Natural Resources Research, Vol. 31, No. 5, October 2022 (© 2022)
https://doi.org/10.1007/s11053-022-10078-z

Original Paper

Sequential Value of Information for Subsurface Exploration Drilling

T. Hall,[1,3] C. Scheidt,[2] L. Wang,[1] Z. Yin,[1] T. Mukerji,[2] and J. Caers[1]

Received 29 October 2021; accepted 21 May 2022
Published online: 29 June 2022

Quantitative methods are needed for systematic decision-making during exploration for subsurface resources, but few methods exist that fully incorporate the interaction of geological, operational, and financial conditions. The sequential nature of planning where to drill for subsurface exploration is not commonly addressed by conventional techniques in a quantitative fashion, despite its foundational relevance to hypothesis testing. Value of information (VOI) can incorporate various aspects of subsurface exploration decision-making as well as sequence. Here, we use VOI to determine the optimal sequence and placement of exploration boreholes when varying conditions such as target resource volume and drilling cost. Using VOI, we show that the optimal placement and selection of exploration boreholes change when planning to drill one borehole at a time compared to planning to drill two boreholes sequentially. A formulation and tutorial explanation of VOI for sequential decision situations are shown using a synthetic case. We demonstrate a test case using data from a real metal deposit.

KEY WORDS: Subsurface exploration, Value of information, Sequential information gathering, Optimal drill placement, Decision-making under uncertainty.

INTRODUCTION

Understanding the nature of the subsurface is critical to an increasing diversity of scientific problems, including the storage of hydrogen (Tarkowski 2019) and captured CO_2 (Wei et al. 2021), the availability of Martian water–ice (Morgan et al. 2021), as well as the incumbent domains of petroleum, minerals, groundwater, and contaminants. The variety of subsurface contexts is driven by technological innovation as well as the need to accommodate a changing climate and a growing population. For example, groundwater resources are increasingly stressed due to over-pumping and climate change (Wu et al., 2020), yet the investigation of deep, more resilient aquifers is relatively new compared to shallower aquifers (Kang & Jackson, 2016; Lall et al., 2020).

When considering metals, the global development of renewable energy technology and infrastructure will increase demand for lithium, cobalt, and graphite by approximately 450% over the next three decades (Hund et al. 2020). For the UK alone, total fleet replacement of internal-combustion vehicles with battery electric vehicles requires 75% of the global annual supply of lithium, 200% of the

[1]Department of Geological Sciences, Stanford University, Stanford, USA.
[2]Department of Energy Resources Engineering, Stanford University, Stanford, USA.
[3]To whom correspondence should be addressed; e-mail: trhall@stanford.edu

1520-7439/22/1000-2413/0

global annual supply of cobalt, and 100% of the global annual supply of neodymium (Herrington 2021). In the future, the global metal demand may be entirely sustained by closed-loop recycling; however, current technological limitations necessitate the growth of the mining industry (Ali et al. 2017).

Discoveries of significant mineral deposits and therefore addition of new resources into the global supply have waned due to limited surface expression of geological indicators of mineralization as well as reduced capital expenditure (Schodde 2017), which some authors believe is a significant constraint on future supply (Ali et al. 2017; Schodde 2017; Hund et al. 2020; Herrington 2021). Other authors find mineral exploration is not a limiting factor and rather it is the environmental, social, and governance of mining (ESG) that limits production (Jowitt et al. 2020). Furthermore, some authors find that the recent decline in discoveries will not be a significant constraint on future supply due to production and extraction innovations (Rötzer and Schmidt 2018; Ericsson et al. 2019). Other authors find that using estimations of extractable global resources in order to guide investment or policy-making is highly uncertain (West 2020).

The significance of ESG restrictions (difficult to open a new mine), production innovation, and increasing commodity price may lead mining companies to re-assess previously unprofitable mines currently in reclamation in addition to pursuing greenfields targets. Efficient asset management requires that the decision to develop or sell a property is made accurately and in a reasonable timeframe—a mistaken or delayed decision may cost decades in upkeep. Therefore, methods that aid exploration for both new and existing properties are needed.

Uncertainty regarding the spatial characterization of mineral resources affects the rate of upstream discovery and profitability of downstream production. Companies perform various tests of the subsurface to reduce this uncertainty and thereby generate geological interpretations and predictions for exploration decision-making (Yousefi et al. 2021; Zuo et al. 2021) and resource estimation, mine planning, and production forecasting (Dimitrakopoulos 2018; Deutsch 2021). These tests include drilling, geophysics, geochemistry, and field-based observations. The interpretation of test results is the responsibility of decision-makers who then must determine the necessity, type, extent, and sequence of continued testing.

Experts use various information metrics to determine the optimal placement, orientation, and depth of mineral exploration and infill drilling. These metrics include resource conversion (Deutsch et al., 2007; Nowak and Leuangthong 2019), kriging variance (Gershon et al. 1988; Barnes 1989; Delmelle and Goovaerts 2009; Isaaks and Srivastava 2010), conditional simulations (Journel 1993; Goovaerts 1997; Jean-Paul Chiles and Pierre Delfiner 1999; Verly and Parker 2021), as well as incorporation of these measures along with other estimations of value (Boucher et al. 2005; Eidsvik and Ellefmo 2013; Dirkx and Dimitrakopoulos 2018; Froyland et al. 2018).

Research on decision-making for information gathering in the subsurface has drawn extensively from decision theory (Raiffa and Schlaifer 1961), value of information (Howard 1966; Raiffa 1968), and optimal experimental design (Le and Zidek 2006; Müller 2007; Cressie and Wikle 2011). However, it is also common that experts do not use any of these information metrics or decision-theoretics to guide decision-making, and instead rely on their interpretation and best judgement.

Value of information has been applied to a variety of subsurface domains. A summary reference is Eidsvik et al. (2015). Value of information has been tested extensively for petroleum applications (Newendorp and Schuyler 2002; Bickel and Bratvold 2008; Bratvold et al. 2009), applied to spatial datasets including seismic data (Bhattacharjya et al. 2010; Dutta et al. 2019; Jreij et al., 2021), and mineral exploration (Soltani and Hezarkhani 2011).

Value of information does not strictly assume a multivariate gaussian distribution, it directly incorporates the data variability as well as the impact of the future data on the decision, and the cost of data collection can be compared to the added value. These are fundamental differences compared to more commonly applied methods such as kriging variance.

In many circumstances, information gathering takes place sequentially (Miller 1975; Kochenderfer et al., 2022). That is, the decision-maker may seek to perform a sequence of tests rather than a single static measurement. Broadly, sequential information gathering in the subsurface is described by Eidsvik et al. (2018). Bickel and Smith (2006) demonstrate that six individually unattractive oil wells comprise a valuable opportunity when drilled in a certain sequence. Given the complexity of the subsurface, updating the prior geological model with all possible data outcomes for long sequences of boreholes becomes an intractable problem (Bickel and Smith 2006; Eidsvik et al. 2018). Addressing this intractability by efficiently searching the problem space is a broad area of research (Kochenderfer et al. 2022). For example, approximate solutions from dynamic programming may be used (Powell 2011). Ultimately, the solution is designed specific to the problem domain (Eidsvik et al. 2015).

The very nature of mineral exploration and appraisal, as with other resources, is that it involves sequential planning of information gathering to reduce uncertainty on the subsurface. An explorer must be able to plan a multi-step campaign where they determine how best to proceed with testing their hypotheses of the subsurface for various outcomes, rather than determining where to drill only after the results have been observed. The questions of where to test, which order to perform the tests, and what methods to test with are all components of a well-planned exploration campaign.

Here, we address this sequential nature of subsurface exploration using value of information in order to determine the optimal exploration plan and demonstrate the difference in decision-making when planning to explore "one step at a time" and planning to explore sequentially. We use the quantitative value of information (Howard 1966), rather than other approaches such as realistic value of information (Soltani and Hezarkhani 2013). Sequential decision-making can be directly encoded in value of information, which is a unique feature of the method. The optimal placement and sequence of exploration boreholes are systematically compared based on drilling cost and volume above a chosen threshold (though other statistics may be used). The forecasted performance of each exploration plan is displayed in dollar values, which may allow for more effective communication, planning, and contract negotiation.

In previous work, spatial dependencies and conditional probabilities required for sequential information gathering are elicited from subject matter experts (Bickel et al. 2006) or expert opinions combined with geoscientific information (Eidsvik et al. 2018; Morosov and Bratvold 2022). In this work, probabilities are modeled directly using geological and geophysical models built by experts.

The paper begins with an introduction of value of information and a formulation of value of information for sequential information gathering. Then, a synthetic case is used to demonstrate an application of value of information for exploration drilling for target volumes. Next, value of information is tested on a real case. Finally, the paper is concluded with summarizing remarks and future work.

METHODOLOGY

Review: Value of Information

The calculation of value of information requires the definition and mathematical formulation of various components of a decision situation. This formulation is for discrete cases, for example, when data take a binary form indicating the presence or absence of a geological feature. For continuous cases, value of information is calculated by integrating over all possible data outcomes (Raiffa 1968; Eidsvik et al. 2015).

The primary uncertainty that a decision-maker seeks to resolve via data collection is defined as the distinction of interest. For the subsurface, distinctions of interest include the presence, absence, or various configurations of geological features with scientific or economic importance (Bhattacharjya et al. 2010; Eidsvik et al. 2015). Here, we denote the distinction of interest as a categorical variable: s. Note that value of information analysis is not limited to categorical data; it is done here to provide a tutorial explanation. The decision-maker's understanding of each possible instantiation of the distinction of interest is modeled by the probability $p(s)$. The decision-maker chooses an action from a set of alternatives a, defined a-priori. The set of alternatives that a decision-maker may select from are specific to the decision situation; examples include drilling a well for oil, developing a mine, performing a geophysical survey, or selling an asset. A scenario is a possible instantiation of the distinction of interest following a decision. The decision-maker uses the value function $v(s, a)$ to quantify the value assigned to each scenario and action. In some cases, a decision-maker may perform testing (information gathering) of the distinction of interest before making a larger commitment of time or money. This testing does not completely reveal the distinction of interest (imperfect information) but may improve the decision-maker's understanding. In order to quantify the added benefit of data collection, the decision-maker calculates the value of information.

The value of information is the difference between the value of a decision situation before (PV) and after (PoV) testing; thus,

$$\mathrm{VOI} = \mathrm{PoV} - \mathrm{PV} \tag{1}$$

The value of a decision situation before any information gathering is defined as the prior value, PV, the maximum over all actions of the expected value over all scenarios:

$$\mathrm{PV} = \max_a \left\{ \sum_s E[v(s,a)] \right\} \tag{2}$$

Risk preferences may vary between decision-makers, and in such cases, utility functions may be used to model those preferences (von Neumann and Morgenstern 1944). Here, we assume a risk-neutral decision-maker with a linear utility function.

To calculate the posterior value, PoV, we introduce the data variable $\boldsymbol{d}$, modeling the uncertainty of future information gathering; $\boldsymbol{d}$ is a vector and potentially high-dimensional (e.g., possible readings from a geophysical sensors, or lithologies observed along a borehole). The relationship between $\boldsymbol{d}$ and s is modeled as the likelihood function $p(\boldsymbol{d}|s)$. The posterior distribution of the distinction of interest is calculated via Bayes' Rule; thus,

$$p(s|\boldsymbol{d}) = \frac{p(\boldsymbol{d}|s)p(s)}{p(\boldsymbol{d})} \tag{3}$$

In the categorical case, $p(\boldsymbol{d})$ is obtained via summation of the marginal probabilities of s; thus,

$$p(\boldsymbol{d}) = \sum_s p(\boldsymbol{d}|s)p(s) \tag{4}$$

The value with information, PoV, is defined as the summation of maximum expected value for all available actions over all possible data outcomes:

$$\mathrm{PoV} = \sum_{\boldsymbol{d}} \max_a \left\{ \sum_s E[v(s,a)|\boldsymbol{d}] \right\} p(\boldsymbol{d}) \tag{5}$$

The VOI is then compared to the cost of data acquisition $\mathrm{VOI} - \mathrm{cost}_d$ to determine whether it is worthwhile to perform information gathering. When this value is negative, there is no benefit to the proposed information gathering.

