a^{*^{\text{fl}}}$ optimal decision considering future information from current decision stage

$a^{*^{\text{fs}}}$ optimal decision considering future information from all remaining decision stages

a^{key} decision alternative providing important information for key uncertainty reduction

b discount rate

b discount rate

C covariance of model parameter

d number of remaining actions

d number of remaining actions

E expected value of objective function

EV^* maximum expected value over current uncertainty state

f evaluation function/estimated maximum net present value

f objective function/evaluation function (depending on context)

f_m first derivatives of objective function with respect to model parameter

f_{mm} second derivatives of objective function with respect to model parameter

g actual economic value from previous actions

g economic value from previous actions

h heuristic function/estimated maximum future value

h heuristic function/estimated maximum future value

h history of past decisions and observations

h^* maximum future value

$IWEI$ injector economic index

J objective function

k log-permeability

L taper length

m model parameter

m model parameter

n_{eff} effective sample size

n_s environment state at decision stage

n_s environment state at decision stage

N_{Ω^b} Number of observation subspaces

N_e Ensemble size

N_e total number of model realizations

N_r number of selected wells

N_s number of remaining wells

N_s number of selected wells

N_w total number of wells

N_w total number of wells

N_x total number of random controls

NPV net present value

NPV net present value

O observation space

o observation obtained from specific decision

o^b best observation subset for key uncertainty reduction

P pressure

$P(\Omega^b \mid \Theta^m)$ probability of observing $o^b \in \Omega^b$ at key uncertainty subregion Θ^m

$P(\Theta^m \mid \Omega^b)$ posterior probability of key uncertainty subregion Θ^m with observation $o^b \in \Omega^b$

$P(\Theta^m)$ prior probability of key uncertainty subregion Θ^m

$P(o \mid h, a)$ probability of observing o from decision a following history h

$PWEI$ producer economic index

q production/injection rate

q production/injection rate

Q^* maximum expected value over all possible future observations from all remaining actions

r reward/cost

r reward/cost

T total number of time steps

T total number of time steps

t time

t time

u uncertainty state

W cost of drilling well

W cost of drilling well

x control variable

x control variable

Δt^{Φ} learning time period

Δt^L learning time period

Φ learning technique index

f_n evaluation function at a specific state

f_n evaluation function at a specific state

h_n heuristic function at a specific state

h_n heuristic function at a specific state

i control index/learning period index/cell index (depending on context)

i key uncertainty subregion index

j decision stage index/model realization index (depending on context)

j model realization index/time step index(depending on context)

j time step index

k key observation subspace index

k well index

L learning technique index

o oil

o oil

s decision stage

s decision stage

w water/well (depending on context)

w water/well(depending on context)

wi water injection

wi water injection

cos   Cosine distance

L_1 Manhattan distance

L_2 Euclidean distance

eff   effective sample size

$\bar{m}$ reservoir mean model

Φ learning

i injection

i injection

L learning

p production

p production

b   best observation subset

lin   linear approximation

m   model parameter

o   observation

quad   quadratic approximation

Acknowledgments

Appendix 1

To go from the second line to the third line in (24), recall that

$$E[\mathbf{b}\mathbf{b}^T] = \text{cov}(\mathbf{b}, \mathbf{b}) + E[\mathbf{b}]E[\mathbf{b}^T] \tag{33}$$

Assuming stationarity, this simplifies to

$$E[\mathbf{b}\mathbf{b}^T] = \text{cov}(\mathbf{b}, \mathbf{b}) + \bar{\alpha}^2 \mathbf{1}\mathbf{1}^T \tag{34}$$

Similarly

$$E[\mathbf{b}\mathbf{b}_0^T] = \text{cov}(\mathbf{b}, \mathbf{b}_0) + \bar{\alpha}^2 \mathbf{1}\mathbf{1}^T \tag{35}$$

The unbiasedness condition gives that

$$\bar{\alpha}^2 \mathbf{1}\mathbf{1}^T \mathbf{w} = \frac{1}{N}\bar{\alpha}^2 \mathbf{1}\mathbf{1}^T \mathbf{1}. \tag{36}$$

Appendix 2

The covariance function C is related to the variance and the semivariogram,

$$C(h) = \sigma^2 - \gamma(h), \tag{37}$$

where σ^2 is the variance, h is the distance between two observations, and $\gamma(h)$ is the semivariogram at distance h.

Since the observation β is obtained from random controls of individual realizations, we need to model the covariance function of β with two terms, i.e., variability in β at a fixed model realization and variability in β at a fixed control variable,

$$C_\beta = C_\beta(h_x) + C_\beta(h_m), \tag{38}$$

where $C_\beta(h_x)$ is the covariance function for β at fixed model realization; h_x is distance between observations; $C_\beta(h_m)$ is the covariance function for β at fixed control variable;

h_m is the distance of realizations corresponding to the observed controls, i.e., $h_m = 0$ if observed values are from the same realization, $h_m = 1$ if observed values are from different realizations.

$C_\beta(h_x)$ can be obtained from the average variogram of β of all ensemble realizations,

$$C_\beta(h_x) = \frac{1}{N} \sum_{j=1}^{N} \sigma_\beta^2(m_j) - \frac{1}{N} \sum_{j=1}^{N} \gamma(h_x, m_j), \tag{39}$$

where $\sigma_\beta^2(m_j)$ and $\gamma(h_x, m_j)$ are the variance and variogram of β at a fixed realization m_j, respectively.

For fixed control variable, the values of β_{ij} and $\beta_{ij'}$ will be correlated for model realizations m_j and $m_{j'}$ at a small distance. Here, we assume that the model realizations are far enough apart such that the observed values of β from different realizations are independent. Then the covariance function for β for fixed control variable can be described as

$$C_\beta(h_m) = \begin{cases} \frac{1}{N_x} \sum_{k=1}^{N_x} \sigma_\beta^2(x_k) & \text{if } h_m = 0 \\ 0 & \text{if } h_m = 1, \end{cases} \tag{40}$$

where $\sigma_\beta^2(x_k)$ is the variance of observed β from different realizations at a fixed control x_k.

$$\text{cov}(\mathbf{b}\mathbf{b}^T) = \begin{bmatrix} \text{cov}(\beta_{11}, \beta_{11}) & \text{cov}(\beta_{11}, \beta_{22}) & \cdots & \text{cov}(\beta_{11}, \beta_{NN}) \\ \text{cov}(\beta_{22}, \beta_{11}) & \text{cov}(\beta_{22}, \beta_{22}) & & \text{cov}(\beta_{22}, \beta_{NN}) \\ \vdots & & \ddots & \vdots \\ \text{cov}(\beta_{NN}, \beta_{11}) & \text{cov}(\beta_{NN}, \beta_{22}) & \cdots & \text{cov}(\beta_{NN}, \beta_{NN}) \end{bmatrix} \tag{41}$$

$$\text{cov}(\mathbf{b}\mathbf{b}_0^T) = \begin{bmatrix} \text{cov}(\beta_{11}, \beta_{01}) & \text{cov}(\beta_{11}, \beta_{02}) & \cdots & \text{cov}(\beta_{11}, \beta_{0N}) \\ \text{cov}(\beta_{22}, \beta_{01}) & \text{cov}(\beta_{22}, \beta_{02}) & & \text{cov}(\beta_{22}, \beta_{0N}) \\ \vdots & & \ddots & \vdots \\ \text{cov}(\beta_{NN}, \beta_{01}) & \text{cov}(\beta_{NN}, \beta_{02}) & \cdots & \text{cov}(\beta_{NN}, \beta_{0N}) \end{bmatrix} \tag{42}$$

Bibliography

[1] A. ALEXANDERIAN, N. PETRA, G. STADLER, AND O. GHATTAS, *Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations*, SIAM/ASA Journal on Uncertainty Quantification, 5 (2017), pp. 1166–1192.

[2] E. G. D. BARROS, S. MACIEL, R. J. MORAES, AND R. M. FONSECA, *Automated clustering based scenario reduction to accelerate robust life-cycle optimization*, in ECMOR XVI-16th European Conference on the Mathematics of Oil Recovery, Barcelona, Spain, 2018, European Association of Geoscientists & Engineers.

[3] H. BEYER AND B. SENDHOFF, *Robust optimization – a comprehensive survey*, Computer Methods in Applied Mechanics and Engineering, 196 (2007), pp. 3190–3218.

[4] M. A. CARDOSO AND L. J. DURLOFSKY, *Linearized reduced-order models for subsurface flow simulation*, Journal of Computational Physics, 229 (2010), pp. 681–700.

[5] M. A. CARDOSO AND L. J. DURLOFSKY, *Use of reduced-order modeling procedures for production optimization*, SPE Journal, 15 (2010), pp. 426–435.

[6] C. CHEN, Y. WANG, G. LI, AND A. C. REYNOLDS, *Closed-loop reservoir management on the Brugge test case*, Computational Geosciences, 14 (2010), pp. 691–703.

[7] P. CHEN, U. VILLA, AND O. GHATTAS, *Taylor approximation and variance reduction for PDE-constrained optimal control under uncertainty*, Journal of Computational Physics, 385 (2019), pp. 163–186.

[8] Y. CHEN, D. S. OLIVER, AND D. ZHANG, *Efficient ensemble-based closed-loop production optimization*, SPE Journal, 14 (2009), pp. 634–645.

[9] J. DARLINGTON, C. C. PANTELIDES, B. RUSTEM, AND B. A. TANYI, *An algorithm for constrained nonlinear optimization under uncertainty*, Automatica, 35 (1999), pp. 217–228.

[10] J. DARLINGTON, C. C. PANTELIDES, B. RUSTEM, AND B. A. TANYI, *Decreasing the sensitivity of open-loop optimal solutions in decision making under uncertainty*, European Journal of Operational Research, 121 (2000), pp. 343–362.

[11] D. DENNEY, *Pros and cons of applying a proxy model as a substitute for full reservoir simulations*, Journal of Petroleum Technology, 62 (2010), pp. 634–645.

[12] A. DOUCET, S. GODSILL, AND C. ANDRIEU, *On sequential Monte Carlo sampling methods for Bayesian filtering*, Statistics and Computing, 10 (2000), pp. 197–208.

[13] L. J. DURLOFSKY, *Upscaling and gridding of fine scale geological models for flow simulation*, in 8th International Forum on Reservoir Simulation Iles Borromees, Stresa, Italy, 2005.

[14] R. E. KORF, *Depth-limited search for real-time problem solving*, Real-Time Systems, 2 (1990), pp. 7–24.

[15] R. M. FONSECA, B. CHEN, J. D. JANSEN, AND A. C. REYNOLDS, *A stochastic simplex approximate gradient (StoSAG) for optimization under uncertainty*, International Journal for Numerical Methods in Engineering, 109 (2017), pp. 1756–1776.

[16] A. T. F. S. GASPAR, E. MUÑOZ MAZO, AND D. SCHIOZER, *Case study of the structure of the process for production strategy selection*, International Journal of Modeling and Simulation for the Petroleum Industry, 4 (2011), pp. 9–15.

[17] G. GASPARI AND S. E. COHN, *Construction of correlation functions in two and three dimensions*, Q. J. R. Meteorol. Soc., 125 (1999), pp. 723–757.

[18] R. HAMMING, *Coding and Information Theory*, Prentice-Hall, 1980.

[19] R. G. HANEA, P. CASANOVA, L. HUSTOFT, R. B. BRATVOLD, R. NAIR, C. HEWSON, O. LEEUWENBURGH, AND R. M. FONSECA, *Drill and learn: A decision making workflow to quantify value of learning*, in SPE Reservoir Simulation Conference, Society of Petroleum Engineers, 2017.

[20] P. E. HART, N. J. NILSSON, AND B. RAPHAEL, *A formal basis for the heuristic determination of minimum cost paths*, IEEE Transactions on Systems Science and Cybernetics, 4(2) (1968), pp. 100–107.

[21] J. D. JANSEN AND L. J. DURLOFSKY, *Use of reduced-order models in well control optimization*, Optimization and Engineering, (2016), pp. 105–132.

[22] M. A. JARO, *Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida*, Journal of the American Statistical Association, 84 (1989), pp. 414–420.

[23] M. JESMANI, B. JAFARPOUR, M. C. BELLOUT, AND B. FOSS, *A reduced random sampling strategy for fast robust well placement optimization*, Journal of Petroleum Science and Engineering, 184 (2020), p. 106414.

[24] S. KIM, R. PASUPATHY, AND S. G. HENDERSON, *A guide to sample average approximation*, in Handbook of Simulation Optimization, Springer, 2015, pp. 207–243.

[25] L. KISH, *Survey sampling*, John Wiley and Sons, Inc, New York, 1965.

[26] A. J. KLEYWEGT, A. SHAPIRO, AND T. HOMEM-DE MELLO, *The sample average approximation method for stochastic discrete optimization*, SIAM Journal on Optimization, 12 (2002), pp. 479–502.

[27] T. KORENIUS, J. LAURIKKALA, AND M. JUHOLA, *On principal component analysis, cosine and euclidean measures in information retrieval*, Information Sciences, 177 (2007), pp. 4893–4905.

[28] L. F. Lamas, V. E. Botechia, D. J. Schiozer, and M. Delshad, *Optimization for drilling schedule of wells in the development of heavy oil reservoirs*, Brazilian Journal of Petroleum and Gas, 11 (2017).

[29] O. Leeuwenburgh, A. G. Chitu, R. Nair, P. J. P. Egberts, L. Ghazaryan, T. Feng, and L. Hustoft, *Ensemble-based methods for well drilling sequence and time optimization under uncertainty*, in ECMOR XV-15th European Conference on the Mathematics of Oil Recovery, Amsterdam, Netherlands, 2016, European Association of Geoscientists & Engineers.

[30] V. I. Levenshtein, *Binary codes capable of correcting deletions, insertions and reversals.*, Soviet Physics Doklady, 10 (1966), pp. 707–710.

[31] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker, Inc., New York, 1992.

[32] S. Rahim and Z. Li, *Well placement optimization with geological uncertainty reduction*, IFAC-PapersOnLine, 48 (2015), pp. 57–62. 9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2015.

[33] C. P. Robert and G. Casella, *Introducing Monte Carlo Methods with R*, Springer, 2010.

[34] S. Ronald, *More distance functions for order-based encodings*, in 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence, 1998, pp. 558–563.

[35] T. Schiavinotto and T. Stützle, *A review of metrics on permutations for search landscape analysis*, Computers & Operations Research, 34 (2007), pp. 3143–3153.

[36] S. Sharma, *Applied Multivariate Techniques*, John Wiley & Sons, Inc., New York, USA, 1996.

[37] V. L. S. Silva, A. A. Emerick, P. Couto, and J. L. D. Alves, *History matching and production optimization under uncertainties–application of closed-loop reservoir management*, Journal of Petroleum Science and Engineering, 157 (2017), pp. 860–874.

[38] R. van Doren, Jorn F. M.and Markovinović and J. D. Jansen, *Reduced-order optimal control of water flooding using proper orthogonal decomposition*, Computational Geosciences, 10 (2006), pp. 137–158.

[39] G. Van Essen, M. Zandvliet, P. Van den Hof, O. Bosgra, and J. D. Jansen, *Robust waterflooding optimization of multiple geological scenarios*, SPE Journal, 14 (2009), pp. 202–210.

[40] L. Wang and D. S. Oliver, *Efficient optimization of well drilling sequence with learned heuristics*, SPE Journal, 24 (2019), pp. 2111–2134.

Paper C

Improving Sequential Decisions – Efficiently Accounting for Future Learning

Lingya Wang, Dean S. Oliver

Improving Sequential Decisions – Efficiently Accounting for Future Learning

Lingya Wang[a,b], Dean S. Oliver[a]

[a] - NORCE Norwegian Research Centre AS, Norway
[b] - University of Bergen, Norway

Abstract

In sequential field development planning, past decisions not only directly affect the maximum achievable expected NPV but also influence the future information that can be used to reduce geological uncertainty. To act optimally, when choosing actions, we must also take into account the opportunities to improve the optimal strategy by reducing future uncertainty. In most applications, however, the effect of future information on the optimal decisions is ignored because it would be computationally intractable to update the reservoir model and re-optimize to account for all possible outcomes of future observations. To efficiently make optimal decisions while considering future possibilities for learning through actions, we developed a flexible workflow built on the key-feature-based value of information (VOI) analysis, which is obtained by identifying key reservoir features for optimization problems and key observations for improving future decisions. Instead of considering future information from all remaining actions, we only consider the important information from key actions to reduce the uncertainty with the largest influence on the optimal strategy – that which would be most helpful in improving future decisions. The efficiency of the method results from the focus on the use of key observations to reduce key uncertainty, rather than using all observations to reduce all uncertainties.

