Value of Information in Optimizing Reservoir Development under Geological Uncertainty

Takashi GODA* , Katsuta OHNO, and Kozo SATO

Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, JAPAN

(Received July 7, 2016)

The value of information (VOI) analysis has been recognized as a useful tool for measuring how much the expected monetary value can be increased as a result of an information-gathering activity. The VOI analysis is usually applied in the context of decision making under geological uncertainty, where only a relatively small number of decision alternatives are taken into account. In this paper we discuss possible applications of the VOI analysis in optimization of reservoir development under geological uncertainty. More precisely, by means of the VOI analysis, we evaluate how much an information-gathering activity provides an increase of the maximized expected monetary value, where the maximization is considered in a probabilistic sense to account for geological uncertainty. In this application we often need to deal with a quite large or even infinite number of candidate solutions within an optimization problem, which may cause trouble for an efficient implementation of the VOI analysis. After introducing a general methodology to estimate the VOI in our context, we validate our methodology through a toy problem, and moreover apply to a simple waterflooding problem. We find out that the VOI analysis can be conducted efficiently even in optimization under geological uncertainty by specializing our methodology properly.

Keywords
Value of information, Optimization considering uncertainty, Decision making, Nested Monte Carlo, Derivative-free optimization

1. Introduction

Optimization of reservoir development under geological uncertainty plays a crucial role in the success of maximizing the monetary value gained by developing a reservoir in a probabilistic sense. Under the risk neutrality assumption, the objective reduces to maximizing the expectation of the monetary value with respect to probability distributions of uncertain geological parameters\(^1,2\). Now suppose that some information-gathering activities can be conducted to reduce some degree of geological uncertainty. The posterior probability distributions of uncertain geological parameters conditional to the result of such information-gathering activities can be calculated according to Bayes’ theorem by multiplying the prior probability distributions by the likelihood function. The difference of the prior and posterior probability distributions yields that of the optimal development strategies found before and after these information-gathering activities are conducted. In general, however, it is unclear how much the maximum expected monetary value can be increased by gathering information. If no increase of monetary value can be expected, such an information-gathering activity is worthless in terms of optimization under geological uncertainty. Otherwise, such an information-gathering activity deserves to be conducted for improving the optimal development strategy. Thus, a prior evaluation on how much the maximum expected monetary value can be increased is useful not only in assessing whether a certain information-gathering activity is worth conducting or not, but also in designing a proper information-gathering activity.

The aim of this paper is to evaluate such a possible increase of the maximum expected monetary value brought by gathering information in a qualitative manner. In fact, our problem here can be linked to the so-called value of information (VOI) analysis, as will be discussed later in Section 3. Since introduced by Howard\(^3\), the concept of the VOI has been extensively studied in the area of decision analysis\(^4-8\), and has also been applied to, for instance, medical decision analysis\(^9-12\) as well as petroleum reservoir engineering\(^13-20\). Bratvold et al.\(^21\) conducted a recent survey on the latter application. Due to the unfortunate facts that evaluating the VOI is computationally demanding and that the (nested) Monte Carlo estimator of the VOI is biased\(^10\), several attempts have been made to develop an efficient algorithm to
estimate the VOI$^{9,11,12,17,20}$. Nevertheless, these attempts were made in the context of decision making with an implicit assumption that the number of decision alternatives is finite and even quite small.

If the concept of VOI is to be applied in optimization under geological uncertainty, as shall be done in this paper, we often need to deal with a quite large or even infinite number of candidate solutions, which may cause further trouble for an efficient evaluation of the VOI. More precisely, if there is only a small number of decision alternatives, the maximum of the expected monetary value among them can be found exactly, although the expected monetary value for each alternative should be approximated by an appropriate numerical integration. If there is a quite large or infinite number of candidate solutions within an optimization problem, on the other hand, even the maximum of the expected monetary value cannot be found exactly in general, so that some optimization algorithm should be applied to search for the maximum. Therefore, an increase of the computational cost is inevitable in our current context. In this paper, we address this computational issue to examine the feasibility of applying the concept of the VOI in optimization of reservoir development under geological uncertainty.

The remainder of this paper is organized as follows. In the next section, we give the definition of the VOI in the context of decision making. In Section 3, we incorporate the concept of the VOI into optimization under geological uncertainty, and then show an illustrative example for which the VOI can be calculated analytically. In order to deal with the case where an analytical calculation of the VOI is not available, we introduce a general methodology to estimate the VOI in Section 4. We validate our methodology through a toy problem in Section 5, and furthermore, we demonstrate a practical applicability of the concept of the VOI in our context through a two-dimensional waterflooding problem in Section 6. Finally we conclude this paper with some remarks.

2. Value of Information

Here we mainly follow the exposition of Nakayasu et al.$^{20}$ to introduce the definition of VOI in the context of decision making, where only a finite number of decision alternatives are taken into account.

Let $N_a$ be the number of decision alternatives with an individual alternative represented by $a_i$ for $i = 1, \ldots, N_a$. The task of a decision maker is to decide which alternative $a_i$ is optimal under uncertainty of $X$, where $X$ is assumed to be a set of continuous random variables defined on $\Omega_X$. For each alternative $a_i$, a monetary value function $f(\cdot, a_i):\Omega_X \to \mathbb{R}$ is assigned.