The formulation above is for static information gathering. In static information gathering, data collection takes place during a single event in time, after which the decision-maker cannot choose to collect additional information.

Review: Sequential Value of Information

The decision-maker may perform sequential information gathering when they can choose to collect additional information after observing the outcome of the first test (Miller 1975). Here, we review the formulation of value of information for sequential problems. The important distinction between static and sequential information gathering is that for sequential cases, the decision-maker has the option to proceed with a second test after observing the outcome of the first.

To represent distinct epochs during sequential information gathering, we introduce a subscript $\boldsymbol{d}_n$ where $n = 1, 2, \ldots, h$. The time horizon h is the maximum number of data collection epochs. The posterior value of sequential information gathering with time horizon $t = 2$ encodes the decision-maker's option to either stop data collection after the first test, or continue to perform a second test. Here, we use the notation $\mathrm{PoV}(\boldsymbol{d}_1)$ to represent the posterior value for all data outcomes of the first test (Eq. 5). The posterior value for the sequential case, $\mathrm{PoV}(\boldsymbol{d}_2|\boldsymbol{d}_1)$, is the maximum of the value after the first data collection, $\mathrm{PoV}(\boldsymbol{d}_1)$, and the value of continuing with the second test, $\mathrm{Cont}(\boldsymbol{d}_2|\boldsymbol{d}_1)$.

$$\mathrm{PoV}(\boldsymbol{d}_2|\boldsymbol{d}_1) = \max \left\{ \begin{array}{c} \mathrm{PoV}(\boldsymbol{d}_1) \\ \mathrm{Cont}(\boldsymbol{d}_2|\boldsymbol{d}_1) \end{array} \right\} \tag{6}$$

$$\mathrm{Cont}(\boldsymbol{d}_2|\boldsymbol{d}_1) = \sum_{\boldsymbol{d}_1} \left\{ \sum_{\boldsymbol{d}_2} \max_a \left\{ \sum_s E[v(s,a)|\boldsymbol{d}_2, \boldsymbol{d}_1] \right\} p(\boldsymbol{d}_2|\boldsymbol{d}_1) - \mathrm{cost}_{\boldsymbol{d}_2} \right\} p(\boldsymbol{d}_1) \tag{7}$$

VOI is the difference between the prior and posterior value: $\mathrm{VOI} = \mathrm{PV} - \mathrm{PoV}(\boldsymbol{d}_2|\boldsymbol{d}_1)$. Similar to the static case, VOI is compared to the cost of collecting $\boldsymbol{d}_1$ in order to determine whether or not to proceed with collecting $\boldsymbol{d}_1$.

Sequential Value of Information for Cases with Spatial Uncertainty: General Formulation

Data collection and modeling in the subsurface involve spatial uncertainty of the distinction of interest. To account for spatial uncertainty of the subsurface models, the distinction of interest is represented by a set of model realizations, $\boldsymbol{m}$. The superscript $\boldsymbol{m}^{\ell}$ is used to represent a single realization of the subsurface, where $\ell = 1, \ldots, L$, and L is the total number of realizations. A decision-maker may use a summary statistic to simplify their decision-making process. For example, they may only be interested in exploration plans that reduce uncertainty on a volume of minerals, petroleum, or groundwater greater than or equal to a threshold. Here, we use t to represent this threshold.

The decision maker selects an action from three possible alternatives: $\{Walk\ Away, Go\ Ahead, Collect\ Data\}$. The decision-maker will choose to *Go Ahead* when they have determined that the prospective area is suitable for more advanced study (i.e., delineation drilling, pre-feasibility study, development), which represents a significant downstream commitment of time and capital. If the project is deemed unsuitable for immediate development, they will choose to *Walk Away* and incur no cost nor reward. *Collect Data* represent the choice to proceed with carrying out a test of the subsurface without fully committing to development (i.e., drilling one or more exploration boreholes).

Success is the reward (here, in dollars) if the true configuration of the subsurface is equal to or exceeds the volume threshold, whereas Failure (also in dollars) is the outcome if the volume is insufficient. We use an indicator function; thus,

$$I_t(\boldsymbol{m}) = \begin{cases} 1 & \text{if} \quad \text{volume}(\boldsymbol{m}) \geq t \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The indicator function is used to calculate the expected value, prior to and following data collection; thus,

$$\begin{aligned} E[v(I_t(\boldsymbol{m}), Go\ Ahead)] = {} & \text{Success} \times p(I_t(\boldsymbol{m}) = 1) \\ & + \text{Failure} \times p(I_t(\boldsymbol{m}) = 0) \end{aligned} \tag{9}$$

$$\begin{aligned} E[v(I_t(\boldsymbol{m}) | \boldsymbol{d}, Go\ Ahead)] = {} & \text{Success} \times p(I_t(\boldsymbol{m}) | \boldsymbol{d} = 1) \\ & + \text{Failure} \times p(I_t(\boldsymbol{m}) | \boldsymbol{d} \\ & = 0) \end{aligned} \tag{10}$$

The prior and posterior value with sequential information is rewritten with the indicator function as:

$$\text{PV} = \max_a \{E[v(I_t(\boldsymbol{m}), a)]\} \tag{11}$$

$$\text{PoV}(\boldsymbol{d}_1) = \sum_{\boldsymbol{d}_1} \max_a \{E[v(I_t(\boldsymbol{m}), a) | \boldsymbol{d}_1]\} p(\boldsymbol{d}_1) \tag{12}$$

$$\text{PoV}(\boldsymbol{d}_2 | \boldsymbol{d}_1) = \max \left\{ \begin{array}{c} PoV(\boldsymbol{d}_1) \\ Cont(\boldsymbol{d}_2 | \boldsymbol{d}_1) \end{array} \right\} \tag{13}$$

$$\text{Cont}(\boldsymbol{d}_2 | \boldsymbol{d}_1) = \sum_{\boldsymbol{d}_1} \left\{ \sum_{\boldsymbol{d}_2} \max_a \{E[v(I_t(\boldsymbol{m}), a) | \boldsymbol{d}_2, \boldsymbol{d}_1]\} p(\boldsymbol{d}_2 | \boldsymbol{d}_1) - \text{cost}_{\boldsymbol{d}_2} \right\} p(\boldsymbol{d}_1) \tag{14}$$

A decision tree for this formulation is shown in Figure 1.

SYNTHETIC CASE

Here, we used a synthetic case to illustrate an application of value of information for exploration planning. We defined all components of a simplified decision problem for subsurface exploration drilling and provide the reader with a walkthrough tutorial. The following case was constructed in order to compare the value of information when planning to drill one borehole at a time, drilling two boreholes sequentially, and drilling two boreholes simultaneously. Moreover, we used this case to demonstrate the sensitivity of VOI to various parameters such as borehole cost, borehole depth, and volume threshold.

To construct this case and apply value of information, we defined the following: the set of actions that the decision-maker may select from, the prior models of the subsurface, the possible observations, the method used to condition the prior with observations, and the definition of value. Finally, we performed the calculation of VOI and discussed results throughout the experiments in order to clearly demonstrate the method.

Definition of Possible Actions

This decision problem used the decision tree from Figure 1. A decision-maker seeks to determine whether or not they should perform an exploration drilling campaign in order to reduce their uncertainty on a subsurface resource. If so, they would like to know where to drill, how deep, and the added benefit (in dollar value) of the drilling campaign. As in section Sequential Value of Information for Cases with Spatial Uncertainty: General Formulation, the

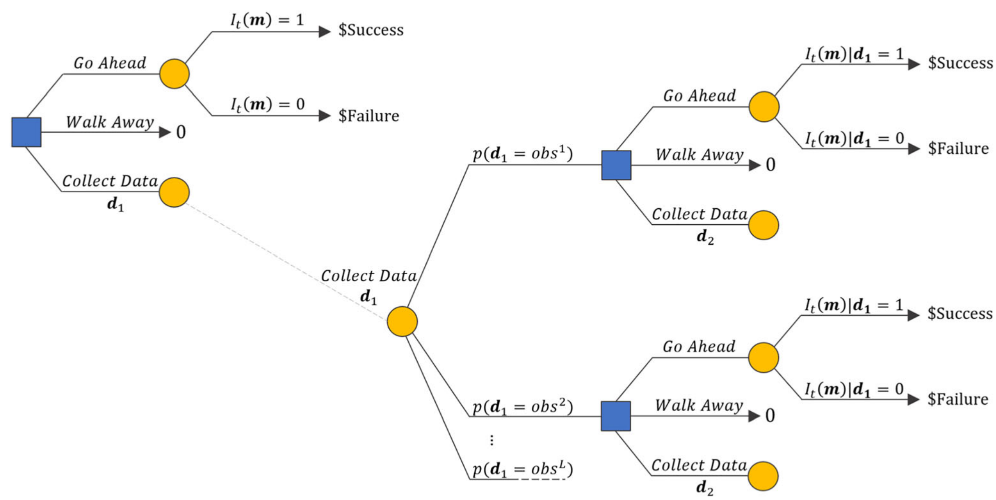


Figure 1. Decision tree for the case of sequentially drilling two exploration boreholes using a volume threshold. The *Collect Data* node is evaluated for every possible observation of $\boldsymbol{d}_1$: $\{obs^1, \ldots, obs^L\}$.

decision-maker may select an action from a set of three alternatives: $\{Walk\ Away,\ Go\ Ahead,\ Collect\ Data\}$.

Prior Modeling of Subsurface

To represent spatial uncertainty of the subsurface, we used sequential indicator simulation, though any ensemble method may be used. We generated $L = 500$ discrete realizations of the subsurface $\{\boldsymbol{m}^1, \ldots, \boldsymbol{m}^{500}\}$; each realization was a two-dimensional 50×30 grid. Each cell in a realization contained a binary value representing economic or non-economic lithology (Fig. 2a). The mean of all subsurface realizations is shown in Figure 2b. The uncertainty on the boundary between the two lithologies is shown in Figure 2c. We assumed no pre-existing drilling information. The prior realizations were generated using sequential indicator simulation.

Observations

For each location x and depth y, unique drilling observations, $\{\boldsymbol{d}^1, \ldots, \boldsymbol{d}^L\}$, were obtained by applying a linear forward model to each subsurface realization, $\boldsymbol{d}^\ell = G\boldsymbol{m}^\ell$. G contained 1's and 0' indicating the presence and absence of a drill location in the gridded model $\boldsymbol{m}$. In this case, all possible boreholes were oriented vertically.

Conditioning With Information

Calculation of VOI requires conditioning the prior realizations to all combinations of borehole data $\boldsymbol{d}^\ell$ (the future information is unknown, but its prior distribution is known). Hence, we needed to calculate L posterior models, each constrained to one realization of the borehole $\boldsymbol{d}^\ell$. Because borehole data were linear data in the sense of a Bayesian inverse problem, we used the Bayes–linear–Gauss equation after performing dimensionality reduction (Scheidt et al. 2018; Caers et al. 2022). A review of this is found in Appendix 1.