In this work, we built supervised-learning algorithms to identify the optimal combination of observations for reducing key uncertainty and simultaneously to estimate the information's reliability. This allows automatic detection of key observations and direct computation of the posterior probability distribution of key uncertainty based on Bayes' rule, avoiding the need for full history matching to re-estimate the uncertainty. Moreover, the entire key observation space is divided into a limited number of disjointed subspaces, such that observations located in the same subspace have almost the same prediction precision for key uncertainty reduction. It is then only necessary to update the reservoir model for each subspace instead of for all distinct sets of observations. Our methods are illustrated by the application of the drilling-order problem in a synthetic field model, for which the drilling sequence of wells is an important contributor to the reservoir's profitability and for which the optimal solution changes significantly with key reservoir features. Results show that using such a simplified VOI analysis based on key actions and key observations can efficiently improve the expected outcome of an optimal strategy with very little performance loss. Although the key actions provide important information for key uncertainty reduction, taking key action rather than the

initial optimal decision for the current uncertainty state is not always worthwhile even if the information is obtained without explicit cost. Since there may be an indirect cost of information caused by taking an action that appears to be sub-optimal based on past information, it is necessary to consider both the possibility of key uncertainty reduction and the possibility of high expected NPV to determine whether it is worth taking the action to improve future decisions.

Keywords: Dynamic sequential-decision making; Value-of-information analysis; Future learning possibility; Key uncertainty reduction; Supervised learning; Robust optimization;

Introduction

Almost all published reservoir management or field development optimization studies have as a goal, the generation of a sequence of actions that is optimal for the current level of knowledge. There is an implicit assumption that the sequence that is delivered should be adhered to, whatever the results of the drilling or the control settings. These strategies, would in fact be optimal if there was no opportunity to later make revisions. In reality, one would, of course, modify the drilling schedule or the operation of wells as soon as one obtained new information that revealed a different picture of the reservoir.

How should we account for the possibility of learning from actions when optimizing field development for expected net present value (NPV)? To account for future learning requires computation of the value of information (VOI), as it may be advantageous to "pay" for information by making a decision that appears to be sub-optimal for the current assessment of uncertainty in reservoir characterization. If the value of the information obtained by taking an action is greater than the loss of expected NPV, then it is beneficial to take the action.

Unfortunately, while the need to account for the sequential nature of the field development problem is well known, it has generally been ignored in reservoir optimization [17, 24, 29, 7]. The key challenge is that in order to rigorously compute the value of information one must consider all possible values of data that might be obtained from an action then solve a history matching problem with uncertainty assessment for each possible outcome of the data [2, 4, 16, 1]. Then optimization must be performed to determine what action should be taken and a value assigned to all possible outcomes. Although several approaches have been proposed to estimate the value of information [11, 6, 9, 15, 1] for problems in which production flow data must be assimilated, the cost of the combined history matching and optimization is prohibitive for realistic problems. Hence most applications that have considered future learning have had very few decision options (for example drill or not to drill a well) or to problems in which the data assimilation is extremely easy and there are few possible data. Even in those cases however the optimization applications or the data simulation were relatively simple [8, 3, 16].

The problem of robust optimization, taking into account the possibility of uncertainty reduction through the acquisition of data, is closely related to the concept of value of information [25, 12, 5]. The application to closed-loop reservoir management (CLRM) is of particular interest. Barros et al. [4] compute the value of information obtained from an optimal CLRM strategy with traditional production observations. The information is then used to re-estimate uncertainty and re-optimize the controls. They showed that it

was possible to compute the value of CLRM in a rigorous, but highly expensive, way, but they did not use the value of information to modify the optimal controls. In an application to the optimization of drilling order, Hanea et al. [13] also investigated the value of information but, like Barros et al. [4], did not use the value of information to improve the expected outcome of optimization. Torrado et al. [27] applied partially observable Monte-Carlo planning algorithm to optimize the drilling schedule considering future uncertainty reduction based on observations through an entire drilling sequence. Their approach is similar to VOI analysis while potentially evaluating only the strategies with high expected values, and the posterior probabilities of uncertainty are estimated by sampling deterministic realizations at given previous observations instead of through a history matching process. Even so, in the case with only two possible observations from each well, many expensive simulations were still needed to compute the optimal solution, since drilling sequences with all distinct sequential observations had to be evaluated and the number of possible combinations was large. A more general application, in which the value of future information was used to optimize bottom-hole pressure controls on wells in a single inverted 5-spot pattern, has been described by Barros et al. [1]. The procedure was shown to increase the expected value of the field although, as in other applications, the computational cost appears to make the method impractical without substantial modification.

In this paper, we consider a realistic problem in which there are many possible decisions at each step, and many possible data, which are determined by the decisions, and but we make the computation manageable by identifying key information that would help in making optimal decisions and key actions that would result in obtaining that information. Through VOI analysis, we aim to obtain a more robust decision considering the opportunities to improve optimal strategy resulting from future uncertainty reduction. Not all decision alternatives, however, may be able to provide information for making better future decisions. Instead of considering the effects of future information from all possible decisions, an efficient and effective way to account for the possibility of future learning is only taking into account the important information from key actions for characterizing key reservoir features for optimization problems. In this way, a standard VOI analysis with extensive form can be simplified with very little performance loss based on key uncertainty with the largest influence on optimal strategy and key observations for improving future decisions. Moreover, the entire key observation space can be divided into a limited number of disjoint subspaces, i.e., observations located in the same subspace have almost the same prediction precision for key uncertainty reduction. In that case, it is only necessary to update the reservoir model for each observation subspace instead of for all distinct sets of observations. Using such a simplified key-feature-based VOI analysis, it is possible to make optimal decisions efficiently considering future learning possibilities. The performance of this approach is illustrated by the application of the drilling-order problem in a synthetic field model. When evaluating the optimal sequence, we neglect the possibility of learning at later times because that information at late time will generally have smaller effect on the optimization of the first few steps in the sequence.

By identifying key uncertainty for the optimization problems, we can identify key actions that would provide the most valuable future information for improving optimal decisions. To efficiently identify key observations, we build supervised-learning algorithms that are able to capture the mapping between observations and key reservoir

features to automatically detect the optimal combination of observations and simultaneously evaluate the information's reliability for each observation subspace. This allows the direct computation of posterior probability of key uncertainty using Bayes' rule, avoiding the need for full history matching to re-estimate the uncertainty. Note that here we are dealing with information content of hypothetical data – data that might be obtained after drilling a well. The actual data that is obtained will be different because the rates schedule will be different, and the wells may be controlled by tubing head pressure (THP) instead of bottom-hole pressure (BHP), etc. When the actual data are obtained, it is feasible to perform an actual history match and update the model, because only one set of data needs to be history matched in that case.

Hong et al. [16] carefully articulated the concept of VOI from the perspective of decision analysis, and demonstrated the value of obtaining saturation information in a 2D waterflooded reservoir for design of a polymer flood. They conclude, however, that VOI analysis plays no role for water, oil, and gas production rate data and well BHP data "because the data have already been or will definitely be gathered." In contrast, our interest is in focused on the value of information that can be obtained from production data, as the actions that we take in the field control the type of information that is obtained and the timing of the acquisition. Although the information obtained from production data may be obtained without explicit cost, it may have a hidden cost if obtaining it requires one to operate a field sub-optimally for the current uncertainty. An obvious case is the running of a pressure shut-in test to obtain an estimate of reservoir pressure or wellbore skin. If the well is already equipped with a downhole gauge, the cost of the information is largely due to deferred production and the information content from the data is not due only to the fact that pressures are recorded, but also to the fact that the control setting has been altered. In our drilling sequence problem, the timing of information acquisition is at the control of the operator and the "cost" is the loss of expected NPV incurred by drilling the wells in a sub-optimal sequence. To determine whether it is worth taking key action earlier in sequence to obtain the information for improving future decisions, we must evaluate the *net* expected value of information with this indirect cost that is associated with changing optimal decisions for current uncertainty state to key action.

This paper is organized as follows. Section 2 introduces the robust decision-making problem under uncertainty and the technologies we use for solving this problem, including key-feature-based VOI analysis for considering future learning possibilities, the supervised-learning algorithm for identifying key observations, the learned heuristic search method for optimizing sequence of discrete actions [30] and bias-correction methods for estimating the expected NPV [31]. Section 3 presents the numerical results of the drilling-order problem in a synthetic model. In this section, we investigate the effects of various geological features on the optimal drilling sequence, the reliability of the key observations identified using supervised learning models with regard to key uncertainty reduction, and VOI analysis performances through key actions with different initial probabilities of key uncertainty. Finally, the conclusions of this study are provided in Section 4.

Methodology

Robust decision making under uncertainty

The general purpose of robust field development optimization is to identify an optimal
strategy that maximizes the expectation of an objective function (e.g., expected NPV) in
an uncertain reservoir model. Geological uncertainty frequently results in large uncer-
tainty in reservoir performance, but that uncertainty can be reduced using observations
obtained from past decisions through history matching or data assimilation. In tra-
ditional CLRM, the model is updated and uncertainty is re-estimated based on past
observations before making the next decision and the optimal decision for each deci-
sion stage is obtained by performing a re-optimization in the currently updated reservoir
model. In other words, the optimal decision is typically determined by maximizing the
expected NPV over the *current assessment of uncertainty*. The decision we make at the
current time will also influence the possibility of obtaining information that might re-
duce the reservoir uncertainty and improve the optimized strategy. Therefore, the true
optimal action for each decision step depends on both the past decisions and the con-
sequences on future uncertainty reduction. In this section, we use the drilling-order
problem (i.e., maximize the expected NPV by optimizing the drilling sequence of wells)
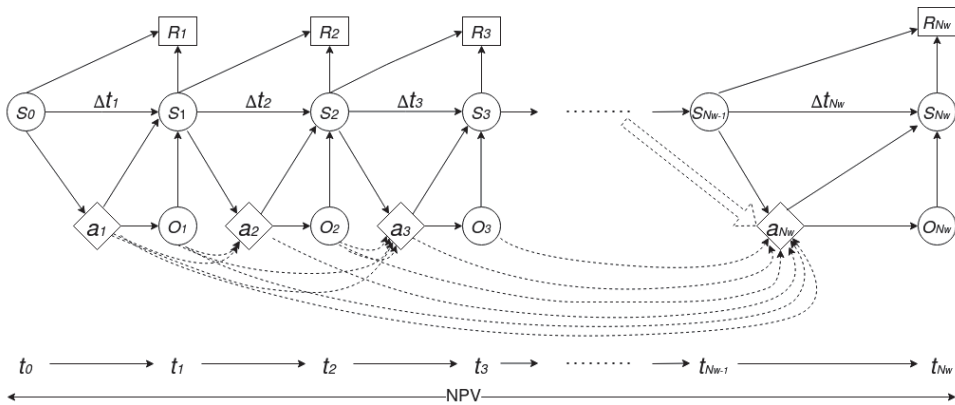as an example to demonstrate the optimal decisions obtained with different concerns.

Figure 1: Example of an observation-based dynamic drilling-sequence planning with N_w wells

Suppose that we need to optimize the drilling schedule of N_w wells and each drilled
well results in observations that can be used to re-estimate uncertainty before choosing
the next well to be drilled. Hence, after drilling each new well, the reservoir model is
updated based on previously obtained observations, before optimizing the next decision.
Figure 1 shows an example path generated by an ordered sequence of N_w drilling actions
with sequential observations from all wells. The set of actions $a_1, a_2, \ldots, a_{N_w}$ represents
the sequence of N_w wells drilled at time $t_0, t_1, \ldots, t_{N_w-1}$. The sequence $o_1, o_2, \ldots, o_{N_w}$
denotes the observations obtained from the drilling of each well. We have assumed that
these observations are immediately available for updating the reservoir model. The state
$s_0, s_1, \ldots, s_{N_w}$ denotes the specific environments at each decision step constrained to the
past decisions and corresponding observations. Uncertainty at each environment state
s_j is re-estimated based on the observations $o_1, o_2, \ldots, o_j$ from the previously j drilled
wells. The expected NPV is the cumulative reward consisting of the sum of rewards
$R_1, R_2, \cdots, R_{N_w}$ over the time periods $\Delta t_1, \Delta t_2, \cdots, \Delta t_{N_w}$. As illustrated in Fig. 1, the

previous and the current decisions affect both the possibility of the future choices of actions and the possibility of future observations. Therefore, the robust optimal decision at each decision step should be determined by considering the possibility of achieving high expected NPV, and the opportunity to improve the optimal strategy based on the future uncertainty state, namely, the future learning possibilities.

Optimization ignoring future learning possibilities

In most applications of CLRM optimization, to reduce the complexity of the sequential decision problem under uncertainty, the effect of future uncertainty reduction on the optimal strategy is ignored when making decisions. In other words, the optimal action for each decision step is computed by performing a re-optimization in the updated reservoir model based on the current assessment of uncertainty, without considering the consequence of this decision on the future uncertainty state [17, e.g.,]. In that case, after completing the drilling of j wells, the next best action a_{j+1}^* is drilling the well that leads to the maximum expected NPV of complete drilling sequences over the current uncertainty state u_j based on the observations in history h_j, i.e.,

$$a_{j+1}^* = \underset{a_{j+1} \in A_{j+1}(h_j)}{\operatorname{argmax}} \operatorname{EV}^* \left(h_j, a_{j+1}, u_j \right), \tag{1}$$

where h_j is an observable history consisting of a sequence of selected actions (i.e., j drilled wells) and observation pairs, $h_j = (a_1, o_1, \ldots, a_j, o_j)$, where observation o_j obtained from each past action a_j might be a single datum (e.g., types of facies) or a collection of data (e.g., production data of various types over a time interval); $A_{j+1}(h_j)$ is the current action space at a given history h_j, which consists of the $(N_w - j)$ remaining wells; u_j is the current assessment of uncertainty based on the past observations from j drilled wells in history h_j; $\operatorname{EV}^* (h_j, a_{j+1}, u_j)$ is the maximum expected NPV for complete drilling sequences over the uncertainty state u_j constrained to history h_j followed by taking a_{j+1} as the next decision. Note that in this approach any possible future information from the remaining $N_w - j$ actions is not considered, including the observation o_{j+1} from the current decision alternatives, a_{j+1}. Thus, the EV^* in Eq. 1 is evaluated over the uncertainty state u_j instead of u_{j+1}. To compute the optimal decision a_{j+1}^*, learned heuristic search [30] is an efficient approach, which allows for optimizing either only the first few decisions or a complete strategy. The key advantage of this approach is that an approximation of the maximum expected value EV^* constrained to the past decisions can be accurately estimated without finding the entire optimal strategy.

Fully structured robust decision making

As discussed in the previous section, when selecting an action that will increase the expected NPV, we should also take into account the possibility of future uncertainty reduction, rather than basing our decision solely on the maximization of expected NPV over current uncertainty. The optimal choice of the next well after sequentially drilling j wells should therefore be based on the expected value over all possible observations from all remaining wells (assuming no explicit cost for collecting information from each drilled well),

$$a_{j+1}^{*^{\text{fs}}} = \underset{a_{j+1} \in A_{j+1}(h_j)}{\operatorname{argmax}} \sum_{o \in O_{a_{j+1}}} p(o|h_j, a_{j+1}) Q_{N_w - (j+1)}^*(h_{j+1}), \tag{2}$$

where $O_{a_{j+1}}$ is the observation space obtained from a_{j+1}; $p(o|h_j, a_{j+1})$ are the marginal probabilities of distinct observations; and $Q^*_{N_w-j}(h_{j+1})$ is the optimal expected value considering all possible future observations constrained to history h_{j+1} including the observation from a_{j+1}. The optimal expected value can be calculated in a backward induction procedure, i.e.,

$$Q^*_{N_w-(j+k)}(h_{j+k}) = \max_{\substack{a_{j+k+1} \in A_{j+k+1}(h_{j+k}) \\ o \in O_{a_{j+k+1}}}} \sum p(o|h_{j+k}, a_{j+k+1}) Q^*_{N_w-(j+k+1)}(h_{j+k+1}),$$

(3)

$$\text{for} \quad k = 1, 2, \ldots, N_w - (j+1),$$

where $Q^*_0(h_{N_w})$ is the expected NPV over the final uncertainty state updated using all sequential observations from a complete drilling sequence in history h_{N_w}. Using backward induction to solve the optimization problem in Eq. 2 is also known as the standard VOI decision analysis process with extensive form [28, 22]. This approach is a fully structured decision tree that considers all possible combinations of the sequences of remaining actions with distinct sequential observations [16].

As a simple illustration, suppose that N_w -2 wells have been drilled sequentially resulting in history h_{N_w-2} and the optimal next well is chosen from the two remaining wells W_a, W_b. Each well can provide two possible distinct observations o_{f_1}, o_{f_2} about the type of facies. Figure 2 shows a simple example of determining the optimal next well from W_a and W_b through the backward induction procedure. In this case, determining the optimal action considering the future information from the two remaining wells, requires consideration of 8 possible combinations of sequences with distinct observations. Because the number of options is small, the optimal action $a^{*\text{fs}}_{j+1}$ based on the expected values over all possible future observations (Eqs. 2-3), is easily determined to drill W_a as the next well. However, as this optimization requires information about the expected NPV from all possible combinations of sequences with observations and about the marginal probabilities of all possible observations from remaining actions, the size of the decision tree is exponential in the number of distinct states related to both the action space and the observation space obtained from each action. Consequently, if there were 8 possible remaining wells while each well provides only two distinct observations, then there would be $8! \times 2^8 \approx 1 \times 10^7$ possible combinations of drilling sequences with distinct sequential observations. The use of such a fully structured decision tree will be computationally intractable even before taking into account the cost of updating the reservoir model.