For a risk-neutral decision maker, the optimal alternative is one which maximizes the expected monetary value

$$\int_{\Omega_X} f(x, a_i)p_X(x)\,dx,$$

where $p_X: \Omega_X \to \mathbb{R}$ denotes a prior probability density of $X$. Thus the expected monetary value (EMV) without (prior to) any information is given by

$$\text{EMV}_{\text{prior}} = \max_{i=1,\ldots,N_a} \int_{\Omega_X} f(x, a_i)p_X(x)\,dx.$$

Now let us consider the situation where the decision maker can decide which alternative $a_i$ is optimal given additional information $Y$. If $Y$ is perfect, i.e., if the decision maker can know $X$ precisely by observing $Y$, he/she can decide the optimal alternative which maximizes the monetary value function $f(x, a_i)$ itself depending on the resulting value of $X = x$. Thus the EMV with (posterior to) perfect information is given by

$$\text{EMV}_{\text{perfect}} = \int_{\Omega_X} \max_{i=1,\ldots,N_a} f(x, a_i)\,p_X(x)\,dx,$$

and thus, the expected value of perfect information (EVPI) is given by

$$\text{EVPI} = \text{EMV}_{\text{perfect}} - \text{EMV}_{\text{prior}}.$$

On the other hand, if $Y$ is not perfect but just a sample from a likelihood function $p_Y$, he/she shall decide the optimal alternative which maximizes the conditional expected monetary value

$$\int_{\Omega_X} f(x, a_i)p_{X|Y}(x|y)\,dx,$$

when the resulting value of $Y$ equals $y$, where $Y$ is again assumed to be a set of continuous random variables defined on $\Omega_Y$. In the above, $p_{X|Y}$ denotes the conditional probability density function of $X$ given $Y$, and from Bayes’ theorem and the chain rule of conditional probability, we have

$$p_{X|Y}(x|y) = \frac{p_{Y|X}(y|x)p_X(x)}{p_Y(y)},$$

where

$$p_Y(y) = \int_{\Omega_X} p_{Y|X}(y|x)p_X(x)\,dx.$$

Thus the EMV with (posterior to) sample information is given by

$$\text{EMV}_{\text{post}} = \int_{\Omega_X} \max_{i=1,\ldots,N_a} \int_{\Omega_Y} f(x, a_i)p_{X|Y}(x|y)\,dx\,p_Y(y)\,dy.$$

Correspondingly, the expected value of sample information (EVSI) is given by

$$\text{EVSI} = \text{EMV}_{\text{post}} - \text{EMV}_{\text{prior}}.$$

Note that we always have $0 \leq \text{EVSI} \leq \text{EVPI}$. The term VOI denotes either EVPI or EVSI depending on whether the information $Y$ is perfect or not. From the above definitions of the VOI, the following two or three

inputs are required to evaluate the VOI:

- $p_x$: a prior probability density of $X,$
- $f(\cdot; a)$: a monetary value function for each decision alternative $a,$ for $i = 1, \ldots, N,$ and
- $p_Y$: a likelihood function of the information $Y.$

It is obvious that the last item is not necessary in case of perfect information. We refer to the recent papers\textsuperscript{15,16,19} for a discussion on how the choice of $p_Y$ affects the value of EVSI. Although we only considered the case where a decision maker is risk neutral, risk attitude can be incorporated in the definition of VOI by using an appropriate utility function\textsuperscript{11,16}.

3. Incorporating into Optimization under Uncertainty

3.1. Value of Information in Optimization under Uncertainty

Under the risk neutrality assumption, a continuous optimization under geological uncertainty can be formulated as follows. Again let $X$ be a set of continuous random variables defined on $\mathcal{Q}_s$ and $p_X : \mathcal{Q}_s \rightarrow \mathbb{R}$ be its probability density. Moreover let $A$ be a set of continuous control parameters defined on $\mathcal{Q}_s$ and $f : \mathcal{Q}_s \times \mathcal{Q}_t \rightarrow \mathbb{R}$ be a function whose maximum is to be searched with respect to $A$ under uncertainty of $X.$ Thus the problem is to maximize the following expectation

$$\int_{\mathcal{Q}_s} f(x, a)p_X(x)dx =: F(a),$$

as a function of $a \in \mathcal{Q}_s.$ The optimal solution is of course given by

$$a^* = \arg \max_{a \in \mathcal{Q}_s} F(a).$$

If the function $f : \mathcal{Q}_s \rightarrow \mathbb{R}$ is bounded above, such an $a^*$ always exists, but may not be unique.

As can be seen from the similarity between Eqs. (1) and (4), it is straightforward to incorporate the concept of the VOI into optimization under geological uncertainty by replacing $a,$ and $\max_{x \in A} f(x)$ by $a$ and $\max_{x \in \mathcal{Q}_s}$ respectively. That is, we have

$$\text{EMV}_{\text{pre}} = \max_{a \in \mathcal{Q}_s} \int_{\mathcal{Q}_t} f(x, a)p_X(x)dx,$$

$$\text{EMV}_{\text{post}} = \int_{\mathcal{Q}_s} \max_{a \in \mathcal{Q}_s} f(x, a)p_X(x)dx,$$

$$\text{EMV}_{\text{sample}} = \int_{\mathcal{Q}_s} \max_{a \in \mathcal{Q}_s} \int_{\mathcal{Q}_t} f(x, a)p_X(x)p_Y(x|\mathcal{Y})dx p_Y(x)dx,$$