Definition of Value of Information

The value of a realization is a function of whether its volume of economic lithology exceeds the chosen threshold, t. We adapted Eqs. 9 and 10 for the synthetic case as follows:

$$
\begin{aligned}
E[v(I_t(\boldsymbol{m}),\ Go\ Ahead)] = p(I_t(\boldsymbol{m}) = 1) \\
\times (\text{Success} - \text{cost}_{\text{total}}) \\
+ p(I_t(\boldsymbol{m}) = 0) \\
\times (\text{Failure} - \text{cost}_{\text{total}})
\end{aligned} \tag{15}
$$

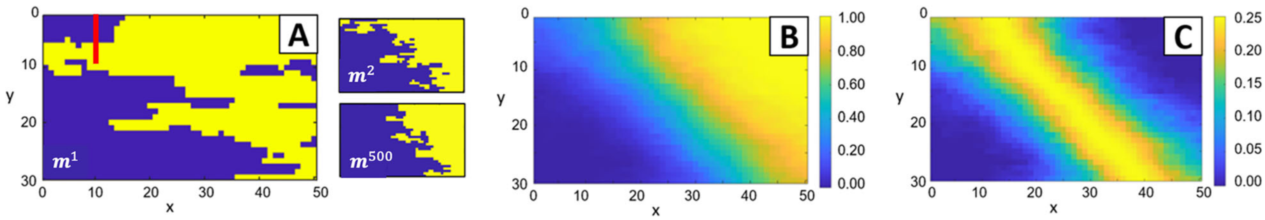


Figure 2. (**a**)Realizations of the subsurface. The top of the figure (y = 0) represents the ground surface. Yellow indicates economic lithology. Blue indicates non-economic lithology. The red line indicates a borehole with a length of ten units. (**b**) Mean of all subsurface realizations, where values closer to one indicate a higher probability of economic lithology. (**c**) Variance of all subsurface realizations.

$$\begin{aligned} E[v(I_t(\boldsymbol{m})|\boldsymbol{d},\ \mathit{Go\ Ahead})] = {} & p(I_t(\boldsymbol{m}) = 1|\boldsymbol{d}) \\ & \times (\text{Success} - \text{cost}_{\text{total}}) \\ & + p(I_t(\boldsymbol{m}) = 0|\boldsymbol{d}) \\ & \times (\text{Failure} - \text{cost}_{\text{total}}) \end{aligned} \quad (16)$$

The volume threshold t, and constants Success, and Failure, (in dollar values) were chosen from the prior distribution of model realizations, which is shown in Figure 3. $\text{cost}_{\text{total}}$ was the cost of extracting the resource from initial development to closure and reclamation, which was incurred whether or not the resource exceeds the volume threshold. We used *Success* and *Failure* to all classify realizations that either exceeded (in the case of Success) or did not exceed (in the case of Failure) the volume threshold.

The decision-maker would like to determine where to drill in order to reduce uncertainty on the resource volumes that exceed the P25 (of the prior). Therefore, we set t to the P25, $t = 673$ (Fig. 3). *Success* was set to t($\mathit{Success} = 673$), which represents the decision-maker's satisfaction for any realization that exceeds the volume threshold. *Failure* was set to the P5 ($\mathit{Failure} = 581$), a low revenue if the volume threshold is not exceeded. *Success* and *Failure* must be chosen as quantiles of the prior, since the decision model depends on Bayes' rule. $\text{cost}_{\text{total}}$ is set to 150.

In Eqs. 17 through 21, we show how VOI is calculated for this case using Monte Carlo simulation. The prior value is calculated as:

$$\text{PV} = \max_a \left\{ \frac{1}{L} \sum_{\boldsymbol{m}^\ell} v(I_t(\boldsymbol{m}), a) \right\} \quad (17)$$

The posterior value is calculated for all possible data outcomes via Monte Carlo simulation for each observation , $\boldsymbol{d}^\ell$; thus,

$$\text{PoV} = \frac{1}{L'} \sum_{\boldsymbol{d}_1^{\ell'}} \max_a \left\{ \frac{1}{L} \sum_{\boldsymbol{m}^\ell} v(I_t(\boldsymbol{m}), a) | \boldsymbol{d}_1^{\ell'} \right\} \quad (18)$$

The value of information is the difference between PoV and PV. (Eq. 1).

Then, we considered a second exploration borehole, $\boldsymbol{d}_2$, for a different location and depth. We created a distinction between the value after the first data collection $\boldsymbol{d}_1$ and the second data collection $\boldsymbol{d}_2$. Equation 18 was modified to represent the value of stopping after the first data collection; thus,

$$\text{PoV}(\boldsymbol{d}_1) = \frac{1}{L'} \sum_{\boldsymbol{d}_1^{\ell'}} \max_a \left\{ \frac{1}{L} \sum_{\boldsymbol{m}^\ell} v(I_t(\boldsymbol{m}), a) | \boldsymbol{d}_1^{\ell'} \right\} \quad (19)$$

The continued value that of drilling the second borehole after observing the results of the first is:

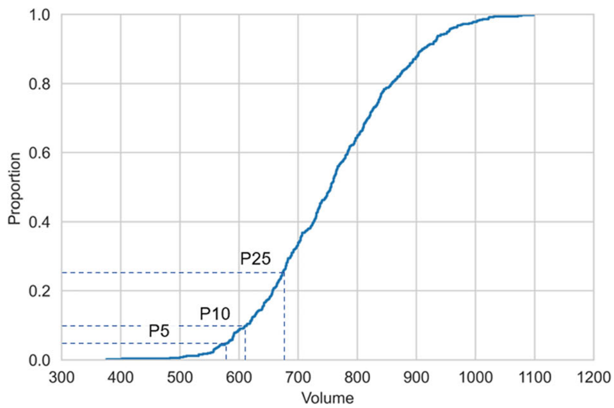
$$\text{ContVal}(\boldsymbol{d}_2|\boldsymbol{d}_1) = \frac{1}{L''} \sum_{\boldsymbol{d}_2^{\ell''}} \left\{ \frac{1}{L'} \sum_{\boldsymbol{d}_1^{\ell'}} \max_a \left\{ \frac{1}{L} \sum_{\boldsymbol{m}^\ell} v(I_t(\boldsymbol{m}), a) | \boldsymbol{d}_1^{\ell'}, \boldsymbol{d}_2^{\ell''} \right\} - \text{cost}_{\boldsymbol{d}_2} \right\} \quad (20)$$

We then use Eqs. 1 and 13 to calculate the value of information. The cost of data collection was subtracted from the VOI to determine the magnitude of the added value. Here, the drilling cost $cost_{\boldsymbol{d}_n}$ was calculated by multiplying a cost per unit depth k and the total depth y of a borehole; thus,

$$\text{cost}_{\boldsymbol{d}_n} = k \times y \quad (21)$$

Results

Here, we carried out the calculation of value of information in the synthetic case to compare decision-making when planning only one borehole, planning a borehole with the option of continuing to



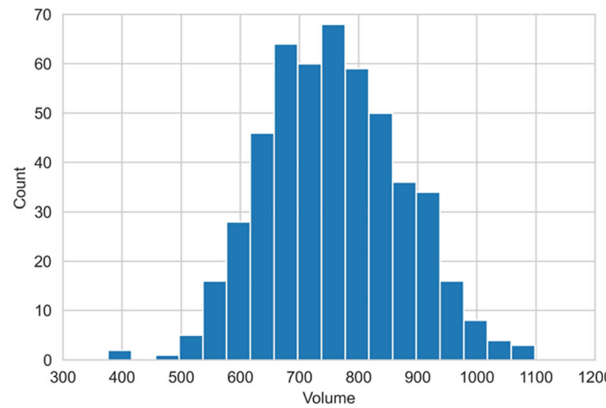


Figure 3. Cumulative distribution (left) and histogram (right) for economic volume of the prior model realizations. The P25 is 673. The P10 is 613. The P5 is 581. The minimum volume in all prior realizations is 376.

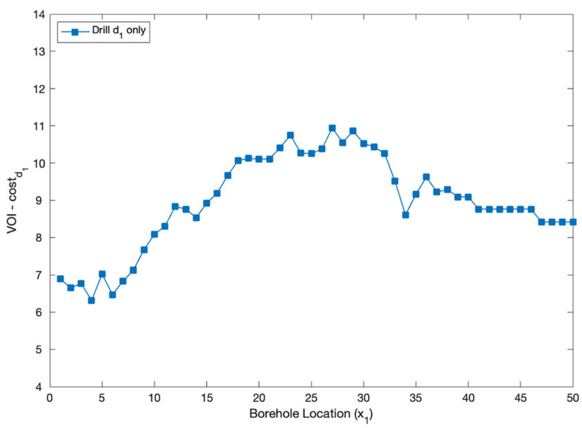


Figure 4. Value of information at each location x_1 for drilling one borehole $\boldsymbol{d}_1$ to depth $y_1 = 6$. Points represent calculated values. Line is interpolation.

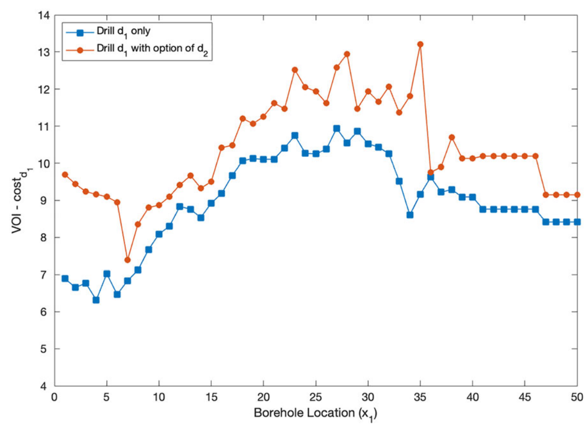


Figure 5. Plot of VOI for static drilling of one borehole only (blue) and sequential drilling of one borehole with the option of continuing to a second (red). The blue line is the value of information minus cost for drilling only the first borehole of fixed depth $y_1 = 6$ at each location. The red line is the value of information minus cost for drilling the first borehole at each location with the option of continuing to the second borehole ($x_2 = 40, y_2 = 20$).

a second, and planning to drill both simultaneously. We fixed the depth of the two proposed boreholes since our goal here is to offer a tutorial explanation of value of information, not to immediately solve for the optimal exploration plan. We then showed the sensitivity of VOI to drilling cost, drilling depth, and volume threshold.

We began by calculating the VOI for planning to drill only one vertical borehole. $\text{VOI} - \text{cost}_{\boldsymbol{d}_n}$ was calculated for each possible drilling location x_1 and is shown in Figure 4. The depth of this first borehole was fixed at $y_1 = 6$. The location x_1 with the maximum VOI ($\text{VOI} = 10.94$) was $x_1 = 27$. The decision-maker always chooses to drill the borehole at the location with the maximum VOI. Therefore, when the decision-maker plans to only drill once, they would choose to drill the borehole at $x_1 = 27$. This location corresponds to the boundary of economic volumes, which exceeded the P25 threshold.

Next, we compared the VOI for drilling only one borehole to drilling one borehole with the option of continuing to a second. We calculated the value of the first borehole at each possible location x_1 if the decision-maker has the option of continuing to a second borehole $\boldsymbol{d}_2$ (Fig. 5). Here, we specified this second borehole with fixed location $x_2 = 40$, fixed depth $y_2 = 20$, oriented vertically.

The location of $\boldsymbol{d}_1$ with the maximum VOI (VOI $= 13.21$) for sequential drilling was $x_1 = 35$. The location of $\boldsymbol{d}_1$ with the maximum VOI for the sequential case was different than the location of $\boldsymbol{d}_1$ with the maximum VOI when drilling only once ($x_1 = 27$). Therefore, if the decision-maker plans on drilling a second borehole at $x_2 = 40$ of depth $y_2 = 20$, they choose to drill the first borehole in a different location than when they plan to drill only once. The position of $\boldsymbol{d}_1$ differed in the sequential case because in this case the VOI for drilling the first borehole with the option of continuing to a second was 21% greater than the VOI for drilling the first borehole alone. Also note that the second borehole may never be drilled, depending on the outcome of the first borehole.

Next, we continued the comparison to a case for drilling two boreholes simultaneously, where the second borehole was drilled without waiting to observe the results from the first borehole. Figure 6 shows a plot of VOI for drilling once, drilling sequentially, and drilling simultaneously.

Figure 6 shows that only in a few circumstances ($x_1 = 3,\ 5,\ 7$ through 12) the VOI for simultaneous drilling exceeded sequential drilling, and overall, sequential drilling was still the maximum VOI where $x_1 = 35$. The difference between VOI for simultaneous and sequential drilling was due to the decision-maker's option of drilling the second borehole

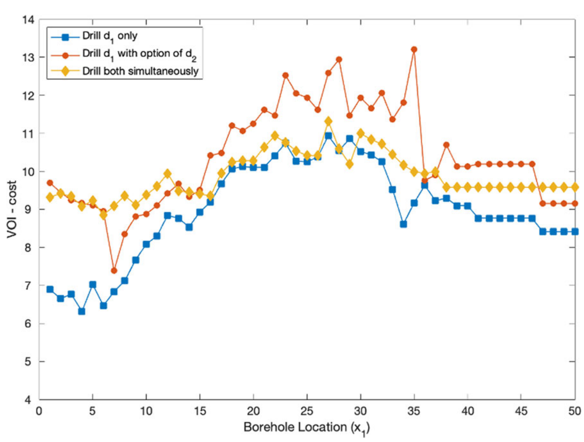


Figure 6. Value of information for static drilling of one borehole only (blue), drilling one borehole with the option of continuing to a second (red), and static drilling of both boreholes simultaneously (yellow). For the simultaneous case, cost is the cost of drilling both the first and second boreholes. For the other cases, cost is the cost of drilling the first borehole.

after observing the first in the sequential case. If the outcome of the first borehole is not suitable, the decision-maker can choose to stop exploring. If the decision-maker chooses to drill simultaneously, they will incur the cost of both boreholes, even if one borehole does not add any information.