Although it is possible to approximately solve Eq. 2 by formulating the problem as a partially observable Markov decision process (POMDP) [32, 26], the cost of solving a POMDP can be prohibitive for reservoir applications [27], since the evaluations of the expected values require many expensive simulations and the number of the states that need to be evaluated in a POMDP can be large, especially when many various combinations are likely to generate high expected values. Hence, computing the optimal decision $a^{*\text{fs}}_{j+1}$ that considers all possible future observations is only applicable to reservoir simulation-based problems with small numbers of distinct actions and distinct observations in practice.
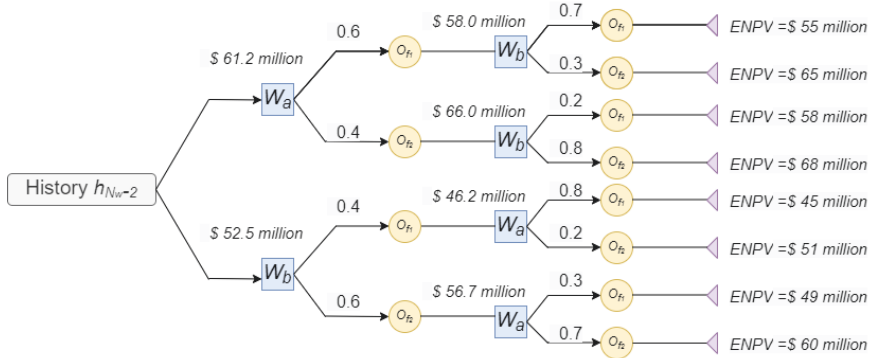
Figure 2: A fully structured decision tree for determining the order of two remaining wells in consideration of all possible future observations

Accounting for future learning through the current decision

Instead of using a fully structured decision tree, a more feasible way to obtain an optimal decision that considers the future learning possibilities is to take into account the effects of future observations resulting from only the current decision step,

$$a_{j+1}^{*\text{fl}} = \operatorname*{argmax}_{a_{j+1}\in A_{j+1}(h_j)} \sum_{o\in O_{a_{j+1}}} p(o|h_j,a_{j+1})\,\text{EV}^*\left(h_j,a_{j+1},u_{j+1}^o\right), \tag{4}$$

where $\text{EV}^*\left(h_j,a_{j+1},u_{j+1}^o\right)$ is the maximum achievable expected NPV constrained to the previous actions in history h_j and the current decision alternative a_{j+1}. This expectation is evaluated over the uncertainty state u_{j+1}^o updated based on the future possible observation o from a_{j+1}. Note that $\text{EV}^*\left(h_j,a_{j+1},u_{j+1}^o\right)$ is different from the expected value $Q_{N_w-(j+1)}^*(h_{j+1})$ in Eq. 2 which accounted for future information from all remaining decisions.

In the terminology of VOI, $a_{j+1}^{*\text{fl}}$ (Eq. 4) is the optimal decision based on the expected value with additional information (EVWI) through one decision point [16],

$$a_{j+1}^{*\text{fl}} = \operatorname*{argmax}_{a_{j+1}\in A_{j+1}(h_j)} \text{EVWI}_{a_{j+1}}, \quad \text{EVWI}_{a_{j+1}} = \sum_{o\in O_{a_{j+1}}} p(o|h_j,a_{j+1})\,\text{EV}^*\left(h_j,a_{j+1},u_{j+1}^o\right),$$
$$\tag{5}$$

while a_{j+1}^* (Eq. 1), that ignores the effects of all possible future observations, is the optimal decision determined by the expected value without additional information (EVWOI),

$$a_{j+1}^* = \operatorname*{argmax}_{a_{j+1}\in A_{j+1}(h_j)} \text{EVWOI}_{a_{j+1}}, \quad \text{EVWOI}_{a_{j+1}} = \text{EV}^*\left(h_j,a_{j+1},u_j\right). \tag{6}$$

We assume that there is no cost for acquiring information from $a_{j+1}\in A_{j+1}(h_j)$. Because $a_{j+1}^{*\text{fl}}$ is obtained considering the possibility of future learning before committing to a decision, $a_{j+1}^{*\text{fl}}$ generally is a more robust decision than a_{j+1}^*, which ignores the effect of all future information.

Figures 3 and 4 show the VOI decision trees from the example of two remaining wells with and without considering the effect of future observations from current decision
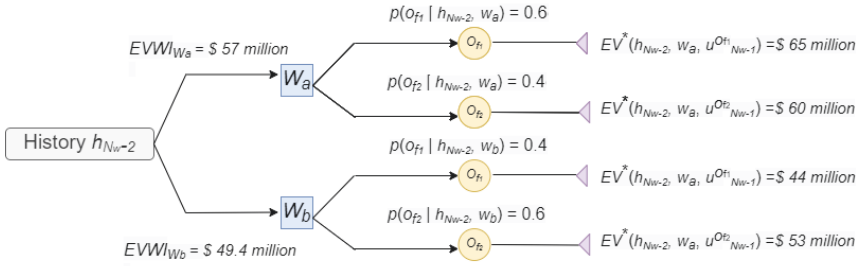
$$p(o_{f_1} \mid h_{Nw\text{-}2}, w_a) = 0.6$$

$$EV^*(h_{Nw\text{-}2}, w_a, u^{Of_1}{}_{Nw\text{-}1}) = \$\,65 \text{ million}$$

$$EVWI_{Wa} = \$\,57 \text{ million}$$

$$p(o_{f_2} \mid h_{Nw\text{-}2}, w_a) = 0.4$$

$$EV^*(h_{Nw\text{-}2}, w_a, u^{Of_2}{}_{Nw\text{-}1}) = \$\,60 \text{ million}$$

$$p(o_{f_1} \mid h_{Nw\text{-}2}, w_b) = 0.4$$

$$EV^*(h_{Nw\text{-}2}, w_a, u^{Of_1}{}_{Nw\text{-}1}) = \$\,44 \text{ million}$$

$$p(o_{f_2} \mid h_{Nw\text{-}2}, w_b) = 0.6$$

$$EV^*(h_{Nw\text{-}2}, w_a, u^{Of_2}{}_{Nw\text{-}1}) = \$\,53 \text{ million}$$

$$EVWI_{Wb} = \$\,49.4 \text{ million}$$

Figure 3: VOI analysis considering the future information obtained from current decision alternatives

$$EVWOI_{Wa} = \$\,46 \text{ million}$$

$$EV^*(h_{Nw\text{-}2}, w_a, u_{Nw\text{-}2}) = \$\,46 \text{ million}$$

$$EV^*(h_{Nw\text{-}2}, w_b, u_{Nw\text{-}2}) = \$\,48 \text{ million}$$
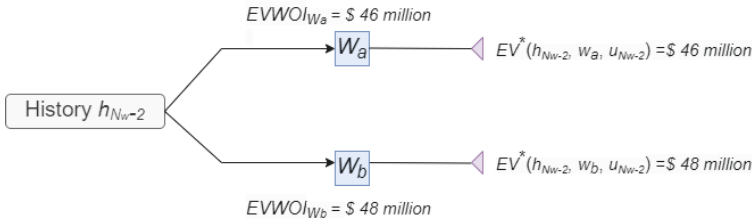
$$EVWOI_{Wb} = \$\,48 \text{ million}$$

Figure 4: A simple example of the decision tree ignoring the future learning possibilities

alternatives. As illustrated in Fig 4, the optimal choice for the next well that ignores the future learning possibilities is W_b, which is obtained by maximizing the expected NPV over the current assessment of uncertainty (Eq. 6). When the effect of future possible observations is considered (Fig. 3), the optimal next well is W_a, which has a higher EVWI and is a more robust decision. For a large problem with many decision alternatives, although the optimal decision $a^{*\text{fl}}_{j+1}$ that is obtained from a simplified VOI decision tree might not be identical to $a^{*\text{fs}}_{j+1}$ that is obtained from a standard VOI analysis with extensive form, the cost of computing $a^{*\text{fl}}_{j+1}$ is much lower than that in $a^{*\text{fs}}_{j+1}$. In general, the information obtained from the later decision stages has a smaller impact on improving the optimal strategy. We expect that simplifying the VOI analysis by only considering the information from the current decision step would not incur much performance loss, i.e., $a^{*\text{fl}}_{j+1}$ is expected to be an approximation solution close to the optimal decision $a^{*\text{fs}}_{j+1}$. In this work, we focus on how to efficiently improve the optimal decision a^*_{j+1} to $a^{*\text{fl}}_{j+1}$.

Although $a^{*\text{fl}}_{j+1}$ does not require a fully structured decision tree, directly solving Eq. 5 is still prohibitive in most reservoir applications for which the costs of history matching and optimization are large. If there are N_d decision alternatives and N_o distinct observations from each decision, it would be necessary to update the reservoir model and perform the optimization $N_d \times N_o$ times to obtain the optimal decision $a^{*\text{fl}}_{j+1}$. Hence, it is desirable to make the computation manageable and design a more practical way to compute the optimal decision $a^{*\text{fl}}_{j+1}$ in consideration of the possibilities of future learning.

In general, we might expect that gained information will reduce the uncertainty, thereby leading to better future decisions. Some decisions, however, may result in little

information or information that is irrelevant to the optimization of the objective. In that case, accounting for the possibility of future uncertainty reduction will only increase the cost of making decisions, while the optimal decision may not be changed (i.e., the optimal strategies for maximizing the expected NPV over uncertainty state u_{j+1}^o and u_j are the same). Hence, because the computational cost of considering many possible observations is high, it is more important to account for the future information that is most likely to improve decisions than to consider as much information as possible from remaining actions. In this work, we use a simplified VOI decision analysis to efficiently account for the possibility of future learning when choosing actions, in which only the key information that would have a large influence on the optimal decisions is taken into account.

Planning for future learning

Learning through key action

Due to limited observations of the reservoir, the properties of the subsurface (e.g., porosity, permeability, fluid contact locations, fault transmissibilities) may be highly uncertain. Uncertainty in some properties may have little effect on the optimal decisions and information on those nonessential properties would not be beneficial for optimally managing the reservoir, even if the uncertainties could be reduced significantly. Hence, when evaluating the desirability of performing an action to learn about the reservoir, we can focus our attention on obtaining information from a few key actions that can be used to reduce key uncertainty that have large impact on optimal decisions. Then, the optimal decision can be made based on the trade-off between the key action that would provide the most important information to reduce key uncertainty and the action that would achieve the maximum expected value over current uncertainties.

Figure 5 shows a feasible workflow that efficiently accounts for the possibility of future learning of key information through key actions that would be most helpful for improving future decisions. By identifying key uncertainty for the optimization problem, we can identify the key information-gathering action that would provide the most important observations for reducing key uncertainty, potentially leading to better future decisions. The key action could provide a large number of observations from various available information sources. To avoid the cost of formal history matching, we select observations for which the connection to uncertainty reduction in key reservoir features is straightforward. In this work, we build supervised-learning models to identify the optimal combination of observations for key uncertainty reduction and simultaneously evaluate the reliability of information. Then, the probability of key uncertainty with the given observations can be computed directly using Bayes' rule instead of using data assimilation algorithms, which typically require many expensive simulations to obtain estimates of the posterior probability distribution. Hence, the workflow is applicable for reducing key uncertainty for optimization problem without the requirement of an expensive history matching process to update the reservoir model.

Although information obtained from a key action is most likely to improve future decisions, taking a key action to acquire important future observations is not necessarily worthwhile, even if these observations may be obtained without an explicit cost. In some cases, the optimal strategy obtained without acquiring future information (i.e., the optimal strategy for current assessment of uncertainty) may have a higher expected value

than that of the optimal strategy considering the possibility of future learning through key action (i.e., the optimal strategy with additional information from key action). The cause of this situation is that there is a hidden cost when taking the key action would lead to a sub-optimal solution, so that value gained by using additional information from key action may not be able to compensate for this hidden cost. To determine whether it is worth taking the key action, we need to assess an implicit *net* expected value of information associated with a change in decisions (i.e., change action obtained from the optimal strategy for the current assessment of uncertainties to key action), which is the difference between the expected value with future information from key action and the expected value without any future information.
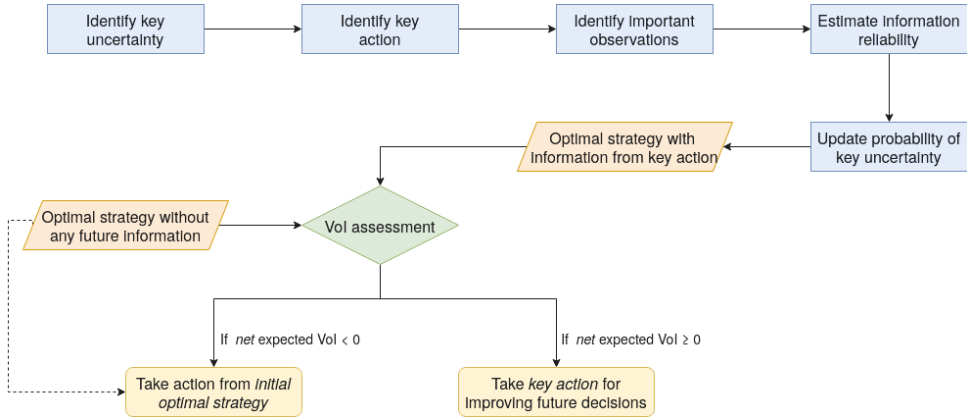

Figure 5: Decision making while considering future learning through key action and key information

Value-of-information analysis through key action

Instead of computing the actual $\text{EVWI}_{a_{j+1}}$ for all possible decision alternatives $a_{j+1} \in A_{j+1}(h_j)$ (Eq. 5) to obtain an optimal decision that considers future learning possibilities, we simplify the VOI decision tree to only two decision alternatives, i.e., a_{j+1}^* and a_{j+1}^{key},
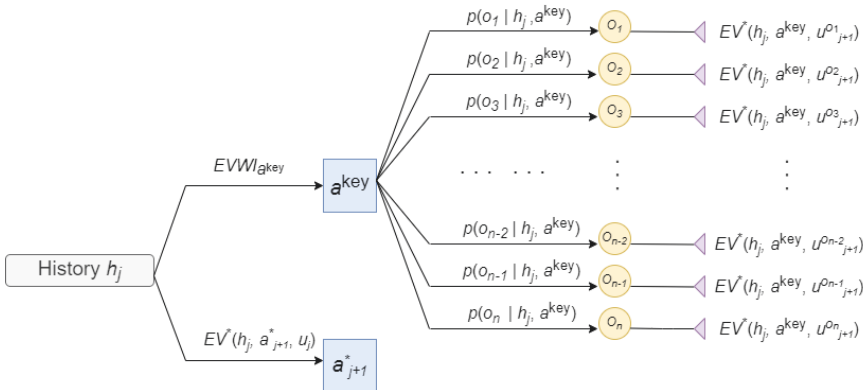

Figure 6: Simplified VOI decision tree with future information only from key action

by considering only future information from the key action (Fig. 6). a_{j+1}^* is the decision obtained from the optimal strategy for the current assessment of uncertainties (Eq. 1), while a_{j+1}^{key} is the key action that would provide important information for improving future decisions. In other words, a_{j+1}^{key} is the decision alternative from the current action space $A_{j+1}(h_j)$ that is expected to result in a high *net* EVOI associated with changing decision a_{j+1}^* to a_{j+1}^{key} defined as,

$$\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}} = \text{EVWI}_{a_{j+1}^{\text{key}}} - \text{EV}^*(h_j, a_{j+1}^*, u_j), \tag{7}$$

where $\text{EV}^*(h_j, a_{j+1}^*, u_j)$ is the expected value from the optimal strategy for current uncertainty without any future information (Eq. 1 and Eq. 6) and $\text{EVWI}_{a_{j+1}^{\text{key}}}$ is the expected value with additional information from a_{j+1}^{key}. The expected value with additional information is computed as

$$\text{EVWI}_{a_{j+1}^{\text{key}}} = \sum_{o \in O_{a_{j+1}^{\text{key}}}} p(o|h_j, a_{j+1}^{\text{key}}) \times \text{EV}^*(h_j, a_{j+1}^{\text{key}}, u_{j+1}^o), \tag{8}$$

where $p(o|h_j, a_{j+1}^{\text{key}})$ is the marginal probability of a specific observation; $O_{a_{j+1}^{\text{key}}}$ is set of all possible distinct observations from a_{j+1}^{key} and $\text{EV}^*(h_j, a_{j+1}^{\text{key}}, s_{j+1}^{o_i})$ is the expected value from the optimal strategy for uncertainty state u_{j+1}^o updated with additional observation from a_{j+1}^{key}.

As shown in Eq. 7, performing the VOI analysis through a_{j+1}^{key} is similar to a simplified VOI analysis with the decision alternative $a_{j+1} \in A_{j+1}(h_j)$ that has a high $\text{EVWI}_{a_{j+1}}$. If a_{j+1}^{key} has the maximum $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$, a_{j+1}^{key} would be identical to the optimal decision a_{j+1}^* from Eq. 5. In that case, simplifying the VOI decision tree to a_{j+1}^{key} would not incur performance loss compared to directly solving Eq. 5. The main advantage of using $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ instead of $\text{EVWI}_{a_{j+1}}$ is that a_{j+1}^{key} with a high $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ can be identified without comparing the actual $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ of all possible decisions.