$$\text{EVPI} = \text{EMV}_{\text{post}} - \text{EMV}_{\text{pre}},$$

$$\text{EVI} = \text{EMV}_{\text{sample}} - \text{EMV}_{\text{pre}}.$$  

As already mentioned, in order for EMV$_{\text{pre}}$ to be well-defined, the function $F$ given in Eq. (4) must be bounded above. Similarly, in order for EMV$_{\text{post}}$ to be well-defined, the function $f(x, \cdot)$ must be bounded above as a function of $a$ for any fixed $x \in \mathcal{Q}_s.$ Furthermore, in order for EMV$_{\text{sample}}$ to be well-defined, the following function $G : \mathcal{Q}_t \times \mathcal{Q}_s \rightarrow \mathbb{R}$ must be bounded above as a function of $a$ for any given $y \in \mathcal{Q}_s$:

$$G(y, a) = \int_{\mathcal{Q}_s} f(x, a)p_X(x|y)dx.$$  

In the subsequent arguments, we shall always assume that the function $f$ satisfies all of these conditions such that the VOI is well-defined.

Note that in the case where the number of candidate solutions is quite large but finite, the definition of the VOI given in Section 2. applies as it is by regarding a set of decision alternative as a set of candidate solutions.

3.2. Illustrative Example

Let us consider an early stage of the reservoir development. Let $X = (X_1, \ldots, X_s)$ be a set of $s$ uncertain geological parameters including the reservoir area, thickness, porosity, water saturation ($S_w$), formation volume factor (FVF), and so on. Under a suitable re-parametrization of those inputs, for instance using $1 - S_w$ as an uncertain parameter instead of $S_w$ or taking the inverse of FVF as an uncertain parameter, the reserve can be estimated by the product $\Pi_{j=1}^s X_j.$\textsuperscript{22} We assume that uncertain parameters are non-negative and independent each other, and that each parameter $X_j$ follows a probability distribution $p_{X_j}$ with mean $\mu_j$ and variance $\sigma_j^2.$

Suppose that we want to optimize an initial investment cost, denoted by a non-negative real $a$, to develop the target reservoir. For low, we cannot gain a monetary value so much even if the reserve is large. For $a$ high, on the other hand, we can gain a large monetary value only if the reserve is large. Since the reserve is uncertain, the initial investment cost should be carefully optimized.

We formulate these observations as follows. Assume that the monetary value gained becomes negative if the reserve is less than $a$, and positive otherwise, and that the slope of the monetary value $f(x_1, \ldots, x_s, a)$ as a function of the reserve $\Pi_{j=1}^s x_j$ depends linearly on $a.$ To satisfy these assumptions, we consider the following function

$$f(x_1, \ldots, x_s, a) = a\left(\prod_{j=1}^s x_j - a\right).$$

In this problem setting, EMV$_{\text{pre}}$ can be calculated analytically as follows:

$$\text{EMV}_{\text{pre}} = \max_{a \geq 0} \int_{\mathcal{Q}_s} f(x_1, \ldots, x_s, a) \prod_{j=1}^s p_{X_j}(x_j) dx_j = \max_{a \geq 0} \left(\prod_{j=1}^s \mu_j - a\right) = \frac{1}{4} \prod_{j=1}^s \mu_j.$$
where the last equality is due to the fact that the maximum with respect to \( a \) is attained when
\[
a = \left( \prod_{j=1}^s \mu_j \right) / 2.
\]

Now let us consider an ideal sample information \( Y \) which gives precise values of \( (X_j)_{\subseteq u} \) for a subset \( u \subseteq \{1, \ldots, s\} \) but has no hint for the remaining parameters. In what follows, we write \( X_u = (X_j)_{\subseteq u} \) and \( X_{\bar{u}} = (X_j)_{\not\subseteq u} \). For such an information, \( \text{EMV}_{\text{post}} \) can be calculated analytically as follows:

\[
\text{EMV}_{\text{post}} = \int_{\Omega_u} \left[ \prod_{j=1}^s p_{X_j}(x_j) \right] \right|_{x_{\bar{u}}} \right|_{x_u} dx_j = \frac{1}{4} \prod_{j=1}^s \left[ \mu_j^2 \left( \sigma_j^2 + \mu_j^2 \right) \right] \left( \prod_{j=1}^s \mu_j^2 \right) - \left( \prod_{j=1}^s \mu_j^2 \right).
\]

Thus, EVSI is given by
\[
\text{EVSI} = \text{EMV}_{\text{post}} - \text{EMV}_{\text{pri}} = \frac{1}{4} \prod_{j=1}^s \left[ \mu_j^2 \left( \sigma_j^2 + \mu_j^2 \right) \right] - \left( \prod_{j=1}^s \mu_j^2 \right).
\]

If we restrict the case where \( |u| = 1 \), i.e., \( Y \) is a partial perfect information only for a single uncertain parameter, the largest EVSI is attained for \( X_1 \) with the largest ratio \( \sigma_j/\mu_j \) in this example. This way the VOI analysis can be used not only in the area of decision making but also in optimization under geological uncertainty. However, since it is not always the case where an analytical calculation of the VOI is possible, we discuss how to estimate the VOI in the following.