Now, we show the sensitivity of VOI to three parameters: cost, drilling depth, and volume threshold. First, we show the sensitivity of value of information to the cost per unit length of drilling k. Figure 7 shows two plots of static and sequential drilling when varying k.

In Figure 7, we observe that the VOI for sequential and simultaneous drilling was lower when using increased drilling costs. When $k = 0.10$, VOI for simultaneous drilling exceeded the VOI for sequential drilling if x_1 was between 8 and 12. When $k = 0.20$, VOI for simultaneous drilling exceeded the VOI for sequential drilling only when $x_1 = 7$ or $x_1 = 8$. This result shows that when drilling cost is high, the decision-maker chooses to drill sequentially and not incur the full cost of drilling both boreholes before observing the outcome of the first.

Next, we show how the optimal drilling placement and depth of a first borehole may change when planning to drill sequentially compared to drilling only once. For drilling only once, we calculated the VOI for all locations and depths of a single borehole. For the sequential case, we calculated the VOI for all locations and depths of the first borehole with the option of continuing to a fixed second borehole ($x_2 = 40,\ y_2 = 20$). In Figure 8, we show two plots of the VOI for all boreholes locations x_1 and borehole depths y_1.

The location and depth of the borehole with the maximum VOI when planning to drill once were $x_1 = 36,\ y_1 = 15,$ VOI $= 11.80$. The location and depth of the borehole with the maximum VOI when planning to drill sequentially were $x_1 = 28,\ y_1 = 7,$ VOI $= 12.59$. When planning to drill sequentially, the decision-maker will place the first borehole at a different location and a different depth than if they had planned to drill only once.

A decision-maker may choose to explore for different volumes of a resource, given their risk preference. Therefore, next we study the sensitivity of VOI when changing the volume threshold. We reformulated the scenario to represent an explorer's preference to reduce uncertainty on the lowest ten percent of economic volumes (P10, Fig. 3). Success was set to the P10 (Success $= 613$). Failure was set

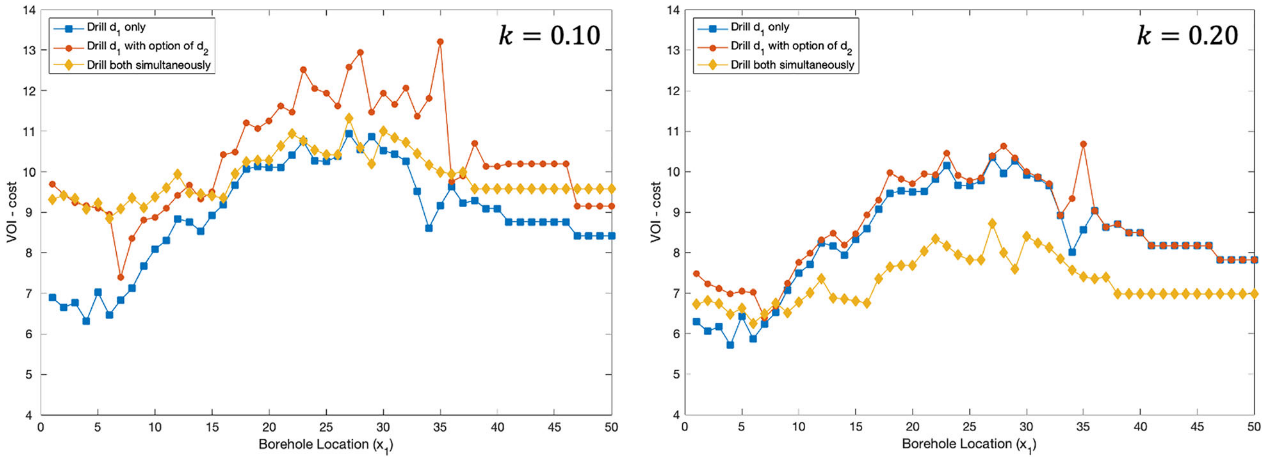


Figure 7. Two plots of VOI – cost for drilling once, drilling sequentially, and drilling two boreholes simultaneously for a cost per unit length $k = 0.10$ (Left) and $k = 0.20$ (Right).

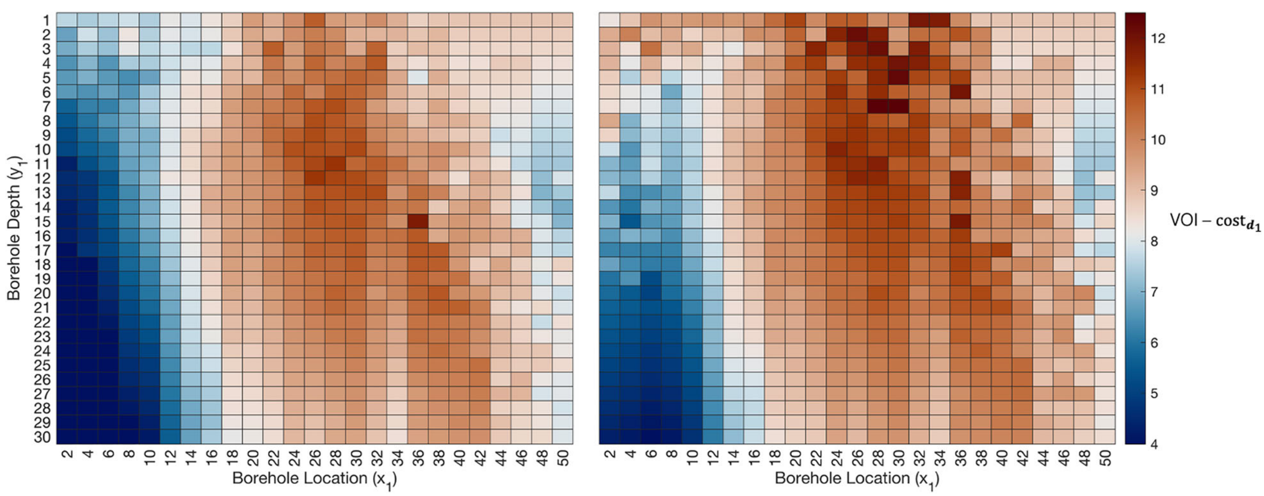


Figure 8. (Left) A heatmap of VOI – cost for all locations and depths (x_1, y_1) when planning to drill only one borehole. (Right) A heatmap of VOI-cost for all locations and depths (x_1, y_1) when planning to drill the first borehole with the option of continuing to a second ($x_2 = 40$, $y_2 = 20$).

to the minimum economic value of all realizations (Failure = 376). cost$_{total}$ was set to 150. Figure 9 shows the VOI for all borehole locations and depths for a single borehole.

The borehole location and depth for this scenario with the maximum VOI were $x_1 = 28$, $y_1 = 2$, VOI = 7.26. This differs from the borehole location and depth when drilling once for volumes greater than the P25 (Fig. 8, left). The difference in depth and placement of boreholes for the two different volume thresholds was due to the shallow depth of the smaller economic volume beneath the surface.

REAL CASE

In this section, we used real data from a well-defined ore deposit to determine the optimal se-

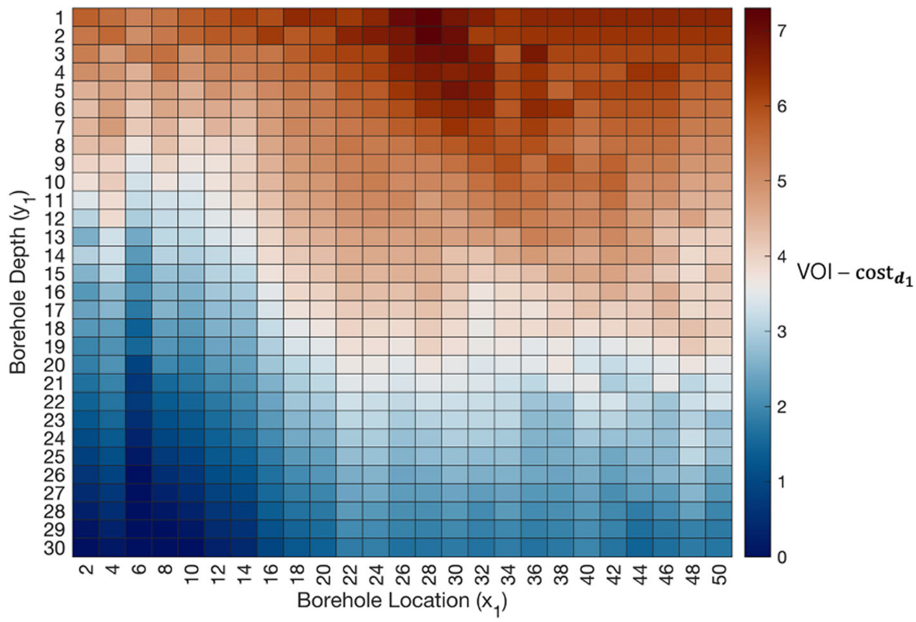


Figure 9. Heatmap of VOI for volume threshold less than the P10 (613), varying borehole locations of $\boldsymbol{d}_1$.

quence of initial exploration drilling. We emulated a real exploration campaign by generating a prior model using airborne geophysics, without using any of the drilling information contained in the database. A subset of boreholes was selected from the drilling database, which was treated as the sets of possible candidates for drilling. We calculated VOI for all single boreholes and possible pairs of two boreholes out of the subset in order to compare and rank the possible exploration plans.

The dataset was from a real metal deposit and consisted of 52 boreholes and comprehensive geochemical assay information as well as airborne geophysics. For proprietary reasons, the deposit and its location cannot be described. A subset of six boreholes was selected from the dataset, see Figure 10. The value of information was calculated for each pair out of the six boreholes as a single value, quantifying the additional value of exploration (in dollars). Calculating VOI requires generating a prior model (section Prior Modeling), a method of updating the prior model with drilling information (section Posterior Modeling), and calculating the value of information based on the prior and posterior realizations, the dollar value of one cell of the massive sulfide unit, and drilling cost (section Value Calculation).

Prior Modeling

The prior model consisted of 517 realizations $\{\boldsymbol{m}^1, \ldots, \boldsymbol{m}^{517}\}$, on a three dimensional $104 \times 120 \times 64$ grid generated using magnetic field data acquired via airborne survey (Caers et al. 2022). Multiple inversions of the data were generated using a sparse norm inversion, which were transformed into realizations that indicate either the presence or absence of the economic lithology. Figure 11 is used to ascertain the distribution of prior model volumes. The maximum volume was approximately 1500 times greater than the minimum volume. Uncertainty in geophysical measurements was high relative to the accuracy of drilling information, which made initial placement of boreholes a challenging and important task.