The $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ from changing decision a_{j+1}^* to a_{j+1}^{key} (Eq. 7) can be rewritten as

$$\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}} = \left(\text{EVWI}_{a_{j+1}^{\text{key}}} - \text{EV}^*(h_j, a_{j+1}^{\text{key}}, u_j) \right) - \left(\text{EV}^*(h_j, a_{j+1}^*, u_j) - \text{EV}^*(h_j, a_{j+1}^{\text{key}}, u_j) \right), \tag{9}$$

where the first group of terms is the standard definition of EVOI for a_{j+1}^{key}, i.e./ it is the difference in the expected values with and without additional information from a_{j+1}^{key},

$$\text{EVOI}_{a_{j+1}^{\text{key}}} = \text{EVWI}_{a_{j+1}^{\text{key}}} - \text{EVWOI}_{a_{j+1}^{\text{key}}} = \text{EVWI}_{a_{j+1}^{\text{key}}} - \text{EV}^*(h_j, a_{j+1}^{\text{key}}, u_j). \tag{10}$$

The second group of terms in Eq. 9 is the expected cost of information (ECOI), or the hidden cost, caused by a sub-optimal solution constrained to a_{j+1}^{key} chosen as the next decision using the current assessment of uncertainty,

$$\text{ECOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}} = \text{EV}^*(h_j, a_{j+1}^*, u_j) - \text{EV}^*(h_j, a_{j+1}^{\text{key}}, u_j). \tag{11}$$

Thus, $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ actually is an implicit *net* $\text{EVOI}_{a_{j+1}^{\text{key}}}$ accounting for the hidden cost $\text{ECOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$,

$$\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}} = \text{EVOI}_{a_{j+1}^{\text{key}}} - \text{ECOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}. \tag{12}$$

If a_{j+1}^{key} is identical to a_{j+1}^*, there is no hidden cost, i.e., $\text{ECOI}_{a_{j+1}^{\text{key}} \to a_{j+1}^{\text{key}}} = 0$. In general, however, a_{j+1}^{key} will not be the same as the optimal decision a_{j+1}^* for the current uncertainty state, especially when the decision space is large.

According to Eq. 12, a_{j+1}^{key} is the decision alternative that is expected to result a large $\text{EVOI}_{a_{j+1}^{\text{key}}}$ but a small $\text{ECOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$. The standard $\text{EVOI}_{a_{j+1}^{\text{key}}}$ (Eq. 10) depends on whether a_{j+1}^{key} is able to provide useful information for making better future decisions, i.e., important information for key uncertainty reduction. The hidden cost $\text{ECOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ depends on whether the optimal strategies constrained to a_{j+1}^{key} can achieve a high expected NPV over current uncertainty state, which is possible to be evaluated when computing the optimal decision a_{j+1}^* that ignores the effects of future possible observations. Therefore, by considering the possibility of obtaining valuable information for reducing key uncertainty and the possibility of achieving a high expected NPV for current assessment of uncertainty, we can identify the a_{j+1}^{key} that is likely to result in a high $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$; that is the decision alternative $a_{j+1} \in A_{j+1}(h_j)$ with a high $\text{EVWI}_{a_{j+1}}$.

To judge whether it is preferable to take action a_{j+1}^{key} over a_{j+1}^*, we need to assess $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ before committing to a decision, noting that $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ may be negative value to the hidden cost. If $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}} > 0$, the value gained by using the information from a_{j+1}^{key} can compensate for the hidden cost caused by a sub-optimal solution, i.e., $\text{EVOI}_{a_{j+1}^{\text{key}}} > \text{ECOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$. It is then worth taking a_{j+1}^{key} to acquire the information that will help in improving future decisions.

The simplified VOI analysis based on $\text{EVOI}_{a_{j+1}^* \to a_{j+1}^{\text{key}}}$ does not take in account the future learning possibilities through a_{j+1}^* (Fig. 6). In some cases, a_{j+1}^* may also be able to provide important future information for reducing key uncertainty. To obtain a more robust decision, we can take into account the possibilities of future learning through both a_{j+1}^{key} and a_{j+1}^*,

$$\hat{a}_{j+1}^{*\text{fl}} = \arg \max_{a \in [a_{j+1}^*, a_{j+1}^{\text{key}}]} \text{EVWI}_a = \arg \max_{a \in [a_{j+1}^*, a_{j+1}^{\text{key}}]} \sum_{o \in O_a} p(o|h_j, a) \times \text{EV}^*(h_j, a, u_{j+1}^o). \quad \text{(13)}$$

However, this approach will increase the computational cost of making a decision since the evaluation of EVWI for each action is based on the maximum expected values, $\text{EV}^*(h_j, a, u_{j+1}^o)$, corresponding to various observations. This requires re-optimization multiple times to obtain all expected values. Consequently, directly solving Eq. 7 or Eq. 13 is likely to be impractical when all distinct observations are accounted for, although the number of decision alternatives in VOI analysis is reduced by identifying a_{j+1}^{key}. To the computation of EVWI manageable, we will present a methodology in the following section for efficiently estimating EVWI by using key observations to reduce key uncertainty, rather than using all observations to reduce all uncertainties.

Key observation selection

In ensemble-based methods, a set of N_e model realizations is used to represent the uncertainties in reservoir properties. To reduce the effects of sampling error, N_e is typically chosen to be on the order of 100. Each model realization is capable of generating a specific set of simulated observations obtained by taking action a_{j+1}^{key}, e.g., production data over a certain period, in which case there would be N_e distinct sets of

observations – one set from each ensemble member. If each set of observations is used in the estimation of $\text{EVWI}_{a^{\text{key}}}$, the computational cost will be high since both history matching and optimization have to be performed N_e times to obtain the maximum expected values from all posterior ensembles. Hence, it is generally infeasible to consider all N_e distinct realizations of future outcomes in standard VOI analysis with history matching for making a decision, especially when the observation space or decision space is large. Moreover, when all observations obtained from a_{j+1}^{key} are used to simultaneously re-estimate uncertainty, the largest decrease in uncertainty may be in properties that are irrelevant to current decisions. All data will, of course, be used to update model uncertainty after an action has been taken.

As mentioned previously, the additional value with information is achieved by the reduction in uncertainty of model parameters that will affect the optimal decisions, rather than all reductions in uncertainties in model parameters. Instead of using all observations to reduce all uncertainties, we can approximately compute $\text{EVWI}_{a^{\text{key}}}$ by reducing key uncertainties of the optimization problem based on key observations that are most helpful in exploring key reservoir properties. Such an approximation of $\text{EVWI}_{a_{j+1}^{\text{key}}}$ can be used to indicate the importance of a_{j+1}^{key}. Because VOI analysis is performed by ranking the importance of decision alternatives based on the expected values and $\text{EVWI}_{a_{j+1}^{\text{key}}}$ deals with the information content of hypothetical data, we expect (without proof) that using the $\text{EVWI}_{a_{j+1}^{\text{key}}}$ computed based on key information would not incur performance loss in VOI analysis and the optimization framework. When the actual observations are obtained from an action that has been executed, an actual history match will be performed with all observations to update various uncertainties in reservoir properties.

Performing the action a_{j+1}^{key} can provide a large number of observations, but the reduction in key uncertainties from some observations may be very small. Accounting for all information in the VOI, including those nonessential observations, will increase the computational effort associated with updating the reservoir model. Hence, we would like to use a reduced set of important observations to update key uncertainties, i.e., key information is defined to be a subset of observations that are most helpful in reducing key uncertainties. Instead of updating key uncertainties for each possible outcome of key observations obtained from an individual realization, we divide the entire observation space R_b^n associated with the best subset, B, into a limited number of disjoint subspaces (e.g., $R_b^n = \Omega_1^b \cup \Omega_2^b$ and $\Omega_1^b \cap \Omega_2^b = \emptyset$). Suppose that observations located in the same subspace have almost the same prediction precision to reduce uncertainties, then the posterior probability distributions of uncertainties conditioned on observations in the same subspace would be similar. In that case, there is no need to compute N_e posterior ensembles considering all distinct sets of observations from individual realizations. $\text{EVWI}_{a_{j+1}^{\text{key}}}$ could be efficiently evaluated by performing the optimization process only in a few posterior ensembles associated with the observation subspaces Ω_k^b,

$$\text{EVWI}_{a_{j+1}^{\text{key}}} = \sum_{k=1}^{N_{\Omega^b}} p(o^b \in \Omega_k^b | h_j, a_{j+1}^{\text{key}}) \times \text{EV}^*(h_j, a_{j+1}^{\text{key}}, u_{j+1}^{\Omega_k^b}), \tag{14}$$

where N_{Ω^b} is the number of observation subspaces, which is much smaller than the ensemble size ($N_{\Omega^b} \ll N_e$); $u_{j+1}^{\Omega_k^b}$ is the updated uncertainty state for observed values

$o^b \in \Omega_k^b$. Figure 7 shows an example of VOI analysis considering only the important future observations through key action for key uncertainty reduction, while the entire key observation space is divided into four disjoint subspaces. We refer to such a simplified VOI decision tree as *key-feature-based* VOI analysis. It is performed only through the key action and the key information that have been identified for exploring reservoir features with large influences on the optimal decisions.

To ensure the usefulness of key observations and their subspaces in reducing key uncertainties, the entire key observation space is divided such that each subspace Ω_k^b will have high a probability $P(\Omega_k^b|\Theta_k^m)$ for indicating a specific key uncertainties sub-region Θ_k^m, while the probability $P(\Omega_k^b|\Theta_i^m)$ for key uncertainties located in other sub-regions Θ_i^m is low. Suppose that distribution of key uncertainties is divided into N_{Ω^b} disjoint subregions $\Theta^m = [\Theta_1^m, \Theta_2^m, \ldots, \Theta_{N_{\Omega^b}}^m]$. The best observation space division $\Omega^b = [\Omega_1^b, \Omega_2^b, \ldots, \Omega_{N_{\Omega^b}}^b]$ can then be described as

$$\Omega^b = \underset{\Omega^b}{\operatorname{argmax}} \sum_{k=1}^{N_{\Omega^b}} \left[P(\Omega_k^b|\Theta_k^m) - \sum_{\substack{i=1 \\ i \neq k}}^{N_{\Omega^b}} P(\Omega_k^b|\Theta_i^m) \right]. \tag{15}$$

For key uncertainties with categorical variables, each category can be set as one specific subregion Θ_i^m. For continuous variables, instead of randomly dividing the distribution of key uncertainties into a limited number of subregions, the division of Θ^m can be optimized based on the performance of the corresponding Ω^b,

$$\Theta^m = \underset{\Theta^m}{\operatorname{argmax}} \sum_{k=1}^{N_{\Omega^b}} P(\Omega_k^b|\Theta_k^m), \tag{16}$$

which is a simplification of Eq. 15 since $\sum_{k=1}^{N_{\Omega^b}} P(\Omega_k^b|\Theta_i^m) = 1$ for a specific key uncertainties subregion Θ_i^m.

Based on the prior probability $P(\Theta_i^m)$ of each subregion Θ_i^m and the information's reliability $P(\Omega_k^b|\Theta_i^m)$, the posterior probability $P(\Theta_i^m|\Omega_k^b)$ can be computed using Bayes' rule,

$$P(\Theta_i^m|\Omega_k^b) = \frac{P(\Omega_k^b|\Theta_i^m) \times P(\Theta_i^m)}{\sum_{i=1}^{N_{\Omega^b}} P(\Omega_k^b|\Theta_i^m) \times P(\Theta_i^m)}, \tag{17}$$

where $\sum_{i=1}^{N_{\Omega^b}} P(\Omega_k^b|\Theta_i^m) \times P(\Theta_i^m)$ is the marginal probability of observing $o^b \in \Omega_k^b$ in the prior ensemble.

Approximations applied to solve Eq. 15 would affect the performance of key uncertainty reduction. To ensure the effectiveness of using Ω^b to reduce key uncertainties, an appropriate approach that can effectively identify the observation subspaces Ω^b with high information's reliability $\sum_{k=1}^{N_{\Omega^b}} P(\Omega_k^b|\Theta_k^m)$ is required. In this work, we apply multiple supervised-learning algorithms to identify the optimal observation subset and the corresponding best space division Ω^b. Meanwhile, the reliability of information $P(\Omega_k^b|\Theta_k^m)$ for each subspace can also be estimated when evaluating the learning algorithm's performance with the optimal observation subset. Then, the posterior probability of key uncertainties can be computed using Bayes' theorem (Eq. 17). Consequently, there is no need to use data assimilation algorithms that update every model parameter for VOI
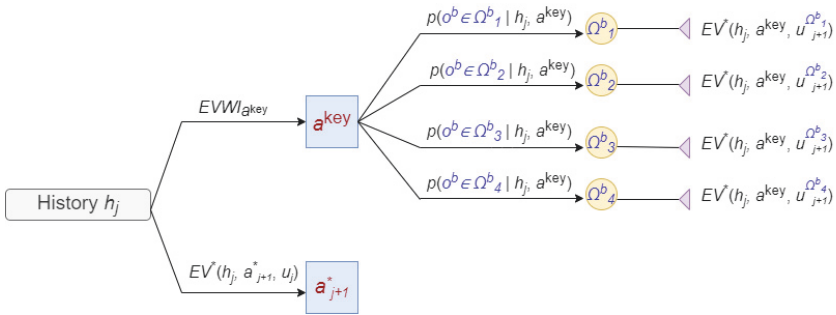
Figure 7: Key-feature-based VOI analysis obtained by identifying key uncertainties for optimization problems and key observations for reducing key uncertainty
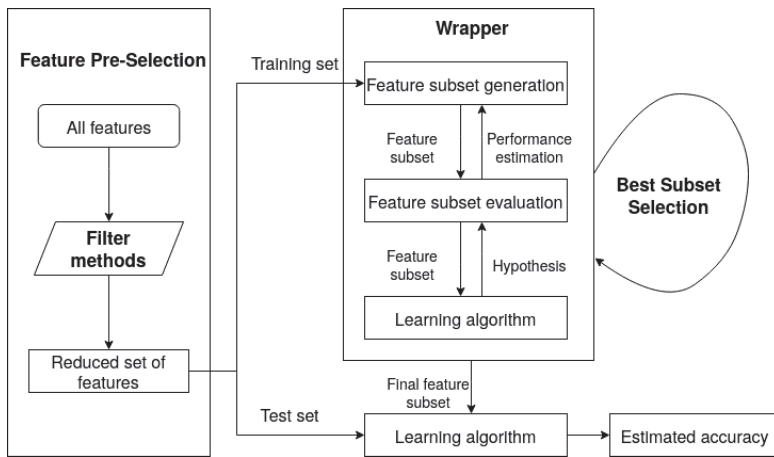

Figure 8: Key feature selection process based on filter and wrapper methods

analysis in our workflow. Building supervised-learning models to identify key observations requires a dataset (also called the original dataset) that includes all possible observations and the corresponding distribution of key uncertainties. This dataset can be obtained by collecting relevant information from a number of individual realizations applied with key action.

Figure 8 shows the process of selecting key observations from the original dataset by using filter and wrapper methods with an inductive learning model that is able to capture the mapping between the inputs (i.e., observations from key action) and the outputs (i.e., key uncertainty). The original dataset may contain hundreds to thousands of observations, which are known as features in learning models. Note that in supervised-learning algorithms, the observations acquired from key action are input variables and are called "features", while the values of key reservoir properties are output variables. A large number of features would make the model more complex and may lead to overfitting due to the curse of dimensionality. To avoid these issues, we first apply filter methods [19] to quickly remove redundant and irrelevant features by ranking the features using some relevance measure, regardless of learning algorithms, obtaining a

subset of features. The selected feature subset after filtering is generally not the optimal feature subset for key uncertainty reduction. Thus, a wrapper method [18] involved in supervised-learning models is used to select the best combination of features that gives the optimal results for learning algorithms, i.e., a feature subset that leads to high prediction accuracy. To avoid overfitting in learning models, we split the original dataset into separate training and test subsets and use resampling methods (e.g., cross-validation) to evaluate learning models' performance with limited data samples. The prediction accuracy (i.e., reliability of information) is estimated from the test error associated with specific learning models.

Suppose that N_{sl} individual realizations are applied to generate the dataset for building supervised-learning models. The cost of updating the key uncertainty for all observation subspaces will be N_{sl} simulations. Using multiple supervised-learning algorithms will not increase the simulation cost since the learning models are built using the same samples. If N_{opt} simulations are required for a single robust optimization, it will require $N_{sl} + N_{opt} \times (N_{\Omega}^{b} + 1)$ simulations to perform the VOI analysis through key action (Eq. 7) and key information (Eq. 14), which is much lower than the cost of solving Eq. 4 with an exhaustive history matching and optimization procedure that requires $N_d \times N_e \times (N_{hm} + N_{opt})$, where N_d is the number of decision alternatives and N_{hm} is the cost of history matching. Even accounting for future learning possibilities through all N_d decision alternatives, performing VOI analysis through key information still requires many fewer simulations than directly solving Eq. 4. In that case, the cost of VOI analysis through key information is $N'_{sl} + N_s \times N_{\Omega}^{b} \times N_{opt}$, in which N'_{sl} simulations are performed to investigate the reliability of information from all N_d possible decision alternatives, and N_s is the number of decisions that are identified with reliable information for key uncertainty reduction and small hidden cost caused by sub-optimal solutions, which is generally smaller than N_d.