4. General Methodology to Estimate Value of Information

In order to deal with the case where an analytical calculation of the VOI is not possible, here we introduce a general methodology to estimate the VOI defined in the previous section. In the following algorithms, the stopping criterion used within each optimization process needs to be properly chosen by the users. Let us assume that drawing random samples of \( X \sim p_X \) is easy. First, \( \text{EMV}_{\text{pri}} \) can be simply estimated as follows.

**Algorithm 1** (Estimation of \( \text{EMV}_{\text{pri}} \)) Let \( p_X \) be a probability density of \( X \). For a positive integer \( N \), draw random samples \( x_1, \ldots, x_N \) from \( p_X \). Then do the following:

1. Choose an initial candidate \( a \in \Omega_1 \) and let \( r^* = -\infty \).
2. Evaluate
\[
r = \frac{1}{N} \sum_{i=1}^N f(x_i, a)
\]
(3) If \( r > r^* \) holds, replace \( r^* \) by \( r \).
(4) If the stopping criterion is satisfied, output \( r^* \) as an estimate of \( \text{EMV}_{\text{pri}} \). Otherwise, choose a new candidate \( a \in \Omega_1 \) by some optimization algorithm and go to Step 2.

In a similar way, \( \text{EMV}_{\text{post}} \) can be estimated as follows.

**Algorithm 2** (Estimation of \( \text{EMV}_{\text{post}} \)) Let \( p_X \) be a probability density of \( X \). For a positive integer \( N \), draw random samples \( x_1, \ldots, x_N \) from \( p_X \). From \( i = 1, \ldots, N \), do the following:

1. Choose an initial candidate \( a \in \Omega_1 \) and let \( r^* = -\infty \).
2. Evaluate \( r_i = f(x_i, a) \). If \( r > r^* \) holds, replace \( r^* \) by \( r_i \).
3. If the stopping criterion is satisfied, output \( r^* \). Otherwise, choose a new candidate \( a \in \Omega_1 \) by some optimization algorithm and go to Step 2.

Finally output
\[
\frac{1}{N} \sum_{i=1}^N r_i^*
\]
as an estimate of \( \text{EMV}_{\text{post}} \).

Then EVPI can be estimated by subtracting the estimate of \( \text{EMV}_{\text{pri}} \) from that of \( \text{EMV}_{\text{post}} \). If at most \( L \) iterations are needed until a stopping criterion is satisfied in Step 3 of these two algorithms, we need \( O(LN) \) function evaluations to estimate EVPI.

Estimating \( \text{EMV}_{\text{post}} \) and EVSI is more computationally demanding as can be seen below. If drawing random samples of \( Y \sim p_Y \) given \( X \) is easy, it follows from the chain rule (3) that drawing random samples of \( Y \sim p_Y \) can be also easily done by randomly sampling \( X \sim p_X \) and then \( Y \sim p_Y \) sequentially. However, computing a probability density \( p_Y \) is not necessarily easy because we need to evaluate an integral of (3). In what follows, we introduce two methods to estimate \( \text{EMV}_{\text{post}} \) without computing \( p_Y \): one is based on Markov chain Monte Carlo (MCMC) sampling and the other is on importance sampling. We refer to the books\(^{23,24}\) for extensive information on these sampling methods.

The first method is actually a straightforward implementation of MCMC sampling for estimation of \( \text{EMV}_{\text{post}} \). The significant advantage of MCMC sampling is that we can draw a chain of samples of \( X \sim p_X \) given \( Y \) directly without any need to compute \( p_Y \).

\[
\]
Although consecutive samples drawn by MCMC sampling are not independently distributed, the chain after a large number of steps can be used as a sequence of random samples from \( p_X \). In this way, \( \text{EMV}_{\text{post}}^{\text{sample}} \) can be estimated as follows.

**Algorithm 3** (Estimation of \( \text{EMV}_{\text{post}}^{\text{sample}} \)) Let \( p_X \) be a probability density of \( X \) and \( p_{Y|x} \) a likelihood function. For a positive integer \( M \), draw random samples \( y_1, \ldots, y_M \) from \( p_Y \). From \( j = 1, \ldots, M \), do the following:

1. For a positive integer \( N \), draw a chain of samples \( x_1, \ldots, x_N \) from \( p_X \) given \( Y = y_j \) by using some MCMC sampling.
2. Choose an initial candidate \( a \in \Omega_4 \) and let \( r_j' = -\infty \).
3. Evaluate
   \[
   r_j = \frac{1}{N} \sum_{i=1}^{N} f(x_i, a)
   \]
   (4) If \( r_j > r_j' \), replace \( r_j' \) by \( r_j \).
5. If the stopping criterion is satisfied, output \( r_j' \).
   Otherwise, choose a new candidate \( a \in \Omega_4 \) by some optimization algorithm and go to Step 3.

Finally output
\[
\frac{1}{M} \sum_{j=1}^{M} r_j'
\]
as an estimate of \( \text{EMV}_{\text{post}}^{\text{sample}} \).