Posterior Modeling

Here, we explain data collection and conditioning the prior with observations. Possible observations for each of the six boreholes vary based on the position of the borehole within the block model. An observation, $\boldsymbol{d}^\ell$, is a vector of binary values corresponding to whether each cell in the borehole

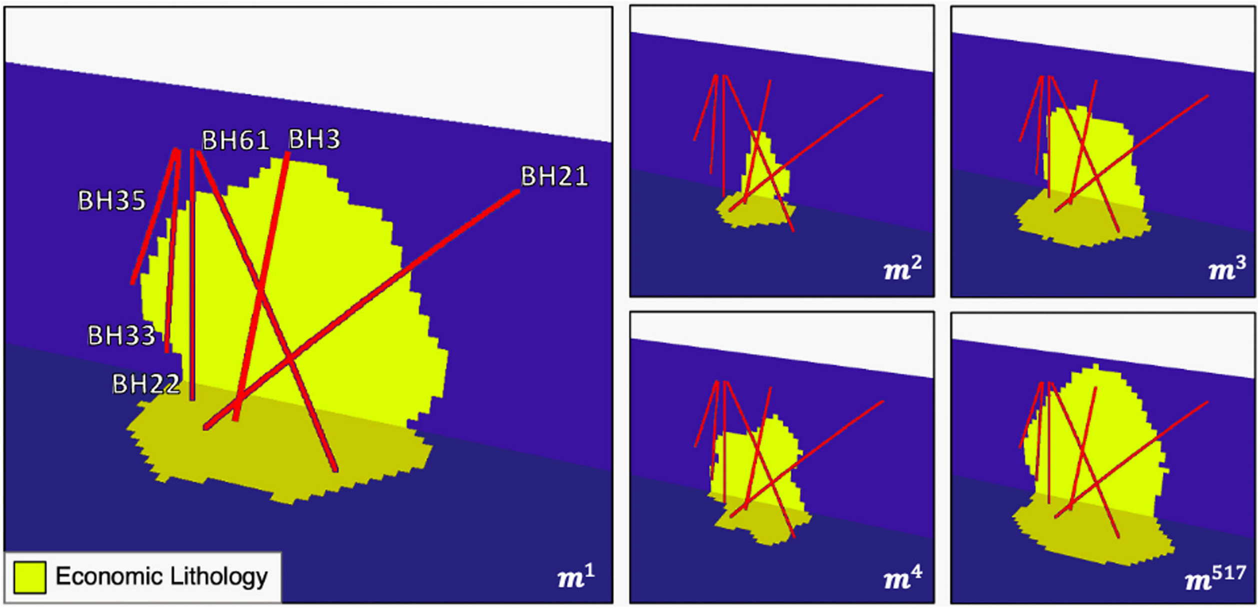


Figure 10. Five realizations of the prior model are shown in cross section, with the subset of six boreholes selected from the original database.

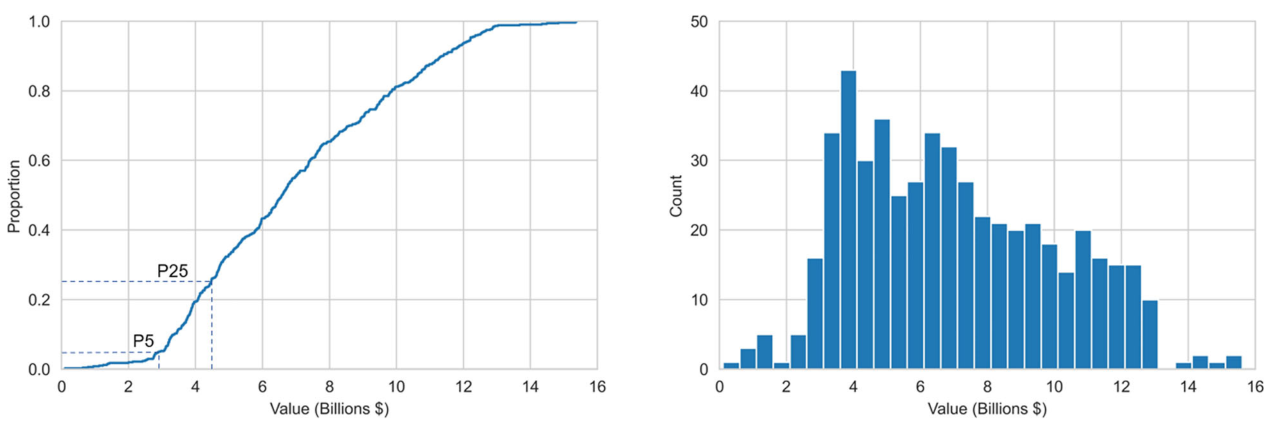


Figure 11. Cumulative distribution (left) and histogram (right) of the economic value of each realization in the prior model, shown in billions of dollars. Success and Failure were determined by selecting the P25 and P5 quantiles of this distribution. The P25 was \$4.46 billion. The P5 was \$2.80 billion.

intercepts the economic lithology. Posterior realizations $p(\boldsymbol{m}|\boldsymbol{d}^{\ell})$ were generated via Monte Carlo simulation using an ensemble smoother method described in Appendix 2, though any method to condition the prior with information could be used.

Value Calculation

The output of the value of information calculation is a dollar value which describes the additional benefit of carrying out the exploration plan. We demonstrated an application of the VOI formulation for volume thresholds described in section Defini-

Table 1. Drilling depth, cost, and percent true deposit intersected for each borehole

Borehole ID	Depth of borehole (feet)	Cost of Borehole cost_d (US Dollars)	Total Intersection of True Deposit(%)
BH21	1204.3	30,108	35
BH22	935.2	23,378	32
BH33	529.3	13,233	64
BH35	360.6	9,015	30
BH61	930.4	23,258	40
BH3	763.8	19,094	96

Table 2. Value of information for drilling each borehole individually. All values except BH3 are presented in millions of dollars, and are the calculated VOI – cost. BH3 is presented in dollars, and is negative since its VOI is zero

Borehole Number	VOI – cost_{d_i}
BH21	80.8
BH22	105.2
BH33	102.2
BH35	93.3
BH61	137.8
BH3	− 19,094

tion of Value of Information (Eqs. 15 – 20) and calculated the value of exploration plans targeting volumes greater than or equal to the 25th percentile (P25). We defined new values for Success, Failure, $\text{cost}_{\text{total}}$, and cost_{d_n}, specific to this real case. Note that these values were intended as approximations, and do not represent a rigorous economic model. For example, the cost of drilling could be determined by calculating cost for assays, drill mobilization, drilling, downhole survey and consumables, but we used a summary of this information as a total dollar cost per foot. Success and Failure were determined by selecting quantiles of the distribution of economic value (in dollars) for all prior realizations (Fig. 11), in the same manner as described in section Definition of Value of Information

The economic value for a realization (Eq. 20) was calculated using the volume (in tons) of economic lithology in the realization, multiplied by the net smelter return (NSR) of one ton of economic lithology. This calculation was an additional step, different than the tutorial case, because now we must incorporate the real economic value and density. For this deposit, the net smelter return, r_{NSR}, was \$142/ton. The volume (in tons) of a cell of economic lithology was calculated by multiplying the mean density of the economic lithology ($\rho_{\text{econ}} = 3.34\ \text{t/m}^3$) by the cell volume ($\text{vol}_{\text{cell}} = 1000\ \text{m}^3$) and the total count of all cells that belong to the economic lithology in the realization.

$$\text{Economic value of } \boldsymbol{m}^{\ell} = r_{\text{NSR}} \times \rho_{\text{econ}} \times \text{vol}_{\text{cell}} \times \sum_{n} \boldsymbol{m}^{\ell}(\text{lith}_n) \tag{22}$$

Note that in Eq. 22, we used $\boldsymbol{m}(\text{lith}_n)$ to determine if the lithology lith at cell n was economic or not; thus,

$$m^{\ell}(\text{lith}_n) = \begin{cases} 1 & \text{if } \ \text{lith}_n \text{ is economic} \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

For this case, we defined Success as the P25 of the distribution of economic value for all prior model realizations, Success = \$4.46 bn. Failure was defined as the P5 of the same distribution, Failure = \$2.8 bn. The volume threshold, t, was the P25 of the distribution of volume for all prior model realizations.

The cost of development, $\text{cost}_{\text{total}}$, was the total amount of money projected to be spent on extracting the resource, which included all costs from permitting, development, operation, through closure and reclamation. Here, we assigned $\text{cost}_{\text{total}} = \$2$ bn.

The cost of drilling a borehole, cost_d, was calculated by multiplying the total depth of the borehole (in feet) by a drilling cost per foot. We specified the drilling cost per foot $k = \$25/\text{ft}$ (\$80.45/m), though any cost may be used.

$$\text{cost}_d = \text{depth}_d \times k \tag{24}$$

The cost of drilling each borehole is shown in Table 1. Value of information was calculated for the real case using Eqs. 13, 15–20.

Table 3. Value of information minus drilling cost for drilling each borehole with the option of continuing to a second. All values are presented in millions of dollars, and are the calculated VOI minus the cost of $\boldsymbol{d}_1$

VOI Sequential Drilling	BH21	BH22	BH33	BH35	BH61	BH3
w/ option: BH21		119.5	109.2	108.2	156.8	80.8
w/ option: BH22	120.8		118.7	110.9	145.5	105.2
w/ option: BH33	123.7	112.4		110.7	145.6	102.2
w/ option: BH35	106.1	110.7	110.1		143.1	93.3
w/ option: BH61	115.9	164.8	183.5	137.6		137.8
w/ option: BH3	80.8	105.6	102.2	93.3	152.4	

185.0 VOI (Millions $) 80.0

Table 4. Value of of information minus drilling cost for drilling a pair of boreholes simultaneously. Values in bold are less than their corresponding sequential value of information (Table 3). Italicized values are less than the corresponding value of information for drilling only one of either borehole (Table 2). Values that appear equal in Tables 3 and 4 (e.g., column BH3) are in fact lower in the simultaneous case due to the greater cost of drilling both boreholes at once. For such cases, the impact of greater cost for simultaneous drilling is obscured by rounding. All values are presented in millions of dollars, and are VOI minus the cost of both boreholes

VOI Simultaneous Drilling	BH21	BH22	BH33	BH35	BH61	BH3
BH21		***79.6***	**87.5**	***70.9***	**90.3**	**80.8**
BH22	***79.6***		**102.7**	**93.7**	**145.5**	**105.2**
BH33	**87.5**	**102.7**		***85.7***	**111.0**	**102.2**
BH35	***70.9***	**93.7**	***85.7***		178.3	**93.3**
BH61	**90.3**	**145.5**	**111.0**	178.3		**137.8**
BH3	**80.8**	**105.2**	**102.2**	**93.3**	**137.8**	

180.0 VOI (Millions $) 70.0

Borehole Descriptions

Here, we offer more detailed description of the boreholes to complement Figure 10. The percentage of each borehole that intersected the true orebody is shown in Table 1. Borehole 21 pierced the orebody diagonally, and directly into the middle of the deposit, in an opposite direction to Borehole 61. Borehole 61 also pierced the orebody diagonally, but was aligned tangentially to the outer edges of some prior models. Borehole 3 was located at the middle of the deposit. Boreholes 22, 33, and 35 were positioned on the outer edge of the deposit.

Results

Our aim was to compare the value of information for drilling only one borehole, drilling two boreholes sequentially, and drilling two boreholes simultaneously. By performing this comparison, we determined the ranking of exploration plans and show how the explorer may choose to drill differently if sequential or simultaneous drilling is possible. Following this comparison, we performed a more detailed investigation of sequential drilling and show whether or not the decision-maker should proceed to collect data based on each outcome of the first borehole, and the added benefit of continuing to drill. The value of information for drilling only one borehole is shown in Table 2. The value of information for drilling one borehole with the option of continuing to a second (sequential drilling) is shown in Table 3. All results are displayed in millions of dollars.

In Table 2, we observe that the borehole with the highest VOI was BH61 and the borehole with the lowest VOI was BH3. All boreholes except BH3 had a value of information greater than zero. In Table 3, we observe that the exploration plan with the maximum value of information was to drill BH33 followed by BH61 (VOI = \$183.5 million.