In this paper, the performance of key-feature-based VOI analysis (Fig. 7) applied with supervised-learning algorithms (Fig. 8) is illustrated by an application of the drilling-order problem in a synthetic model (REEK field). The key uncertainty that has the largest influence on the optimal drilling sequence is whether one fault is completely sealing or not, for which the output variable in learning models is a category. In that case, there is no need to optimize the division of key uncertainty since each subregion Θ_i^m corresponds to a specific category. We first use the Minimum Redundancy Maximum Relevance (mRMR) method [21, 10, 23] to eliminate some less important features, then further reduce the number of features by using the area under the receiver operating characteristic (ROC) curve [14], which measures the classification performance at various thresholds. These two steps are independent of any learning algorithms. To obtain the optimal combination of observations, we investigate four different classification models (i.e., k-Nearest Neighbor, Logistic Regression, Support Vector Machine, and Random Forest). We then use the best-performing learning algorithm to identify the best observation subset and evaluate each subspace's prediction accuracy based on the test dataset.

Robust optimization of well drilling schedule

In order to perform the optimization efficiently, we require two additional technologies to deal with the search for an optimal sequential solution, and to deal with uncertainty

in the reservoir characteristics. Learned heuristic search is an effective and efficient search method for solving the optimization problems with discrete actions [30]. This approach allows for optimizing only the first few actions by limiting the search depth so that the optimal well for each decision step could be obtained at a reduced cost without finding the entire optimal drilling sequence [31]. The key point of this method is that an accurate approximation of the maximum achievable expected NPV constrained to previous wells, i.e., $\mathrm{EV}^*(h_j, a_{j+1}, u_j)$ in Eq. 1, can be evaluated by first using a crude heuristic function to estimate the maximized value. The accuracy of the heuristic is then improved by learning the errors of the initial approximate values obtained from previous decision steps. In this way the search direction can be guided toward the optimal solution quickly and effectively. In this paper, we apply two different online-learning techniques (i.e., single-step adjustment and multiple-time-periods learning) to improve the initial imprecise heuristic values, which can be inexpensively obtained by assuming that all remaining wells are drilled simultaneously at the next step and then put on production immediately after completing the drilling of wells.

The second technology is the application of bias-correction methods to the estimate of NPV obtained from the mean reservoir model to efficiently compute a good approximation of the expected NPV over an ensemble of reservoir model realizations [31]. Although the mean model $\bar{m}$ generally provides a poor estimate of the expected value when the objective function $J(x, m)$ at control x is nonlinear in the uncertain model parameter m, this approximation can be significantly improved by estimating a multiplicative bias correction factor $\alpha(x)$. The estimation only requires information from individual simulations with distinct controls and model realizations, i.e.,

$$E[J(x,m)] \approx \hat{\alpha}(x)J(x,\bar{m}), \quad \hat{\alpha}(x) = G(\beta_1, \beta_2, \cdots, \beta_n, x), \tag{18}$$

where $\hat{\alpha}(x)$ is the estimate of bias correction factor between the ensemble average value and the value obtained from the mean model, i.e., $\hat{\alpha}(x) \approx \frac{\sum_{j=1}^{N_e} J(x,m_i)}{N_e J(x,\bar{m})}$. G is an estimating function for $\alpha(x)$ based on a set of observations β obtained by applying n randomly sampled controls to individual realizations and the mean reservoir model, where β_j is the partial correction factor at a random control x_j of a random individual realization m_j, i.e., $\beta_j = \frac{J(x_j,m_j)}{J(x_j,\bar{m})}$. When estimating $\alpha(x)$, high weights would be assigned to the observed values of β from similar controls because they are expected to provide more useful information. Using such a bias-correction method, robust optimization requires additional simulations only from the mean model during the optimization process, in which case the robust optimal solution could be obtained at a much lower cost compared to using the ensemble average of simulation results to estimate the expected value.

There are three different ways to estimate $\alpha(x)$: distance-based localization, regularized localization, and optimal weights based on the covariance of correction factors. In this work, we estimate the expected NPV by applying distance-based localization to correct the bias in $J(x, \bar{m})$ since the other two methods require additional information such as the variance of the bias correction factor, which is generally unknown. To efficiently identify similar drilling sequences, a well-position based distance metric is used to measure the similarity of drilling sequences in terms of the bias correction factor.
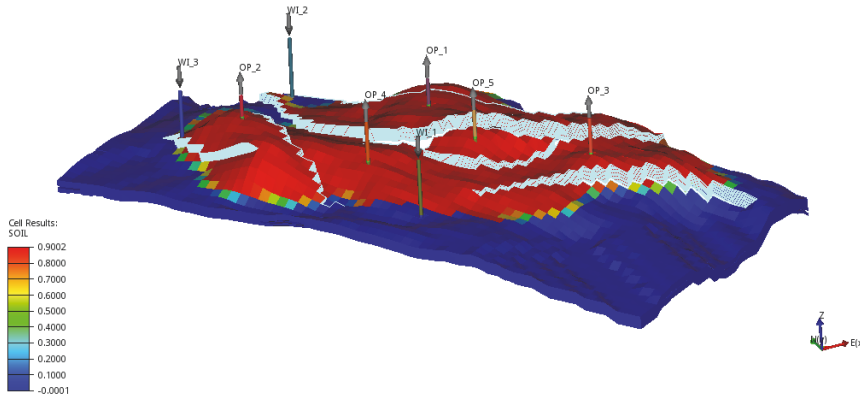
Results and Discussion

Figure 9: Well and fault locations in the REEK Field and initial oil saturation in one realization

Reservoir model

REEK Field is a three-phase black-oil reservoir model with specified locations for eight vertical wells (five producers and three injectors) that need to be drilled sequentially (Fig. 9). It consists of $40 \times 64 \times 14$ grid cells, of which 34,770 are active. The maximum rates of production and injection wells are 6,000 m^3/day and 10,000 m^3/day, respectively. The minimum BHP of the producers is 250 bar, while the BHP of the injectors cannot exceed 320 bar. The porosity field, permeability field, and fault transmissibility multipliers are all uncertain. Recent studies have used this model for optimizing the drilling order of wells, in which the control variables are discrete[31, 30, 13, 20]. However, these studies do not consider the possibilities of future learning through action in the optimization process, i.e., the optimal well for each decision step is based on the current assessment of uncertainty (Eq. 1). To obtain a more robust optimal solution, we will apply a simplified VOI decision analysis based on key reservoir features and key observations (Fig. 7). This method will enable us to make the optimal decision more efficiently while considering the future reduction of key uncertainty in the drilling-order problem.

In the original REEK model, the drilling schedule of wells is an important contributor to the reservoir profitability, i.e., the expected NPV of the optimized drilling sequence could be as much as 25 % higher than random drilling schedules. However, the deterministic optimal drilling sequence does not change significantly with the geological uncertainty [31]. To increase the effect of geological properties on the drilling-order problem, we modified the original REEK Field by extending fault F5 (Fig. 10) so that the reservoir model could be separated into two compartments, and we assume this extended fault F5 would either be non-sealing or sealing in individual model realizations. When fault F5 is completely sealing, the injector WI_1 will be totally isolated from all the other wells, and there is no benefit from drilling WI_1. In that case, the optimal drilling sequence of the remaining wells may change significantly. Table 1 shows the economic parameters used for computing NPV (10 years), the reservoir properties, and control variables in this modified REEK Field. In this work, we have used the same values for economic parameters as in Hanea et al. [13], which investigated the impact of history matching well data on creating value after re-optimizing the drilling schedule of wells
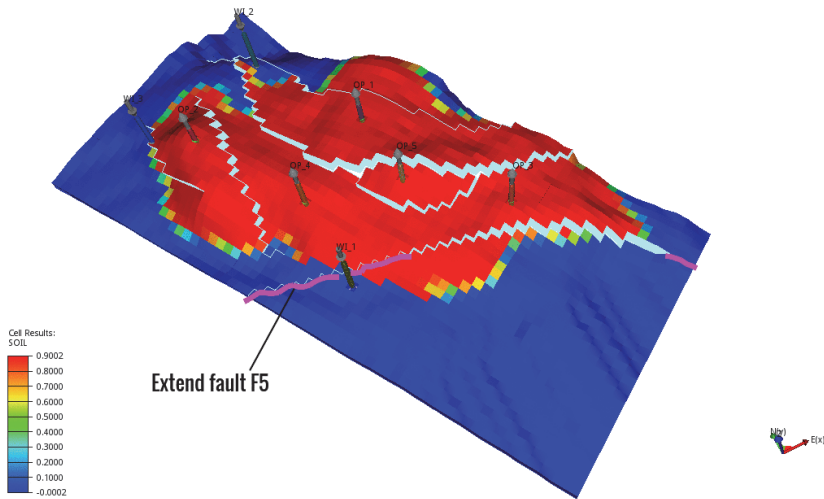
Figure 10: Modified REEK model with extended fault F5 near Injector WI_1

Field	REEK model
Start time	1/12/1999
Time period of NPV (years)	10
Discount factor	0.08
Produced-oil price (USD/m^3)	60
Water-separation cost (USD/m^3)	5
Water-injection cost (USD/m^3)	1
Drilling cost of each well (USD)	1 million
Number of grid blocks	$40 \times 64 \times 14$ (34,770 active cells)
Number of faults	6 (fault F5 is extended)
Number of wells (all vertical wells)	8 (5 producers and 3 injectors)
Drilling period of wells(months)	6
Maximum production rate (m^3/day)	6000
Minimum BHP of producers (bars)	250
Maximum injection rate (m^3/day)	10000
Maximum BHP of injectors (bars)	320
Number of geological realizations	100
Permeability (md)	0 to 3500 (average 733)
Porosity	0 to 0.45 (average 0.159)
Fault transmissibility multiplier of other faults	0 to 1 (average 0.105)
Fault transmissibility multiplier of fault F5	0 (sealing) or $>$ 0 (non-sealing)

Table 1: Economic parameters for NPV and reservoir properties in modified REEK model

in the REEK model. Although the oil price has consequently been set to a low value, it does not affect the conclusions from our experiment results.

To obtain useful information for improving optimal decisions, we first identify the key uncertainty of the drilling-order problem in this modified REEK Field, i.e., illustrating whether a non-sealing/ sealing fault F5 has the largest influence on the optimal drilling sequence. Then, we identify the key action that would provide the most important observations for reducing the key uncertainty. To compute a robust, optimal, and complete solution, we could apply the workflow that accounts for future learning possibilities through key actions (Fig. 5) at each decision step before drilling a new well. Since information obtained at the early stage is most effective for improving the future decisions and thereby increasing the expected NPV after re-estimating the geological uncertainty[13], we simplify the drilling-order problem in our application and neglect the possibilities of future learning at later stages when evaluating the optimal complete drilling sequence, i.e., only identify the key action for the *first decision step* and update the reservoir model based on the *key observations* from that action before drilling the second well. In this paper, we did not reduce the key uncertainty through a history matching process, although that would naturally occur after production data is obtained. Instead, we updated the reservoir model directly using Bayes' rule because the reliability of information collected for the reduction of key uncertainties can be evaluated simultaneously when identifying the best observation subset using a supervised learning model. Because the VOI assessment depends both on the accuracy of information and the prior probabilities, we will study the performances of VOI analysis through key action for the drilling-order problem with different initial probabilities of key uncertainty.

Key uncertainty for drilling-order problem

To identify the key uncertainty of the drilling-order problem, we performed a simple Monte Carlo experiment in which all variables were perturbed simultaneously, and then optimized for each realization for studying the sensitivity of the optimal solution to different geological features. In this work, we use the Manhattan distance, i.e., the sum of the absolute differences between positions of wells in the drilling sequence, to measure the similarity between optimal drilling sequences [31]. Figure 11a compares the distributions of Manhattan distance between deterministic optimal drilling sequences of individual realizations in two cases, one in which fault F5 is always non-sealing (histogram in yellow) and another in which fault F5 alternates between being non-sealing and sealing (histogram in blue). To obtain reliable results, we repeated the experiment 100 times for each case. The optimal drilling sequence clearly varies more significantly when fault F5 is changed from non-sealing to sealing. For the individual realizations with various porosity fields, permeability fields, and fault transmissibility multipliers while fault F5 is always non-sealing, the average Manhattan distance between the deterministic solutions was substantially smaller than when F5 is not always non-sealing, and the optimal well order for some positions in the drilling sequence were almost independent of the geological uncertainty. In this modified REEK model, it seems that whether fault F5 is completely sealing or not has a relatively larger influence than other geological features on the optimal drilling sequence.

In some cases, optimal solutions with a large Manhattan distance might have similar

(a) Distribution of Manhattan distance between the optimal drilling sequences from individual realizations

(b) Distribution of the relative changed in NPV in the reference model with different solutions
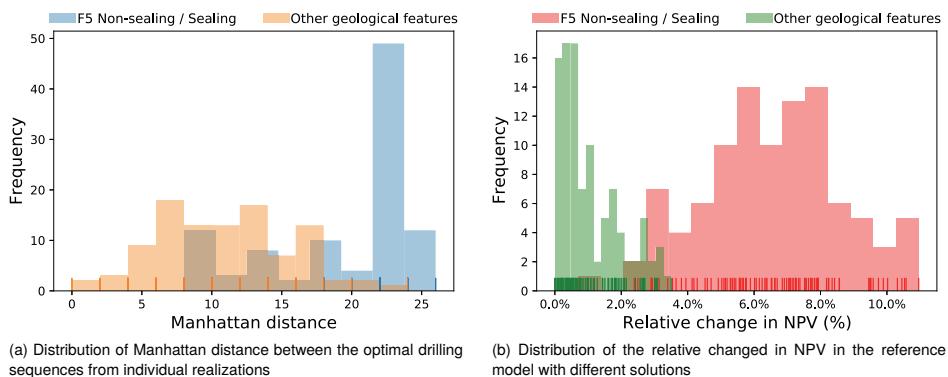
Figure 11: Variability in the optimal drilling sequences and variability in the NPV in the reference model with different optimized drilling sequences

economic values, for example, if only the optimal wells for the later decision stages are changed or their positions are swapped. Hence, we also investigated the variability in NPV obtained in a fixed reference model applied with different deterministic optimal solutions to further illustrate the importance of the sealing properties of fault F5 as a key feature for the drilling-order problem (Figure 11b). The results show that the NPV changes by less than 2% in most cases when various deterministic drilling sequence solutions for a non-sealing fault F5 are applied to a fixed reference model (histogram in green), even if the deterministic optimized drilling sequence might vary with a Manhattan distance larger than 22 in several cases. In cases with solutions obtained respectively from individual realizations with a non-sealing (i.e., reference model) and a sealing fault F5 (histogram in red), the relative change in NPV could be as much as 10% in the reference model. Thus, in terms of either a change in optimal decisions or a potential improvement in NPV, we observed that the key uncertainty for the drilling-order problem in this modified REEK field is whether F5 is completely sealing or not. To improve future decisions by using additional information, the acquired information should be able to provide useful observations for exploring this key reservoir feature.

Identifying key action and collecting information

In this paper, to reduce key uncertainties in the drilling-order problem, we use information from production and pressure data that can be obtained from standard oil-field monitoring. In the modified REEK Field (Fig. 10), Producers OP_3, OP_4, OP_5 and WI_1 are located near fault F5. We expect that, compared to the other wells, observations from these wells may be potentially more useful for predicting whether fault F5 is completely sealing or not. In this case, although OP_5 is close to fault F5, an examination of the information from OP_5 showed that it was less reliable as a source of information than what could be obtained from OP_3 or OP_4. When WI_1 or OP_5 is drilled as the first well, the hidden cost of information caused by sub-optimal solutions (i.e., it can be estimated when computing the optimal solution over the current uncertainty state) is larger than that from OP_3 and OP_4. Therefore, based on the possibility of obtaining valuable information for key uncertainty reduction and the possibility of

achieving high expected NPV, only OP_3 and OP_4 are considered as possible key actions for the first decision step.

Figure 12 shows the oil production rate and BHP in the first 6 months (i.e., the assumed drilling period for each well) obtained from 100 individual realizations with a non-sealing and sealing fault F5, where OP_3 and OP_4 are drilled as the first well, respectively. In almost all cases, production is first constrained to a maximum rate of 6,000 m^3/day and then decreased to hold the producer at a minimum BHP of 250 bar. When fault F5 is sealing, the production rate (red curves in Fig 12a and 12c) decreases more rapidly than when the fault F5 is non-sealing while maintaining the pressure at 250 bar, and the pressure (magenta curves in Fig 12b and 12d) drops faster while maintaining a production rate of 6,000 m^3/day. It seems that both OP_3 and OP_4 can potentially provide useful information for reducing the uncertainty about whether fault F5 is non-sealing or sealing, which would influence the rates of decline in both production and pressure.