For the second method based on importance sampling, we introduce an auxiliary probability density \( q_Y : \Omega_4 \rightarrow \mathbb{R} \) such that \( q_Y \) is easy to compute and drawing random samples from \( q_Y \) is also easy. Using (2) and multiplying and dividing the right-hand side of (5) by \( q_Y \), we have
\[
\text{EMV}_{\text{post}}^{\text{sample}} = \int_{\Omega_4} \left[ \max_{a \in \Omega_4} \frac{f(x, a)}{q_Y(y)} \frac{p_{Y|x}(y|x) p_X(x)}{p_Y(y)} \right] p_Y(y) dy
\]
\[
= \int_{\Omega_4} \max_{a \in \Omega_4} \frac{f(x, a)}{q_Y(y)} \frac{p_{Y|x}(y|x) p_X(x)}{q_Y(y)} \right] q_Y(y) dy.
\]

The last expression of \( \text{EMV}_{\text{post}}^{\text{sample}} \) can be exploited to provide the second estimation algorithm as shown below.

**Algorithm 4** (Estimation of \( \text{EMV}_{\text{post}}^{\text{sample}} \)) Let \( p_X \) be a probability density of \( X \), \( p_{Y|x} \) a likelihood function, and \( q_Y \) an auxiliary probability density of \( Y \). For a positive integer \( M \), draw random samples \( y_1, \ldots, y_M \) from \( q_Y \). From \( j = 1, \ldots, M \), do the following:

1. For a positive integer \( N \), draw random samples \( x_1, \ldots, x_N \) from \( p_X \).
2. Choose an initial candidate \( a \in \Omega_4 \) and let \( r_j' = -\infty \).
3. Evaluate
   \[
   r_j = \frac{1}{N} \sum_{i=1}^{N} f(x_i, a) \frac{p_{Y|x}(y_j|x_i)}{q_Y(y_j)}
   \]
   (4) If \( r_j > r_j' \), replace \( r_j' \) by \( r_j \).
5. If the stopping criterion is satisfied, output \( r_j' \).
   Otherwise, choose a new candidate \( a \in \Omega_4 \) by some optimization algorithm and go to Step 3.

Finally output
\[
\frac{1}{M} \sum_{j=1}^{M} r_j'
\]
as an estimate of \( \text{EMV}_{\text{post}}^{\text{sample}} \).

This algorithm does work as long as \( q_Y(y) > 0 \) whenever \( p_{Y|x}(y|x) > 0 \) for any \( x \in \Omega_X \). EVSI can be estimated by subtracting the estimate of \( \text{EMV}_{\text{post}}^{\text{sample}} \) obtained by **Algorithm 1** from that of \( \text{EMV}_{\text{post}}^{\text{sample}} \) obtained by either **Algorithm 3** or 4. If at most \( L \) iterations are required until a stopping criterion is satisfied in Step 4 of either **Algorithm 3** or 4, we need \( O(LMN) \) function evaluations to estimate EVSI. Therefore, it is clear that estimating EVSI is more computationally demanding than EVPI.

Here we give some remarks on the above algorithms.

- We need to implement some optimization algorithm to search for a new candidate \( a \in \Omega_4 \). Although the choice of an optimization algorithm is arbitrary, derivative-based methods are hard to apply in general, so that we recommend to use derivative-free methods which can be either adaptive or non-adaptive. These include grid search, Nelder-Mead simplex search, and evolutionary computation algorithms as well as many others.

- We can replace random samples with stratified samples, Latin hypercube samples or quasi-Monte Carlo samples if possible. For the low-dimensional case, it is even possible to apply some numerical integration method such as mid-point formula and Gaussian quadrature formula instead of drawing random samples.

- There are several MCMC sampling methods which can be used in **Algorithm 3**, such as Metropolis-Hastings sampling, Gibbs sampling, and slice sampling. In practice, we need to choose an appropriate sampling method depending on the problem at hand. Similarly, the auxiliary probability density \( q_Y \) in **Algorithm 4** should be designed appropriately depending on the problem. Empirically, we recommend to use \( q_Y \) which is close to \( p_X \).

5. **Toy Problem**

5.1. **Problem Setting**

The toy problem considered here is a simplification of that used in **Subsection 3.2**. As a geological
uncertainty, let $X$ denote the reserve itself instead of a set of multiple geological parameters. We assume that $X$ is log-normally distributed with given two parameters $\mu \in \mathbb{R}$ and $\sigma > 0$, that is, the prior probability density is given by

$$p_X(x) = \frac{1}{\sqrt{2\pi} \sigma x} \exp \left\{ -\frac{(\ln x - \mu)^2}{2\sigma^2} \right\}. \quad (8)$$

The prior mean and standard deviation of $X$ are denoted by $\mu_X$ and $\sigma_X$, respectively. Here we have

$$\pi = 2 \ln \mu_X - \ln \left( \mu_X^2 + \sigma_X^2 \right) / 2 \quad \text{and} \quad \sigma^2 = \ln \left( 1 + \sigma_X^2 / \mu_X^2 \right).$$

In the following argument, we normalize the uncertainty of $X$ by using the coefficient of variation $CV = \sigma / \mu_X$, i.e., the value of $CV$ represents the uncertainty of our prior knowledge. In fact, such a normalization was considered previously by the authors$^{16,19}$. Similarly to Eq. (7), the monetary value function is now given by

$$f(x, a) = \frac{a(x - a)}{u},$$

for a given scaling factor $u > 0$. In this problem setting, $EMV_{\text{pri}}$ and $EMV_{\text{perfect}}$ can be obtained analytically as:

$$EMV_{\text{pri}} = \frac{\mu_X^2}{4u} \quad \text{and} \quad EMV_{\text{perfect}} = \frac{\mu_X^2 + \sigma_X^2}{4u}.$$