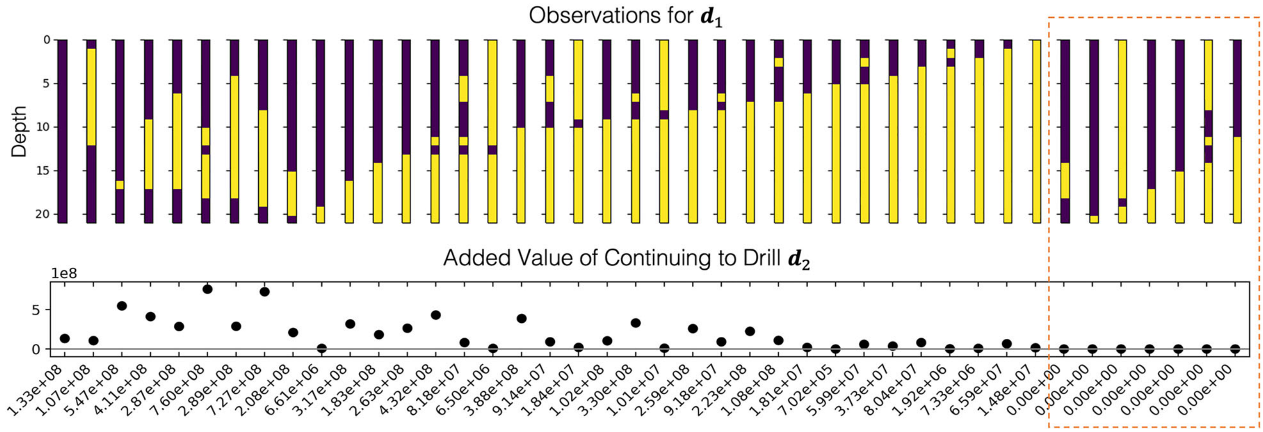


Figure 12. Added value (in dollars) of continuing to drill d_2 = BH61 for each possible unique outcome of d_1 = BH33. Yellow represents intersections of economic lithology. Purple represents intersections of non-economic lithology. The decision-maker will not continue to drill BH61 for any observations of BH33 with zero added value, indicated by the orange box. There are 20 instances of these unique outcomes out of all 517 possible outcomes.

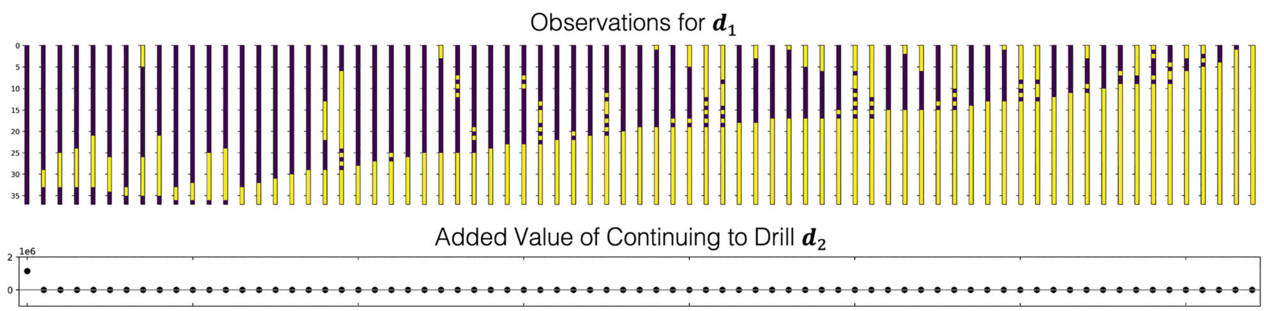


Figure 13. Added value of continuing to drill d_2 = BH3 for each possible unique outcome of d_1 = BH21. There are 74 possible unique outcomes of BH21. The added value of BH3 is greater than zero ($1.15 million) for only one outcome of BH21. There are two instances of this unique outcome out of all 517 possible outcomes.

The two sequential exploration plans with the highest VOI did not involve drilling BH61 first. The VOI for all sequential plans was greater than or equal to the value of information for drilling only the first borehole. Drilling BH61 with the option of continuing to drill BH3 was the fourth highest VOI for sequential drilling. In Table 4, we observe that the exploration plan with the maximum value of information was to drill BH35 and BH61 simultaneously (VOI = $178.3 million). For all pairs of boreholes except BH35 and BH61, the VOI for simultaneous drilling was less than that of the VOI for drilling the same boreholes sequentially. In three cases, the VOI for simultaneously drilling two boreholes was less than that of drilling either borehole individually (BH21 and BH22, BH21 and BH35, BH33 and BH35).

Figures 12 and 13 show the added value of continuing to drill a second borehole for each possible outcome of a first borehole. We show an example for the sequence of boreholes with the greatest value of information (BH33 then BH61, Figure 12), and the least (BH21 then BH3, Figure 13). In Figure 12, the added value of continuing to drill d_2 (BH61) was greater than zero for 35 unique outcomes of d_1 (BH33). The probability of an outcome of d_1 belonging to this set of 35 unique outcomes was 0.961. For the remaining seven unique outcomes of d_1, the added value of continuing to drill d_2 was zero. The outcome of d_1 with the maximum added value of d_2 was $760 million. In Figure 13, the added value of continuing to drill d_2 (BH3) was greater than zero for one unique outcome of d_1 (BH21) with probability 3.89×10^{-3}. The added value of continuing to drill $\boldsymbol{d}_2$ for this observation was \$1.15 million.

DISCUSSION

The results from the real case show that the decision-maker will change their selection of a first exploration borehole if they plan to drill sequentially. In the real case, if the decision-maker plans to drill only once, they would choose to drill the borehole with the highest VOI: BH61. The location of BH61 is not directly in the center of the deposit (BH3) or far to the exterior (BH35). However, if the decision-maker plans to drill two boreholes sequentially, they would choose to drill BH33 followed by BH61 (at an increased value of 75%). BH33 is located more toward the exterior of the deposit than BH61 and intersects less of the economic lithology. This demonstrates that choosing to explore in a ‘‘greedy’’ manner, that is, choosing the borehole with the immediate maximum VOI, is suboptimal when compared to planning exploration drilling sequentially. Furthermore, in this case, planning to drill sequentially also reduces exploration cost. BH33's depth is 529.3 feet compared to BH61's depth of 930.4 feet (Table 1)—BH33 will be completed in less time and at lower cost. If BH61 were not part of the original set, the decision-maker would choose to drill BH22 when planning to drill once (VOI = \$105.2 million), but would not choose to drill BH22 if they plan to drill sequentially (BH21 then BH33, VOI = \$123.7 million).

Planning for sequential exploration also leads the decision-maker to select boreholes that they otherwise might never consider. BH3 has negative VOI when drilled alone, but in the sequence where BH61 is drilled with the option of continuing to BH3, it has the fifth-highest VOI of all fifty possible exploration plans. Drilling BH3 alone has negative value of information because it does not reduce uncertainty on the chosen volume threshold. Furthermore, there is no value improvement in drilling BH3 first followed by any borehole, compared to the second borehole's VOI when drilled alone. This is due to the central location of BH3, since drilling at the center first does not reduce uncertainty on the prior model or volume. The decision-maker may just as drill the second borehole alone, since BH3 offers no additional information.

In all cases except for one, sequential drilling has a higher value of information than simultaneous drilling. This demonstrates two unique features encoded in calculating sequential value of information: waiting to observe the outcome of the first borehole before committing to drill the second and walking away from further exploration if the outcome is unsuitable (volume threshold is not satisfied, cost of continued drilling is not justified).

When exploring to reduce uncertainty on a volume, it does not make sense to drill into the center of the deposit immediately, but rather at its edges where variance is highest. Exploration plans with BH3 first have low VOI since BH3 is located toward center of the deposit. The three exploration plans with the highest VOI all involve drilling a borehole which is located on the periphery of the economic lithology. Once the first borehole is drilled at the edge, then a borehole toward the interior provides more valuable information, since the prior has been updated with the first observation.

In one case, simultaneous drilling exceeds sequential drilling. Drilling BH35 and BH61 simultaneously yields the second highest VOI and exceeds the corresponding sequential VOI of the same boreholes. As demonstrated in the synthetic case, the VOI for a simultaneous exploration plan may exceed the VOI of a sequential exploration plan when drilling cost is low (Fig. 7). The cost of drilling BH35 with BH61 is the fifth lowest cost of any pair of boreholes, which likely contributes to the high VOI. BH35 is also located on the periphery of the deposit, rather than in its center (Fig. 10).

When comparing the VOI for the optimal static drilling plan (BH61 VOI = 137.8) to the VOI for the optimal sequential plan (BH33 with the option of BH61 VOI = 183.5), there is a difference of \$45.7 million dollars. Explaining the added benefit of sequential drilling in a dollar value may lead to more impactful communication and planning with non-technical management and staff.

In a real setting, Table 2 is used to quickly determine the optimal alternative if a drillhole is no longer available due to operational constraints. This allows for decision-making that is consistent and can be quickly audited. Similarly, Figures 12 and 13 show a more detailed analysis which can be used to determine the value and optimal decision as drilling proceeds. In Figure 12, the added value of drilling BH61 is highest for observations of BH33 which intercept relatively less economic lithology, since additional drilling indicates a larger deposit than drilling BH33 alone. In a real setting, a driller or geologist can reference this quantitative information in order to support their decision of whether to continue drilling a second borehole. In Figure 13, the outcome of BH21 does not contain any economic lithology, and given the central location of BH3, BH3 reduces uncertainty on models that BH21 missed.

Due to the double-loop Monte Carlo simulation, the calculation of sequential value of information (Eq. 20) is very computationally demanding. Calculation of sequential value of information for BH21 and BH22 took approximately 35 h on an Intel Core i7-8750H (2.20 GHz) with 32 GB RAM. We selected BH21 and BH22 since they are the pair of boreholes with the highest number of total observations. This computational problem will become more severe when considering more complex geological models and observations, as well as increasing the time horizon of an exploration plan.

The Net Present Value (NPV) is another metric used in resource evaluation, and has been used to measure the added benefit of infill drilling in mining (Froyland et al. 2018), as well as petroleum (Onwunalu and Durlofsky, 2010). We chose to use value of information as it directly encodes the decision problem of whether or not to collect information. Net present value has been used in conjunction with value of information in Dutta et al. (2019).

Another application of sequential value of information with this dataset is to incorporate all existing borehole information from the database and determine the optimal sequence of additional infill drilling, including optimization for position and depth. We have not addressed this topic in this paper, and focus instead on the selection of the initial exploration boreholes chosen from a set of candidates.

CONCLUSIONS

Optimal borehole placement and the importance of sequential planning are demonstrated by focusing on the sequential nature of subsurface exploration and applying VOI. The optimal placement of boreholes changes depending on whether the explorer can consider drilling a second borehole after observing the outcome of the first. In some cases, boreholes that would not be drilled alone are highly valuable when drilled in combination with others. Using VOI, an explorer can frame the decision problem in terms of a summary statistic in order to show optimal exploration plans that reduce uncertainty on chosen thresholds, such as volume. Options for sequential exploration drilling are directly encoded in value of information and monetary values are used, which enables the incorporation of cost of data collection and clear communication of prospective value with management. We incorporate spatially varying prior models using Monte Carlo simulation and update with information sampled from these models. Thus, we do not rely on an expert's potentially subjective assessment of correlation between boreholes. The method is computationally intensive, and future work will focus on approximate techniques in order to address more complex decision situations.

FUNDING

We thank the TomKat Center for Sustainable Energy and the affiliates of the Stanford Center for Earth Resources Forecasting for funding and supporting this research.

APPENDIX 1

CONDITIONING THE PRIOR MODEL WITH DATA

The prior model consists of a set of L discrete realizations ($\boldsymbol{m}$), drawn from the same geological scenario. These prior realizations are unconditional and need to be conditioned to drilling observations (hard data). Scheidt et al. (2018) showed that if the conditioning problem can be expressed a linear inverse problem (Eq. 1), then there exist techniques such as the Bayes-linear-Gauss equation that can efficiently solve the conditioning problem. We start with a forward model of the data:

$$\boldsymbol{d} = G\boldsymbol{m} \text{ where } \boldsymbol{m} \text{ is the disctete model and } G \text{ is a linear operator} \tag{25}$$

The linear operator G is a matrix containing as many columns as grid cells in the model. The number of rows in G is the number of hard data points

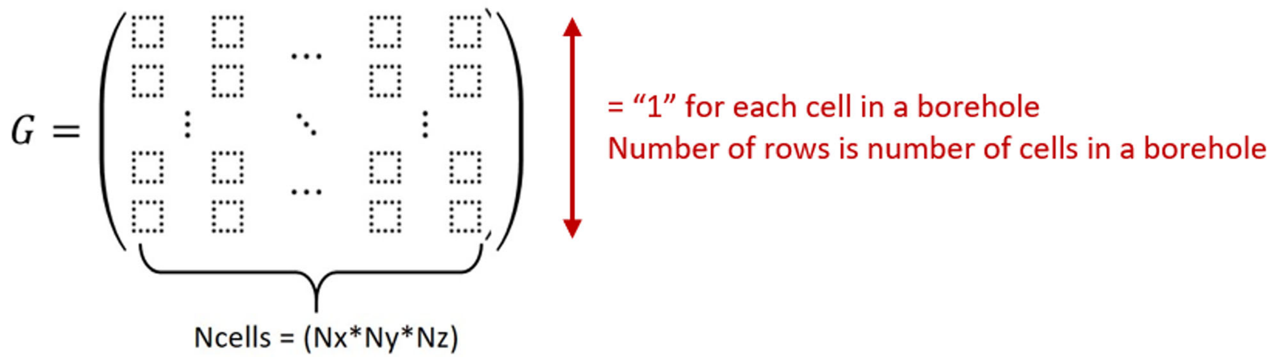


Figure 14. Linear operator G.