Figure 12: Oil production rate and pressure in first 6 months obtained from individual realizations with non-sealing or sealing fault F5 when OP_3 and OP_4 are drilled as the first well respectively

Figure 13 shows the derivative of production rate when the producer is held at the minimum BHP of 250 bar and the pressure derivative when the producer is controlled by the maximum rate 6,000 m^3/day. Note that in Fig. 13a and 13c the derivative of the production rate is shown with time starting from the first day when BHP =250 bar, while the x-axis in Fig. 13b and 13d for the pressure derivative displays the time starting

(a) Derivative of oil production rate from OP_3 drilled as the first well

(b) Normalized pressure derivative from OP_3 drilled as the first well

(c) Derivative of oil production rate from OP_4 drilled as the first well

(d) Normalized pressure derivative from OP_4 drilled as the first well

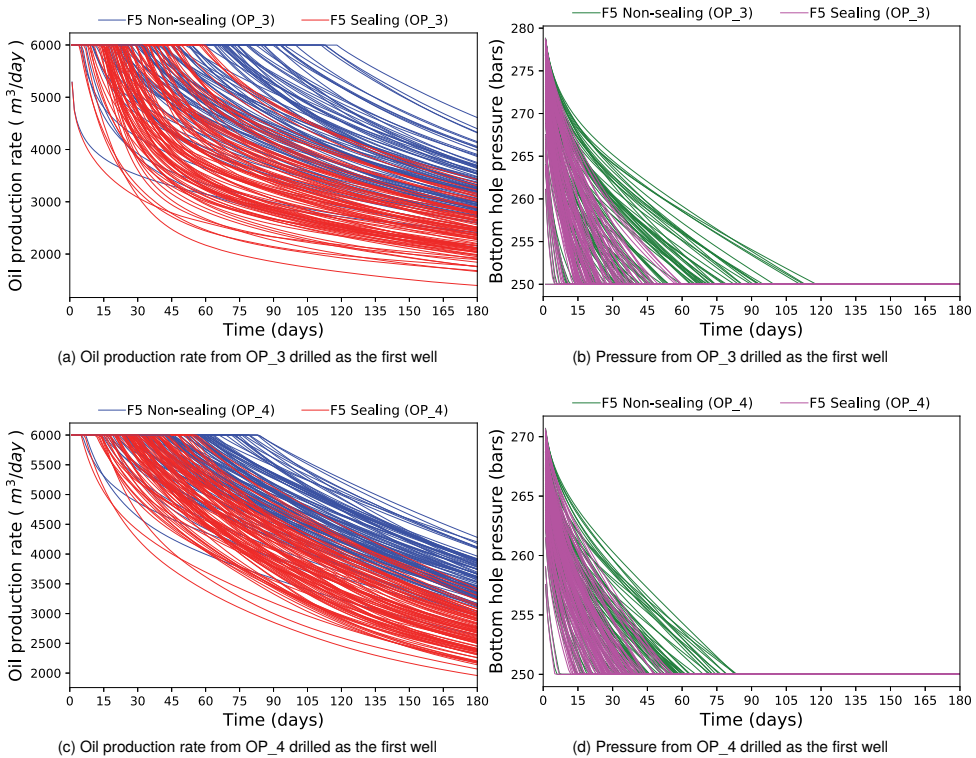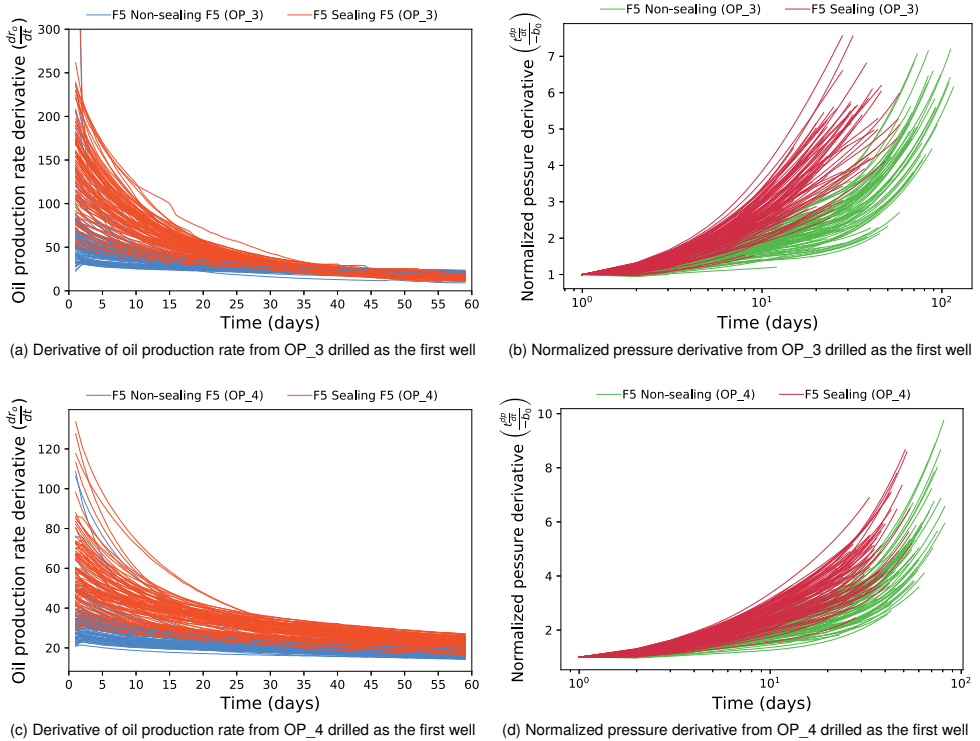*Figure 13: Derivative of production rate and pressure derivative obtained from individual realizations with non-sealing or sealing fault F5 when OP_3 and OP_4 are drilled as the first well respectively*

from the first day of production. Here we use the normalized logarithmic derivative of pressure to compute the pressure derivative, i.e., $\frac{t\frac{dp}{dt}}{-b_0}$, where $-b_0$ is the initial value of logarithmic pressure derivative at the beginning of production. Compared to the results obtained from the production rate and pressure (Fig. 12), it seems that the separation between model realizations with a non-sealing and sealing fault F5 is better when using the derivative information, especially with regard to observations of the pressure derivative when OP_3 is drilled as the first well (Fig. 13b). Note that computation of the normalized pressure derivative requires evaluation of 4 pressure values, so it may not be surprising that it is more informative for identifying potential barriers than a pressure measurement. Although we may intuitively find a good observation (e.g., pressure derivative at day 10) for predicting whether F5 is non-sealing or sealing in this example, that might not be the case in other situations. The information obtained from a key action could contain hundreds or thousands of observations. Manual identification of the important observations from such a large dataset is laborious and time-consuming. Moreover, to obtain highly reliable information for key uncertainty reduction, we usually need to combine multiple observations. Therefore, it is necessary to apply a practical method for the automatic defection of key observations.

In this work, we build supervise-learning models that are able to capture the mapping between the inputs (observations) and the outputs (non-sealing/sealing fault F5) to select the best observation subset with high prediction accuracy. Because the best subset might contain observations from different sources, we consider all information related to production rate, pressure, and their derivatives in the process of identifying key observations. In our application, it takes only a few minutes to identify the optimal combination of observations from the original information dataset with more than 700 observations using supervised-learning algorithms.

Selecting the best observation subset

In the six-month period after drilling the first well a large amount of production and pressure data are recorded. Some of the data are apparently unaffected by whether fault F5 is non-sealing or sealing (e.g., BHP of OP_3 is 250 bar in the last 2 months production before drilling the next well). To identify the important observations from such a large original dataset, we first use the mRMR feature selection method to remove irrelevant and redundant observations and obtain a small observation subset that may provide useful information. This is followed by a ROC curve analysis to further reduce the size of the subset, leaving only the observations with relatively good classification performance. In this way, the dimension of the observation dataset can be reduced quickly without incurring the loss of important information. Finally, a supervised-learning model can be applied with a small observation subset (i.e., few input variables) containing most of the useful information needed to efficiently identify the optimal combination of observations with highly reliable information for key uncertainty reduction. Although we did not consider the effect of observation error in the following analysis, the only change required in methodology would be to add random noise to the modeled observations.

Figures 14a and 14b show the ROC curves of 20 best observations obtained from the production/pressure data when OP_3 and OP_4 are drilled as the first well, respectively. In these figures the true positive rate (TPR, y-axis) is plotted against the false
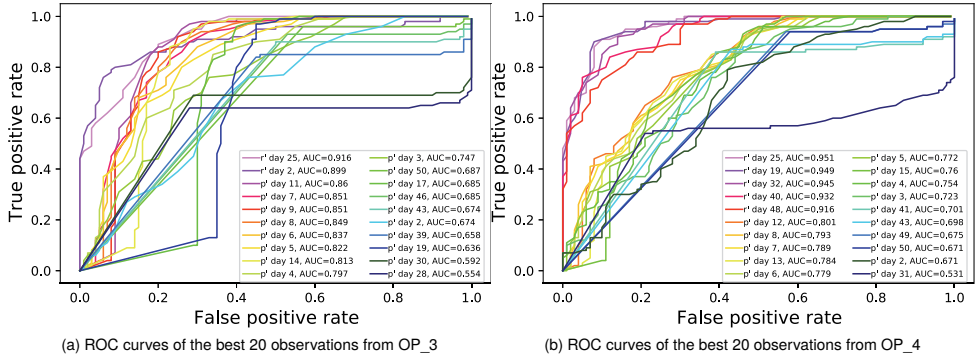
Figure 14: Comparison of the receiver operating characteristic (ROC) plots for the 20 top-ranked observations when OP_3 and OP_4 are drilled as the first well respectively

positive rate (FPR, x-axis) at various thresholds. The TPR represents the proportion of positive samples (i.e., individual realizations with a non-sealing fault F5) that are correctly identified, while the FPR is the proportion of negatives samples (i.e., individual realizations with a sealing fault F5) that are incorrectly identified as positive cases. The classification performance of a single observation is quantified by using the area under the ROC curve (AUC), which measures the entire two-dimensional area underneath the ROC curve from (0,0) to (1,1). A large AUC score indicates that the single observation has a good ability to distinguish between different classes. All of the 20 best observations are from r' or p', and this shows that information obtained from pressure and rate derivatives (Fig.13) provides more useful observations for predicting whether or not fault F5 is sealing than directly using production rates or pressure (Fig.12). Of the 20 top-ranked observations, most are associated with the pressure derivative, however, observations from the production rate derivative have higher AUC scores. In this case, it seems that p' provides more observations with useful information, while observations from r' provide better predictive performance in distinguishing between realizations with a non-sealing or sealing fault F5. If we use only the pressure data for reducing key uncertainty, the drilling of OP_3 as the first well would yield more important observations than the drilling of OP_4; seven observations of p' from OP_3 have AUC scores larger than 0.8, while only one observation of p' with an AUC > 0.8 emerges from OP_4. If only considering the information from rate data, drilling OP_4 as the first well would provide more key observations with higher AUC scores than those from OP_3.

Figure 15 shows the distributions of observed values obtained from individual realizations at the best single observation r'_b and p'_b (largest AUC score) identified from the derivatives of production rate and pressure, respectively. The red vertical lines represent the best cutting point that maximizes the difference between the TPR and FPR. The probabilities of the individual realizations with observed values located in each region determined by the optimal threshold are summarized in Tab. 2. When OP_4 is drilled as the first well, the prediction accuracy reaches as high as 90% based on r'_b, but the accuracy rate based on p' is less than 75%. When OP_3 is drilled as the first well, r'_b also provides information with a higher reliability than that of p'_b, but using p'_b to identify the models with a non-sealing fault F5 would be more effective, i.e., 94 of 100 individual realizations with a non-sealing F5 have $p'_b < \delta_{p'_b}$. In this example, although it is
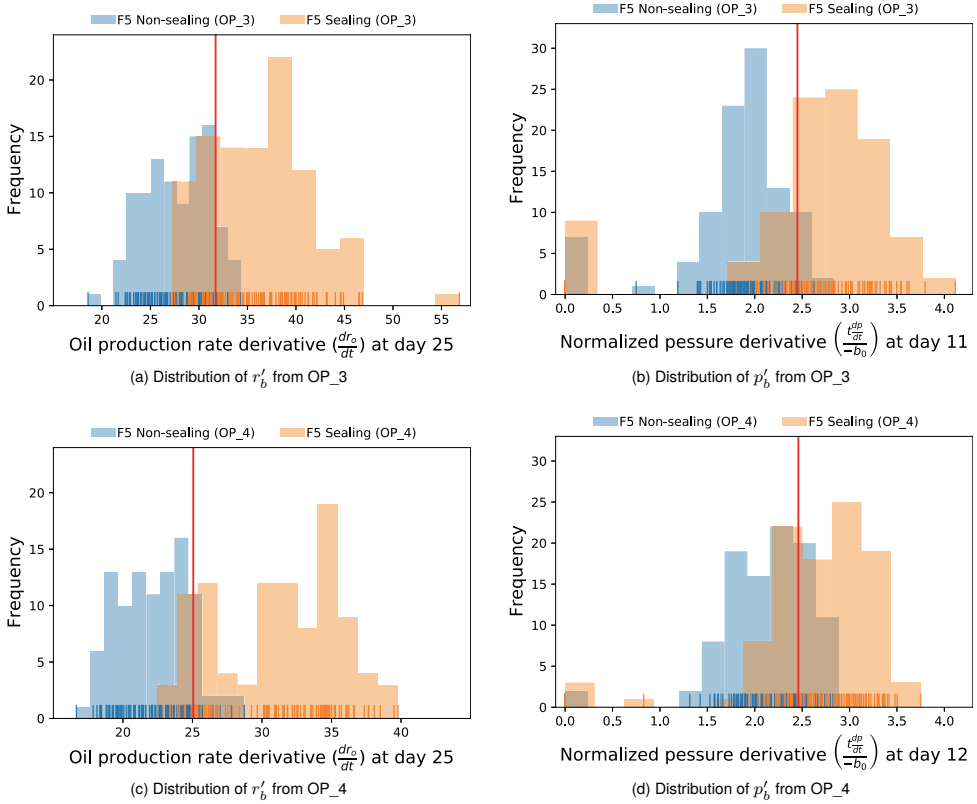
Figure 15: *Distributions of the observed values of r'_b and p'_b obtained from individual model realizations with non-sealing or sealing fault F5*

possible to reduce key uncertainty using only a single observation, this might not always be the case in other problems. In most applications, it is generally necessary to use multiple observations to obtain reliable information and thereby reduce key uncertainty. A quick, simple way is the direct use of a combination of the observations with high AUC scores. However, when the classification performances of these observations are similar, the prediction accuracy will most likely not improve.

By using supervised-learning algorithms, we can efficiently identify the optimal observation subset with high prediction accuracy. We should expect that a good observation subset consists of observations with useful information (e.g., AUC > 0.5). Based on the 20 top-ranked observations identified through the ROC curve analysis, we apply supervised-learning algorithms to identify the best combination of observations for predicting whether F5 is non-sealing or sealing. To avoid overfitting of the learning models, the original observation dataset obtained from 200 samples (100 individual realizations with non-sealing and sealing fault F5 respectively) is split into a training set (80%) and a test set (20%). The learning model is built based on the training set while the performance of the model is evaluated in the test set. To acquire a more statistically reliable estimate of performance, we use k-fold cross-validation resampling method to evaluate the learning model on the limited dataset.

Best single observation	Information from OP_3	Information from OP_4
$P(r_b' < \delta_{r_b'}|\mathsf{F5}_{\mathrm{non-sealing}})$	0.89	0.90
$P(r_b' \geq \delta_{r_b'}|\mathsf{F5}_{\mathrm{sealing}})$	0.80	0.87
$P(p_b' < \delta_{p_b'}|\mathsf{F5}_{\mathrm{non-sealing}})$	0.94	0.76
$P(p_b' \geq \delta_{p_b'}|\mathsf{F5}_{\mathrm{sealing}})$	0.74	0.72

Table 2: Probabilities of the individual realizations with the observed values located in each region determined by the optimal threshold of the best single observation r_b' or p_b'
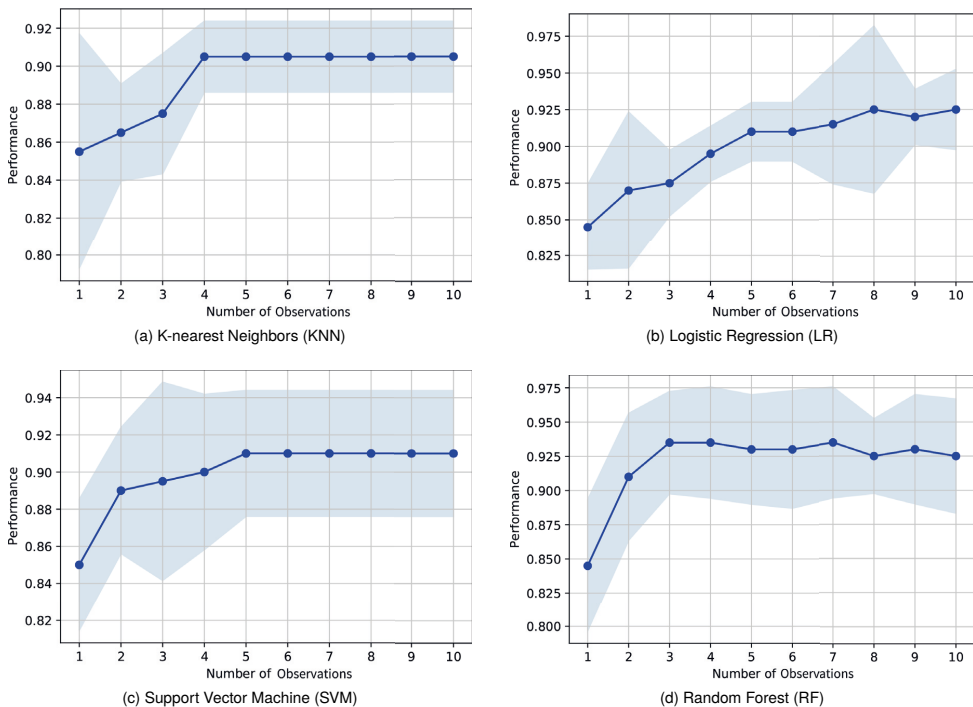

Figure 16: Performances of four supervised-learning models for predicting whether F5 is sealing or not using the optimal observation subset based on information from OP_3

Figure 16 shows the performances of four classification models applied to predict whether or not fault F5 is sealing using optimized observation subsets with different size based on the information from OP_3. The performance score on the y-axis represents the prediction accuracy measured in the test set. The light blue area indicates the standard deviation of prediction accuracy estimated through 5-fold cross validation. Results show that the Random Forest model (Fig. 16d), which is an ensemble machine learning algorithm based on bootstrap aggregation(bagging), performed better than the other three supervised-learning algorithms (i.e., k-Nearest Neighbors, Logistic Regression, Support Vector Machine). Compared with the case involving only one single observation, prediction accuracy improves using an optimized combination of two or three observations. However, when using more than three observations, the performance of the learning model does not improve with the addition of new observations; rather, as the number of input variables increases, the model becomes more complex, making it more prone to overfitting the training set. In this case, we can use only three key observations to predict whether or not F5 is sealing with high accuracy. When OP_4 is drilled as the first well, we observe similar properties in the process of identifying the optimal observations subset: the Random Forest algorithm extracts an observation subset with relatively higher prediction accuracy than the subsets produced by the other algorithms and an optimal number of approximately three observations.