Therefore, EVPI is given by

$$EVPI = EMV_{\text{perfect}} - EMV_{\text{pri}} = \frac{\sigma_X^2}{4u} - \frac{\mu_X^2 CV^2}{4u}.$$  

In case of sample information $Y$ which is now assumed to be obtained by conducting some test to estimate the reservoir extent, we consider the following Gaussian likelihood function

$$p_Y(y|x) = \frac{1}{\sqrt{2\pi} \sigma_y} \exp \left\{ -\frac{(y - x)^2}{2\sigma_y^2} \right\}. \quad (9)$$

for a given $\sigma_y > 0$. Note that $Y$ can take a negative value. Again for normalization we introduce the parameter $b = \sigma_Y / \mu_Y$ which represents the reliability of $Y$. In this case, it is no longer possible to calculate $EMV_{\text{post}}$ and $EVSI$ analytically. It can be shown at least, however, that the information $Y$ is perfect. Figure 1 shows prior probability densities $p_X$ with different values of $CV$ on the left and monetary value functions $f(\cdot, a)$ with different values of $a$ on the right. In order to give a flavor of how the information $Y$ affects the optimal initial investment cost $a$, we plot the function $a^{\text{max}}(y)$ as a function of $y$ within the range $[0, 120]$ in Figure 2. The left shows $a^{\text{max}}$ with different values of $b$ for the case $CV = 0.3$, whereas the right does with different values of $CV$ for the case $b = 0.2$. In the absence of information, the optimal initial investment cost equals $\mu_Y / 2 = 60$. In both the plots, all the curves intersect at around $y = \mu_Y (\approx 120)$, which gives $a^{\text{max}} = 60$. This implies that the optimal initial investment cost does not change so much when the observation result corresponds to the prior mean. On the contrary, it can be seen that the optimal initial investment cost after gathering the information $Y$ changes significantly when the observation result $y$ is far from the prior mean.

5.2. Numerical Results

Here we estimate EVPI and EVSI by using our algorithms presented in Section 3. We always apply a

Fig. 1 Prior Probability Densities $p_X$ with Various Values of $CV$ (left) and Monetary Value Functions $f(\cdot, a)$ with Various Values of $a$ (right)
mid-point formula instead of drawing random samples with \( M'N' = 10^3 \) and use Algorithm 4 to estimate \( \text{EMV}_{\text{post}} \). Furthermore, we set the upper limit of an initial investment cost to 200, which should be more practical than leaving it unbounded, and thus, the search range of an initial investment cost is restricted to the interval \([0, 200]\). As an optimization algorithm, we use the so-called golden section search \(^{25,26}\), which is one of the classical adaptive algorithms. After \( T' \) steps, the length of the search range can be reduced to \( 200 \times w'^{T'} \) with \( w' = (\sqrt{5} - 1)/2 \). For any unimodal function, the sequence of function evaluations given by the golden section search converges to the true maximum.

Let us consider the case of perfect information first. Since we know the exact EVPI in this case, it is possible to validate our methodology by comparing our estimates of EVPI with the exact EVPI. Figure 3 shows the EVPI estimates for \( CV = 0.1, 0.11, \ldots, 0.49, 0.5 \) obtained by the golden section search with different maximum numbers of steps \( T = 3, 5, 7, 10 \). The estimates are in good agreement with the exact ones already for \( T = 5 \). In order to quantify the rate of convergence, we computed the root mean square error (RMSE) of the estimates for each \( T' = 3, \ldots, 10 \). As shown in Fig. 4, the RMSE converges exponentially fast up to \( T = 7 \), beyond which the RMSE does not converge anymore since the numerical error which comes from the approximate evaluation of the integral cannot be reduced without increasing \( N \) in Algorithms 1 and 2. Thus, we can say that \( T' = 7 \) is enough for the golden section search to obtain a satisfactory convergence here, which means that the computational difficulty of this problem is equivalent to that of evaluating EVPI in the context of decision making under uncertainty with 7 decision alternatives.

Let us move on to the case of sample information. We recall that in this case we no longer know the exact EVSI, although the function \( G(y, a) \) given in Eq. (6) is known to be concave and unimodal as a function of \( a \) for any fixed \( y \). Therefore, if we neglect the approximation error which comes from Step 3 of Algorithm 3 or 4, the golden section search can asymptotically find the exact peak \( a_{\text{max}}(y) \) for any fixed \( y \), which leads to an
accurate estimate of EVSI for large $T$. Figure 5 (left) shows the EVSI estimates for a fixed $b = 0.2$ and $CV = 0.1, 0.11, \ldots, 0.49, 0.5$ obtained by the golden section search with different maximum numbers of steps $T = 3, 5, 7, 10$, whereas Fig. 5 (right) does those for a fixed $CV = 0.3$ and $b = 0, 0.01, \ldots, 0.25$. In order to quantify the convergence behavior, we consider the root mean square difference (RMSD) of consecutive estimates with maximum numbers of steps $T$ and $T+1$. As shown in Fig. 6, the RMSD converges exponentially fast, which is quite similar to that can be seen in Fig. 4. It can be inferred from the fitted curves that we only need around $T \approx 10$ to make RMSD less than $10^{-2}$. Thus, a suitable specialization of our general methodology can provide a reliable estimate of VOI with fewer function evaluations.