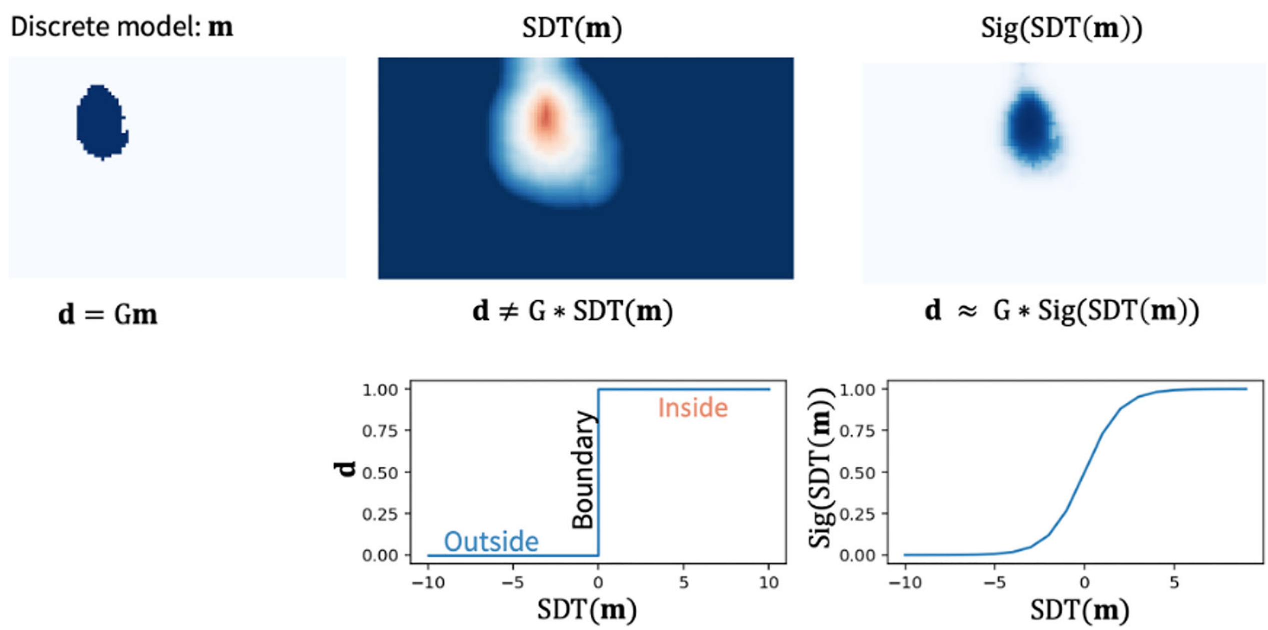


Figure 15. (top) Example of a prior model (left) and its transformations: $SDT(\boldsymbol{m})$ (middle) and $Sig(SDT(\boldsymbol{m}))$ (right) with their relation to the data and the relation between $SDT(\boldsymbol{m})$ and $\boldsymbol{d}$ and $SDT(\boldsymbol{m})$ and $Sig(SDT(\boldsymbol{m}))$**,** respectively.

(Fig. 14). The values in each cell of G are whether the cell contains hard data—not the values of the hard data itself.

The high-dimensionality of the model $\boldsymbol{m}$ necessitates the use of a dimensionality reduction technique. Any technique can be used in this approach, such as principal component analysis (PCA), or discrete cosine transform (DCT). These dimensionality reduction techniques express the model $\boldsymbol{m}$ as:

$$\boldsymbol{m} = \varphi\alpha$$

where φ is a matrix containing the set of basis functions and α, a matrix of coefficients.

The linear relationship from Eq. 25 becomes:

$$\boldsymbol{d} = G_2\alpha \text{ with } G_2 = G_\varphi$$

Assuming that α are normally distributed, Bayes-linear-Gauss equation can be applied to do the conditioning. If $\alpha \sim N(\mu, \Sigma)$, then $\alpha_{\text{post}} = \alpha|\boldsymbol{d} \sim N(\overline{\mu}, \Sigma)$ with:

- $\overline{\mu} = \mu + \Sigma G_2'\left(G_2\Sigma G_2'\right)^{-1}(\boldsymbol{d}_{obs} - G_2\mu)$, with $\boldsymbol{d}_{obs}$ the observed drilling data
- $\overline{\Sigma} = \Sigma - \Sigma G_2'(G_2\Sigma G_2')^{-1}G_2\Sigma$

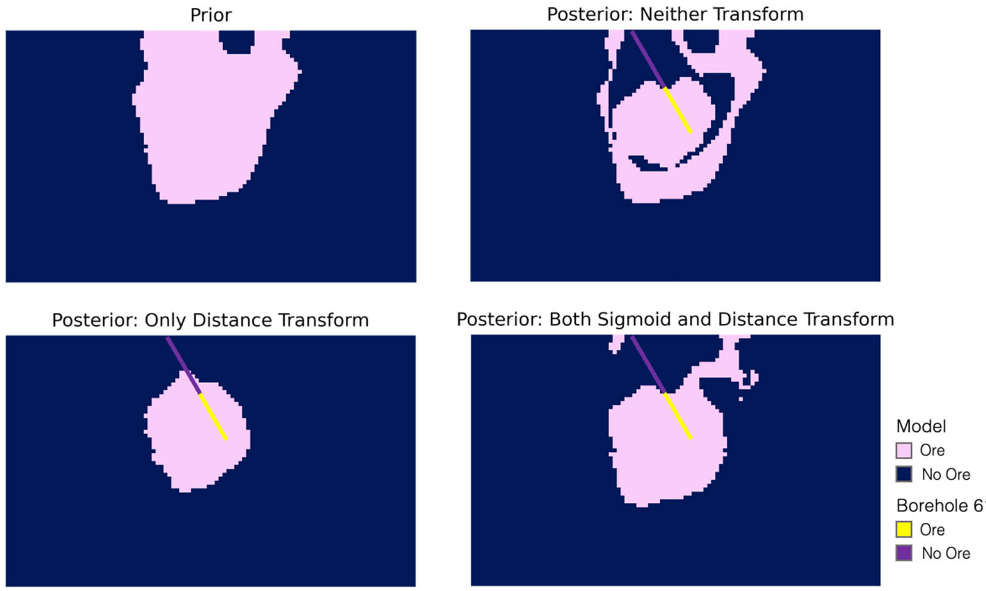


Figure 16. A comparison of one realization of the prior model (top left) with a realization of the posterior model without using any transform (top right), only using the signed distance transform (bottom left), and using both the sigmoid and signed distance transforms (bottom right). The prior model was conditioned on one outcome of Borehole 61 for this demonstration.

A set of posterior coefficients α_{post} are then sampled using the above multi-variate normal distribution ($\bar{\mu}$, $\bar{\Sigma}$). ‘‘Pre-posterior’’ realizations are then obtained by mapping back the posterior coefficients into the original space (by inverting the linear dimensionality reduction technique, PCA, DCT, etc.). Next, a threshold is applied to the pre-posterior realizations to generate the final, discrete posterior models:

$$\boldsymbol{m}_{\text{post}} = \boldsymbol{m}_{\text{prepost}} > 0.5$$

This suite of discrete posterior models is then used as an input to the value of information calculation.

APPENDIX 2

CONDITIONING THE PRIOR MODEL WITH DATA WITH ENSEMBLE SMOOTHER

Ensemble smoother (van Leeuwen and Evensen, 1996; Emerick and Reynolds, 2013) is a data assimilation technique that uses an ensemble of models to perform model updating. In the case of linear forward modeling ($\boldsymbol{d} = G\boldsymbol{m}$) and under the Gaussian assumption: $\boldsymbol{m} \sim \mathrm{N}(\boldsymbol{\mu}_m, \mathrm{C_m})$, the ensemble smoother updates each model of the prior ensemble using the following equation:

$$\boldsymbol{m}^{\ell}_{\text{pos}} = \boldsymbol{m}^{\ell} + C_{\text{md}} * C^{-1}_{\text{dd}} \left(\boldsymbol{d}_{\text{obs}} - \boldsymbol{d}^{\ell}\right)$$

where $\mathrm{C_{md}}$ represents the cross-covariance matrix between $\boldsymbol{m}$ and $\boldsymbol{d}$, and C_{dd} the covariance matrix of $\boldsymbol{d}$. Here, $\boldsymbol{d}_{obs}$ is the observed data, and $\boldsymbol{d}$ is sampled from the prior model $\boldsymbol{m}$.

In the case of discrete models, additional prior transformations are needed to better meet the Gaussian assumption. One such idea is to apply a signed distance transform (SDT, Osher and Fedkiw, 2002; Hakim-Elahi and Jafarpour, 2017). The SDT is defined at grid cell each location by the signed distance to the boundary of the closest object $d(x)$:

$$SDT(x) = \begin{cases} -d(x) & \text{if } x \text{ is in the object} \\ d(x) & \text{if } x \text{ is outside the object} \\ 0 & \text{if } x \text{ is at the boundary} \end{cases}$$

Because the SDT transformation is not linear, the linear forward model no longer applies ($\boldsymbol{d} \neq G * SDT(\boldsymbol{m})$), hence we propose an additional sigmoid transformation making the relationship be-

Table 5. VOI and ranking for all single borehole candidates when performing both sigmoid and signed distance transformations, neither transformation, and the signed distance transform only

Borehole	Fully Transformed		Neither Transform		Signed Distance Transform Only	
	VOI	Rank	VOI	Rank	VOI	Rank
BH21	80.8	5	113.8	5	0	–
BH22	105.2	2	160.1	3	0	–
BH33	102.2	3	133.9	4	0	–
BH35	93.3	4	165.9	2	0	–
BH61	137.8	1	196.5	1	0	–
BH3	0	6	0	6	0	–

tween data and transformed models $Sig(SDT(\boldsymbol{m}))$ closer to linear (see Fig. 15):

$$Sig(x) = \frac{1}{1 + \exp(-x)}$$

The ensemble smoother is then applied on the non-linear transformation $Sig(SDT(\boldsymbol{m}))$ Because these transformations are bijective, they can be undone to obtain the discrete ore body model.

Here, we demonstrate the necessity of the signed distance transform and sigmoid function for preparing the prior models before input to the ensemble smoother. In Figure 16, we show posterior models when the transforms are not performed to the posterior model using fully transformed data. When neither transform is performed, the posterior model closely matches the borehole data, but the posterior model does not appear realistic due to the outer ''shell'' of ore lithology which isn't fully updated from the prior model. If only the signed distance transform is performed, the posterior model does not match the borehole data as closely, and appears over smoothed. When using both the sigmoid and distance transforms, the borehole data is closely matched and the posterior model maintains features from the prior model without over smoothing.

Table 5 shows the VOI of single boreholes when performing both the sigmoid and distance transformations, neither transformation, and the signed distance transform only. When performing neither transform, VOI is higher than the fully transformed case for all boreholes except BH3, where the VOI is equal in both cases. As shown in Figure 16, this is due to the improper updating of the prior model which leads to a much greater estimate of cells containing ore. Nevertheless, the ranking of boreholes BH61, BH21, and BH3 is preserved between both cases. When performing only the signed distance transform, the VOI for all boreholes is zero. This low VOI is due to the over smoothing shown in Figure 16.

REFERENCES

Ali, S. H., Giurco, D., Arndt, N., Nickless, E., Brown, G., Demetriades, A., et al. (2017). Mineral supply for sustainable development requires resource governance. *Nature. Nature Publishing Group*. https://doi.org/10.1038/nature21359.