Table 3 shows the optimal observation subset with three key observations identified using the Random Forest model, as well as the probabilities of individual realizations with observed values located in the two best disjoint subspaces Ω_1^b, Ω_2^b for predicting whether fault F5 is non-sealing and sealing, respectively. We note that, although the best single observation r_b' and p_b' have good classification performances (Fig. 15), the optimal combination of observations does not necessarily include r_b' or p_b'. The optimal observation subset from OP_3 has one observation of pressure derivative p' obtained at day 19, and this is only the 18th-best single observation with AUC =0.636 (Fig. 14a). However, after additional two observations from r' are combined, the accuracy rate for identifying models with non-sealing and sealing fault could reach 92% and 94%, respectively. When OP_4 is drilled as the first well, the optimal observation subset consists of three observations from r', whereas the best single observation r_b' at day 25 with AUC=0.951 is not in the optimal subset. However, the prediction accuracy is increased to 95% when fault F5 is non-sealing. The above results show that drilling either OP_3 or OP_4 at the first decision step will result in obtaining highly reliable information for reducing key uncertainty in the drilling-order problem (i.e., whether F5 is completely sealing or not). By using a supervised-learning algorithm, we efficiently identified key observations from both OP_3 and OP_4 as well as the best division of space for prediction purposes, and we also simultaneously estimated the the reliability of information (i.e., prediction accuracy) in each subspace (Tab. 3). This allows the direct computation of the posterior probability of key uncertainty using Bayes' theorem, thereby avoiding the need for history matching to re-estimate uncertainty.

Assessing value of information through key action

By identifying the key reservoir feature that has the largest influence on the optimal drilling sequence, we obtained two possible key actions at the first step (i.e., OP_3 and OP_4) that are more likely to provide useful information for reducing the key uncer-

$N_f^b = 3$	Information from OP_3	Information from OP_4
Optimal observation subset	$o^b = (r'_{\text{day2}}, r'_{\text{day25}}, p'_{\text{day19}})$	$o^b = (r'_{\text{day19}}, r'_{\text{day32}}, r'_{\text{day48}})$
$P(o^b \in \Omega_1^b \| \text{F5}_{\text{non-sealing}})$	0.916	0.945
$P(o^b \in \Omega_2^b \| \text{F5}_{\text{sealing}})$	0.941	0.878

Table 3: Best observation subset and the reliability of the information in each observation subspace obtained from Random Forest model

(a) Key uncertainty reduction based on information from OP_3

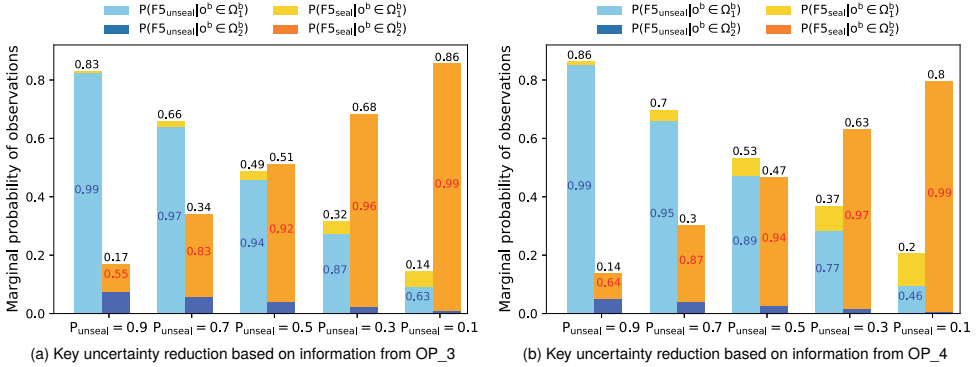(b) Key uncertainty reduction based on information from OP_4

Figure 17: Posterior probabilities and marginal probabilities of observations obtained from different prior probability distributions of key uncertainty

tainty in the drilling-order problem: whether fault F5 is sealing or not. Key observations obtained from both OP_3 and OP_4 are demonstrated to have high predictive accuracy for indicating if fault F5 is sealing or non-sealing. Although reducing key uncertainty would potentially lead to better future decisions, taking a key action to acquire the useful information for key uncertainty reduction is not always worthwhile, since there may be a high hidden cost of obtaining information caused by the sub-optimality of the solution which uses the key action. To judge whether taking the key action increases the maximum expected NPV in current uncertainty state, we must evaluate the EVOI associated with this hidden cost (Eq.7). To compute the expected value of information, we need to obtain the expected NPV from the initial optimal solution obtained utilizing the prior probability of key uncertainty and the optimal solutions (i.e., constrained to the selected key action) in the posterior probability, and the marginal probabilities of observations. All of these values are related to the prior probability. Hence, instead of evaluating the EVOI only in one specific prior probability, we investigate the performances of the VOI analysis through key action with different prior probabilities of key uncertainty.

Figure 17 shows effects of using the best observation subset o^b from OP_3 or OP_4 to evaluate the posterior probabilities of a non-sealing and sealing fault F5 with different initial probabilities, i.e., $P_{\text{unseal}} = (0.9, 0.7, 0.5, 0.3, 0.1)$. The black values on the tops of bars represent the marginal probabilities of observations located in the best two disjoint subspaces Ω_1^b and Ω_2^b, i.e., $P(o^b \in \Omega_1^b)$ and $P(o^b \in \Omega_2^b)$, irrespective of the reservoir features. Since the prediction accuracy for observations located in the same subspace is almost the same, we only need to compute the posterior probabilities at $o^b \in \Omega_1^b$ and $o^b \in \Omega_2^b$ (marked with different colors in the bars) considering all possible observations associated with o^b. Note that the values in blue indicate the
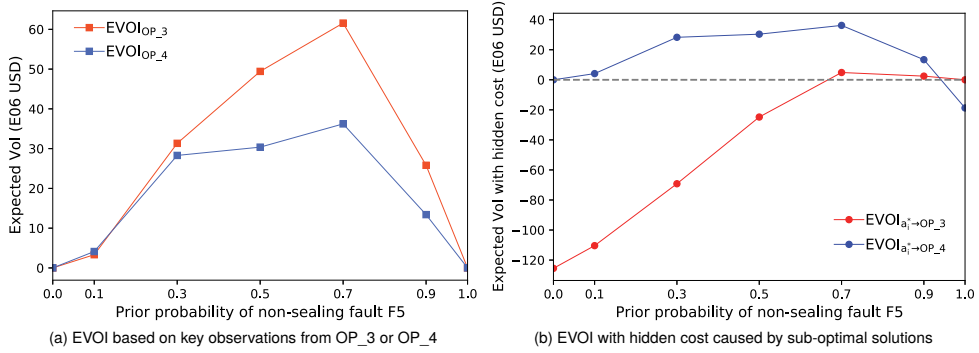
Figure 18: Comparison of standard EVOI and EVOI with the hidden cost of obtaining information from OP_3 or OP_4

posterior probabilities $P(\text{F5}_{\text{unseal}}|o^b \in \Omega_1^b)$ while the values in red represent $P(\text{F5}_{\text{seal}}|o^b \in \Omega_2^b)$. Overall, using key observations from both OP_3 and OP_4 can significantly reduce the uncertainty about whether fault F5 is completely sealing or not, but the posterior probabilities are strongly influenced by the prior probabilities. The posterior probabilities with observations $o^b \in \Omega_1^b$ and $o^b \in \Omega_2^b$ change considerably especially when $0.3 < P_{\text{F5}_{\text{unseal}}} < 0.7$. When $P_{\text{F5}_{\text{unseal}}} = 0.5$, the posterior probabilities $P(\text{F5}_{\text{unseal}}|o^b \in \Omega_1^b)$ and $P(\text{F5}_{\text{seal}}|o^b \in \Omega_2^b)$ based on the key observations obtained from OP_3 could reach 0.94 and 0.92, respectively. When $P_{\text{F5}_{\text{unseal}}} > 0.9$ or < 0.1, $o^b \in \Omega_1^b$ and $o^b \in \Omega_2^b$ are almost perfect observations for indicating a non-sealing and sealing fault F5, respectively. At a fixed prior probability, both marginal probabilities of observations and posterior probability distributions change slightly when key observations respectively from OP_3 and OP_4 are used, although the reliability of information is different (Tab. 3). In this case, drilling either OP_3 or OP_4 as the first well can result in obtaining information with similar effectiveness in terms of the reduction of key uncertainty for the drilling-order problem, i.e., whether F5 is sealing or not.

The purpose of collecting information for key uncertainty reduction is to obtain a better optimal drilling order solution with higher expected value. To study whether key observations identified from OP_3 and OP_4 are useful in increasing expected profitability, we computed the EVOI (Eq. 10) to evaluate their potential for increasing the expected NPV from improved optimal solutions. Figure 18a compares the $\text{EVOI}_{\text{OP}_3}$ and $\text{EVOI}_{\text{OP}_4}$ obtained from the difference between the expected values of optimal drilling sequences with and without use of additional information (i.e., key observations) to reduce the uncertainty about whether F5 is sealing or not. Results show that both $\text{EVOI}_{\text{OP}_3}$ and $\text{EVOI}_{\text{OP}_4}$ are positive when $0 < P_{\text{F5}_{\text{unseal}}} < 1$. This indicates that the optimal drilling sequence of the remaining wells is improved after reducing the key uncertainty in the drilling-order problem. Also, the key observations from either OP_3 or OP_4 are always helpful in making better future decisions with different prior probabilities. When the initial probability of a non-sealing fault F5 is near 0.7, both $\text{EVOI}_{\text{OP}_3}$ and $\text{EVOI}_{\text{OP}_4}$ reach their maximum values. Although the effects of using information from OP_3 and OP_4 to reduce key uncertainty are similar (Fig.17), $\text{EVOI}_{\text{OP}_3}$ is always higher than $\text{EVOI}_{\text{OP}_4}$, i.e., more *additional* value can be created with the information from OP_3. If we only consider the EVOI for choosing a key

action, it seems that OP_3 should be preferred to OP_4. However, there might be a large hidden cost of information caused by a sub-optimal solution when OP_3 or OP_4 is drilled as the first well. Hence, we should evaluate the net EVOI with this hidden cost to determine whether it is worth drilling OP_3 or OP_4 first to obtain information for improving future decisions, rather than taking the optimal decision for achieving the maximum expected NPV over the current uncertainty state.

Figure 18b shows the net EVOI including the hidden cost of information when OP_3 or OP_4 is drilled as key action for first decision step. Note that there is no need to compute the EVOI (Eq. 10) and the cost of information ECOI (Eq. 11) separately for obtaining the net EVOI through a key action. This net EVOI can be computed based on the $\text{EVOI}_{a_i^* \to a_{j+1}^{\text{key}}}$ obtained from changing the initial optimal decision a_i^* for current uncertainty state to key action a_{j+1}^{key} (Eq. 7). Although $\text{EVOI}_{\text{OP}_3} > \text{EVOI}_{\text{OP}_4}$, the hidden cost of obtaining information from OP_3 is much larger than the cost from OP_4 especially when there is a high probability that fault F5 is sealing. When the initial probability of a non-sealing fault F5 is less than 0.7, the $\text{EVOI}_{a_i^* \to \text{OP}_3}$ is negative, which means that the additional value created by using the information from OP_3 to improve the optimal drilling sequence of the remaining wells is not enough to compensate for the hidden cost of obtaining information using a sub-optimal drilling sequence when OP_3 drilled as the first well, i.e., $\text{EVOI}_{\text{OP}_3} < \text{ECOI}_{a_i^* \to \text{OP}_3}$. In that case, there is no benefit to drilling OP_3 first even if it can provide highly reliable information for reducing key uncertainty for the drilling-order problem, and OP_3 is preferred to a_i^* only when $P_{\text{F5}_{\text{unseal}}} \geq 0.7$. However, $\text{EVOI}_{a_i^* \to \text{OP}_4}$ is always positive and larger than $\text{EVOI}_{a_i^* \to \text{OP}_3}$ unless the prior probability of non-sealing fault F5 is extremely high. After consideration of the hidden cost of information, when $P_{\text{F5}_{\text{unseal}}} < 0.94$, OP_4 will be a better choice of key action than OP_3 while OP_3 will be preferred to OP_4 if $P_{\text{F5}_{\text{unseal}}} \geq 0.94$. In this case, performing the key-feature based VOI analysis (Fig. 7), which considers only the future learning possibility through OP_3 or OP_4 that has a higher $\text{EVOI}_{a_i^* \to a_{j+1}^{\text{key}}}$, leads to the same optimal decision with consideration of the future information from all possible decision alternatives, which illustrates that by taking into account future information from key action, we are able to make optimal decisions which account for the possibilities of further learning without sacrificing the quality of solution.

Instead of computing all $\text{EVOI}_{a_i^* \to a_{j+1}^{\text{key}}}$ to identify the preferred a_{j+1}^{key}, we can study the initial optimal drilling sequence to quickly obtain a good a_{j+1}^{key} with a potentially higher $\text{EVOI}_{a_i^* \to a_{j+1}^{\text{key}}}$. Since the hidden cost of information is caused by a sub-optimal solution and the wells that have important contributions to increase the expected NPV are generally preferable for drilling at an early stage, we expect that the cost $\text{ECOI}_{a_i^* \to a_{j+1}^{\text{key}}}$ from a_{j+1}^{key} that is drilled at a later stage along the initial optimal complete drilling sequence will be potentially larger than the hidden cost of obtaining information from a_{j+1}^{key} drilled at an early stage. When there are several possible a_{j+1}^{key} with similar reliability of information, we can choose the one that is supposed to be drilled earlier for maximizing the expected NPV in the current uncertainty state to avoid a high hidden cost of information. In this work, we used learned heuristic search with mean model bias-correction methods to efficiently obtain robust optimal drilling sequence under uncertainty. During the search process, the maximum expected NPV constrained to

different selected wells is estimated without finding the actual optimal solution. Hence, we could also obtain an approximation of $\mathrm{ECOI}_{a_i^* \to a_{j+1}^{\mathrm{key}}}$ when computing the initial optimal decision a^* and without incurring additional costs.

For the drilling-order problem, we observe that the optimal drilling sequence always starts with OP_4 when $P_{\mathrm{F5_{unseal}}} < 0.94$, and then changes to OP_3 drilled as the first well when $P_{\mathrm{F5_{unseal}}} \geq 0.94$, which indicates that one of the possible key actions (e.g., OP_3 or OP_4) has no hidden cost of information caused by a sub-optimal solution since it is identical to the initial optimal decision a^*, i.e., $\mathrm{ECOI}_{a^* \to a^*} = 0$. When a^* is also able to provide useful information for key uncertainty reduction, we can consider the future information from both a^* and a_{j+1}^{key} that is identified from the other decision alternatives to make a more robust decision in consideration of future learning possibilities (Eq. 13). In this example, the optimal decision obtained after taking into account the future information from both OP_3 and OP_4 is still the initial optimal decision a^* since taking a^* is able to both generate highly reliable information for key uncertainty reduction (i.e., valuable information for improving future decisions) and maximize the expected NPV for the current assessment of uncertainty (i.e., no hidden cost of information), although this might not always be the case in other problems. Note that here, we only investigated the future learning possibilities at the first decision step. After drilling a new well, the reservoir model will be updated through history matching based on the actual data. The key-feature-based VOI analysis (Fig. 7) could then be performed again to determine the next optimal well. At the second and later decision stages, however, information from the remaining actions will generally have a smaller potential for improving the optimal strategy, and there may be no clear key uncertainties for the optimization problem. In that case, there would be no need to consider the effect of future information when making the optimal decision, and one could simply use a standard robust optimization method [31].