6. Two-dimensional Waterflooding Problem

6.1. Problem Setting

We conduct numerical experiments for a two-dimensional waterflooding problem to demonstrate a practical applicability of the concept of the VOI in optimization under geological uncertainty. We use the software IMEX for reservoir simulations. Let us consider a virtual oil reservoir of the size $1 \text{ km} \times 1 \text{ km} \times 20 \text{ m}$ with a no-flow boundary, which is discretized into $25 \times 25$ equal-sized grid cells for reservoir simulations. In order to model one quarter of a repeated five-spot pattern, one water injector is located at the $(1,1)$-th grid cell and one producer is at the $(25,25)$-th grid cell. For simplicity, we assume that the reservoir is homogeneous and that all the geological parameters except the permeability are already known. To unify the notation, we denote the permeability by $X [\text{md}]$ instead of the usual convention $k$. The prior probability density for $X$ is given by a log-normal density function (8) with the mean $\mu_X = 120$ and $CV = 0.3$, see also Fig. 1 (left). Other necessary inputs and model parameters are given as in Table 1. Note that we consider different four values of the discount rate, denoted by $r$.

Suppose that a risk-neutral developer is trying to optimize the maximum water-injection rate $a [\text{m}^3/\text{day}]$ from the injector such that the expected monetary value
is maximized under uncertainty of permeability \( X \). The production of oil is done by setting the bottomhole pressure (BHP) at the producer constantly to 20 MPa. Given \( x \) and \( a \), the monetary value is given by

\[
f(x, a) = \sum_{\text{year } t} \frac{\sum_{\text{year } t} q_0(t|x, a) - c_W q_0(t|x, a) - c_0 q_0(t|x, a) - \lambda a}{1 + r} \Delta t, \tag{11}
\]

where \( N_t \), \( r \), \( p_o \), \( c_w \), \( c_0 \) denote the production duration, the discount rate, the oil price and the costs of disposing and injecting water, respectively. Recall that \( N_t = 10 \), \( p_o = 50 \), \( c_w = 3 \), \( c_0 = 3 \) in this paper, where the units of the latter three are $/bbl. Furthermore, \( q_0(t|x, a) \) and \( q_0(t|x, a) \) are the production rates at the year \( t \) for oil and water, respectively, and \( q_0(t|x, a) \) is the injection rate at the year \( t \). Note that \( q_0(t|x, a) \) can be possibly reduced from the maximum water-injection rate \( a \) if the BHP at the injector is increased up to 35 MPa. The second term in Eq. (11) is included to represent CAPEX.

The problem here is to evaluate the value of collecting the information \( Y \) by conducting a test to measure \( X \) with some degree of accuracy. The likelihood function of \( Y \) is given by the form (9) with the parameter \( b \in [0, 0.25] \). When \( b = 0 \), we assume that information \( Y \) brought from the test is perfect.

6.2. Numerical Results

Before estimating EVPI and EVSI, we first show how the uncertainty of \( X \) affects the optimal value of \( a \). Figure 7 shows the profiles of \( f(x, a) \) as functions of \( a \) for \( x = 50, 100, 150, 200, 250 \), where \( f(x, a) \) is evaluated by running the reservoir simulator for every hundred from \( a = 0 \) to \( a = 2000 \). For any value of the discount rate \( r \), the profile of \( f(x, a) \) is a concave function of \( a \) with a single peak when \( x \) is fixed. The monotone increase of \( f(x, a) \) for \( a \) below the peak stems from an increased oil production due to water injection, whereas the monotone decrease for \( a \) beyond the peak stems from overwhelming of the costs of disposing and injecting water and the necessary CAPEX. For the case \( r = 0 \), the function \( f(x, a) \) reaches its maximum when \( a \) is around 300 for the low-permeability case \( x = 50 \), whereas it does when \( a \) is around 700 for all of the other higher-permeability cases. As \( r \) increases, the dependence of the peak location on \( x \) becomes more evident. For the case \( r = 0.2 \), for instance, the function \( f(x, a) \) reaches its maximum when \( a \) is again around 300 for the case \( x = 50 \), whereas it does \( a \) is around 1000 for the case \( x = 250 \).

Now we estimate EVPI and EVSI by using our algorithms presented in Section 3. Similarly to the previous section, we apply a mid-point formula with \( M = N = 10^2 \) and use Algorithm 4 to estimate \( \text{EMV}_{\text{post}}^{\text{sample}} \).

We restrict the search range of the maximum water-injection rate \( a \) to the interval \([0, 1000]\). As an optimization algorithm, we always use the golden section search with the maximum number of steps \( T = 22 \). Since we no longer know whether functions to be maximized are unimodal or not, the golden section search may fail to find true optimum. Nevertheless, the exponential convergence of the golden section search enables fast evaluations of EVPI and EVSI with few maximum number of steps \( T \). We set \( T = 22 \) because the length of search range can be reduced from 1000 to less than 0.1, which assures a sufficient convergence. Note that if we search for a maximum with the precision 0.1 uniformly within the interval \([0, 1000]\), we need \( T = 10^4 \) steps. Since one reservoir simulation run takes one second or less for our field model, evaluating EVSI takes less than 3 days by using the golden section search, while it is expected to take about 3 years by using such a uniform search. We can easily see the importance of choosing a suitable optimization algorithm for estimating the VOI in the current context.