Barnes, R. J. (1989). *Sample Design for Geologic Site Characterization* (pp. 809–822). Dordrecht: Springer. https://doi.org/10.1007/978-94-015-6844-9_64.

Bhattacharjya, D., Eidsvik, J., & Mukerji, T. (2010). The value of information in spatial decision making. *Mathematical Geosciences, 42*(2), 141–163.

Bickel, J. E., & Smith, J. E. (2006). Optimal sequential exploration: A binary learning model. *Decision Analysis, 3*, 16–32.

Bickel, J. E., Smith, J. E., & Meyer, J. L. (2006). Modeling Dependence Among Geologic Risks in Sequential Exploration Decisions. *Society of Petroleum Engineers (SPE)*. https://doi.org/10.2118/102369-ms.

Bickel, J. E., & Bratvold, R. B. (2008). From uncertainty quantification to decision making in the oil and gas industry. *Energy Exploration and Exploitation, 26*(5), 311–325.

Boucher, A., Dimitrakopoulos, R., & Vargas-Guzmán, J. A. (2005). *Joint Simulations, Optimal Drillhole Spacing and the Role of the Stockpile*. Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-3610-1_4.

Bratvold, R. B., Bickel, J. E., & Lohne, H. P. (2009). Value of information in the oil and gas industry: Past, present and future. *SPE Reservoir Evaluation and Engineering, 12*(4), 630–638.

Caers, J., Scheidt, C., Yin, Z., Wang, L., Mukerji, T., & House, K. (2022). Efficacy of information in mineral exploration drilling. *Natural Resources Research*. https://doi.org/10.1007/s11053-022-10030-1.

Cressie N., & Wikle C. (2011). Statistics for Spatio-Temporal Data. Wiley. https://www.wiley.com/en-us/Statistics+for+Spatio+Temporal+Data-p-9780471692744.

Delmelle, E. M., & Goovaerts, P. (2009). Second-phase sampling designs for non-stationary spatial variables. *Geoderma, 153*(1–2), 205–216.

Deutsch, C. V. (2021). Implementation of geostatistical algorithms. *Mathematical Geosciences, 53*(2), 227–237.

Deutsch C.V, Leuangthong O., & Ortiz J. (2007). A case for geometric criteria in resources and reserves classification.

Transactions: Society for Mining Metallurgy and Exploration, *322*(1), 1–11.
Dimitrakopoulos R. (2018). Stochastic mine planning-methods, examples and value in an uncertain world. In *Advances in Applied Strategic Mine Planning* (pp. 101–115), Springer International Publishing, https://doi.org/10.1007/978-3-319-69320-0_9.
Dirkx, R., & Dimitrakopoulos, R. (2018). Optimizing infill drilling decisions using multi-armed bandits: application in a long-term multi-element stockpile. *Mathematical Geosciences, 50*(1), 35–52.
Dutta, G., Mukerji, T., & Eidsvik, J. (2019). Value of information of time-lapse seismic data by simulation-regression: Comparison with double-loop Monte Carlo. *Computational Geosciences, 23*(5), 1049–1064.
Eidsvik, J., & Ellefmo, S. L. (2013). the value of information in mineral exploration within a multi-gaussian framework. *Mathematical Geosciences, 45*(7), 777–798.
Eidsvik, J., Martinelli, G., & Bhattacharjya, D. (2018). Sequential information gathering schemes for spatial risk and decision analysis applications. *Stochastic Environmental Research and Risk Assessment, 32*(4), 1163–1177.
Eidsvik, J., Mukerji, T., & Bhattacharjya, D. (2015). *Value of information in the earth sciences. Value of information in the earth sciences*. Cambridge University Press. https://doi.org/10.1017/cbo9781139628785.
Emerick, A. A., & Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers and Geosciences*, *55*, 3–15.
Ericsson, M., Drielsma, J., Humphreys, D., Storm, P., & Weihed, P. (2019). Why current assessments of 'future efforts' are no basis for establishing policies on material use—a response to research on ore grades. *Mineral Economics*. https://doi.org/10.1007/s13563-019-00175-6.
Froyland, G., Menabde, M., Stone, P., & Hodson, D. (2018). The value of additional drilling to open pit mining projects. In R. Dimitrakopoulos (Ed.), *Advances in Applied Strategic Mine Planning.* Cham: Springer. https://doi.org/10.1007/978-3-319-69320-0_10.
Gershon, M., Allen, L. E., & Manley, G. (1988). Application of a new approach for drillholes location optimization. *International Journal of Surface Mining, Reclamation and Environment, 2*(1), 27–31.
Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. Oxford University Press.
Hakim-Elahi, S., & Jafarpour, B. (2017). A distance transform forcontinuous parameterization of discrete geologic facies for subsurface flow model calibration. Water Resources Research, *53*, 8226–8249. https://doi.org/10.1002/2016WR019853.
Herrington, R. (2021). Mining our green future. *Nature Reviews Materials, 6*(6), 456–458.
Howard, R. A. (1966). Information value theory. *IEEE Transactions on Systems Science and Cybernetics, 2*(1), 22–26.
Hund K., Porta D. la, Fabregas T.P., Laing T., & Drexhage J. (2020). Climate Smart Mining Facility. Minerals for Climate Action: The Mineral Intensity of the Clean Energy Transition. Online: https://pubdocs.worldbank.org/en/961711588875536384/Minerals-for-Climate-Action-The-Mineral-Intensity-of-the-Clean-Energy-Transition.pdf.
Isaaks, E. H., & Srivastava, R. M. (2010). An Introduction to Applied Geostatistics, by E. H. Isaaks and R. M. Srivastava. *Geographical Analysis, 26*(3), 282–283.
Chiles, J. –P. & Delfiner, P. (1999). Geostatistics: Modelling Spatial Uncertainty. In: J.-P. Chils & P. Delfiner (Eds.), Wiley, https://doi.org/10.1002/9780470316993.
Journel, A. G. (1993). *Geostatistics: Roadblocks and Challenges* (pp. 213–224). Dordrecht: Springer. https://doi.org/10.1007/978-94-011-1739-5_18.
Jowitt, S. M., Mudd, G. M., & Thompson, J. F. H. (2020). Future availability of non-renewable metal resources and the influence of environmental, social, and governance conflicts on metal production. *Communications Earth & Environment, 1*(1), 1–8.
Jreij, S. F., Trainor-Guitton, W. J., Morphew, M., & Chen Ning, I. L. (2021). The value of information from horizontal distributed acoustic sensing compared to multicomponent geophones via machine learning. *Journal of Energy Resources Technology, 143*(1), 010902-1 https://doi.org/10.1115/1.4048051.
Kang, M., & Jackson, R. B. (2016). Salinity of deep groundwater in California: Water quantity, quality, and protection. *Proceedings of the National Academy of Sciences of the United States of America, 113*(28), 7768–7773.
Kochenderfer, M. J., Wheeler, T. A., & Wray, K. H. (2022). *Algorithms for Decision Making*. MIT Press.
Lall, U., Josset, L., & Russo, T. (2020). A snapshot of the world's groundwater challenges. *Annual Review of Environment and Resources, 45*, 171–194.
Le, N. D., & Zidek, J. V. (2006). *Statistical analysis of environmental space-time processes*. New York: Springer. https://doi.org/10.1007/0-387-35429-8.
Morosov, A. L., & Bratvold, R. B. (2022). Appraisal campaign selection based on the maximum value of sequential information. *Journal of Petroleum Science and Engineering, 208*(Part B), 109473.
Miller, A. C. (1975). Value of sequential information. *Management Science, 22*(1), 1–11.
Morgan, G. A., Putzig, N. E., Perry, M. R., Sizemore, H. G., Bramson, A. M., Petersen, E. I., et al. (2021). Availability of subsurface water-ice resources in the northern mid-latitudes of Mars. *Nature Astronomy, 5*(3), 230–236.
Müller, W. G. (2007). *Collecting spatial data: Optimum design of experiments for random fields*. Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-540-31175-1.
Newendorp, P. D., & Schuyler, J. (2002). *Decision analysis for petroleum exploration* (2nd ed.). Planning Press.
Nowak M., & Leuangthong O. (2019). Optimal drill hole spacing for resource classification. In *Mining Goes Digital - Proceedings of the 39th international symposium on Application of Computers and Operations Research in the Mineral Industry, APCOM 2019* (pp. 115–124). CRC Press/Balkema. https://doi.org/10.1201/9780429320774-14.
Onwunalu, J. E., & Durlofsky, L. J. (2010). Application of a particle swarm optimization algorithm for determining optimum well location and type. *Computers & Geosciences, 14*, 183–198.
Osher, S., & Fedkiw, R. (2002). Level set methods and dynamic implicit surfaces. Applied mathematical sciences. New York:Springer.
Powell W. B. (2011). Approximate Dynamic Programming: Solving the Curses of Dimensionality: Second Edition. Wiley Blackwell, https://doi.org/10.1002/9781118029176.
Raiffa. (1968). Decision analysis: introductory lectures on choices under uncertainty. https://psycnet.apa.org/record/1968-35027-000.
Raiffa H., & Schlaifer R. (1961). Applied Statistical Decision Theory. Wiley, https://www.wiley.com/en-us/Applied+Statistical+Decision+Theory-p-9780471383499.
Rötzer, N., & Schmidt, M. (2018). Decreasing metal ore grades-Is the fear of resource depletion justified? *Resources, 7*(4), 88.
Scheidt, C., Li, L., & Caers, J. (2018). Quantifying Uncertainty in Subsurface Systems. *Wiley*. https://doi.org/10.1002/9781119325888.
Schodde. (2017). Long term trends in global exploration – are we finding enough metal? Minex Consulting. Online: http://minexconsulting.com/long-term-trends-in-global-exploration-are-we-finding-enough-metal/.

Soltani, S., & Hezarkhani, A. (2011). Determination of realistic and statistical value of the information gathered from exploratory drilling. *Natural Resources Research, 20*, 207–216.

Soltani, S., & Hezarkhani, A. (2013). A simulated annealing-based algorithm to locate additional drillholes for maximizing the realistic value of information. *Natural Resources Research, 22*, 229–237.

Tarkowski, R. (2019). Underground hydrogen storage: Characteristics and prospects. *Renewable and Sustainable Energy Reviews, 105*, 86–94.

van Leeuwen, P. J., & Evensen, G. (1996). Data assimilation and inverse methods in terms of a probabilistic formulation. *Monthly Weather Review*, *124*(12), 2898–2913.

Verly, G., & Parker, H. M. (2021). Conditional simulation for mineral resource classification and mining dilution assessment from the early 1990s to Now. *Mathematical Geosciences, 53*(2), 279–300.

von Neumann J., & Morgenstern O. (1944). *Theory of Games and Economic Behavior,* 2nd rev. Princeton University Press.

Wei, Y.-M., Kang, J.-N., Liu, L.-C., Li, Q., Wang, P.-T., Hou, J.-J., et al. (2021). A proposed global layout of carbon capture and storage in line with a 2 °C climate target. *Nature Climate Change, 11*(2), 112–118.

West, J. (2020). Extractable global resources and the future availability of metal stocks: ''Known Unknowns'' for the foreseeable future. *Resources Policy, 65*, 101574.

Wu, W.-Y., Lo, M.-H., Wada, Y., Famiglietti, J. S., Reager, J. T., Yeh, P.J.-F., et al. (2020). Divergent effects of climate change on future groundwater availability in key mid-latitude aquifers. *Nature Communications, 11*(1), 1–9.

Yousefi, M., Carranza, E. J. M., Kreuzer, O. P., Nykänen, V., Hronsky, J. M. A., & Mihalasky, M. J. (2021). Data analysis methods for prospectivity modelling as applied to mineral exploration targeting: State-of-the-art and outlook. *Journal of Geochemical Exploration, 229*, 106839.

Zuo, R., Kreuzer, O. P., Wang, J., Xiong, Y., Zhang, Z., & Wang, Z. (2021). Uncertainties in GIS-based mineral prospectivity mapping: Key types, potential impacts and possible solutions. *Natural Resources Research. Springer, 30*(5), 3059–3079.