Conclusion

In this paper, we proposed a flexible workflow built on a key-feature-based value of information analysis to make optimal decisions efficiently while accounting for the possibilities for future learning through actions. Taking into account the effects of future information before committing to a decision allows improvement of the optimal strategy. However, it is infeasible and unnecessary to account for all possible future observations from remaining actions (i.e., a standard VOI analysis with extensive form). In our approach, the VOI analysis is only performed on a small number of key actions that will provide key information for reducing the most important uncertainties in the optimization problem. i.e., information that will be for making better future decisions. Then, the optimal decision is made based on the trade-off between the key actions and the initial optimal decision obtained without considering any future information. The simplified VOI analysis based only on key actions and key information might not result in the same optimal solution as the complete VOI analysis, but it offers a practical way to obtain a near-optimal decision that accounts for the possibilities of future learning (i.e., the opportunities to improve optimal strategy resulting from future uncertainty reduction). The key actions can be identified by considering the possibility of obtaining valuable information for reducing key uncertainties and the possibility of achieving high expected NPVs for the current uncertainty state, so that there is no need

to compare the actual expected values of all possible decisions. The focus on the use of key information to reduce key uncertainties avoids the need for full history matching to re-estimate all uncertainties in the optimization. Instead of considering all distinct sets of observations obtained from all ensemble members when updating the reservoir model, we divide the entire key observation space into a limited number of disjointed subspaces, i.e., each subspace will have high information reliability for indicating a specific key uncertainties subregion, and observations located in the same subspace have similar prediction precision for key uncertainty reduction. Consequently, we only have to re-estimate key uncertainties for each observation subspace and perform the optimization process in a few posterior ensembles for computing the expected value with information. The following conclusions can be drawn from the present study:

- Although many uncertainties arise in reservoir characterization, some of them have little influence on the optimal decisions, even if they might be reduced significantly by assimilation of acquired observations. By identifying key uncertainties for the optimization problems, we can identify key actions that would provide the most useful information for improving future decisions.

- When all observations are used to simultaneously re-estimate uncertainty, the largest decrease in uncertainty may be in properties that are irrelevant to current decisions, and the reduction in key uncertainties from some observations might be very small. However, the computational penalty of including those nonessential observations in updating the reservoir model can be large. Thus, instead of only reducing the decision space by identifying key actions, we also need to identify the most important observations for reducing key uncertainties to make the computation manageable.

- Performing a key action to acquire information for reducing key uncertainty is not necessarily worthwhile, even if there is no explicit cost in obtaining the information and future decisions could be improved. When taking key action to obtain information leads to a sub-optimal solution, there is a hidden cost in obtaining the information. Instead of using the additional value that could be created with information to judge whether it is worth taking action, the criteria should be the *net* expected value of information, including the hidden cost associated with changing the optimal decision for the current uncertainty state to the key action.

- The initial action in the optimal sequence of actions based on current information might, in some cases, be a possible key action. To obtain a robust decision, it may be necessary to consider the possibility of future learning through both the initial optimal decision and alternative key actions identified from the remaining decision alternatives.

- The expected value of information attributed to key actions will depend on the prior probability distribution of key uncertainties. Hence, changing the prior probabilities will not only affect the standard (naïve) computation of the optimal solution, but will also affect the optimal decision obtained from VOI analysis.

Although the methodology is illustrated by the application of drilling-order problems, it can be extended to general sequential decision-making problems under uncertainty while considering the effect of future information. The key point is to effectively identify

key actions and key observations that are associated with the key reservoir features for optimization problems. In our example application, *key uncertainties* are identified by studying the sensitivity of deterministic optimization solutions to different individual uncertainties. *Key actions* are identified by evaluating the reliability of information for reducing key uncertainties and the hidden cost of obtaining information from key actions. For large problems, it may be necessary to explore more generalized and efficient approaches for identifying key uncertainties and key actions. To efficiently identify *key observations*, we built supervised-learning algorithms that can automatically detect the optimal combination of observations as well as the best division of space for reducing key uncertainty. At the same time, we estimate the prediction accuracy (i.e., the information's reliability) for observations located in each subspace. This allows directly computing the posterior probability of key uncertainty based on Bayes' theorem, avoiding the necessity of expensive data assimilation algorithms to update the entire reservoir model. Using learning algorithms to identify the important observations is applicable for optimization problems with multiple key uncertainties that are continuous or categorical variables. For continuous variables, the distribution of key uncertainties could be divided into a set of optimized subregions based on the performance of the observation subspaces in reducing key uncertainties. Our simplified VOI analysis considers the future information only resulting from the current decision step. If the key information can only be obtained by taking at least two actions (i.e., individual decision alternatives are shown to be unreliable as a source of information for reducing key uncertainty), we could extend the VOI analysis by considering the possibility of future learning through the following two decision steps, i.e., a combination of information from two actions. The simultaneous consideration of information from two actions would increase the complexity of VOI analysis, computational cost of expected value with information, and the hidden cost of obtaining key information caused by sub-optimal solution, which will be constrained to more past decisions.

Nomenclature

f^*	actual expected objective function value of optimal control
f^*	maximum net present value of optimal control
α	bias correction factor
α	bias correction factor
$\bar{\alpha}$	average value of bias correction factor
$\bar{\varepsilon}$	mean observed single-step error of estimated value
$\bar{m}$	expected value of model parameter
$\bar{m}$	expected value of model parameter
$\bar{\gamma}$	mean observed single-step ratio of estimated value
$\bar{\mu}$	mean observed single-step ratio of variability in estimated value
$\bar{\mu}_{n_s}$	mean single-step ratio of variability in estimated value
β	partial correction factor for individual realizations
β	partial correction factor
Δt	time interval
Δt	time interval
Δx	length interval
δ	distance between controls
γ	single-step ratio of estimated value
γ	single-step ratio of estimated value
$\hat{\alpha}_{\mathrm{loc}}$	local estimate
$\hat{\alpha}_r$	regularized estimate
$\hat{\gamma}$	forecast error of initial evaluation function value
$\hat{f}$	learned evaluation function
$\hat{f}$	learned evaluation function
$\hat{h}$	learned heuristic function

$\hat{\gamma}$ estimated single-step ratio of estimated value

$\hat{\mu}$ estimated single-step ratio of variability in estimated value

λ regularization parameter/Lagrangian parameter (depending on context)

$\mathbf{b_0}$ vector of β at a fixed control x_0

$\mathbf{b}$ vector of β from random controls and realizations

$\mathbf{P}$ vector of drilling sequence

$\mathbf{w}$ weight vector of β

$\text{ECOI}_{a^* \to a^{\text{key}}}$ hidden cost of information caused by sub-optimal solutions constrained to decision a^{key}

EVOI expected value of information

$\text{EVOI}_{a^* \to a^{\text{key}}}$ expected value of information changing decision a^* to a^{key}

EVWI expected value with information

EVWOI expected value without information

μ single-step ratio of variability in estimated value

μ viscosity of fluid

ω weight

Ω^b key observation subspace

ρ distance-based weight

σ_α^2 variance of bias correction factor α

σ_β^2 variance of partial correction factor β

θ permeability

Θ^m subregion of key uncertainty

ε single-step error of estimated value

A cross-sectional area

A decision space

a decision alternative

a^* optimal decision over the current assessment of uncertainty

$a^{*^{\text{fl}}}$ optimal decision considering future information from current decision stage

$a^{*^{\text{fs}}}$ optimal decision considering future information from all remaining decision stages

a^{key} decision alternative providing important information for key uncertainty reduction

b discount rate

b discount rate

C covariance of model parameter

d number of remaining actions

d number of remaining actions

E expected value of objective function

EV^* maximum expected value over current uncertainty state

f evaluation function/estimated maximum net present value

f objective function/evaluation function (depending on context)

f_m first derivatives of objective function with respect to model parameter

f_{mm} second derivatives of objective function with respect to model parameter

g actual economic value from previous actions

g economic value from previous actions

h heuristic function/estimated maximum future value

h heuristic function/estimated maximum future value

h history of past decisions and observations

h^* maximum future value

$IWEI$ injector economic index

J objective function

k log-permeability

L taper length

m model parameter

m model parameter

n_{eff} effective sample size

n_s environment state at decision stage

n_s environment state at decision stage

N_{Ω^b} Number of observation subspaces

N_e Ensemble size

N_e total number of model realizations

N_r number of selected wells

N_s number of remaining wells

N_s number of selected wells

N_w total number of wells

N_w total number of wells

N_x total number of random controls

NPV net present value

NPV net present value

O observation space

o observation obtained from specific decision

o^b best observation subset for key uncertainty reduction

P pressure

$P(\Omega^b \mid \Theta^m)$ probability of observing $o^b \in \Omega^b$ at key uncertainty subregion Θ^m

$P(\Theta^m \mid \Omega^b)$ posterior probability of key uncertainty subregion Θ^m with observation $o^b \in \Omega^b$

$P(\Theta^m)$ prior probability of key uncertainty subregion Θ^m

$P(o \mid h, a)$ probability of observing o from decision a following history h

$PWEI$ producer economic index

q production/injection rate

q production/injection rate

Q^* maximum expected value over all possible future observations from all remaining actions

r reward/cost

r reward/cost

T total number of time steps

T total number of time steps

t time

t time

u uncertainty state

W cost of drilling well

W cost of drilling well

x control variable

x control variable

Δt^{Φ} learning time period

Δt^L	learning time period
Φ	learning technique index
f_n	evaluation function at a specific state
f_n	evaluation function at a specific state
h_n	heuristic function at a specific state
h_n	heuristic function at a specific state
i	control index/learning period index/cell index (depending on context)
i	key uncertainty subregion index
j	decision stage index/model realization index (depending on context)
j	model realization index/time step index(depending on context)
j	time step index
k	key observation subspace index
k	well index
L	learning technique index
o	oil
o	oil
s	decision stage
s	decision stage
w	water/well (depending on context)
w	water/well(depending on context)
wi	water injection
wi	water injection
cos	Cosine distance
L_1	Manhattan distance
L_2	Euclidean distance
eff	effective sample size
$\bar{m}$	reservoir mean model
Φ	learning
i	injection
i	injection
L	learning

p production

p production

b   best observation subset

lin   linear approximation

m   model parameter

o   observation

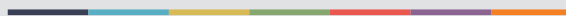quad   quadratic approximation

Acknowledgments

Bibliography

[1] E. G. D. BARROS, P. M. J. V. DEN HOF, AND J. D. JANSEN, *Informed production optimization in hydrocarbon reservoirs*, Optimization and Engineering, 21 (2020), pp. 25–48.

[2] E. G. D. BARROS, J. D. JANSEN, AND P. M. J. VAN DEN HOF, *Value of information in parameter identification and optimization of hydrocarbon reservoirs*, IFAC-PapersOnLine, 48 (2015), pp. 229–235.

[3] E. G. D. BARROS, O. LEEUWENBURGH, P. M. J. VAN DEN HOF, AND J. D. JANSEN, *Value of multiple production measurements and water front tracking in closed-loop reservoir management*, in SPE Reservoir Characterisation and Simulation Conference and Exhibition, Abu Dhabi, UAE, 2015, Society of Petroleum Engineers.

[4] E. G. D. BARROS, P. M. J. VAN DEN HOF, AND J. D. JANSEN, *Value of information in closed-loop reservoir management*, Computational Geosciences, 20 (2016), pp. 737–749.

[5] R. B. BRATVOLD, J. E. BICKEL, AND H. P. LOHNE, *Value of information in the oil and gas industry: Past, present, and future*, SPE Reservoir Evaluation & Engineering, 12 (2009), pp. 630–638.

[6] B. CHEN, J. HE, X.-H. WEN, W. CHEN, AND A. C. REYNOLDS, *Uncertainty quantification and value of information assessment using proxies and Markov chain Monte Carlo method for a pilot project*, Journal of Petroleum Science and Engineering, 157 (2017), pp. 328–339.

[7] Y. CHEN, D. S. OLIVER, AND D. ZHANG, *Efficient ensemble-based closed-loop production optimization*, SPE Journal, 14 (2009), pp. 634–645.

[8] P. CUNNINGHAM AND S. BEGG, *Using the value of information to determine optimal well order in a sequential drilling program*, The American Association of Petroleum Geologists, 92 (2008), pp. 1393–1402.

[9] J. EIDSVIK, G. DUTTA, T. MUKERJI, AND D. BHATTACHARJYA, *Simulation-regression approximations for value of information analysis of geophysical data*, Mathematical Geosciences, 49 (2017), pp. 467–491.

[10] P. A. ESTEVEZ, M. TESMER, C. A. PEREZ, AND J. M. ZURADA, *Normalized mutual information feature selection*, IEEE Transactions on Neural Networks, 20 (2009), pp. 189–201.

[11] T. Goda, K. Ohno, and K. Sato, *Value of information in optimizing reservoir development under geological uncertainty*, Journal of the Japan Petroleum Institute, 60 (2017), pp. 41–52.

[12] C. J. Grayson, *Decisions under uncertainty: Drilling decisions by oil and gas operators*, Harvard University Press, Boston, Massachusetts, 1960.

[13] R. Hanea, P. Casanova, L. Hustoft, R. Bratvold, R. Nair, C. W. Hewson, O. Leeuwenburgh, and R.-M. Fonseca, *Drill and learn: A decision-making work flow to quantify value of learning*, SPE Reservoir Evaluation & Engineering, 22 (2019), pp. 1131–1143.

[14] J. Hanley and B. McNeil, *The meaning and use of the area under a receiver operating characteristic (roc) curve*, Radiology, 143 (1982), pp. 29–36.

[15] J. He, P. Sarma, E. Bhark, S. Tanaka, B. Chen, X.-H. Wen, and J. Kamath, *Quantifying expected uncertainty reduction and value of information using ensemble-variance analysis*, SPE Journal, 23 (2018), pp. 428–448.

[16] A. Hong, R. Bratvold, P. Thomas, and R. Hanea, *Value-of-information for model parameter updating through history matching*, Journal of Petroleum Science and Engineering, 165 (2018), pp. 253–268.

[17] J. D. Jansen, D. R. Brouwer, G. Nævdal, and C. P. J. W. Van Kruijsdijk, *Closed-loop reservoir management*, First Break, 23 (2005), pp. 43–48.

[18] R. Kohavi and G. H. John, *Wrapper for feature subset selection*, Artificial Intelligence, 97 (1997), pp. 273–324.

[19] C. Lazar, J. Taminau, S. Meganck, D. Steenhoff, A. Coletta, C. Molter, V. De Schaetzen, R. Duque, H. Bersini, and A. Nowé, *A survey on filter techniques for feature selection in gene expression microarray analysis*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 9 (2012), pp. 1106–1119.

[20] O. Leeuwenburgh, A. G. Chitu, R. Nair, P. J. P. Egberts, L. Ghazaryan, T. Feng, and L. Hustoft, *Ensemble-based methods for well drilling sequence and time optimization under uncertainty*, in ECMOR XV-15th European Conference on the Mathematics of Oil Recovery, Amsterdam, Netherlands, 2016, European Association of Geoscientists & Engineers.

[21] H. Peng, F. Long, and C. Ding, *Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy*, IEEE transactions on pattern analysis and machine intelligence, 27 (2005), pp. 1226–1238.

[22] H. Raiffa and R. Schlaifer, *Applied statistical decision theory*, Harward University Press, Boston, USA, 1961.

[23] S. Ramírez-Gallego, I. Lastra, D. Martínez-Rego, V. Bolón-Canedo, J. M. Benítez, F. Herrera, and A. Alonso-Betanzos, *Fast-mrmr: Fast minimum redundancy maximum relevance algorithm for high-dimensional big data*, International Journal of Intelligent Systems, 32 (2017), pp. 134–152.

[24] P. Sarma, L. J. Durlofsky, K. Aziz, and W. H. Chen, *Efficient real-time reservoir management using adjoint-based optimal control and model updating*, Computational Geosciences, 10 (2006), pp. 3–36.

[25] R. Schlaifer, *Probability and statistics for business decisions : an introduction to managerial economics under uncertainty*, McGraw-Hill Book Company, New York, 1959.

[26] E. J. Sondik, *The optimal control of partially observable Markov processes*, PhD thesis, Stanford University, Stanford, CA, USA, 1971.

[27] R. R. Torrado, J. Rios, and G. Tesauro, *Optimal sequential drilling for hydrocarbon field development planning*, in Proceedings of the 29th AAAI Conference on Innovative Applications (IAAI-17), 2017, pp. 4734–4739.

[28] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, USA, 1944.

[29] C. Wang, G. Li, and A. C. Reynolds, *Production optimization in closed-loop reservoir management*, SPE Journal, 14 (2009), pp. 506–523.

[30] L. Wang and D. S. Oliver, *Efficient optimization of well drilling sequence with learned heuristics*, SPE Journal, 24 (2019), pp. 2111–2134.

[31] L. Wang and D. S. Oliver, *Fast robust optimization using bias correction applied to the mean model*, Computational Geosciences, 25 (2021), pp. 475–501.

[32] K. Åström, *Optimal control of markov processes with incomplete state information*, Journal of Mathematical Analysis and Applications, 10 (1965), pp. 174–205.

uib.no