Similarly to Fig. 2, we plot the function \( a_{\text{max}} \), which is defined as in Eq. (10), with different values of \( b \) in Fig. 8. For the perfect information case \( b = 0 \), \( y \) should be read as \( x \). For the case \( r = 0 \), it can be seen that the function \( a_{\text{max}} \) has an upper bound near 650 regardless of the value of \( b \), and that the profiles of \( a_{\text{max}} \) significantly differ from each other depending on the value of \( b \) for \( y \) less than 100. As \( r \) increases, an upper bound on \( a_{\text{max}} \) also increases and the profiles of \( a_{\text{max}} \) significantly differ from each other depending on the value of \( b \) for wider range of \( y \). This means that the optimal choice of the maximum water-injection rate under uncertainty becomes more sensitive to the reliability of information when \( r \) is large.

Finally, the resulting values of EVPI and EVSI as functions of \( b \) are shown in Fig. 9 for all the cases \( r = 0, 0.05, 0.1, 0.2 \). It is obvious that, for any value of \( r \), EVSI decreases as \( b \) increases. As \( r \) increases from 0
to 0.1, EVSI also increases for any $b$, although we cannot see a clear difference of the EVSI profiles between the cases $r' = 0.1$ and $r' = 0.2$. The increase of EVSI up to $r' = 0.1$ can be explained by the fact that the variability of $a^\text{max}$ becomes more significant for larger $r$. However, as $r$ becomes larger, the expected monetary values (both prior and posterior) themselves become lower, and thus, it can be expected that EVSI cannot increase more when $r$ goes beyond 0.1. This is why the EVSI profiles for the cases $r = 0.1$ and $r = 0.2$ are similar.

In order for the information $Y$ to be valuable in optimizing under uncertainty of $X$, EVSI must be positive. Considering the extrapolations of the resulting values of EVSI by using the fitted linear functions of $b$, we see that $b$ must be less than 0.27 and 0.35 for the cases $r = 0$ and $r = 0.2$, respectively. Therefore, if $b$ is larger than these values, the reliability of the information $Y$ is not enough and the optimal value of $a$ can be determined without the need of gathering the information. In this way our proposed methodology enables not only to evaluate how much the information-gathering activity provides an increase of the maximized expected monetary value, but also to identify the necessary condition for the information to be valuable in optimizing reservoir development under geological uncertainty.

7. Concluding Remarks

In this paper we first incorporated the concept of VOI into optimization of reservoir development under geological uncertainty together with an illustrative example. We then introduced a general methodology to estimate the VOI in our context. In practice, one should implement appropriate optimization algorithm and sampling method depending on a problem at hand. In order to validate our methodology, we conducted numerical experiments for the toy problem and the simple waterflooding problem. In the toy problem, we confirmed that by using a suitable search algorithm the computational difficulty can be reduced to that of estimating the VOI in the usual setting, i.e., in the context of decision making under uncertainty with a small number of decision alternatives. Moreover, we demonstrated through the simple waterflooding problem that our methodology is of practical use for measuring how much the maximum expected monetary value can be increased as a result of an information-gathering activity.
Acknowledgment

The author would like to thank two anonymous referees for their careful reading of the manuscript and their helpful comments and suggestions. This work is supported by JX Nippon Oil & Gas Exploration Corporation. The work of T. G. is supported by Arai Science and Technology Foundation. The authors are thankful to CMG Ltd. for the software license.

References

17) Bickel, J. E., SPE Econ. & Mgmt., 4, 198 (2012).

Fig. 8 Profiles of $a^{\text{max}}$ with Various Values of $b$ for the Waterflooding Problem for the Cases $r = 0$ (left top), $r = 0.05$ (right top), $r = 0.1$ (left bottom) and $r = 0.2$ (right bottom), Respectively

Fig. 9 EVSI Estimates and the Fitted Regression Lines as Functions of $b$ with Various Values of $r$ for the Waterflooding Problem
Geosci., 20, 737 (2016).

요 旨

地質的不確実性下での油層開発最適化における情報の価値

合田 隆, 大野 克太, 佐藤 光三
東京大学大学院工学系研究科, 113-8656 東京都文京区本郷7-3-1

情報の価値（VOI: value of information）分析は、ある情報収集を行うことにより期待利潤がどの程度増加するのかを測る有用な手段として知られている。通常、VOI分析は地質的不確実性下における意思決定問題、とりわけ選択肢が比較的少数の意思決定問題に適用される。本研究では地質的不確実性下における油層開発最適化問題への VOI分析応用可能性について議論する。すなわち、開発最適化によって最大化される期待利潤がある情報収集によってどの程度増加させられるのかを VOI分析を通じて定量評価する。最適化問題では非常に膨大な数の候補解から有効な解を探索する必要があるため、従来の VOI分析に比べて効率的な実装が難しいと予想される。“不確実性下での最適化問題”の文脈において VOIを推定する一般的な計算アルゴリズムを導入したのち、単純な例題ならびに水攻法を模倣した油層開発最適化問題に応用する。VOI分析に適切な最適化手法を組み込むことによって、効率的に VOIが評価できることを示す。